arXiv:2303.01797v1 [math.OC] 3 Mar 2023

凸多目标优化中的随机逼近

卡洛·阿尔贝托·德·贝尔纳迪、恩里科·米列里纳、埃莱娜·莫尔霍和雅各布·索马利亚

Abstract. 给定一个严格凸的多目标优化问题，其目标函数为 f_1、.....、f_N，记其解为 x_0。该解是线性标量化问题的最小点，其中目标函数是 f_1、.....、f_N的凸组合，权重为 t_1、.....、t_N。本文的主要结果给出了平均误差的估计：当我们用从 f_1、.....、f_N中选取的 n 个函数的凸组合的最小点来逼近x_0 时，所产生的平均误差；这些函数分别以概率 t_1、.....、t_N选取，并采用相同的系数 $1/n$进行加权。特别地，我们证明上述平均误差在 n趋于 ∞时收敛至 0，且关于权重 t_1、.....、t_N一致。随机逼近定理证明中的关键工具是一种几何性质，我们称之为小直径性质；它确保函数 f_1、.....、f_N的凸组合的最小点连续依赖于该凸组合的系数。

1. Introduction

本文的主要目的是，按照 [2,定理 1（第28]页）的思路，提出一个关于凸多目标优化问题解的逼近结果。在该文中，作者给出了一个结果，并由他们自己解释了其含义（原文为法语，以下引文译自英文，参见第 30]页的 [2, 备注）：

“该定理意味着，要计算函数f_1、.....、f_N的一个帕累托最优，只需近似地计算其中 n 个函数的帕累托最优；更进一步，从 N 中选取 n 个元素的大多数选择都能得到良好的结果。”

为阐明上述评论，我们简要说明什么是多目标优化问题。关于这一领域的完整综述，请读者参阅如今已成为经典的专著 [9] 和 [8]。设 $f_1, \dots, f_N$ 是从赋范空间 X 到 $\mathbb{R}$的 N函数。当不存在满足以下条件的 $x \in X$时，点 $x_0 \in X$ 称为函数 $f_1, \dots, f_N$的帕累托最优点（或有效点）：对于$k = 1, \dots, N$，有 $f_k(x) \leqslant f_k(x_0)$；并且对于某个 $i \in \{1, \dots, N\}$，有 $f_i(x) < f_i(x_0)$。最常见的方法之一

2020 *Mathematics Subject Classification.* 主要分类：90C29，46N10；次要分类：90C25。
Key words and phrases. 多目标优化，解映射的连续性，凸函数的凸组合，小直径性质。作者的研究部分由 GNAMPA（INdAM——意大利高等数学研究所）资助。C.A. De Bernardi、E. Miglierina 和 E. Molho 的研究部分由西班牙科学与创新部、西班牙国家研究署通过项目 PID2020-112491GB-I00 资助。

是寻找帕累托最优点的所谓线性标量化，即求解一族标量优化问题，其中目标函数由函数 $f_1, \ldots, f_N$ 的凸组合给出（例如，参见 [9, 第4]章和 [8, 第5]章）。本文讨论多目标优化问题中一种特殊的线性标量化情形：在对函数 $f_1, \ldots, f_N$ 作出凸性假设后，每个帕累托最优点都可以完全表征为函数 $f_1, \ldots, f_N$ 的凸组合的最小点。具体而言，若函数 $f_1, \ldots, f_N$ 严格凸，则点 x_0 是函数 $f_1, \ldots, f_N$ 的帕累托最优点，当且仅当存在 $t = (t_1, \ldots, t_N) \in \Sigma_N := \{t \in [0,1]^N : \sum_{i=1}^N t_i = 1\}$，使得 x_0 是标量函数 $u_t(x) = \sum_{k=1}^N t_k f_k(x)$ 的最小点（参见 [2,引理1，第 27]页）。

线性标量化在多目标优化理论与实践中的重要性，促使我们研究函数凸组合的性质。此外，正如我们将要看到的，要获得具有[2], 精神的结果，一个关键工具将是这样一种几何性质：凸组合的最小点能够关于凸组合的系数连续地依赖于连续严格凸函数 f_1、……、 f_N。

Enflo 和 Beauzamy 在 [2] 中提出的方法，对于多目标优化领域而言具有深刻的原创性，并为解集的逼近提供了不同寻常的视角。尽管具有这些有趣的特点，论文 [2] 似乎并未在多目标优化文献中得到广泛关注。因此，我们希望研究这一结果，并尝试削弱其假设。 [2] 方法的关键在于研究具有目标函数 f_1、……、 f_N 的凸多目标优化问题的帕累托最优点，将这些点与 f_1、……、 f_N 的凸组合的最小点解对应起来。特别地，我们需要确保映射 $\varphi(t) = \operatorname{argmin}(u_t)$ 的连续性，其中 $t \in \Sigma_N$。此外，鉴于线性标量化过程，这类结果本身就很有意义。

