arXiv:2302.02689v1 [math.OC] 6 Feb 2023

Bregman函数的本质

Edouard Pauwels*

2023年2月7日草稿

摘要

设 $C \subset \mathbb{R}^n$ 为凸紧集，且内点非空； h 为定义域为 C、在 C上连续的Legendre函数。我们证明， h 是Bregman函数，当且仅当它在 C 上严格凸，且 C 是一个多胞体。这为许多基于Bregman散度的算法的序列收敛提供了洞见：Bregman拓扑与欧几里得拓扑之间的抽象相容性条件，可以等价地替换为关于 h 和 C的显式条件。这也强调，对于这些方法（定义域超出多面体定义域）的普遍收敛理论，需要比Bregman条件更加细致的完善。

1 引言

令 $C \subset \mathbb{R}^n$ 为具有非空内部的凸紧集， $h\colon\ C \mapsto \mathbb{R}$ 为凸函数。本文始终假设 h 是 Legendre 函数：

- h 是本质光滑的：在 int$C$ 上连续可微，且当趋近于 C 的边界时， $\|\nabla h\|$ 趋于$+\infty$ 。
- h 在 int$C$ 上严格凸。

我们还假设全文中 h 在 C 上连续。因此，对于所有 $y \in C$ 和 $x \in$int$C$，与 h 相关的Bregman散度为：

$$D_h(y,x) = h(y) - h(x) - \langle \nabla h(x), y - x \rangle .$$

1.1 Fej´erian 序列：

设 $f\colon\ C \to \mathbb{R}$ 为凸的适当下半连续函数。考虑问题

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

*IR法国国家科学研究中心（CNRS），图卢兹大学，法国图卢兹。法国大学研究院（IUF） .

并记 $S \subset C$ 为（1）的解集。下面的定义改编自欧氏情形 [11]。

定义 1. 设 $(x_k)_{k\in\mathbb{N}}$ 是int$C$中的一个序列，若满足以下条件，则称其为问题(1) 的 ***Fej´ér*** 序列，或 ***Fej´er*** 序列：

- 所有聚点都属于 S。
- 对于所有 $y \in S$， $D_h(y, x_k)$ 当 $k \to \infty$ 时存在有限极限。

实际上，定义 1 可针对一般的闭凸集 $S \subset C$ 给出（参见例如 [11]，其内容可直接改写）。为简洁起见，我们采用优化问题 (1)，因为它已经涵盖了范围广泛的算法。

1.2 算法示例

下面列出一些在凸优化中生成 Fej´ér 序列的常见算法。

镜像下降： 该算法由 Nemirovsky [13]提出。假设 f 是 Lipschitz 连续的。初始化 $x_0 \in$ int$C$，给定正步长序列 $(\alpha_k)_{k\in\mathbb{N}}$，迭代如下

$$x_{k+1} = \nabla h^*(\nabla h(x_k) - \alpha_k v_k), \tag{2}$$

$$v_k \in \partial f(x_k). \tag{3}$$

假设 $\sum_{k=0}^{+\infty} \alpha_k = +\infty$ 和 $\sum_{k=0}^{+\infty} \alpha_k^2 < +\infty$，所得算法序列具有 Fej´er 性质。在这种情况下，[6]中的方程 (4.22) 确保聚点属于 S，而方程 (4.21) 结合第 2.2[14] 节的 [引理 2，确保对于所有$y \in S$， $D(y,\ x_k)$ 都存在极限。

Bregman梯度与 NoLips算法： 假设 f 在包含 C 的开集上是 C^1，且 $h - \alpha f$ 对于某个 $\alpha > 0$是凸的。初始化 $x_0 \in$ int$C$ 并迭代如下

$$x_{k+1} = \nabla h^*(\nabla h(x_k) - \alpha \nabla f(x_k)). \tag{4}$$

在这种情况下，对于所有 $y \in S$，当 k 增长时， $D_h(y,\ x_k)$ 单调不增 [4, 引理 5]。这将推广到 [4]中针对复合目标提出的 NoLips算法。同样， [1]中的工作在 Hessian-Riemann梯度流的背景下，提出了 Fej´er 性质的连续时间变体。

