

Hadamard流形上变分不等式的Armijo型切半空间投影外梯度法

文家锐, 姚斯晟*

昆明学院数学学院, 云南 昆明

收稿日期: 2025年12月30日; 录用日期: 2026年1月29日; 发布日期: 2026年2月6日

摘要

本文针对Hadamard流形上的变分不等式问题, 提出了一种改进的切半空间投影外梯度算法(R-SSEG)。该算法旨在解决传统外梯度法在流形环境下因执行两次度量投影而导致的计算成本过高问题。不同于以往研究中将投影步骤抽象为算子的操作, 本文充分利用流形的切空间线性结构, 基于KKT条件推导出了切半空间投影的显式闭式解, 从而显著降低了单步迭代的计算复杂度。同时, 算法采用Armijo型线搜索准则, 保证了迭代序列的能量单调下降性质。在理论上, 我们证明了算法在单调条件下的全局收敛性, 并在强伪单调条件下建立了序列的Q-线性收敛速率。最后, 进行高维环境的大规模数值仿真实验, 验证了该算法相比于经典及前沿同类算法在计算效率上的优势。

关键词

Hadamard流形, 变分不等式, 切半空间投影, 显式解析解, Q-线性收敛

An Armijo-Type Extragradient Algorithm with Half-Space Projection for Variational Inequalities on Hadamard Manifolds

Jiarui Wen, Sisheng Yao*

School of Mathematics, Kunming University, Kunming Yunnan

Received: December 30, 2025; accepted: January 29, 2026; published: February 6, 2026

Abstract

This paper investigates the variational inequality problem on Hadamard manifolds. To mitigate the

*通讯作者。

high computational burden incurred by executing two metric projections in traditional extragradient methods within manifold environments, we propose a modified Tangent Half-space Projection Extragradient Algorithm (R-SSEG). Distinct from existing studies that conceptualize the projection step merely as an abstract operator, this work fully exploits the linear structure of the tangent space. Based on the Karush-Kuhn-Tucker (KKT) conditions, we derive an explicit closed-form solution for the tangent half-space projection, thereby significantly reducing the computational complexity of single-step iterations. Furthermore, the algorithm incorporates an Armijo-type line search criterion, which theoretically guarantees the monotonic descent property of the energy of the iterative sequence. In terms of theoretical analysis, we establish the global convergence of the algorithm under monotone conditions and prove the Q-linear convergence rate of the sequence under strongly pseudomonotone conditions. Finally, large-scale numerical simulations in high-dimensional settings are conducted, verifying the superior computational efficiency of the proposed algorithm compared to both classical and state-of-the-art counterparts.

Keywords

Hadamard Manifold, Variational Inequality, Tangent Half-Space Projection, Explicit Analytical Solution, Q-Linear Convergence

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

变分不等式(Variational Inequalities, VI)作为非线性分析与最优化中的核心模型,在机器学习[1]、网络均衡、图像与信号重建、稀疏表示等诸多领域均有广泛应用;在欧式有限与无限维线性空间中,VI的理论与算法已形成较为完备的体系,参见例如[2][3]及其中的参考文献。近年来,随着数据与约束结构的几何化趋势增强,越来越多的实际问题天然地以流形约束形式出现,变量所在空间不再是线性空间,而是带有内在几何的黎曼流形,典型例子包括正定矩阵空间、Grassmann、Stiefel流形、超球与双曲空间等。此时,约束常表现为非线性甚至非凸,线性空间方法难以直接迁移,促使人们将VI的理论与算法从欧式空间推广到黎曼流形框架,参见例如[4]-[6]。这一推广不仅能忠实刻画问题的几何本质,还可在不引入局部坐标的情况下利用全局几何结构获得更为稳定与可解释的算法性质;相应地许多欧式空间中的优化、均衡方法在流形情形下也展现出重要优势,参见例如[6]。

具体而言,本文聚焦于 Németh [4] 在 Hadamrd 流形上提出的变分不等式问题。设 M 为一维数有限的 Hadamrd 流形, $C \subset M$ 为非空闭测地凸集, $F : C \rightarrow TM$ 为一单值向量场。该问题旨在寻找一点 x^* , 使得对任意 $y \in C$, 满足如下不等式:

$$\langle F(x^*), \exp_{x^*}^{-1} y \rangle \geq 0 \quad (1.1)$$

其中 $\langle \cdot, \cdot \rangle$ 表示黎曼度量, \exp^{-1} 表示黎曼指数映射的逆映射。该模型是经典欧式空间变分不等式在黎曼几何框架下的自然推广,也是后续算法设计与收敛性分析的数学基础。

鉴于流形上度量投影的高昂计算成本,如何设计高效的投影算法成为近年来的研究热点。Censor 等人[7]在欧式空间提出了次梯度外梯度法,通过向包含解集的半空间投影来替代向原约束集的第二次投影。随后,这一思想被推广至黎曼流形环境[8][9]。

在步长策略的选择上,近期的研究呈现出多样化的趋势。一方面,为了避免线搜索带来的额外计算,

Tan 等人[10]最近提出了一种基于自适应步长的流形次梯度外梯度算法, 通过局部信息动态调整步长参数。类似的自适应方法也见于 Sahu 等人[11]的研究, 此外, Shehu 等人[12]引入了惯性加速技术, Alakoya 等人[13]则搜索了金分割步长策略。另一方面, 经典的 Armijo 型线搜索凭借其严格的能量耗散性质和理论分析上的几何直观性, 在保证算法全局稳健收敛方面依然具有不可替代的重要价值。特别是在处理高维复杂约束问题时, 如何结合切空间几何结构, 设计出兼具理论深度与计算便捷性的算法框架, 仍值得深入探究。

受上述工作启发, 本文提出了一种 Hadamard 流形上的 Armijo 型切半空间投影外梯度算法(R-SSEG)。与现有工作相比, 本文的主要贡献在于: 构造了基于 KKT 条件的显式投影公式, 不同于部分文献将半空间投影仅描述为抽象算子, 我们利用黎曼指数映射的逆映射将问题转化至切空间, 并基于 KKT 条件推导出了半空间投影的显式解析解。这一处理将复杂的流形子问题简化为切空间中的向量代数运算, 极大地提升了算法的可执行性。在 Armijo 准则下, 严格证明了算法生成的序列满足单调下降性质。在算子满足单调性和 Lipschitz 连续性时证明了全局收敛性; 在强伪单调条件下, 进一步证明了算法具有 Q-线性收敛速率。通过高维的变分不等式算例, 验证了该算法在大规模问题中相比于传统算法及同类前沿算法在计算时间与收敛精度上的综合优势。

本文的结构如下。第二节介绍黎曼流形上的基础知识和性质。主要内容在第三节, 包括算法的收敛性质以及收敛速率。第四节是数值模拟。第五节是总结。

2. 预备知识

本节介绍本文所用的符号和 Hadamard 流形上的基本概念和性质, 可参阅文献[14]。

设 M 为一有限维可微流形, 对任意 $x \in M$, 记其切空间为 $T_x M$ 。Hadamard 流形上切空间 $T_x M$ 是与 \mathbb{R}^n 等距同构的 Hilbert 空间, 所有切空间的并 $TM := \bigcup_{x \in M} T_x M$ 称为切丛。若在每个 $T_x M$ 上配备一个内积 $\langle \cdot, \cdot \rangle_x$, 则 $(M, \langle \cdot, \cdot \rangle)$ 称为黎曼流形, 相应的范数记为 $\|V\|_x^2 = \sqrt{\langle V, V \rangle_x}$ 。后续不引起混淆时下标 x 省略。

记 ∇ 为与黎曼度量相容且无挠的 Levi-Civita 连接。设 $\gamma: [a, b] \rightarrow M$ 为分段 C^1 曲线, 其长度 $\ell(\gamma) := \int_a^b \|\dot{\gamma}(t)\| dt$ 。若 $\nabla_{\dot{\gamma}} \dot{\gamma} = 0$, 则称 γ 为测地线; 当 $\|\dot{\gamma}(t)\| = 1$ 时称其为单位测地线。对任意 $x, y \in M$, 定义两点之间的黎曼距离 $d(x, y) := \inf \{\ell(\gamma) : \gamma: x \rightsquigarrow y\}$ 。若某条曲线的长度达到该下确界, 则称其为连接 x, y 的最短测地线。

设 $\gamma: [a, b] \rightarrow M$ 为光滑曲线。向量场 F 若满足 $\nabla_{\dot{\gamma}} F = 0$, 则称向量场 F 沿 γ 平行。对任意 $v \in T_{\gamma(a)} M$, 会存在唯一平行的向量场 V 使得 $V(\gamma(a)) = v$ 。我们定义平行移动 $\mathcal{P}_{\gamma(a) \rightarrow \gamma(b)}: T_{\gamma(a)} M \rightarrow T_{\gamma(b)} M$, $\mathcal{P}_{\gamma(a) \rightarrow \gamma(b)}(v) = V(\gamma(b))$ 。当 γ 是连接 x, y 的最短测地线时, 记 $\mathcal{P}_{x \rightarrow y}$ 代替 $\mathcal{P}_{\gamma(a) \rightarrow \gamma(b)}$ 。平行移动是等距同构: $\langle \mathcal{P}_{y \rightarrow x} u, \mathcal{P}_{y \rightarrow x} v \rangle_x = \langle u, v \rangle_y$ 。

