

哈达玛流形上的近端梯度法收敛性分析

宋乐乐

上海大学, 上海

Email: 1149902098@qq.com

收稿日期: 2021年4月2日; 录用日期: 2021年5月3日; 发布日期: 2021年5月11日

摘要

本文在哈达玛流形上提出了近端梯度算法, 并给出了收敛性分析。具体地说, 我们将非凸非光滑问题的近端梯度法从欧氏空间推广到哈达玛流形上。在黎曼流形上存在着非线性的困难, 我们根据哈达玛流形的特殊结构, 给出了理论证明。

关键词

近端梯度法, 哈达玛流形, 内蕴算法

Analysis of Convergence of Proximal Gradient Method on Hadamard Manifold

Lele Song

Shanghai University, Shanghai

Email: 1149902098@qq.com

Received: Apr. 2nd, 2021; accepted: May 3rd, 2021; published: May 11th, 2021

Abstract

In this paper, a proximal gradient algorithm for Hadamard manifolds is presented and its convergence is analyzed. Specifically, we extend the proximal gradient method for Nonconvex nonsmooth problems from Euclidan space to Hadamard manifolds. The difficulty of nonlinearity on Bernhard Riemann manifolds is proved theoretically according to the special structure of Hadamard manifolds. In this paper, a proximal gradient algorithm for Hadamard manifolds is presented and its convergence is analyzed. Specifically, we extend the proximal gradient method for nonconvex nonsmooth problems from Euclidean space to Hadamard manifolds. The difficulty of nonlinearity on Bernhard Riemann manifolds is proved theoretically according to the special structure of Ha-

damard manifolds.

Keywords

Near-End Gradient Method, Hadamard Manifold, Intrinsic Algorithm

Copyright © 2021 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

流形是一般几何对象的总称，是欧式空间中的曲线，曲面等概念的推广。流形上的优化不仅是近年来信息科学领域发展起来的一门新兴学科，同时也是优化理论与方法的重要组成部分。在优化领域中，通过利用流形的几何结构，一些欧式空间的非凸问题可以转化为流形上的凸优化问题，一些约束优化问题可以视为流形上的无约束优化问题。因此，近年来这个领域的研究受到越来越多的关注。

哈达玛流形是一个完备的具有非正曲率[1] [2]的简单连通有限维黎曼流形，包括大量重要的矩阵流形和模型。在机器学习[3]，图像处理[4]中都有着广泛的应用。由于其优良的几何结构，哈达玛流形上任意两点间测地线的唯一性假设，避免了算法设计的奇异性。近年来，国内外研究了各种非正曲率流形，如对称正定流形和 n 阶旋转群，已经出现在机器学习和计算机视觉中[5] [6] [7]。考虑流形的几何结构，设计内在的迭代结构保持算法[8] [9]，为一些任务的研究提供了更准确，更有效的途径。

作为非正曲率流形，哈达玛流形具有无共轭点的优良性质，因此其上任意两点之间存在唯一的测地线。它使算法的迭代在整个流形上有了很好的定义，并在这样的流形上发展了许多保持结构的算法。例如，Colao 等人考虑了非线性哈达玛流形上的平衡问题，证明了一类双函数平衡点的存在性[1]和定义在哈达玛流形上的集值映射的不动点定理，并讨论了平衡点的近似问题。近几年，Ruizgarzon 等人提出了哈达玛流形的约束向量优化问题弱有效 Pareto 点的经典充要条件[10]，Ferreia 和 Louzeiro 将梯度方法推广到截面曲率有界的流形上，并证明了一阶收敛速度[11]。另一方面，Ferreira 和 Oliveira 提出了一种求解哈达玛形上光滑凸函数的邻近点方法[12]，并说明了 Riemann 流形中的凸分析，可以用来解决欧式空间中的非凸约束问题，并估计了 $P(1/D)$ 的收敛速度[9] [13]。Wang 等人将求解多值向量场奇异点的非精确邻近点方法推广到哈达玛流形上[14]。Bento 等人还将一种近似点法推广到一般 Riemann 流形，并进行了收敛性分析[15]以及进一步的矢量优化[16]。Ansari 等人提出了一种正则化的非精确邻近点算法，用于寻找哈达玛流形上最大单调集值向量场的奇点[17]。Baygorrea 等人建立了哈达玛流形上拟凸极小化的非精确邻近点方法[18]并估计其收敛速度[19]。Tang 和 Huang 还估计了向量函数在哈达玛流形上的邻近点算法的收敛速度[20]。近几年，Chen 等人提出了 Stiefel 流形上的黎曼近似梯度(RPG)方法[21]，而 Torres Almeida 等人则修改了邻近点算法来优化哈达玛流形上的 DC 函数[22]。结果表明，基于梯度的方法和哈达玛流形上的邻近点算法极大地丰富了流形优化的研究。