采用 D-函数进行近端最小化： 假设 f 是真下半连续的。初始化 $x_0 \in$int$C$，并给定正步长序列 $(\alpha_k)_{k\in\mathbb{N}}$，进行迭代

$$x_{k+1} = \arg\min_{x\in C} \alpha_k f(x) + D_h(x, x_k). \tag{5}$$

在这种情况下，对于所有 $y \in S$, D_h（$y,\ x_k$）都随 k 的增大而单调不增 [10, 引理 3。3] 并且若 $\sum_{k=0}^{+\infty} \alpha_k = +\infty$，则所得序列的所有聚点都在 S [10,定理 3.4] 中。

交替投影： 除了优化问题(1)之外，Bregman 在其奠基性论文[7], 中考虑了一种交替投影算法，该算法生成一个 Fej´er 序列，以求得凸集交集中的一个元素。

1.3 序列收敛分析

与欧几里得情形类似，Fej´erian 性质可用于证明该序列收敛，这与 Opial 引理有关（概述见 [11] ）。只要由 D_h 编码的拓扑等价于欧几里得拓扑，这一论证就成立。由于严格凸性，这在 C 的内部总是成立，但要确保在边界处也成立，还需要作出进一步假设。第一个条件是

$$(x_k)_{k\in\mathbb{N}} \subset \mathrm{int}C,\, y \in C \qquad D_h(y, x_k) \underset{k\to\infty}{\to} 0 \quad \Rightarrow \quad x_k \underset{k\to\infty}{\to} y \tag{A}$$

根据 h 在其定义域上的连续性可知， h 在整个定义域 C 上（而不仅仅是在内部）具有严格凸性，足以推出 (A) [12, 引理 2.16]。若假设相反的蕴含关系

$$(x_k)_{k\in\mathbb{N}} \subset \mathrm{int}C,\, y \in C, \qquad x_k \underset{k\to\infty}{\to} y \quad \Rightarrow \quad D_h(y, x_k) \underset{k\to\infty}{\to} 0 \tag{B}$$

那么我们有如下 F´ejer 论证：假设 (A) 和 (B) 均成立，则任意 Fej´erian 序列 $(x_k)_{k\in\mathbb{N}}$（定义 1）都收敛。事实上，由紧致性可知存在一个聚点 $\bar{x}$，根据定义 1，它必定属于 S；由性质 (A)，取一个子列后，$D_h(\bar{x},\ x_k) \to 0$，但定义 1 保证 $D_h(\bar{x},\ x_k)$ 收敛，因此其极限必为 0，而性质 (B) 保证 $x_k \to \bar{x}$。

注意，由于 h 在其定义域上连续，条件(B)等价于当 $x_k \to y$时，$\langle \nabla h(x_k)\ y - x_k\rangle \to 0$。条件(A)和条件(B)可追溯至 Bregman [7]，并已在文献 [8, 9, 10, 5, 12, 1, 2, 4]中以相同的抽象形式，或经调整后在比此处所考虑范围更广的情形下得到广泛研究。给定一个在其紧定义域 C上连续的 Legendre 函数 h，若条件(A)和条件(B)成立，则称 h 为 Bregman函数。

1.4 主要结果

我们关注以下问题：条件(A)和条件(B)的限制有多严格？换言之，一个连续的 Legendre 函数要成为 Bregman函数，需要满足多少条件？我们给出如下答案。

定理1. 设 $C \subset \mathbb{R}^n$ 为凸紧集，且内部非空； h 为 *Legendre* 函数，其定义域为 C，并在 C 上连续。则：

- （A）成立，当且仅当 h 在 C上严格凸。
- *(B)</style}*成立，当且仅当 C 是一个多胞体。

定理1给出了一个关于 h 的显式充分条件，以确保第1.2节中所述所有算法都能收敛。事实上，若 C 是一个多胞体，且 h 在C上严格凸，则任何 Fej´er 序列都收敛。此外，定理1说明，现有用于证明序列收敛的 Fej´er 论证仅适用于多胞体域，而对于更一般的定义域，收敛分析则需要采用不同的论证。