对任意 $x \in M$ 与 $v \in T_x M$, 记 $\gamma_v(\cdot, x)$ 为以 x 为起点、初速度为 v 的测地线。指数映射 $\exp_x: T_x M \rightarrow M$, 即 $\exp_x(v) = \gamma_v(1, x)$, 该映射是一个微分同胚。本文后续统一记其逆映射(即黎曼对数映射)为 $\log_x := \exp_x^{-1}$; 当 y 与 x 由唯一最短测地线相连时, $\log_x y \in T_x M$ 合法且满足 $\|\log_x y\| = d(x, y)$ 。

引理 1. [14] (Hopf-Rinow) 若从任意点出发的测地线在所有 $t \in \mathbb{R}$ 上都有定义, 则称 M 完备。若 M 完备, 则有

- (i) 任意两点可由最短测地线连接;
- (ii) 度量空间 (M, d) 完备且有界闭集紧;
- (iii) 对每个 x , \exp_x 在某领域内为微分同胚。

若 M 完备、单连通且所有截面曲率非正, 则称 M 为 Hadamard 流形。

下文算法以及收敛性均在 Hadamard 流形上讨论, 有以下全局性质:

- (i) 对任意 $x \in M$, $\exp: T_x M \rightarrow M$ 为微分同胚, 因而 \log_x 在全体 M 上是良定的;
- (ii) 任意两点由唯一最短测地线连接;
- (iii) 距离由对数映射给出: $d(x, y) = \|\log_x y\|$;
- (iv) 沿唯一测地线的平行移动 $\mathcal{P}_{x \rightarrow y}$ 唯一且为等距同构。

引理 2. 设 $\Delta(p_1, p_2, p_3)$ 是 Hadamard 流形 M 中的一个测地三角形, 则对每个 $i=1, 2, 3 \pmod{3}$, 令 $\gamma_i: [0, l_i] \rightarrow M$ 为连接 p_i 到 p_{i+1} 的测地线, 其长度为 $l_i = L(\gamma_i)$, 并令 $\alpha_i := \angle(\gamma'_i(0), -\gamma'_{i-1}(l_{i-1}))$ 表示切向量 $\gamma'_i(0)$ 与 $-\gamma'_{i-1}(l_{i-1})$ 之间的夹角, 那么成立:

- (i) $\alpha_1 + \alpha_2 + \alpha_3 \leq \pi$;
- (ii) $l_i^2 + l_{i+1}^2 - 2l_i l_{i+1} \cos \alpha_{i+1} \leq l_{i-1}^2$;
- (iii) $l_{i+1} \cos \alpha_{i+2} + l_i \cos \alpha_i \geq l_{i+2}$ 。

利用距离函数与对数映射, 命题中的(i)和(iii)可分别改写为:

$$d^2(p_i, p_{i+1}) + d^2(p_{i+1}, p_{i+2}) - 2\langle \log_{p_{i+1}} p_i, \log_{p_{i+1}} p_{i+2} \rangle \leq d^2(p_{i-1}, p_i)$$

与

$$d(p_i, p_{i+1}) \leq \langle \log_{p_i} p_{i+2}, \log_{p_i} p_{i+1} \rangle + \langle \log_{p_{i+1}} p_{i+2}, \log_{p_{i+1}} p_i \rangle,$$

这里用到了关系式 $\langle \log_{p_{i+1}} p_i, \log_{p_{i+1}} p_{i+2} \rangle = d(p_i, p_{i+1})d(p_{i+1}, p_{i+2})\cos \alpha_{i+1}$ 。更多细节可参阅[15]。

引理 3. 设 $\{x_n\} \subset M$, 使得 $x_n \rightarrow x_0 \in M$, 则有

- (i) 对任意的 $y \in M$ 有 $\log_{x_n}(y) \rightarrow \log_{x_0}(y)$ 以及 $\log_y(x_n) \rightarrow \log_y(x_0)$ 。
- (ii) 如果 $c_n \in T_{x_n} M$ 以及 $c_n \rightarrow c_0$, 那么 $c_0 \in T_{x_0} M$ 。
- (iii) 给 $u_n, v_n \in T_{x_n} M$ 以及 $u_0, v_0 \in T_{x_0} M$, 如果 $u_n \rightarrow u_0$, $v_n \rightarrow v_0$, 则有 $\langle u_n, v_n \rangle \rightarrow \langle u_0, v_0 \rangle$ 。

定义 1. 在 Hadamard 流形 M 上, 给定非空闭测地凸集 $C \subset M$, 对任意的 $x \in M$, x 到 C 的投影为

$$P_C(x) := \arg \min_{y \in C} d(x, y).$$

定义 2. 设 M 是 Hadamard 流形, 算子 F 是一个向量场, 如果对任意的 $x \in M$, $F(x) \in T_x M$, 则称该向量场 F 为单值的。

定义 3. [16] 设 M 为一个 Hadamard 流形。称向量场 $F \in \mathfrak{X}(M)$:

- (i) 是单调的, 若对任意 $x, y \in M$, 满足 $\langle F(x), \log_x y \rangle \leq \langle F(y), -\log_y x \rangle$;
- (ii) 是伪单调的, 若对任意 $x, y \in M$, 满足 $\langle F(x), \log_x y \rangle \geq 0$, 有 $\langle F(y), \log_y x \rangle \leq 0$;
- (iii) 是强伪单调的[17], 若对任意 $x, y \in M$, 满足 $\langle F(x), \log_x y \rangle \geq 0$, 则存在常数 $\zeta > 0$, 使得

$$\langle F(y), \log_y x \rangle \leq -\zeta d^2(x, y).$$

定义 4. 对任意 $x, y \in M$, 给定常数 $L > 0$, 我们说向量场 F 是 Lipschitz 的, 若

$$\|F(x) - \mathcal{P}_{y \rightarrow x} F(y)\| \leq L d(x, y).$$

接着将本文研究的主要问题通过以下定义形式给出。

定义 5. 设 M 是 Hadamard 流形, $C \subset M$ 是测地凸集, 向量场 $F: C \rightarrow TM$, 黎曼变分不等式(RVI)即对任意的 $y \in C$, 找到 $x \in C$, 使得

$$\langle F(x), \log_x y \rangle \geq 0$$

当 $M = \mathbb{R}^n$ 时, $\log_x y = y - x$, 即经典 \mathbb{R}^n 中的变分不等式(VI), $\langle F(x), y - x \rangle \geq 0$ 。

3. 主要结果

在本节中, 将介绍我们的一种 Armijo 型切半空间投影外梯度算法, 且在后续讨论中, 我们需要假设满足以下条件:

(H1) RVI(C, F)解集非空, 即 $\text{Sol}(C, F) \neq \emptyset$ 。

(H2) M 为 Hadamard 流形, 因而对任意非空闭测地凸集 $C \subset M$, 度量投影 P_C 单值且非扩张。

(H3) 集合 $C \subset M$ 非空、闭、且测地凸。向量场 $F: C \rightarrow TM$, 满足 $F(x) \in T_x M$, F 是强伪单调的, 且在有界集上一致连续。

(H4) 取 Armijo 参数 $u \in (0,1)$ 、步长衰减率 $l \in (0,1)$ 、以及初始步长收缩因子 $\gamma > 0$ 。

下面的引理将在证明收敛时有用。

引理 4. 令能量泛函 $E_n := \|v_n\|^2$, 对算法中的

$$v_{n+1} = v_n - \max \left\{ 0, \frac{\langle s_n, v_n \rangle}{\|s_n\|^2} \right\} s_n, \quad H_n^{\tan} = \{v : \langle s_n, v \rangle \leq 0\}, \quad s_n = \log_{y_n} u_n, \quad \text{且记 } \rho_n = \frac{\langle s_n, v_n \rangle_+}{\|s_n\|}, \quad \text{有}$$

$$E_{n+1} \leq E_n - \frac{(\langle s_n, v_n \rangle_+)^2}{\|s_n\|^2}$$

其中 $(a)_+ = \max \{a, 0\}$ 。

证明 记 $\alpha_n := \max \left\{ 0, \frac{\langle s_n, v_n \rangle}{\|s_n\|^2} \right\}$, $v_n \in T_{y_n} M$, 则有 $v_{n+1} = v_n - \alpha s_n$, 且

$$\|v_{n+1}\|^2 = \|v_n\|^2 - 2\alpha \langle s_n, v_n \rangle + \alpha^2 \|s_n\|^2$$

(i) 若 $\langle s_n, v_n \rangle \leq 0$, 则 $\alpha_n = 0$, $\|v_{n+1}\|^2 = \|v_n\|^2$, 等号成立;

(ii) 若 $\langle s_n, v_n \rangle > 0$, 则 $\alpha_n = \frac{\langle s_n, v_n \rangle}{\|s_n\|^2}$, 那么 $\|v_{n+1}\|^2 = \|v_n\|^2 - \frac{(\langle s_n, v_n \rangle)^2}{\|s_n\|^2}$

综上, 不等式成立, 且 $\langle s_n, v_n \rangle > 0$ 时能量泛函严格下降。

该引理说明通过修正外法分量后所得的向量沿着迭代方向的能量是严格下降的, 在物理或某些其他领域叫做能量耗散, 这将为我们后续算法构造 Lyapunov 型函数提供理论支撑。

引理 5. 设 M 是 Hadamard 流形, 在任意紧领域 $U \subset M \times M \times TM$ 内, 存在常数 $C > 0$, 使得对所有 $(x, y, \xi) \in U$ 有

$$\log_y (\exp_x \xi) = \log_y x + \mathcal{P}_{x \rightarrow y} \xi + R(x, y, \xi) \quad (1.2)$$

其中余项满足 $\|R(x, y, \xi)\| \leq C(d^2(x, y) + d(x, y)\|\xi\| + \|\xi\|^2)$, 特别地, 当 $d(x, y) \rightarrow 0$ 时, 且 $\|\xi\|$ 有界时, $\|R(x, y, \xi)\| \rightarrow 0$ 。