在本文中，我们将以另一种方式加速求解哈达玛流形上一类非凸非光滑问题的内蕴近端梯度算法。也就是说，我们考虑下面的最小化问题：

$$\min_{x \in M} F(x) = f(x) + g(x) \quad (1)$$

其中 M 是有限维哈达玛流形(具有非正截面曲率的流形), f 是测地光滑函数, g 是下半连续函数, 可以不连续。

2. 预备知识

我们在本节提出了一些关于哈达玛流形的概念。更详细的内容, 请参阅[2] [13] [18]。

哈达玛流形是一个具有非正截面曲率的完全单连通黎曼流形。在哈达玛流形上任意两点之间存在唯一的测地线。设 M 是哈达玛流形, 给定 $x \in M$, $T_x M$ 表示 M 上在 x 点的切空间。对任意 $y \in M$, 指数映射定义为: $\exp_{x_k} : T_x M \rightarrow M$, 指数映射定义为 $\exp_{x_k}^{-1} : M \rightarrow T_x M$ 满足 $\|\exp_x^{-1} y\| = \|\exp_y^{-1} x\| = d(x, y)$, 其中, $\|\cdot\|$ 为切向量在对应黎曼度量 $\langle \cdot, \cdot \rangle$ 下的范数。 $d(x, y)$ 是 x 和 y 两点间的测地距离。在哈达玛流形上, 对于任意两点 $x, y \in M$, 我们有 $\text{grad} \frac{1}{2} d^2(x, y) = -\exp_x^{-1} y$ 。设 $\gamma : [a, b] \rightarrow M$ 为一条曲线, ∇ 表示 M 上的 Levi-Civita 联络。一个沿着曲线 γ 的向量场 V 是平行向量场, 当且仅当 $\nabla_{\gamma'} V = 0$, 当 $\nabla_{\gamma'} \gamma' = 0$ 时, γ 是一条测地线。测地线 $\gamma(\bullet) = \gamma_v(\bullet, x)$ 表示 $\gamma(0) = x$, $\gamma'(0) = v$ 以及 $\gamma(1) = \exp_x v$ 。 $P_{(x,y)} : T_x M \rightarrow T_x M$ 将向量 $v \in T_x M$ 映射到 $P_{(x,y)} v \in T_x M$ 并保持向量内积不变, 称 $P_{(x,y)}$ 为 M 上的平行移动, 即 $\langle v, w \rangle_x = \langle P_{(x,y)} v, P_{(x,y)} w \rangle_y$, 特别地, $\|v\|_x = \|P_{(x,y)} v\|_y$ 。流形上的一个测地三角 $\Delta(xyz)$ 是由 x, y, z 及连接这些点的三条最小测地线组成的集合。余弦定理是哈达玛流形上一个重要的性质, 任意给定 M 上的测地三角 $\Delta(xyz)$, 有 $d^2(x, z) + d^2(z, y) \leq d^2(x, y) + 2 \langle \exp_z^{-1} x, \exp_z^{-1} y \rangle$ 。