定理1可由命题1、命题 2、引理5和引理6结合得出。我们关键地用到了 [15, 引理 1]。需要指出的是，定理1及本文的所有论述仅限于紧致 C，但大多数证明论证都涉及可能无界的闭 C，因此中间结果对于这一更一般的情形也具有独立的参考价值。严格凸性与 (A) 之间的联系很自然，其中一个方向的结论归功于 Kiwiel [12]。(B) 与 C 的多面体性质之间的联系是我们结果中最有意思的部分，我们将从这里开始讨论。

2 条件(B)

我们先通过一个简单的例子来说明 (B) 不成立的情形及其与曲率的关系；该例子曾由不同作者独立地在 [3, 注记 2]中描述。

2.1 直觉：与曲率不相容

设 $h\colon\ x \mapsto -\sqrt{1-\|x\|^2}$ 是 $C \subset \mathbb{R}^2$上的单位欧几里得球。我们有$\nabla h\colon\ x \mapsto x\frac{1}{\sqrt{1-\|x\|^2}}$。在平面上考虑极坐标，$x =$（$r\cos$（$\theta$） $r\sin$（θ））$,$ 其中 e_1 是第一个基向量（1，0），于是有

$$h(e_1) - h(x) - \langle \nabla h(x), e_1 - x\rangle = -\sqrt{1-r^2} + \frac{r(r-\cos(\theta))}{\sqrt{1-r^2}}$$

选择 $\theta(r) = \arccos\left(r - \sqrt{1-r^2}\right)$时，对于 $r \geq 0$，当 $r \to 1, x(r) = (r\cos(\theta(r))\ r\sin(\theta(r)))$ 趋于 e_1 且 $D_h(e_1,\ x) \to 1$。

实际上，曲线 $r \to x(r)$ 收敛到 e_1，并在此处具有垂直切线，如图1所示。这说明了条件 B 失效的主要机制。由于曲率，弦都与 C 的内部相交，因此沿弦的 h 方向导数在弦的端点处趋于无穷大。在这里，曲线仍处于一段弦上，而沿弦的方向导数在该区域为负。

从这个例子中可以得到一个直观认识：Bregman函数的紧定义域 C 的边界不应具有过多曲率。更准确地说，曲率不应在任何地方聚集。这个直觉实际上是正确的，其关键机制在于极点附近存在大量曲率，因此能够产生类似于上文圆形例子的行为，见图1。这种情况不可能过于密集，否则就会与 (B) 矛盾。因此，C 的极点必须是孤立的，这一点在命题1中表述。另一方面，多面体具有非常强的结构，这会确保 (B) 成立；这就是命题 2。

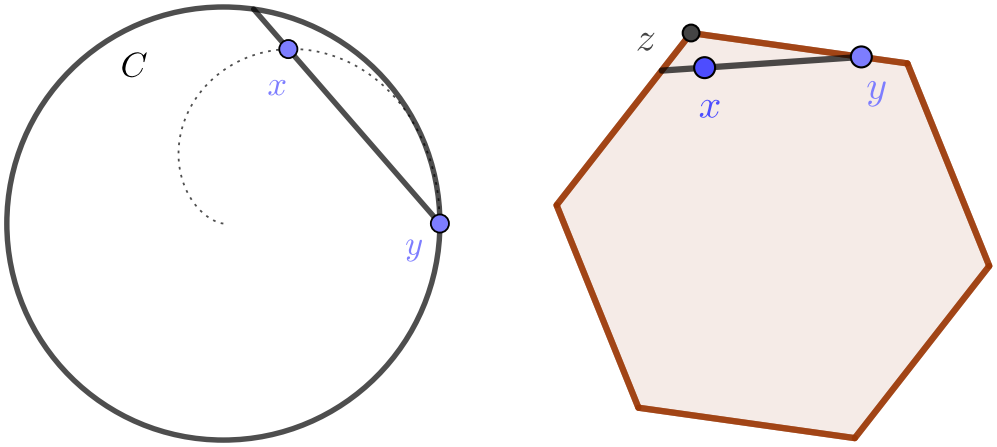


图1：左：切向现象的示意图，即靠近边界以及弦的相对端点。右：一般而言，极点构成一个高曲率区域，并且可以找到构型相似的邻近点。在条件(B)下，这种情况不能过多发生；特别是，极点不应在任何地方聚集。