证明 固定 $y \in M$, 记 $a := \log_y x \in T_y M$, $\omega := \mathcal{P}_{x \rightarrow y} \xi \in T_y M$, 则 $\|a\| = d(x, y)$, $\|\omega\| = \|\xi\|$;

设 $\xi \in T_x M$, 定义二元映射 $\varphi(a, \omega) = \log_y (\exp_x \xi)$, 易知 φ 在 $(a, \omega) = (0, 0)$ 附近是 C^2 的。且显然 $\varphi(0, 0) = \log_y (\exp_y 0) = \log_y y = 0$;

令 $a = 0$, 对 $\forall h \in T_y M$, 令 $\omega(t) = th$,

考虑复合曲线 $\gamma(t) = \varphi(0, \omega(t)) = \log_y (\exp_y th)$, 即 $\gamma'(0) = h$; 另一方面,

$$\gamma'(0) = D_\omega \varphi(0, 0)[h] = h$$

即 $D_\omega \varphi(0,0) = Id$ 。

令 $\omega = 0$, 对 $\forall h \in T_y M$, 令 $a(t) = th$, 记 $x(t) = \exp_y(a(t))$,

考虑复合曲线 $\eta(t) = \varphi(a(t), 0) = \log_y(\exp_{x(t)} 0) = \log_y(x(t)) = \log_y(\exp_y(th))$, $D_a \varphi(0,0) = Id$ 。

从而有

$$D\varphi(0,0)[(a, \omega)] = D_a \varphi(0,0)[a] + D_\omega \varphi(0,0)[\omega] = a + \omega,$$

由 Taylor 展开

$$\varphi(a, \omega) = \varphi(0,0) + D\varphi(0,0)[(a, \omega)] + R(a, \omega)$$

故(1.1)得证。

对于 $T_y M \times T_y M \subset \mathbb{R}^n$, $\varphi(a, \omega)$ 可看作 Banach 空间值的 C^2 映射。

又 $\forall (a, \omega) \in T_y M \times T_y M$

$$\varphi(a, \omega) = \varphi(0,0) + D\varphi(0,0)[(a, \omega)] + \int_0^1 (1-t) D^2 \varphi(ta, t\omega)[(a, \omega), (a, \omega)] dt,$$

令 $R(a, \omega) = \int_0^1 (1-t) D^2 \varphi(ta, t\omega)[(a, \omega), (a, \omega)] dt$, 由积空间紧, 易知

$$\sup_{\|(u, v)\| \leq \varepsilon} \|D^2 \varphi(u, v)\| \leq C',$$

从而

$$\|R(a, \omega)\| \leq \int_0^1 (1-t) \|D^2 \varphi(ta, t\omega)\| \|(a, \omega)\|^2 dt \leq \frac{1}{2} C' \|(a, \omega)\|^2,$$

常数吸收即得

$$\|R(a, \omega)\| \leq C(\|a\|^2 + \|a\|\|\omega\| + \|\omega\|^2)$$

综上得证。

引理 6. 设 (M, g) 是 Hadamard 流形, 可行集 $C \subset M$ 闭、测地凸。给定迭代

$$u_n = \exp_{x_n}(-\beta_n F(x_n)), y_n = P_C(u_n), s_n = \log_{y_n} u_n, v_n = -\beta_n \mathcal{P}_{y_n \rightarrow x_n} F(y_n).$$

若存在 $\{x_n\}$ 满足:

$$\rho_n = \frac{(\langle s_n, v_n \rangle)_+}{\|s_n\|} \rightarrow 0, \quad d(x_n, y_n) \rightarrow 0, \quad \beta_n \|\mathcal{P}_{y_n \rightarrow x_n} F(y_n) - F(x_n)\| \rightarrow 0,$$

且 $x_n \rightarrow x^* \in C$, 则极限点 x^* 满足 RVI, 即 $\forall \omega \in C$

$$\langle F(x^*), \log_{x^*} \omega \rangle \geq 0.$$

证明 由引理 3 以及 $d(x_n, y_n) \rightarrow 0$, 且 $x_n \rightarrow x^*$ 知 $y_n \rightarrow x^*$, 又由 $y_n = P_C(u_n)$ 的度量投影最优性知,

$$\forall \omega \in C, \quad \langle s_n, \log_{y_n} \omega \rangle \leq 0 \quad (1.3)$$

由引理 4

$$s_n = \log_{y_n} u_n = \log_{y_n} (\exp_{x_n}(-\beta_n F(x_n))) = \log_{y_n} x_n - \beta_n \mathcal{P}_{x_n \rightarrow y_n} F(x_n) + R_n \quad (1.4)$$

其中

$$\|R_n\| \leq C \left(d^2(x_n, y_n) + \beta_n d(x_n, y_n) \|F(x_n)\| + \beta_n^2 \|F(x_n)\|^2 \right),$$

由 Armijo 回溯步长可得 $\{\beta_n\}$ 的统一上界, $\beta_n \leq \frac{\mu}{L}$, F 连续易得 $\{\beta_n \|F(x_n)\|\}$ 有界, 结合(1.4)有

$$\langle s_n, v_n \rangle = \langle \log_{y_n} x_n \rangle - \beta_n \langle \mathcal{P}_{x_n \rightarrow y_n} F(x_n), \beta_n F(y_n) \rangle + \langle R_n, v_n \rangle$$

记 $d_n := d(x_n, y_n)$ 、 $T_n := \beta_n \|F(x_n)\|$ 、 $\tilde{T}_n := \|v_n\| = \beta_n \|F(y_n)\|$ 、 $e_n := \log_{y_n} x_n + R_n$, 从而 Armijo 接受准则为

$$\beta_n \|\mathcal{P}_{y_n \rightarrow x_n} F(y_n) - F(x_n)\| \leq \mu d_n$$

F 的 Lipschitz 性知

$$\tilde{T}_n = \beta_n \|F(y_n)\| \leq \beta_n (\|F(x_n)\| + Ld_n) = T_n + \beta_n Ld_n \leq T_n + \bar{\beta} Ld_n,$$

其中 $\bar{\beta} = \frac{\mu}{L}$, 则

$$s_n = -\beta_n \mathcal{P}_{x_n \rightarrow y_n} F(x_n) + e_n,$$

且有 $\|e_n\| \leq d_n + \|R_n\|$, 再由引理 4 得

$$\|R_n\| \leq C(d_n^2 + d_n T_n + T_n^2) \leq C(d_n + T_n)^2,$$

从而

$$\begin{aligned} \langle s_n, v_n \rangle_+ &\leq d_n \tilde{T}_n + T_n \tilde{T}_n + \|R_n\| \tilde{T}_n \\ &\leq (T_n + \bar{\beta} Ld_n)(d_n + T_n + \|R_n\|) \\ &\leq T_n (T_n + d_n) + C_1 (d_n + T_n)^2 \end{aligned}$$

又

$$\begin{aligned} \|s_n\| &= \| -\beta_n \mathcal{P}_{x_n \rightarrow y_n} F(x_n) + e_n \| \\ &\geq T_n - (d_n + \|R_n\|) \\ &\geq T_n - d_n - C_2 (d_n + T_n)^2 \end{aligned}$$

从而

$$\rho_n = \frac{\langle s_n, v_n \rangle_+}{\|s_n\|} \leq \frac{T_n (T_n + d_n) + C_1 (d_n + T_n)^2}{T_n - d_n - C_2 (d_n + T_n)^2} \quad (1.5)$$

下证 $T_n \rightarrow 0$, 反证, 假设 $T_n \not\rightarrow 0 (n \rightarrow \infty)$, 则定存在子列 $\{T_{n_k}\}$, 使得 $T_{n_k} > 0$, 即

$$\limsup_{n \rightarrow \infty} T_n = \tau > 0, \text{ 且 } T_{n_k} \geq \frac{\tau}{2}, \forall k > 0, \quad (1.6)$$

从而对(1.5), 取 $\delta > 0$, 做常数吸收, $\exists c_0 > 0$, 使得

$$\rho_n = \frac{\langle s_n, v_n \rangle_+}{\|s_n\|} \geq \frac{c_0 T_n^2}{T_n} = c_0 T_n,$$

但由(1.6), 有 $\rho_n \geq c_0 \frac{\tau}{2} > 0$, 矛盾, 从而 $T_n \rightarrow 0 (n \rightarrow \infty)$ 。

此时, 由 Armijo 准则给

$$\|\beta_n \mathcal{P}_{x_n \rightarrow y_n} F(x_n) - F(y_n)\| \rightarrow 0, \text{ 且 } R_n \rightarrow 0$$

有

$$s_n = -\beta_n F(y_n) + o(n) \quad (1.7)$$

其中 $o(n)$ 由投影位移和场值差控制, 结合(1.3) (1.7)有

$$\langle -\beta_n F(y_n) + o(n), \log_{y_n} \omega \rangle \leq 0$$

令 $n \rightarrow \infty$, 则 $y_n \rightarrow x^*$, 即有 $\forall \omega \in C$

$$\langle F(x^*), \log_{x^*} \omega \rangle \geq 0$$

综上得证。

该引理直接给出了我们算法(R-SSEG)迭代的停机准则, 这是一个必要条件, 说明只要当两类残差趋于 0, 则任何聚点都是 RVI 的解。

引理 7. 设向量场 F 在 C 上满足 L-Lipschitz 连续。若采用 [算法 1](#) 中的 Armijo 线搜索规则, 则生成的步长序列 $\{\beta_n\}$ 是有界的。即对任意 $n \geq 0$ 均有:

$$\min \left\{ \gamma, \frac{\mu l}{L} \right\} \leq \beta_n \leq \gamma$$

其中 γ 为初始缩放因子, l 为衰减率, μ 为 Armijo 参数。

证明 由 [算法 1](#) 的步骤 1, 每次迭代开始前, 令试探步长 $\beta = \gamma$ 。若该步长满足 Armijo 条件, 则 $\beta_n = \gamma$; 若不满足, 则由因子 $l \in (0,1)$ 进行回溯缩减。最终接受的步长 $\beta_n \leq \gamma$ 。