定义 2.1 M 上的可微函数 f 的梯度是 L-Lipschitz 连续的, 那么称 f 是测地 L-光滑(G-L 光滑)函数。即 $\|P_{(x,y)} \text{grad} f(x) - \text{grad} f(y)\| \leq Ld(x, y)$ 。

注: 定义 2.1 等价于

$$\|f(y) - f(x) - \langle \text{grad} f(x), \exp_x^{-1} y \rangle\| \leq \frac{L}{2} \|\exp_x^{-1} y\|^2 \tag{2}$$

定义 2.2 如果函数 $F : M \rightarrow R$ 在 $x \in M$ 满足

$$\lim_{d(x,y) \rightarrow \infty} \frac{F(y)}{d(x,y)} = +\infty$$

则称 F 在 x 点上是 1-强制的。

定义 2.3 令 $g : M \rightarrow R \cup \{+\infty\}$ 是一个下半连续的恰当函数, g 在 $x \in M$ 上的 Fréchet 次微分定义为

$$\partial_F g(x) := \{v \in T_x M \mid \exists (x_m, v_m) \rightarrow (x, v), s.t. g(x_m) \rightarrow g(x), v_m \in \partial_F g(x_m)\}$$

3. 哈达玛流形上的近端梯度法

在欧式空间中近端算法可以看作是解决一些非光滑、有约束、大规模问题的常规方法。本节中, 我们将介绍哈达玛流形上的近端梯度法。

定义 3.1 黎曼流形上的近端算子定义为:

$$\begin{aligned} x_{k+1} &:= \text{prox}_{\lambda_k g}(x'_k) \\ &= \arg \min_{x \in M} \left(g(x) + \frac{1}{2\lambda_k} d^2(x, x'_k) \right), \end{aligned}$$

其中, $x'_k = \exp_{x_k}(-\lambda_k \text{grad} f(x_k))$ 。

我们接下来给出下列假设:

假设 3.1 f 是一个恰当的, 并且在 M 上是 G-L 光滑函数, g 是一个恰当的, 在 M 上是下半连续函

数。

假设 3.2 函数 $F = f + g$ 是 1-强制的，并且在 M 上有下界。

引理 3.1 设 Ω 在 M 上是一个紧集，则存在一个常数 $C(0 < C < +\infty)$ 使得任意 x, y 和 $z \in M$ ，都有

$$\left\| \exp_x^{-1} z - P_{(x,y)} \exp_y^{-1} z \right\| \leq Cd(x, y) \tag{3}$$

证明：由于 M 是一个有限维哈达玛流形，我们可以把 $\exp_x^{-1} z$ 看作 x 点处的光滑函数，此时，结论得证。

注 实际上，由引理 3.1 我们可以得出 $C \geq 1$ 。梯度向量场在 M 上，当 $\lambda = 1$ 时，梯度向量场 $\text{grad} \frac{1}{2} d^2(\cdot, z)$ 是严格单调的。因此，对任意 x, y 和 $z \in M$ ，我们有

$$\left\langle P_{(x,y)} \text{grad} \frac{1}{2} d^2(\cdot, z) - \text{grad} \frac{1}{2} d^2(\cdot, z), \exp_x^{-1} y \right\rangle \geq d^2(x, y)$$

即，

$$\left\langle P_{(x,y)} (-\exp_y^{-1} z) - \exp_x^{-1} z, \exp_x^{-1} y \right\rangle \geq d^2(x, y)$$

因此，

$$\left\| \exp_x^{-1} z - P_{(x,y)} \exp_y^{-1} z \right\| \geq d(x, y)$$

此时，在引理 3.2 中 $C \geq 1$ 。

接下来，我们给出哈达玛流形上满足上述假设的定步长近端梯度算法(如下算法 1)。

Algorithm 1. Upper proximal gradient method

算法 1. M 上近端梯度法

1) 初始化: $x_1 = x_0, \alpha_k < \frac{1}{L}$

2) for $k = 1, 2, 3 \dots$ do

3)
$$x_{k+1} = \text{prox}_{\alpha_k g} \left(\exp_{x_k} (-\alpha_k \text{grad} f(x_k)) \right)$$
 (4)