2.2 第一蕴含

本节将证明以下结论

命题1. 设 $C \subset \mathbb{R}^n$ 为具有非空内部的凸闭集，且 h 在 C上为 *Legendre* 函数，在 C 上连续，并满足(B)。则 C 的极点局部有限。特别地，若 C 有界，则它是一个多胞体。

我们先给出两个引理

引理 1. 在命题*1*的假设下，设 $x \in \mathrm{bd}C$，且 $y \in C$ 满足$^{x+y}_{2} \in \mathrm{int}C$。则 $D_h(y,(1-\lambda)x \mp \lambda y) \to +\infty$，当 $\lambda \to 0$时。

证明： 注意，在这种情况下，开线段 (x, y) 包含于 C 的内部。对于任意 $\lambda \in (0, 1)$，令 $z_\lambda = (1-\lambda)x + \lambda y$，我们有

$$D_h(y,(1-\lambda)x+\lambda y) = h(y) - h(z_\lambda) - (1-\lambda)\langle \nabla h(z_\lambda), y - x\rangle .$$

令 $\lambda \to 0$，考虑 $^{\nabla h}_{(z_\lambda)\|\nabla h(z_\lambda)\|}$ 的任意聚点 $d \in \mathbb{R}^n$。根据引理 4，我们有 $d \in N_C(x)$ [1, 。2] 并且 $y - x \in \mathrm{int}T_C(x)$（因为 $^{x+y}_{2} \in \mathrm{int}C$），因此 $\langle d, y - x\rangle < 0$。所以，取一个子列使得当 $\lambda \to 0$ 时，

$$D_h(y,(1-\lambda)x+\lambda y) = h(y) - h(z_\lambda) - (1-\lambda)\|\nabla h(z_\lambda)\| \left\langle \frac{\nabla h(z_\lambda)}{\|\nabla h(z_\lambda)\|}, y - x \right\rangle ,$$
$$\to +\infty$$

由于 $h(z_\lambda) \to h(x)$，$\|\nabla h(z_\lambda)\| \to \infty$ 和 $\langle d, y - x\rangle < 0$。由于 d 是任意的，结论成立。证明至此结束。□

引理 2. 在命题*1*的假设下，设 $a \in$边界C 是一个极点，且$y \in C$与 a不同。那么对于任意 ϵ, $K > 0$，存在 $c \in$ 内部C，使得

$$\|c - a\| \leq \epsilon$$
$$D_h(y, c) \geq K.$$

证明： 首先，如果 $[a, y] \cap \text{int}C \neq \emptyset$，那么我们满足引理 1 的条件，由此可得所需结果（取 $x = a$）。因此，我们可以假设 $[a, y] \subset \text{bd}C$。

断言 1. 存在 $\epsilon_1 > 0$，使得对于任意 $b \in$ 内点$C \cap B_{\epsilon_1}(a)$，从 y 到 b的直线会与 C 的边界相交于 x，且满足 $\|x - a\| \leq \epsilon$。

断言的证明。我们可以假设 $\epsilon_1 < \|y - a\|$，因此，由于 $a \neq y$，对于任意 $b \in B_{\epsilon_1}(a)$，点 y 和 b 确定一条唯一的直线。我们还要求 $\epsilon_1 < \epsilon$。反设，对于所有这样的 $\epsilon_1 > 0$，都存在 $b \in \text{int}C \cap B_{\epsilon_1}(a)$，使得从 y 到 b 的直线要么不与 C 的边界相交，要么与其相交于距离大于 ϵ的点。在这两种情况下，由于 $b \in B_\epsilon(a)$（如我们所假设的 $\epsilon_1 < \epsilon$），从 y 到 b 的直线会在一点 $x \in C$处离开 $B_\epsilon(a)$。此时，利用 ϵ_1的一列趋零值，我们可以构造一个收敛到 a 的序列 $(b_k)_{k \in \mathbb{N}}$，使得每条线段 $[y, b_k]$都能延伸至 $x_k \in C$，其中 $b_k \in [x_k, y]$ 且 $\|x_k - a\| = \epsilon$。取极限后，除去子列 $b_k \to a$ 和 $x_k \to x \in C$，我们有 $a \in [x, y]$，其中 $a \neq y$ 且 $a \neq x$，这与 a 是 C的极点这一事实矛盾。断言得证。