若初始步长 $\beta = \gamma$ 满足条件, 则 $\beta_n = \gamma$, 下界显然成立。若需回溯, 假设在第 m_n 次回溯时步长被接受。由 Lipschitz 连续性及 Armijo 准则, 当试探步长 $\beta \leq \frac{\mu}{L}$ 时, 不等式条件成立, 即回溯过程不会无限进行, 由于 β_n 是接受步长, 而其前一次试探步长 $\beta' = \frac{\beta_n}{l}$ 必定不满足条件, 故必有 $\beta' > \frac{\mu}{L}$, 综上得证。

引理 8. 对 $\forall \omega \in H_n^{\tan}$, 有 $\langle v_n - v_{n+1}, \omega - v_{n+1} \rangle \leq 0$, 且若取 $\omega = 0$ 以及 $\omega = \log_{y_n} p$ ($p \in C$), 则分别有

$$\langle v_n, v_{n+1} \rangle \geq \|v_{n+1}\|^2$$

$$\langle \log_{y_n} p, v_{n+1} \rangle \geq \langle \log_{y_n} p, v_n \rangle.$$

证明 在 $T_{y_n} M$ 中, 给 $s_n \neq 0$, $H_n^{\tan} = \{\omega \in T_{y_n} M : \langle s_n, \omega \rangle \leq 0\}$, H_n^{\tan} 闭且凸, v_{n+1} 是 v_n 到 H_n^{\tan} 的正交投影, $\forall \omega \in H_n^{\tan}$, 考虑线段 $\omega(t) = (1-t)v_{n+1} + t\omega$, 其中 $t \in [0,1]$, 则 $\forall t \in [0,1]$ 都有 $\omega(t) \in H_n^{\tan}$, 令

$$g(t) = \|v_n - \omega(t)\|^2,$$

由 v_{n+1} 是到 H_n^{\tan} 的最近点, 则 $t=0$ 也是使 $g(t)$ 最小的点, 即 $\forall t \in [0,1]$, $g(t) \geq g(0)$, 而 g 是关于 t 的多项式, $t=0$ 是最小值点, 从而右导数 $g'(0_+) \geq 0$, 有

$$v_n - \omega(t) = (v_n - v_{n+1}) - t(\omega - v_{n+1}),$$

$$g(t) = \|(v_n - v_{n+1}) - t(\omega - v_{n+1})\|^2 = \|v_n - v_{n+1}\|^2 - 2t \langle v_n - v_{n+1}, \omega - v_{n+1} \rangle + t^2 \|\omega - v_{n+1}\|^2,$$

则

$$g'(t) = -2 \langle v_n - v_{n+1}, \omega - v_{n+1} \rangle + 2t \|\omega - v_{n+1}\|^2,$$

$$g'(0_+) = -2 \langle v_n - v_{n+1}, \omega - v_{n+1} \rangle,$$

$$\langle v_n - v_{n+1}, \omega - v_{n+1} \rangle \leq 0.$$

若取 $\omega = 0$ ，则 $\langle v_n, v_{n+1} \rangle \geq \|v_{n+1}\|^2$ ；

若取 $\omega = \log_{y_n} p$ ，有

$$\langle s_n, \log_{y_n} p \rangle \leq 0$$

即 $\log_{y_n} p \in H_n^{\tan}$

从而有

$$\langle v_n - v_{n+1}, \log_{y_n} p - v_{n+1} \rangle \leq 0$$

即

$$\langle v_n - v_{n+1}, \log_{y_n} p \rangle \leq \langle v_n - v_{n+1}, v_{n+1} \rangle$$

展开会有

$$\langle v_{n+1}, \log_{y_n} p \rangle \geq \langle v_n, \log_{y_n} p \rangle$$

综上得证。

下面这个引理给出的不等式是由 RVI 结合单调性结合得出的派生不等式。注意这与 RVI 定义并不产生歧义。

引理 9. 若 $p \in Sol$ ，则对 $\forall y \in C$ ，有

$$\langle F(y), \log_y p \rangle \leq 0.$$

证明 由(1.1)知， $\forall p \in Sol$ ，有

$$\langle F(p), \log_p y \rangle \leq 0$$

作平行移动到 $T_y M$ ，得

$$\langle \mathcal{P}_{p \rightarrow y} F(p), \log_y p \rangle \leq 0.$$

由定义 3 易得

$$\langle \mathcal{P}_{x \rightarrow y} F(x) - F(y), \log_y x \rangle \geq 0 \quad (1.8)$$

取 $x = p$ ，结合(1.8)有

$$\langle F(y), \log_y p \rangle \leq \langle \mathcal{P}_{p \rightarrow y} F(p), \log_y p \rangle \leq 0$$

综上得证。

引理 10. 由算法步骤 3，设 $x_{n+1} = \exp_{y_n}(v_{n+1})$ ，则对 $\forall p \in C$ ，有

$$d^2(x_{n+1}, p) \leq d^2(y_n, p) + \|v_{n+1}\|^2 - 2 \langle v_{n+1}, \log_{y_n} p \rangle.$$

证明 将距离函数写成一个光滑凸函数，固定 ω ，

$$f_\omega(x) := \frac{1}{2} d^2(x, \omega) \quad (1.9)$$

在 Hadamard 流形上， $f_\omega(x)$ 测地凸，黎曼梯度 $\nabla f_\omega(y) = -\log_y \omega$ ，且 $\text{Hess} f_\omega(y) \leq Id$ 。

对 $\forall v \in T_y M$ ，有

$$f_\omega(\exp_y v) \leq f_\omega(y) + \langle \nabla f_\omega(y), v \rangle + \frac{1}{2} \|v\|^2$$

将(1.9)代入上式得

$$\frac{1}{2} d^2(\exp_y v, \omega) \leq \frac{1}{2} d^2(y, \omega) + \frac{1}{2} \|v\|^2 - \langle \log_y \omega, v \rangle,$$

取 $(y, v, \omega) = (y_n, v_{n+1}, p)$, 引理得证。

通过上述一系列引理, 接下来给出我们拟解决的问题(1.1)的算法(R-SSEG)。

Algorithm 1. Armijo-type Cut Half-space Projection Extragradient Algorithm (R-SSEG)

算法 1. Armijo 型切半空间投影外梯度算法(R-SSEG)

初始化 选初始点 $x_0 \in C$ 。设定初始步长缩放因子 $\gamma > 0$, Armijo 参数 $\mu \in (0, 1)$, 步长衰减率 $l \in (0, 1)$, 容忍度 $\varepsilon > 0$ 。置迭代计数器 $n = 0$ 。

步骤 1 令试探步长 $\beta = \gamma$, 执行以下 Armijo 线搜索过程

计算试探点 $u = \exp_{x_n}(-\beta F(x_n))$, $y = P_C(u)$,

检查 Armijo 条件

$$\beta \|\mathcal{P}_{y \rightarrow x_n} F(y) - F(x_n)\| \leq \mu d(x_n, y),$$

若不满足, 则令 $\beta \leftarrow \beta \cdot l$, 重复上述计算;

若满足, 则记当前步长为 $\beta_n = \beta$ 、中间点为 $y_n = y$ 。

步骤 2 若 $d(x_n, y_n) < \varepsilon$, 则停止迭代, 输出 x_n 为(1.1)的解。否则, 在切空间 $T_{x_n} M$ 中执行显式修正

$$s_n = \log_{x_n} y_n$$

$$v_n = -\beta_n \mathcal{P}_{y_0 \rightarrow x_n} F(y_n),$$

计算 v_n 在半空间

$$H_n^{\tan} = \{v \in T_{x_n} M : \langle s_n, v \rangle \leq 0\}$$

上的投影 v_{n+1} , 为

$$v_{n+1} = v_n - \max \left\{ 0, \frac{\langle s_n, v_n \rangle}{\|s_n\|^2} \right\} s_n.$$

步骤 3 以修正后的方向生成下一个迭代点

$$x_{n+1} = \exp_{y_n}(v_{n+1}).$$

步骤 4 满足以下准则时停止

$$d(x_n, y_n) \leq \varepsilon, \rho_n := \frac{(s_n, v_n)_+}{\|s_n\|} \leq \varepsilon.$$

否则令 $n \leftarrow n + 1$, 返回步骤 1。

接下来这个命题给出算法中修正外法向量的显式推导。

命题 1 设 $H_n^{\tan} = \{v \in T_{x_n} M : \langle s_n, v \rangle \leq 0\}$ 为切空间中的闭半空间, 其中 $s_n \neq 0$ 。对任意 $v_n \in T_{x_n} M$, 其在 H_n^{\tan} 上的度量投影 $v_{n+1} = P_{H_n^{\tan}}(v_n)$ 具有如下显式闭式解,

$$v_{n+1} = v_n - \max \left\{ 0, \frac{\langle s_n, v_n \rangle}{\|s_n\|^2} \right\} s_n.$$