收敛性分析

定理 3.1 当假设 3.1, 3.2 成立时，算法 1 生成的迭代序列 $\{x_k\}$ 在一个有界区域中，设 x^* 是 $\{x_k\}$ 的任意聚点，则 $0 \in \partial F(x^*)$ 。

证明：

$$\begin{aligned} x_{k+1} &= \text{prox}_{\alpha_k g} \left(\exp_{x_k} (-\alpha_k \text{grad} f(x_k)) \right) \\ &= \arg \min_{x \in M} \frac{1}{2\alpha} d^2 \left(x, \exp_{x_k} (-\alpha_k \text{grad} f(x_k)) \right) + g(x) \end{aligned} \tag{5}$$

令 $\omega_k = \exp_{x_k} (-\alpha_k \text{grad} f(x_k))$ ，则

$$g(x_{k+1}) + \frac{1}{2\alpha} d^2(x_{k+1}, \omega_k) \leq g(x_k) + \frac{1}{2\alpha} d^2(x_k, \omega_k)$$

由 f 的 G-L 光滑性质，我们有，

$$\begin{aligned}
F(x_{k+1}) &= f(x_{k+1}) + g(x_{k+1}) \\
&\leq f(x_k) + \langle \text{grad}f(x_k), \exp_{x_k}^{-1} x_{k+1} \rangle + \frac{L}{2} \|\exp_{x_k}^{-1} x_{k+1}\|^2 \\
&\quad + g(x_k) + \frac{1}{2\alpha} (\text{d}^2(x_k, \omega_k) - \text{d}^2(x_{k+1}, \omega_k))
\end{aligned} \tag{8}$$

由余弦定理, 我们有,

$$\begin{aligned}
&\text{d}^2(x_k, \omega_k) - \text{d}^2(x_{k+1}, \omega_k) \\
&\leq -\text{d}^2(x_{k+1}, x_k) + 2 \langle \text{grad}f(x_k), \exp_{x_k}^{-1} x_{k+1} \rangle \\
&= -\|\exp_{x_k}^{-1} x_{k+1}\|^2 - 2\alpha \langle \text{grad}f(x_k), \exp_{x_k}^{-1} x_{k+1} \rangle
\end{aligned} \tag{9}$$

由(8), (9)我们可得,

$$\begin{aligned}
F(x_{k+1}) &\leq F(x_k) + \frac{1}{2\alpha} \left(-\|\exp_{x_k}^{-1} x_{k+1}\|^2 - 2\alpha \langle \text{grad}f(x_k), \exp_{x_k}^{-1} x_{k+1} \rangle \right) \\
&\quad + \langle \text{grad}f(x_k), \exp_{x_k}^{-1} x_{k+1} \rangle + \frac{L}{2} \|\exp_{x_k}^{-1} x_{k+1}\|^2 \\
&= F(x_k) - \left(\frac{1}{2\alpha} - \frac{L}{2} \right) \|\exp_{x_k}^{-1} x_{k+1}\|^2.
\end{aligned} \tag{10}$$

因为 $\alpha \leq \frac{1}{L}$, 所以, $F(x_{k+1}) \leq F(x_k)$. 从而可得 $F(x_k)$ 是非增的, 因此, 我们有

$$F(x_k) \leq F(x_1), \quad \forall k > 0.$$

由于 f 是 1-强制的, 我们可以得到 $\{x_k\}$ 是有界的, 因此, $\{x_k\}$ 有聚点. 因为 $F(x_k)$ 是非增的且有界, 因此, F 在 $\{x_k\}$ 所有的聚点上都有相同的值, 记为 F^* .