我们可以选择 $\epsilon_1 \leq \epsilon$，现在考虑任意一点 $b \in \text{int}C \cap B_{\epsilon_1}(a)$，$b$ 属于某条形如 $[x, y]$ 的线段，其中 $x \in \text{bd}C \cap B_\epsilon(a)$。注意到 $[b, x] \subset B_\epsilon(a)$，因此我们可以应用引理 1，得到 $c \in [b, x]$，使得 $D_h(y,\ c)$ 可以任意大。这就证明了所需结果。

□

命题1的证明： 假设 C 的极点集不是局部有限的。这意味着我们可以找到一个由两两不同的极点组成的有界序列 $(z_k)_{k \in \mathbb{N}}$。该序列有一个收敛子列，设 y 为其极限。由于序列中的元素两两不同，至多存在一个 $k \in \mathbb{N}$ 使得 $y = z_k$，因此我们可以将其从序列中删去。换言之，$z_k \to y \in C$ 当 $k \to \infty$，且对所有 $k \in \mathbb{N}$，都有 $z_k \neq y$。我们将证明在这种情况下，条件 B 不成立。

固定 $k \in \mathbb{N}$，利用引理 2 和 $a = z_k$，我们可以找到 $x_k \in \text{int}C$，使得 $\|x_k - z_k\| \trianglelefteq \|z_k - y\|$ 且 $D_h(y, x_k) \geq k$。现在，由于 $k \to \infty$，我们有 $x_k \to y$ 且 $D_h(y, x_k) \to +\infty$，这与条件 B 矛盾。证明至此结束。□

2.3 逆向蕴含

命题 2. 设 $C \subset \mathbb{R}^n$ 为具有非空内部且局部多面体状（局部可表示为有限个仿射不等式，例如一个多胞体）的凸闭集，h在 C上为 ***Legendre*** 函数，在 C上连续，则 (B) 成立。

该结果由以下结论（[15, 引理 1]）得出。

引理 3 (Tseng 和 Bertsekas)。 设 $h\colon \mathbb{R}^p \mapsto \mathbb{R} \cup +\infty$ 是下半连续的，且在其定义域 C 上连续。那么

- 对于任意 $y \in C$，存在一个以 y 为中心的非退化闭球，使得 $C \cap B$ 是闭集。

- 对于任意 $y \in C$ 和满足 $y + d \in C$ 的 d，以及满足 $x_k + d_k \in C$ （对所有 k 均成立）的序列 $x_k \to y$ 和 $d_k \to d$，我们有

$$\limsup_{k\to\infty} h'(x_k, d_k) \leq h'(y, d).$$

引理 4. 设 C 是一个多面体，且 $y \in C$。那么存在 $\epsilon > 0$，使得对于任意 $x \in C \cap B_\epsilon(y)$，都有 $2x - y \in C$。

证明： 设 $a_1, \ldots, a_m \in \mathbb{R}^n$ 和 $b_1, \ldots, b_m \in \mathbb{R}$ 满足 $C = \{x | \langle a_i, x\rangle \leq b_i, i = 1, \ldots, m\}$。固定 $y \in C$，并设 $I \subset \{1, \ldots m\}$ 为活跃指标集合，使得当且仅当 $i \in I$ 时，$\langle a_i, y\rangle = b_i$ 成立。由线性函数的连续性，存在 $\epsilon > 0$，使得对所有满足 $\|d\| \leq 2\epsilon$ 的 d 以及所有 $i \notin I$，都有 $\langle a_i, y + d\rangle < b_i$。现在，对于任意满足 $\|x - y\| \leq \epsilon$ 的 $x \in C$，我们有

$$\begin{aligned} \langle a_i, 2x - y\rangle &= \langle a_i, y + 2(x - y)\rangle < b_i && \forall i \notin I, \\ \langle a_i, 2x - y\rangle &= 2\langle a_i, x\rangle - \langle a_i, y\rangle \leq 2b_i - \langle a_i, y\rangle = 2b_i - b_i = b_i && \forall i \in I, \end{aligned}$$