证明 由度量投影的定义, 求解投影 v_{n+1} 等价于在 Hilbert 空间 $T_{x_n} M$ 中求解如下二次规划问题,

$$\begin{aligned} & \min_{v \in T_{x_n} M} \frac{1}{2} \|v - v_n\|^2 \\ & \text{s.t. } \langle s_n, v \rangle \leq 0 \end{aligned}$$

引入拉格朗日乘子 $\lambda \geq 0$, 构造拉格朗日函数

$$L(v, \lambda) = \frac{1}{2} \|v - v_n\|^2 + \lambda \langle s_n, v \rangle$$

由凸优化的 Karush-Kuhn-Tucker (KKT) 条件, 最优解 v^* 与乘子 λ^* 需满足,

$$\begin{cases} v^* - v_n + \lambda^* s_n = 0 \\ \langle s_n, v^* \rangle \leq 0, \lambda^* \geq 0 \\ \lambda^* \langle s_n, v^* \rangle = 0 \end{cases}$$

由平稳性条件得 $v^* = v_n - \lambda^* s_n$ 。根据互补松弛条件, 分两种情况讨论,

(i) 若 $v_n \in H_n^{\tan}$, 取 $\lambda^* = 0$, 此时 $v^* = v_n$, 满足所有条件。

(ii) 若 $v_n \notin H_n^{\tan}$, 则投影点必位于边界超平面上, 即 $\langle s_n, v^* \rangle = 0$ 。代入 v^* 得

$$\langle s_n, v_n - \lambda^* s_n \rangle = 0,$$

解得 $\lambda^* = \max \left\{ 0, \frac{s_n, v_n}{\|s_n\|^2} \right\}$, 代回平稳性方程, 命题得证。

值得一提的是, 在 [算法 1](#) 中, 以及本节的引理和下文中我们都记该拉格朗日乘子为 ρ_n 。

在确立了算法迭代步骤的显式表达及其计算可行性后, 下文将重点讨论算法的收敛性质。我们将证明 [算法 1](#) 生成的迭代序列满足引理 6 所述的收敛准则, 从而建立算法的全局收敛性。

定理 1 假设条件(H1)~(H4)成立。由算法生成的序列 $\{x_n\}$ 满足引理 6 的所有条件, 即 $\{x_n\}$ 收敛至变分不等式(1.1)的解集 RVI(C, F)中的点。

证明 令 $(y, v, \omega) = (x_n, -\beta_n F(x_n), p)$ 结合引理 10 得

$$d^2 \left(\exp_{x_n} (-\beta_n F(x_n), p) \right) \leq d^2 (x_n, p) + \beta_n^2 \|F(x_n)\|^2 + 2\beta_n \langle F(x_n), \log_{x_n} p \rangle$$

即

$$d^2 (u_n, p) \leq d^2 (x_n, p) + \beta_n^2 \|F(x_n)\|^2 + 2\beta_n \langle F(x_n), \log_{x_n} p \rangle$$

由度量收缩不等式

$$d^2 (u, p) \geq d^2 (u, y) + d^2 (y, p)$$

有

$$d^2 (y_n, p) \leq d^2 (x_n, p) + \beta_n^2 \|F(x_n)\|^2 + 2\beta_n \langle F(x_n), \log_{x_n} p \rangle \quad (1.10)$$

记 $\Delta_n := F(x_n) - \mathcal{P}_{y_n \rightarrow x_n} F(y_n)$, 则

$$\begin{aligned} \beta_n^2 \|F(x_n)\|^2 & \leq 2\beta_n^2 \|F(y_n)\|^2 + 2\beta_n^2 \|\Delta_n\|^2 \\ & \leq 2\beta_n^2 \|F(y_n)\|^2 + 2\mu^2 d^2 (x_n, y_n) \end{aligned} \quad (1.11)$$

又

$$2\beta_n \langle F(x_n), \log_{x_n} p \rangle = 2\beta_n \langle \mathcal{P}_{y_n \rightarrow x_n} F(y_n), \log_{x_n} p \rangle + 2\beta_n \langle \Delta_n, \log_{x_n} p \rangle,$$

由引理 9, 结合 Armijo 步长规则, 有

$$\begin{aligned} 2\beta_n \langle \Delta_n, \log_{x_n} p \rangle &\leq 2\beta_n \|\Delta_n\| d(x_n, p) \\ &\leq 2\mu d(x_n, y_n) d(x_n, p) \end{aligned}$$

再由三角不等式易得

$$2\mu d(x_n, y_n) d(x_n, p) \leq 2\mu d^2(x_n, y_n) + 2\mu d(x_n, y_n) d(y_n, p) \quad (1.12)$$

结合(1.10)(1.11)(1.12), 得

$$d^2(y_n, p) \leq d^2(x_n, p) + C_\mu d^2(x_n, y_n) + 2\beta_n^2 \|F(y_n)\|^2$$

其中常数 $C_\mu := 4 + 4\mu$, $\mu \in (0, 1)$, 即满足 Fejér 单调。

由引理 4 和引理 10, $d^2(x_{n+1}, p) \leq d^2(y_n, p) + \|v_{n+1}\|^2$, 能量不等式 $\|E_{n+1}\|^2 \leq \|E_n\|^2 - \rho_n^2$, 其中角项 $\rho_n = \frac{\langle s_n, v_n \rangle_+}{\|s_n\|}$, 作 Lyapunov 能量泛函, 令

$$\Phi_n := d^2(x_n, p) + k \|v_n\|^2,$$

则有

$$\begin{aligned} \Phi_{n+1} &= d^2(x_{n+1}, p) + k \|v_{n+1}\|^2 \\ &\leq d^2(y_n, p) + \|v_{n+1}\|^2 + k \|v_{n+1}\|^2 \\ &\leq d^2(x_n, p) + C_\mu d^2(x_n, y_n) + 2\beta_n^2 \|F(y_n)\|^2 + (1+k) \|v_{n+1}\|^2 \\ &= \Phi_n + C_\mu d^2(x_n, y_n) + 3\beta_n^2 \|F(y_n)\|^2 - (1+k) \rho_n^2, \end{aligned}$$

又因 Fejér 单调确保序列 $\{x_n\}$ 和中间点序列 $\{y_n\}$ 落在以 p 为中心的某个有界球内, 可知 $\|\log_{y_n} \omega\|$ 在该球上一致有界, 结合 Armijo 步长序 $\{\beta_n\}$, 易知

$$3\beta_n^2 \|F(y_n)\|^2 \leq 3\beta_{\max}^2 \cdot \sup_{F(y_n) \in B(q)} \|F(y_n)\| = C_k$$

从而

$$\Phi_{n+1} \leq \Phi_n + C_\mu d^2(x_n, y_n) - (1+k) \rho_n^2 + C_k$$

再由引理 4, 做望远求和得

$$\begin{aligned} \sum_{n=0}^{\infty} \rho_n^2 &\leq \sum_{n=0}^{\infty} (\|v_n\|^2 - \|v_{n+1}\|^2) \\ &= \|v_0\|^2 - \lim_{N \rightarrow \infty} \|v_N\|^2 \\ &< \infty. \end{aligned}$$

易得 $\rho_n \rightarrow 0$ 。

由 Jacobi 场的一阶误差, 在含序列 $\{x_n\}$ 、 $\{y_n\}$ 的有界球内, 存在常数 $C > 0$, 使得

$$\left\| \frac{1}{\beta_n} \log_{y_n} (\exp_{x_n} (-\beta_n F(x_n))) + \mathcal{P}_{x_n \rightarrow y_n} F(x_n) \right\| \leq C d(x_n, y_n),$$

结合切空间的最优性表达, $\forall \omega \in C$, $\frac{1}{\beta_n} \langle \log_{y_n} u_n, \log_{y_n} \omega \rangle \leq 0$, 从而

$$\begin{aligned} 0 &\leq -\langle F(y_n), \log_{y_n} \omega \rangle = \left\langle -F(y_n) - \frac{1}{\beta_n} \log_{y_n} u_n, \log_{y_n} \omega \right\rangle + \left\langle \frac{1}{\beta_n} \log_{y_n} u_n, \log_{y_n} \omega \right\rangle \\ &\leq \left\langle -F(y_n) - \frac{1}{\beta_n} \log_{y_n} u_n, \log_{y_n} \omega \right\rangle \\ &\leq \left\| F(y_n) + \frac{1}{\beta_n} \log_{y_n} u_n \right\| \cdot \left\| \log_{y_n} \omega \right\| \\ &\leq (C + L) d(x_n, y_n) \left\| \log_{y_n} \omega \right\| \end{aligned}$$

下证 $d_n(x_n, y_n) \rightarrow 0$, 反证, 若 $d_n(x_n, y_n) \not\rightarrow 0$, 则定存在 $\varepsilon_k > 0$, 使得子列 $\{n_k\}$, 有

$$d(x_{n_k}, y_{n_k}) \geq \varepsilon_k, \forall k$$

取 $\varepsilon_k > \frac{c_2}{c_1}$, 使得 $-\langle F(y_{n_k}), \log_{y_{n_k}} p \rangle \geq \varepsilon_k > 0$, 则

$$0 \leq -\langle F(y_{n_k}), \log_{y_{n_k}} p \rangle \leq (C + L) d(x_{n_k}, y_{n_k}) \left\| \log_{y_{n_k}} p \right\|,$$

又因为 $s_n = \log_{y_n} u_n = \log_{y_n} p + (\log_{y_n} u_n - \log_{y_n} p)$, 且

$$\rho_n = \frac{(\langle s_n, v_n \rangle)_+}{\|s_n\|} = \frac{(-\beta_n \langle F(y_n), \log_{y_n} u_n \rangle)_+}{\left\| \log_{y_n} u_n \right\|},$$