接下来, 我们将证明 $\{x_k\}$ 在一个有界域中. 由(10), 我们可得

$$\left(\frac{1}{2\alpha} - \frac{L}{2} \right) \|\exp_{x_k}^{-1} x_{k+1}\|^2 \leq F(x_k) - F(x_{k+1})$$

因此,

$$\left(\frac{1}{2\alpha} - \frac{L}{2} \right) \sum_{k=1}^{\infty} \|\exp_{x_k}^{-1} x_{k+1}\|^2 \leq F(x_1) - F^* < \infty.$$

进而,

$$\|\exp_{x_k}^{-1} x_{k+1}\|^2 \rightarrow 0 \quad k \rightarrow \infty$$

由(5)的最优性可得,

$$\begin{aligned}
0 &\in \partial g(x_{k+1}) + \frac{1}{2\alpha} \text{grad} \text{d}^2(\cdot, \omega_k) \\
&= \text{grad}f(x_{k+1}) + \partial g(x_{k+1}) - \text{grad}f(x_{k+1}) + \frac{1}{2\alpha} \text{grad} \text{d}^2(\cdot, \omega_k),
\end{aligned}$$

之后, 我们有,

$$\text{grad}f(x_{k+1}) - \frac{1}{2\alpha} \text{grad} \text{d}^2(\cdot, \omega_k) \in \partial F(x_{k+1})$$

因此,

$$\begin{aligned} & \left\| \text{grad} f(x_{k+1}) - \frac{1}{2\alpha} \text{grad} d^2(\cdot, \omega_k) \right\| \\ &= \left\| \text{grad} f(x_{k+1}) - \frac{1}{\alpha} (-\exp_{x_{k+1}}^{-1} \omega_k) \right\| \\ &= \left\| \text{grad} f(x_{k+1}) - P_{(x_k, x_{k+1})} \text{grad} f(x_k) + P_{(x_k, x_{k+1})} \text{grad} f(x_k) + \frac{1}{\alpha} \exp_{x_{k+1}}^{-1} \omega_k \right\| \\ &\leq L \left\| \exp_{x_k}^{-1} x_{k+1} \right\| + \frac{1}{\alpha} \left\| P_{(x_k, x_{k+1})} \alpha \text{grad} f(x_k) + \exp_{x_{k+1}}^{-1} \omega_k \right\|. \end{aligned}$$

因为 f 是 G-L 光滑的, 且 $\{x_k\}$ 在一个有界域上, 存在一个紧集 $\Omega_1 \supset \{x_k\}$ 使得 $\text{grad} f(x)$ 的范数对于任意 $x \in \Omega_1$ 是有限的, 那么, $\{\omega_k\}$ 自然在一个有界区域 Ω_2 中. 除此之外, 我们还可得到 $\{x_k\}$, $\{\omega_k\}$ 都在一个紧集 Ω 中. 因此, 由引理可得, 存在一个常数 C , 使得,

$$\begin{aligned} & \left\| P_{(x_k, x_{k+1})} \alpha \text{grad} f(x_k) + \exp_{x_{k+1}}^{-1} \omega_k \right\| \\ &= \left\| P_{(x_k, x_{k+1})} (\exp_{x_k}^{-1} \omega_k) + \exp_{x_{k+1}}^{-1} \omega_k \right\| \\ &\leq C \left\| \exp_{x_k}^{-1} x_{k+1} \right\|. \end{aligned} \tag{16}$$

由(14), (15), (16)我们可以得出

$$\begin{aligned} & \left\| \text{grad} f(x_{k+1}) - \frac{1}{2\alpha} \text{grad} d^2(\cdot, \omega_k) \right\| \\ &\leq \left(L + \frac{C}{\alpha} \right) \left\| \exp_{x_k}^{-1} x_{k+1} \right\| \rightarrow 0, \quad k \rightarrow \infty. \end{aligned} \tag{17}$$

设 x^* 是 $\{x_k\}$ 的任意聚点, 当 $j \rightarrow \infty$ 时, $\{x_k\}$ 的子列 $x_{k_j} \rightarrow x^*$. 我们现在证明 $0 \in \partial F(x^*)$. 由(5)我们知道

$$g(x_{k_j+1}) + \frac{1}{2\alpha} \text{grad} d^2(x_{k_j+1}, \omega_{k_j}) \leq g(x^*) + \frac{1}{2\alpha} \text{grad} d^2(x^*, \omega_{k_j}),$$