这表明 $2x - y \in C$，从而证明得证。 □

命题 2 的证明： 由 h 在 C 上的连续性，只需证明

$$\lim_{x\to y, x\in \mathrm{int} C} \langle \nabla h(x), y - x\rangle = 0$$

对所有 $y \in C$ 均成立。

固定 $y \in C$，并考虑收敛于 y 的 $(x_k)_{k\in\mathbb{N}}$。 C 可局部表示为一个多面体，因此可以将引理 4 应用于 y 和 C。令 $\epsilon > 0$ 由引理 4 给出，并且不失一般性地假设对所有 $k \in \mathbb{N}$ 都有 $\|y - x_k\| \leq \epsilon$。对于所有 $k \in \mathbb{N}$， $x_k + (y - x_k) = y \in C$，因此将引理 4 应用于 $d_k = y - x_k$ 后，其收敛于 0

$$\limsup_{k\to\infty} \langle \nabla h(x_k), y - x_k\rangle \leq h'(y, 0) = 0.$$

此外，根据引理 4，对所有 $k \in \mathbb{N}$，我们有 $2x_k - y = x_k + (x_k - y) \in C$。因此，可以将引理 4 应用于 $d_k = x_k - y$，其同样收敛于 0，从而得到

$$\limsup_{k\to\infty} \langle \nabla h(x_k), x_k - y\rangle \leq h'(y, 0) = 0.$$

这等价于

$$\liminf_{k\to\infty} \langle \nabla h(x_k), y - x_k\rangle \geq 0.$$

我们已经证明

$$0 \leq \liminf_{k\to\infty} \langle \nabla h(x_k), y - x_k\rangle \leq \limsup_{k\to\infty} \langle \nabla h(x_k), x_k - y\rangle \leq 0,$$

因此极限为 0。证明到此结束。 □

3 条件(A)

其中一个推论由 [12,引理 216]给出。

引理5 (Kiwiel). 设 $C \subset \mathbb{R}^n$ 为凸闭且具有非空内部，设 h 在 C上为 ***Legendre*** 函数，在 C 上连续且在 C上严格凸，则(A) 成立。

反向蕴含可由引理 3 再次推出。

引理6. 设 $C \subset \mathbb{R}^n$ 为凸闭且具有非空内部，设 h 在C上为 ***Legendre*** 函数，在 C 上连续且满足(A)，则 h 在 C上严格凸。

证明： 反设 h 不是严格凸的。这意味着存在 $x, y \in C$，使得

$$h\left(\frac{x+y}{2}\right) = \frac{h(x)+h(y)}{2}, \tag{6}$$

这意味着 h 沿线段 $[x, y]$为仿射函数，即

$$h\left(\frac{x+y}{2} + t(x-y)\right) = h\left(x\left(t+\frac{1}{2}\right) + y\left(\frac{1}{2}-t\right)\right) = \frac{h(x)+h(y)}{2} + t(h(x)-h(y))$$

对所有 $t \in [-1/2, 1/2]$ 都成立。特别地，令 $z = \frac{x+y}{2}$，可得

$$h'(z, x-z) = -h'(z, y-z) = \frac{h(x)-h(y)}{2}. \tag{7}$$

现在考虑 $z_0 \in \text{int}C$ 以及序列 $(z_k)_{k\in\mathbb{N}}$，其中元素属于 int$C$，且对所有 $k \in \mathbb{N}, k \geq 1$, $z_k = \frac{1}{k}z_0 + \frac{k-1}{k}z \in C$ 都成立。
对所有 $k \in \mathbb{N}, k \geq 1$, $y_k = \frac{2}{k+1}z_0 + \frac{k-1}{k+1}y \in C$，我们令其为.....。对所有 $k \geq 1$，有

$$\begin{aligned}
z_k - x &= \frac{1}{k}z_0 + \frac{k-1}{2k}y + \left(\frac{k-1}{2k} - \frac{2k}{2k}\right)x = \frac{2}{2k}z_0 + \frac{k-1}{2k}y - \frac{k+1}{2k}x \\
y_k - x &= \frac{2}{k+1}z_0 + \frac{k-1}{k+1}y - \frac{k+1}{k+1}x = (z_k - x)\frac{2k}{k+1} \\
y_k - z_k &= y_k - x + x - z_k = (z_k - x)\left(\frac{2k}{k+1} - 1\right) = \frac{k-1}{k+1}(z_k - x).
\end{aligned}$$