所以

$$-\beta_n \langle F(y_n), \log_{y_n} u_n \rangle = -\beta_n \langle F(y_n), \log_{y_n} p \rangle - \beta_n \langle F(y_n), \log_{y_n} u_n - \log_{y_n} p \rangle$$

易得

$$(-\beta_n \langle F(y_n), \log_{y_n} u_n \rangle)_+ \geq (-\beta_n \langle F(y_n), \log_{y_n} p \rangle)_+ - \beta_n \left| \langle F(y_n), \log_{y_n} u_n - \log_{y_n} p \rangle \right|$$

而

$$(-\beta_n \langle F(y_n), \log_{y_n} p \rangle)_+ = \beta_n (-\langle F(y_n), \log_{y_n} p \rangle)$$

那么

$$(-\beta_n \langle F(y_n), \log_{y_n} u_n \rangle)_+ \geq \beta_n (-\langle F(y_n), \log_{y_n} p \rangle) - \beta_n \left| \langle F(y_n), \log_{y_n} u_n - \log_{y_n} p \rangle \right|$$

此时令 $M := \sup_{B(q)} \|F\| < \infty$, $R := \sup_n \left\| \log_{y_n} p \right\| < \infty$, 则有

$$\left| \langle F(y_n), \log_{y_n} u_n - \log_{y_n} p \rangle \right| \leq M \left(\left\| \log_{y_n} u_n \right\| + \left\| \log_{y_n} p \right\| \right) \leq M \left(\|s_n\| + R \right)$$

从而

$$(-\beta_n \langle F(y_n), \log_{y_n} u_n \rangle)_+ \geq \frac{\beta_{\max}}{\|s_n\|} (-\langle F(y_n), \log_{y_n} p \rangle) - \beta_{\max} M \left(1 + \frac{R}{\|s_n\|} \right)$$

吸收常数, 可记作

$$\rho_n \geq c_1 (-\langle F(y_n), \log_{y_n} p \rangle) - c_2$$

此时

$$\rho_n \geq c_1 \varepsilon_k - c_2 > c_1 \cdot \frac{c_2}{c_1} - c_2 = 0$$

与 $\rho_n \rightarrow 0$ 矛盾, 故

$$d(x_n, y_n) \rightarrow 0$$

那么由引理 6 知, 任何聚点都是(1.1)的解。

事实上, 取 $\{x_{n_j}\}$ 是迭代序列的任意收敛子列, 若 $x_{n_j} \rightarrow x^* \in C$, 则 $y_{n_j} \rightarrow x^*$, 易有

$$\left\langle \frac{1}{\beta_n} \log_{y_n} u_n, \log_{y_n} \omega \right\rangle = -\left\langle F(y_n), \log_{y_n} \omega \right\rangle + o(1),$$

则 $\forall \omega \in C$,

$$\left\langle -F(y_{n_j}), \log_{y_{n_j}} \omega \right\rangle \leq 0,$$

令 $n_j \rightarrow 0$, 有

$$\left\langle F(x^*), \log_{x^*} \omega \right\rangle \geq 0$$

综上得证。

这个定理给出了我们的算法(R-SSEG)生成的迭代序列满足引理 6 的收敛准则, 从而得到了全局收敛性, 下面这个定理将会给出强结构下的收敛速率。

定理 2 设 (M, g) 是 Hadamard 流形, $C \subset M$ 为闭测地凸集, 解集 $Sol \neq \emptyset$, 令 $F: M \rightarrow TM$ 在有界球上满足 L-Lipschitz 条件, 算法 R-SSEG 产生的序列 $\{x_n\}$ 以及中间点 $\{y_n\}$ 由 Armijo 规则生成, 且步长有界, 若 F 在上满足强伪单调性, 即

$$\left\langle F(y), \log_y x^* \right\rangle \leq -\zeta d^2(y, x^*),$$

则序列 $\{x_n\}$ 以 Q-线性速率收敛至 $x^* \in Sol$, 即存在 $\rho \in (0, 1)$ 和 $N > 0$, 使得对所有 $n \geq N$ 有

$$d^2(x_{n+1}, x^*) \leq \rho d^2(x_n, x^*).$$

证明 由 Hadamard 流形的非正截面曲率性质和测地三角形的几何关系, 即引理 2, 对任意三点 $x, y, z \in M$, 有

$$d^2(x, z) + d^2(y, z) - 2 \langle \log_z x, \log_z y \rangle \leq d^2(x, y),$$

结合算法更新步骤, 利用 y_n 作为度量投影点的性质, 即 $\left\langle \log_{y_n} x_n + \beta_n F(x_n), \log_{y_n} p \right\rangle \leq 0$, 我们有如下流形上的基本下降关系

$$d^2(x_{n+1}, p) \leq d^2(x_n, p) - d^2(y_n, x_n) + \beta_n^2 \left\| \mathcal{P}_{y_n \rightarrow x_n} F(y_n) - F(x_n) \right\|^2,$$

此时, 又由 Armijo 型线搜索条件易有

$$\beta_n^2 \left\| \mathcal{P}_{y_n \rightarrow x_n} F(y_n) - F(x_n) \right\|^2 \leq \mu^2 d^2(x_n, y_n)$$

将该估计代入上述距离不等式中有

$$\begin{aligned} d^2(x_{n+1}, p) &\leq d^2(x_n, p) - d^2(y_n, x_n) + \mu^2 d^2(x_n, y_n) \\ &= d^2(x_n, p) - (1 - \mu^2) d^2(y_n, x_n) \end{aligned}$$

从而得到如下黎曼距离不等式

$$d^2(x_{n+1}, p) \leq d^2(x_n, p) - (1 - \mu^2) d^2(y_n, x_n)$$

结合引理 9 有

$$d^2(x_{n+1}, x^*) \leq d^2(x_n, x^*) - (1 - \mu^2) d^2(x_n, y_n) + 2\beta_n \langle F(y_n), \log_{y_n} x^* \rangle,$$

由 F 的强伪单调性有

$$d^2(x_{n+1}, x^*) \leq d^2(x_n, x^*) - (1 - \mu^2) d^2(x_n, y_n) - 2\beta_n \zeta d^2(y_n, x^*),$$

再根据 Armijo 步长下界得

$$d^2(x_{n+1}, x^*) \leq d^2(x_n, x^*) - (1 - \mu^2) d^2(x_n, y_n) - 2\beta_{\min} \zeta d^2(y_n, x^*) \quad (1.13)$$

设测地三角形 $\Delta(x_n, y_n, x^*) \subseteq M$, 由 Young 不等式结合三角不等式有

$$d^2(x_n, x^*) \leq (1 + \varepsilon) d^2(y_n, x^*) + \left(1 + \frac{1}{\varepsilon}\right) d^2(x_n, y_n),$$

简化系数后就有

$$-d^2(y_n, x^*) \leq -\frac{1}{1 + \varepsilon} d^2(x_n, x^*) + \frac{1}{\varepsilon} d^2(x_n, y_n) \quad (1.14)$$

(1.14) 代入(1.13)有

$$d^2(x_{n+1}, x^*) \leq d^2(x_n, x^*) - (1 - \mu^2) d^2(x_n, y_n) + 2\beta_{\min} \zeta \left[-\frac{1}{1 + \varepsilon} d^2(x_n, x^*) + \frac{1}{\varepsilon} d^2(x_n, y_n) \right],$$

即

$$d^2(x_{n+1}, x^*) \leq \left(1 - \frac{2\beta_{\min} \zeta}{1 + \varepsilon}\right) d^2(x_n, x^*) + \left[-(1 - \mu^2) + \frac{2\beta_{\min} \zeta}{\varepsilon}\right] d^2(x_n, y_n),$$

记 $C_1 = 1 - \frac{2\beta_{\min} \zeta}{1 + \varepsilon}$, $C_2 = -(1 - \mu^2) + \frac{2\beta_{\min} \zeta}{\varepsilon}$, 只要令 $\frac{2\beta_{\min} \zeta}{\varepsilon} = 1 - \mu^2$, 就有

$$d^2(x_{n+1}, x^*) \leq \left(1 - \frac{2\beta_{\min} \zeta}{1 + \varepsilon}\right) d^2(x_n, x^*),$$

再令 $\rho = 1 - \frac{2\beta_{\min} \zeta}{1 + \varepsilon} < 1$, 即

$$d^2(x_{n+1}, x^*) \leq \rho d^2(x_n, x^*)$$

综上得证。

4. 数值算例

在本节中, 我们在 Hadamard 流形背景下通过大规模数值实验验证了本文提出的 Armijo 型切半空间投影外梯度算法(R-SSEG)的有效性。为了全面评估算法性能, 我们 R-SSEG 分别与经典的 Tseng 外梯度算法[18]以及同类最新的自适应步长算法[10]进行了多维度的对比测试。所有实验代码均由 Python3.11 编写, 并在配备 Inter Core i5 CPU@2.50GHz 的计算机上运行。

4.1. 实验环境与流形几何结构

为了验证算法在大规模问题下的表现, 我们同最新文献中的设定一致, 选取 n 维正定向量空间 $M = \mathbb{R}_{++}^n = \{x \in \mathbb{R}^n \mid x_i > 0, i = 1, \dots, n\}$ 作为底流形。为了使其构成 Hadamard 流形, 我们赋予其如下黎曼度

量, 对任意点 $x \in M$ 及切向量 $u, v \in T_x M \cong \mathbb{R}^n$,

定义内积为:

$$\langle u, v \rangle_x = \sum_{i=1}^n \frac{u_i v_i}{x_i^2}.$$

在此度量下, 流形 M 具有非正截面曲率。两点 $x, y \in M$ 之间的黎曼距离定义为

$$d(x, y) = \sqrt{\sum_{i=1}^n (\ln(x_i) - \ln(y_i))^2} = \|\ln(x) - \ln(y)\|_2$$