那么,

$$g(x_{k_j+1}) \leq g(x^*) + \frac{1}{2\alpha} (\text{grad} d^2(x^*, \omega_{k_j}) - \text{grad} d^2(x_{k_j+1}, \omega_{k_j})).$$

有余弦定理, 我们可以得到

$$g(x_{k_j+1}) \leq g(x^*) - \frac{1}{2\alpha} \text{grad} d^2(x_{k_j+1}, x^*) + \frac{1}{\alpha} \langle \exp_{x^*}^{-1} x_{k_j+1}, \exp_{x^*}^{-1} \omega_{k_j} \rangle.$$

由于 $\{\omega_k\}$ 在一个有界区域中, 对于任意 $j > 0$, $\exp_{x^*}^{-1} x_{k_j+1}$ 自然也是有界的. 因此, 我们有

$$\limsup_{j \rightarrow \infty} g(x_{k_j+1}) \leq g(x^*).$$

此外, 由 $g(x)$ 下半连续, 我们有

$$\liminf_{j \rightarrow \infty} g(x_{k_j+1}) \geq g(x^*).$$

所以, 我们可以得到

$$\lim_{j \rightarrow \infty} g(x_{k_j+1}) = g(x^*). \tag{18}$$

由 $f(x)$ 的连续性和(18), 可知当 $j \rightarrow \infty$,

$$F(x_{k_{j+1}}) \rightarrow F(x^*).$$

由(14), (17)—(19), 以及次微分的定义可得

$$0 \in \partial F(x^*).$$

4. 结论

本文针对于极小化非凸非光滑函数问题, 主要研究了哈达玛流形上的近端梯度算法, 我们首先将欧式空间的近端梯度算法推广到哈达玛流形上, 再根据该流形的优良结构, 给出流形上近端梯度法的收敛性证明。理论上, 我们可以得出我们提出的哈达玛流形上的近端梯度算法是收敛的。由于流形上非线性性质, 收敛速度的理论研究存在一定的困难, 还需要进一步研究。