利用引理 3，我们得到

$$\begin{aligned}
\limsup_{k\to\infty} \langle \nabla h(z_k), y_k - z_k \rangle &= \limsup_{k\to\infty} \frac{k-1}{k+1} \langle \nabla h(z_k), z_k - x \rangle \\
&= -\liminf_{k\to\infty} \langle \nabla h(z_k), x - z_k \rangle \\
&\leq h'(z, y-z) = -h'(z, x-z),
\end{aligned}$$

因此

$$\limsup_{k\to\infty} \langle \nabla h(z_k), x - z_k \rangle \leq h'(z, x - z) \leq \liminf_{k\to\infty} \langle \nabla h(z_k), x - z_k \rangle$$

并且 $\langle \nabla h(z_k)\ x - z_k \rangle \to h'(z,\ x - z)$ 为 $k \to \infty$。利用 (6) 和 (7)，由 h 的连续性可推得

$$\lim_{k\to\infty} h(x) - h(z_k) - \langle \nabla h(z_k), x - z_k \rangle = h(x) - \frac{h(x) + h(y)}{2} - \frac{h(x) - h(y)}{2} = 0$$

因此，我们有 $z_k \to z \neq x$，但 $D_h(x,\ z_k) \to 0$，这表明条件 (A) 不成立。通过逆否命题即可证明该结论。□

致谢

作者谨向 J´erˆome Bolte 致谢，感谢他一直以来的支持与富有成效的交流。本研究得到 AI Interdisciplinary Institute ANITI（ANR-19- PI3A-0004）、美国空军科学研究办公室、美国空军装备司令部、美国空军（FA9550-19-1-7026、FA8655-22-1-7012）以及 ANR MaSDOL（19-CE23-0017-01）的资助。

参考文献

[1] Alvarez, F., Bolte, J. 和 Brahic, O.（2004）。凸规划中的Hessian黎曼梯度流。《SIAM控制与优化杂志》，43(2)，477–501。[2] Auslender, A., Teboulle, M.（2006）。凸优化与锥优化的内点梯度法和近端方法。《SIAM优化杂志》，16(3)，697-725。[3] Azizian, W., Iutzeler, F., Malick, J. 和 Mertikopoulos, P.（2022）。约束变分不等式中Bregman近端方法的收敛速度。arXiv预印本 arXiv:2211.08043。[4] Bauschke, H. H., Bolte, J. 和 Teboulle, M.（2016）。超越Lipschitz梯度连续性的下降引理：重新审视一阶方法及其应用。《运筹学数学》，42(2)，330–348。[5] Bauschke, H. H. 和 Borwein, J. M.（1997）。Legendre函数与随机Bregman投影法。《凸分析杂志》，4(1)，27-67。[6] Beck, A. 和 Teboulle, M.（2003）。凸优化的镜像下降与非线性投影次梯度方法。《运筹学快报》，31(3)，167–175。[7] Bregman, L. M.（1967）。求凸集公共点的松弛方法及其在求解凸规划问题中的应用。《苏联计算数学与数理物理》，7(3)，200-217。

[8] Censor, Y., Lent, A.（1981）。区间凸规划的迭代行作用法。《优化理论与应用杂志》，34(3)，321-353。[9]Censor, Y., Zenios, S. A.（1992）。使用d-函数的近端最小化算法。《优化理论与应用杂志》，73(3)，451-464。[10] Chen, G., Teboulle, M.（1993）。使用Bregman函数的类近端最小化算法的收敛分析。《SIAM优化杂志》，3(3)，538-543。[11] Combettes, P. L.（2001）。凸优化中的Fej´er单调性。《优化百科全书》，2，106-114。[12] Kiwiel, K. C.（1997）。严格凸成本和线性约束问题的自由引导松弛法。《运筹学数学》，22(2)，326-349。[13] A. S. Nemirovsky 和 D. B. Yudin（1983）。《优化中的问题复杂性与方法效率》。Wiley-Interscience，纽约。[14] Polyak, B. T.（1987）。《优化导论》。Optimization Software Inc.，出版部，纽约，1，32。[15] Tseng, P., Bertsekas, D. P.（1991）。严格凸成本和线性约束问题的松弛法。《运筹学数学》，16(3)，462-481。