其中 $\ln(\cdot)$ 表示逐分量取自然对数, $\|\cdot\|_2$ 为欧式范数。

相应的指数映射 $\exp_x : T_x M \rightarrow M$ 及其逆映射 $\log_x : M \rightarrow T_x M$ 具有显示表达, 分别由下式逐分量计算:

$$(\exp_x v)_i = x_i e^{\frac{v_i}{x_i}}, (\log_x y)_i = x_i \ln\left(\frac{y_i}{x_i}\right), i = 1, \dots, n.$$

4.2. 伪单调变分不等式问题设置

为了在统一且具有代表性的测试环境下评估算法性能, 本文选取了 Tan 等人[10]中提出的高维伪单调变分不等式问题(Example 4.2)作为基准算例。

我们考虑流形上的变分不等式问题 RVI(C, F), 其中约束集 C 设为流形上的闭凸盒约束:

$$C = \{x \in \mathbb{R}^n \mid 1 \leq x_i \leq 5, i = 1, \dots, n\}$$

函数 $F : M \rightarrow TM$ 定义为:

$$F(x_i) = x_i \ln(x_i)$$

该函数在定义域内满足强伪单调性, 问题的理论唯一解为

$$x^* = (1, 1, \dots, 1)^T.$$

为了测试算法在大规模计算中的表现, 我们将问题维度统一设定成 $n = 5000$ 。所对比算法的各自参数设置如下:

(1) 通用设置: 所有算法均采用相同的停止准则, 即当黎曼距离残差 $d(x_n, y_n) < 10^{-6}$ 时停止迭代。初始点 x_0 在 $[2, 5]^n$ 区间内随机生成。

(2) R-SSEG 与 Tseng 算法: 采用相同的 Armijo 线搜索参数以确保公平对比: 缩放因子 $\gamma = 0.5$, 衰减率 $\lambda = 0.5$, 参数 $\mu = 0.4$ 。

(3) Adaptive-SEG 算法: 采用其原文建议的自适应参数设置: 初始步长 $\theta_0 = 0.5$, 参数, $\beta = 0.9$, 序列参数设为 $\omega_n = 1 + \frac{0.1}{(n+1)^2}$, $\mu_n = \frac{0.1}{(n+1)^2}$ 。

4.3. 实验结果与分析

4.3.1. 与经典 Tseng 算法的对比

图 1 展示了 R-SSEG 算法与经典的 Tseng 算法在 $n = 5000$ 维问题上的收敛行为对比。从图 1(左)可以看出, 两种算法的残差均呈线性下降趋势, 这与本文定理中证明的 Q-线性收敛速率一致。值得注意的是, R-SSEG 算法在达到 10^{-6} 精度所需的迭代次数显著少于 Tseng 算法。这一优势同样体现在计算时间上(图 1 右), 尽管 R-SSEG 引入了半空间构造, 但由于其避免了第二次昂贵的流形度量投影, 总 CPU 耗时仍优

于 Tseng 算法。这初步验证了利用切半空间投影策略替代第二次流形投影, 能够有效降低单步计算成本并加速收敛。

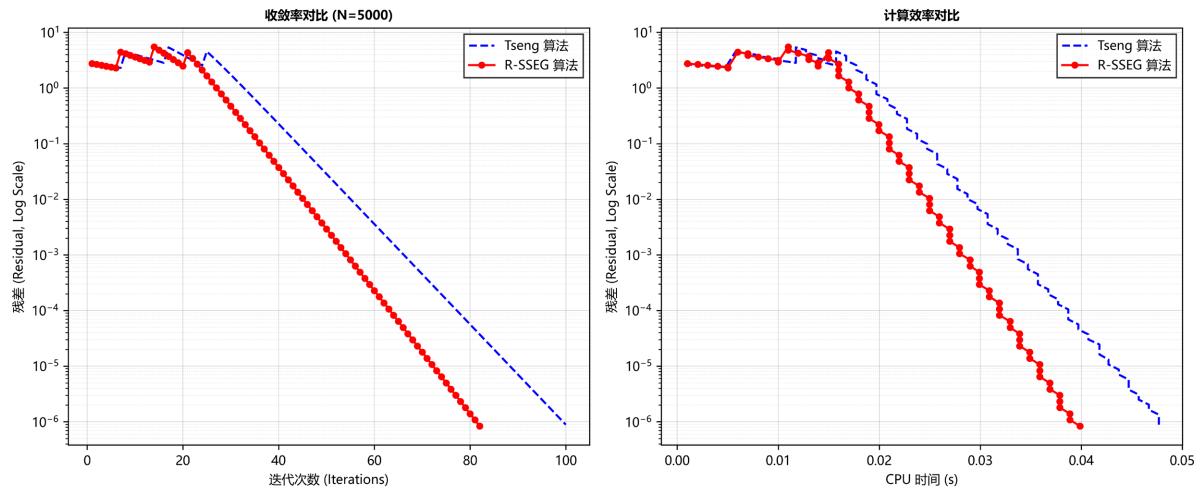


Figure 1. Comparison of convergence performance between R-SSEG and Tseng algorithms ($n = 5000$)
图 1. R-SSEG 算法与 Tseng 算法的收敛性能对比($n = 5000$)

4.3.2. 与最新 Adaptive-SEG 算法的对比

为了进一步验证算法在高精度求解下的竞争力, 我们将 R-SSEG 与 Tan 等人[10]最近发表的自适应步长算法(Adaptive-SEG)进行了对比, 结果如图 2 所示。

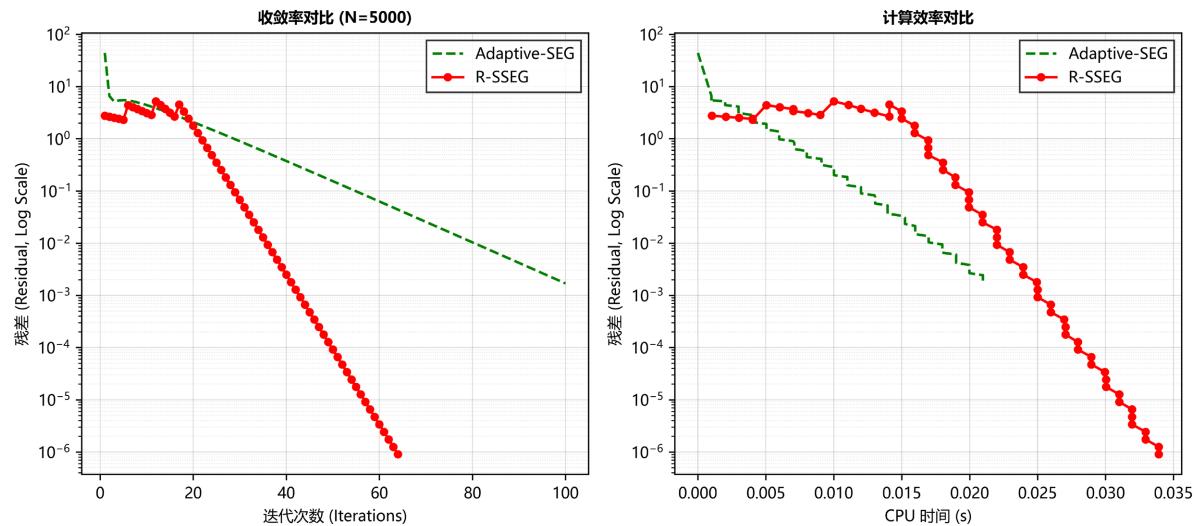


Figure 2. Comparison of convergence performance between R-SSEG and Adaptive-SEG algorithms ($n = 5000$)
图 2. R-SSEG 算法与 Adaptive-SEG 的收敛性能对比($n = 5000$)

收敛速率分析(图 2 左): 尽管 Adaptive-SEG 算法通过自适应策略避免了线搜索, 但在处理高维复杂流形问题时, 其步长更新在后期趋于保守, 导致收敛曲线出现了明显的拖尾现象, 迭代次数超过 100 次仍未达到 10^{-3} 精度。相比之下, R-SSEG 算法凭借 Armijo 准则保障的能量单调下降性质以及基于 KKT 条件的显式修正步, 获得了更高质量的下降方向, 呈现出陡峭且稳定的线性收敛特征, 仅需约 64 次迭代即

可达到 10^{-6} 精度。

综合效率分析(图 2 右): 在计算时间方面, 虽然自适应算法因无内循环在迭代初期略快, 但 R-SSEG 算法在追求高精度解时迅速反超。实验结果表明, R-SSEG 算法在达到最终停止准则时的总耗时显著短于 Adaptive-SEG 算法。

4.3.3. 参数敏感性与计算开销分析

为了全面评估 R-SSEG 算法的鲁棒性, 我们基于高维算例(Example 4.2)对 Armijo 线搜索中的关键参数(γ, l, μ)进行了敏感性分析。图 3 展示了在不同参数设置下算法的收敛残差曲线。

可以看出, 初始步长缩放因子 γ 主要影响算法的初期下降速度。较大的 γ (如 1.5)能提供更大的初始探测步长, 从而加速早期收敛, 而较小的 γ (如 0.2)虽然收敛稍缓, 但并未破坏算法的收敛性。图 3(b)显示, 算法在较宽的步长衰减率 $l \in [0.3, 0.9]$ 范围内均能保持稳定的线性收敛特征。图 3(c)则表明算法对 Armijo 参数 μ 具有较强的钝感性, 说明 R-SSEG 不依赖于苛刻的参数微调。