参考文献

- [1] 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>
- [2] Sakai, T. (1996) Riemannian Geometry (Translations of Mathematical Monographs). American Mathematical Society, Providence, 262-272. <https://doi.org/10.1090/mmono/149>
- [3] Zhang, H.C. and Hager, W.W. (2004) A Nonmonotone Line Search Technique and Its Application to Unconstrained Optimization. *SIAM Journal on Imaging Sciences*, **14**, 1043-1056. <https://doi.org/10.1137/S1052623403428208>
- [4] Beck, A. and Teboulle, M. (2009) A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. *SIAM Journal on Imaging Sciences*, **2**, 183-202. <https://doi.org/10.1137/080716542>
- [5] Ying, S., Wen, Z., Shi, J., Peng, Y., Peng, J. and Qiao, H. (2018) Manifold Preserving: An Intrinsic Approach for Semi-supervised Distance Metric Learning. *IEEE Transactions on Neural Networks and Learning Systems*, **29**, 2731-2742.
- [6] Li, H., Fang, C. and Lin, Z. (2020) Accelerated First-Order Optimization Algorithms for Machine Learning. *Proceedings of the IEEE*, **108**, 2067-2082. <https://doi.org/10.1109/JPROC.2020.3007634>
- [7] Wang, Q., Yuen, P.C. and Feng, G. (2013) Semi-Supervised Metric Learning via Topology Preserving Multiple Semi-Supervised Assumptions. *Pattern Recognition*, **46**, 2576-2587. <https://doi.org/10.1016/j.patcog.2013.02.015>
- [8] Absil, P.A., Mahony, R. and Sepulchre, R. (2009) Optimization Algorithms on Matrix Manifolds. Princeton University Press, Princeton. <https://doi.org/10.1515/9781400830244>
- [9] Bacák, M. (2014) Convex Analysis and Optimization in Hadamard Spaces, Vol. 22. Walter de Gruyter GmbH & Co KG. <https://doi.org/10.1515/9783110361629>
- [10] Ruizgarzon, G., Osunagomez, R. and Ruizzapatero, J. (2019) Necessary and Sufficient Optimality Conditions for Vector Equilibrium Problems on Hadamard Manifolds. *Symmetry*, **11**, 1037. <https://doi.org/10.3390/sym11081037>
- [11] Ferreira, O.P., Louzeiro, M.S. and Prudente, L.F. (2019) Gradient Method for Optimization on Riemannian Manifolds with Lower Bounded Curvature. *SIAM Journal on Optimization*, **29**, 2517-2541. <https://doi.org/10.1137/18M1180633>
- [12] Ferreira, O.P. and Oliveira, P.R. (2002) Proximal Point Algorithm on Riemannian Manifolds. *Optimization*, **51**, 257-270. <https://doi.org/10.1080/02331930290019413>
- [13] Bento, G.C., Ferreira, O.P. and Melo, J.G. (2017) Iteration-Complexity of Gradient, Subgradient and Proximal Point Methods on Riemannian Manifolds. *Journal of Optimization Theory and Applications*, **173**, 548-562. <https://doi.org/10.1007/s10957-017-1093-4>
- [14] Wang, J., Li, C., Lopez, G. and Yao, J.-C. (2015) Convergence Analysis of Inexact Proximal Point Algorithms on Hadamard Manifolds. *Journal of Global Optimization*, **61**, 553-573. <https://doi.org/10.1007/s10898-014-0182-2>
- [15] Bento, G.C., Cruz Neto, J.X. and Oliveira, P.R. (2016) A New Approach to the Proximal Point Method: Convergence on General Riemannian Manifolds. *Journal of Optimization Theory and Applications*, **168**, 743-755. <https://doi.org/10.1007/s10957-015-0861-2>
- [16] Bento, G.C., Ferreira, O.P. and Oliveira, P.R. (2018) Proximal Point Method for Vector Optimization on Hadamard Manifolds. *Operations Research Letters*, **46**, 13-18. <https://doi.org/10.1016/j.orl.2017.10.017>
- [17] Ansari, Q.H., Babu, F. and Yao, J.-C. (2019) Regularization of Proximal Point Algorithms in Hadamard Manifolds.

-
- Journal of Fixed Point Theory and Applications*, **21**, Article No. 25. <https://doi.org/10.1007/s11784-019-0658-2>
- [18] Baygorrea, N., Papa Quiroz, E.A. and Maculan, N. (2016) Inexact Proximal Point Methods for Quasiconvex Minimization on Hadamard Manifolds. *Journal of the Operations Research Society of China*, **4**, 397-424. <https://doi.org/10.1007/s40305-016-0133-3>
- [19] Baygorrea, N., Papa Quiroz, E.A. and Maculan, N. (2017) On the Convergence Rate of an Inexact Proximal Point Algorithm for Quasiconvex Minimization on Hadamard Manifolds. *Journal of the Operations Research Society of China*, **5**, 457-467. <https://doi.org/10.1007/s40305-016-0129-z>
- [20] Tang, F.M. and Huang, P.L. (2017) On the Convergence Rate of a Proximal Point Algorithm for Vector Function on Hadamard Manifolds. *Journal of the Operations Research Society of China*, **5**, 405-417. <https://doi.org/10.1007/s40305-016-0146-y>
- [21] Chen, S., Ma, S., Man-Cho So, A. and Zhang, T. (2020) Proximal Gradient Method for Nonsmooth Optimization over the Stiefel Manifold. *SIAM Journal on Optimization*, **30**, 210-239. <https://doi.org/10.1137/18M122457X>
- [22] Torres Almeida, Y., Cruz Neto, J.X., Oliveira, P.R. and Oliveira Souza, J.C. (2020) A Modified Proximal Point Method for DC Functions on Hadamard Manifolds. *Computational Optimization and Applications*, **76**, 649-673. <https://doi.org/10.1007/s10589-020-00173-3>