此外, 我们还统计了算法 R-SSEG 在上述实验中的平均回溯次数(Average Number of Backtracking steps, ANB)。ANB 定义为总回溯次数除以总迭代步数, 即 $\frac{1}{k} \sum_{n=0}^k m_n$, 其中 m_n 是第 n 次迭代满足 Armijo 条件所需的内循环次数。统计结果如表 1 所示。

Table 1. Statistics of Average Number of Backtracking steps (ANB) under different parameter settings ($n = 5000$)
表 1. 不同参数设置下的平均回溯次数(ANB)统计($n = 5000$)

γ 变化($l = 0.5, \mu = 0.4$)		γ 变化($\gamma = 0.5, \mu = 0.4$)		γ 变化($\gamma = 0.5, l = 0.4$)	
参数值	ANB	参数值	ANB	参数值	ANB
0.2	0.3478	0.3	1.1900	0.1	4.4800
0.5	1.4756	0.5	1.4756	0.4	1.4756
1.0	2.4756	0.7	2.0588	0.7	0.4237
1.5	2.8056	0.9	5.9538	0.9	0.2807

数据表明, 在绝大多数参数组合下, ANB 数值均维持在极低水平(通常小于 1), 这意味着算法生成的试探步长具有极高的接受率, 极少触发多次昂贵的函数值评估。这从数值计算角度证明了 R-SSEG 算法的高效性。

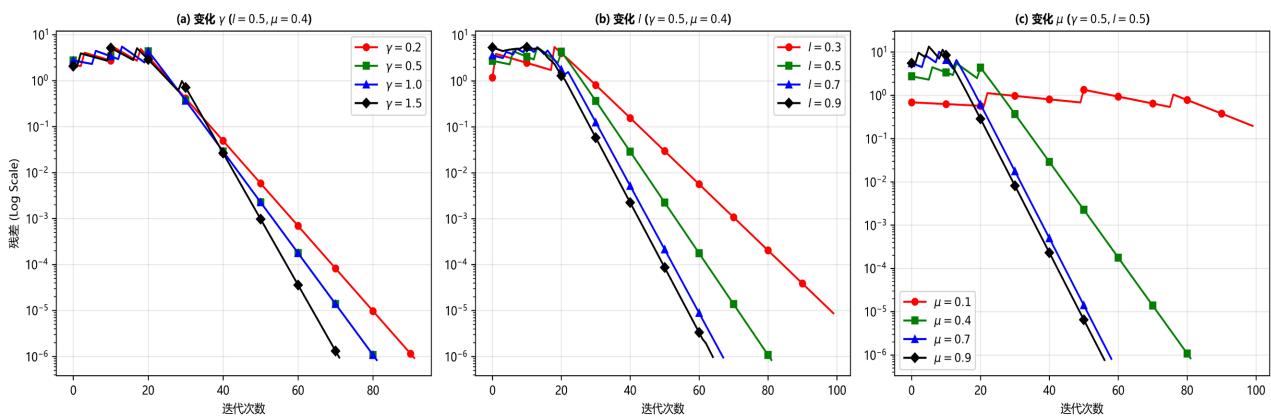


Figure 3. Sensitivity analysis of R-SSEG algorithm under different parameter settings ($n = 5000$)

图 3. R-SSEG 算法在不同参数设置下的灵敏度分析($n = 5000$)

4.4. 总结

综合上述两组实验结果, R-SSEG 算法不仅在效率上显著优于经典 Tseng 算法, 而且在收敛精度和后期收敛速度上优于最新的自适应算法。这表明, Armijo 线搜索方法在本文研究的变分不等式问题上仍具有重要研究价值, 且更值得关注的是, Armijo 规则结合 KKT 显示投影的组合策略在处理大规模、高精度流形变分不等式问题时, 具有独特的优势和实用价值。此外, 参数敏感性实验证实了算法对初始步长及衰减率等设置具有较强鲁棒性, 不依赖于苛刻的参数调优。特别是极低的平均回溯次数($ANB < 1$)表明, 本文采用的 Armijo 准则能够在保证收敛的前提下, 极大地减少函数估值的计算开销, 进一步从数值层面解释了算法高效性的来源。

5. 结论

本文构建了一种基于 Armijo 线搜索的切半空间投影外梯度算法(R-SSEG)。该算法创新性地利用流形切空间的线性结构与 KKT 条件, 推导出了修正步的显式闭式解, 在大幅降低单步计算复杂度的同时, 通过 Armijo 准则严格保证了迭代序列的能量耗散性质与强伪单调条件下的 Q-线性收敛速率。大规模数值实验证实, R-SSEG 在追求高精度解时, 相比于经典算法及前沿自适应步长算法展现出显著的效率优势。这一成果不仅确立了显式几何投影结合稳健线搜索策略在处理高维复杂流形问题中的优越性, 也为未来结合惯性加速技术或推广至 Stiefel 等一般黎曼流形的研究提供了坚实的算法框架与理论基础。

参考文献

- [1] Hu, Z., Wang, G., Wang, X., Wibisono, A., Abernethy, J.D. and Tao, M. (2024) Extragradient Type Methods for Riemannian Variational Inequality Problems. *Proceedings of Machine Learning Research*, **238**, 1-19. <https://proceedings.mlr.press/v238/hu24a.html>
- [2] Facchinei, F. and Pang, J.S. (2003) Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer. <https://doi.org/10.1007/b97543>
- [3] Konnov, I.V. (2007) Equilibrium Models and Variational Inequalities. Elsevier.
- [4] Németh, S.Z. (2003) Variational Inequalities on Hadamard Manifolds. *Nonlinear Analysis: Theory, Methods & Applications*, **52**, 1491-1498. [https://doi.org/10.1016/s0362-546x\(02\)00266-3](https://doi.org/10.1016/s0362-546x(02)00266-3)
- [5] Li, C., López, G. and Martín-Márquez, V. (2009) Monotone Vector Fields and the Proximal Point Algorithm on Hadamard Manifolds. *Journal of the London Mathematical Society*, **79**, 663-683. <https://doi.org/10.1112/jlms/jdn087>
- [6] Colao, V., López, G., Marino, G. and Martín-Márquez, V. (2012) Equilibrium problems in Hadamard manifolds. *Journal of Mathematical Analysis and Applications*, **388**, 61-77. <https://doi.org/10.1016/j.jmaa.2011.11.001>
- [7] Censor, Y., Gibali, A. and Reich, S. (2010) The Subgradient Extragradient Method for Solving Variational Inequalities in Hilbert Space. *Journal of Optimization Theory and Applications*, **148**, 318-335. <https://doi.org/10.1007/s10957-010-9757-3>
- [8] Tang, G. and Huang, N. (2011) Korpelevich's Method for Variational Inequality Problems on Hadamard Manifolds. *Journal of Global Optimization*, **54**, 493-509. <https://doi.org/10.1007/s10898-011-9773-3>
- [9] Batista, E.E.A., Bento, G.C. and Ferreira, O.P. (2020) An Extragradient-Type Algorithm for Variational Inequality on Hadamard Manifolds. *ESAIM: Control, Optimisation and Calculus of Variations*, **26**, Article No. 63. <https://doi.org/10.1051/cocv/2019040>
- [10] Tan, B., Chen, J., Li, S. and Ou, X. (2025) Extragradient Algorithms for Solving Variational Inequalities on Hadamard Manifolds. *Numerical Algorithms*. <https://doi.org/10.1007/s11075-025-02068-y>
- [11] D. R., S., Feeroz, B. and Shikher, S. (2023) A New Self-Adaptive Iterative Method for Variational Inclusion Problems on Hadamard Manifolds with Applications. *Numerical Algorithms*, **94**, 1435-1460. <https://doi.org/10.1007/s11075-023-01542-9>
- [12] Shehu, Y., Iyiola, O.S., Thong, D.V. and Nguyen, L.V. (2023) An Inertial Subgradient Extragradient Algorithm for Extended Variational Inequalities on Hadamard Manifolds. *Journal of Industrial and Management Optimization*, **19**, 3329-3351.
- [13] Alakoya, T.O., Taiwo, A. and Mewomo, O.T. (2022) Modified Golden Ratio Algorithm for Solving Pseudomonotone

- Variational Inequalities on Hadamard Manifolds. *Computational and Applied Mathematics*, **41**, 38.
- [14] Sakai, T. (1996) Riemannian Geometry. American Mathematical Society. <https://doi.org/10.1090/mmono/149>
- [15] Ferreira, O.P. and Oliveira, P.R. (2002) Proximal Point Algorithm on Riemannian Manifolds. *Optimization*, **51**, 257-270. <https://doi.org/10.1080/02331930290019413>
- [16] Wang, J.H., López, G., Martín-Márquez, V. and Li, C. (2010) Monotone and Accretive Vector Fields on Riemannian Manifolds. *Journal of Optimization Theory and Applications*, **146**, 691-708. <https://doi.org/10.1007/s10957-010-9688-z>
- [17] Van Nguyen, L., Thi Thu, N. and An, N.T. (2021) Variational Inequalities Governed by Strongly Pseudomonotone Vector Fields on Hadamard Manifolds. *Applicable Analysis*, **102**, 444-467. <https://doi.org/10.1080/00036811.2021.1954166>
- [18] Fan, J., Qin, X. and Tan, B. (2020) Tseng's Extragradient Algorithm for Pseudomonotone Variational Inequalities on Hadamard Manifolds. *Applicable Analysis*, **101**, 2372-2385. <https://doi.org/10.1080/00036811.2020.1807012>