本文将注意力限制在如下情形：函数 f_1，…， f_N 满足，对于每个 $k = 1$，…， N：

(a) f_k 是一个在有界集上有界的连续严格凸函数；(b) f_k 是强制的（即对于每个 $C \in \mathbb{R}$，集合 $\{x \in X : f(x) \leqslant C\}$ 是有界的）。
在这些假设下，很容易证明，只要 X 是有限维赋范空间，函数 φ 就是连续的。另一方面，当 X 是无限维空间时，并非总能证明映射 φ 的连续性。事实上，我们给出了定义在希尔伯特空间 ℓ_2 上的两个函数的一个例子：即使满足上述假设 (a)-(b)，相应的函数 φ 仍不连续（见例 3.7）。通过增添一种性质，可以得到 φ 的连续性。这种性质已在 [3] 和 [11] 中以不同术语引入，本文称之为小直径性质。第2节主要研究小直径性质。值得指出的是，在命题2.4中，我们证明了一个结果，可以将 x 的范数所具有的强暴露点提升到函数 $f : X \to \mathbb{R}$ 上，从而为寻找满足小直径性质的函数实例提供了有效工具。第3节将证明，如果

函数 f_1、……、f_N 满足小直径性质，并且还满足上述性质 (a)-(b)，那么 就是连续的。最后，第4节给出了针对 f_1、……、f_N 的帕累托最优点的“随机”逼近，该逼近不依赖于与每个帕累托最优点相关联的取值 t（定理 4.4）。正如注4.8所证明的，与 [2], 的结果相比，该定理适用范围更广；在注4.8中，我们利用在 [4], 上构造的特殊范数证明，存在一组函数 f_1、……、f_N，它们不满足 [2, 定理 1，第 28] 页中的假设，却满足我们结果所要求的条件。此外，一个例子表明，小直径性质这一假设不可省略。

2. Notation and preliminaries

本节介绍本文后续部分所需的一些概念和基本结果。本节分为两部分：前一部分研究凸集和函数的几何性质，后一部分简要介绍多目标优化。我们只考虑非平凡的实赋范线性空间。若 X 是一个赋范空间，其拓扑对偶为 X^*，则 B_X、U_X 和 S_X 分别是 X 的闭单位球、开单位球和单位球面。关于此处未作解释的符号和术语，请参见 [1], [7] 和 [10] 。

2.1. 凸集和函数的几何性质。设 D 是赋范空间 X 的一个凸子集。若存在线性泛函 $x^* \in X^* \setminus \{0\}$ 使得 $x^*(x) = \sup_D x^*$，则称点 $x \in D$ 为支撑点。若点 $x \in D$ 满足：x 由 $x^* \in X^*$ 支撑，且 $\{y \in D\colon\ x^*(y) = x^*(x)\} = \{x\}$，则称 $x \in D$ 为 D 的暴露点。若点 $x \in D$ 由 $x^* \in X^*$ 强暴露，即它由 x^* 支撑，且对于所有满足 $\lim_{n\to\infty} x^*(x_n) = \sup_D x^*$ 的序列 $\{x_n\} \subset D$，都有 $x_n \to x$ 。在这种情况下，对于任意 $\lambda > 0$，x 也由 λx^* 强暴露。此外，点 $x \in D$ 由 $x^* \in X^* \setminus \{0\}$ 强暴露，当且仅当 x^* 在 D 上有界，且当 $\delta \to 0$ 时，$\operatorname{diam} S(x^*, \delta, D) \to 0$，其中

$$S(x^*, \delta, D) := \{y \in D\colon x^*(s) > \sup_D x^* - \delta\}.$$

设 $f\colon D \to \mathbb{R}$ 为凸连续函数。根据

$$\operatorname{epi}(f) := \{(x,t) \in D \times \mathbb{R}\colon t \geqslant f(x)\}$$

我们将函数 f的上图集记为。再者， $\partial f(x)$表示 f在 x处的次微分。点$(x, f(x)) \in \operatorname{epi}(f)$由$(x^*, -1)$支撑，当且仅当 $x^* \in \partial f(x)$。最后，我们用 $\mathbb{R}^+$ 表示区间$(0, \infty)$。

定义 2.1. 设 D 是 X, $x \in D$的凸子集，且 $f\colon\ D \to \mathbb{R}$。

- 设 $x^* \in \partial f(x)$，若每个满足 $x^*(x_n) - f(x_n) \to x^*(x) - f(x)$的序列 $\{x_n\} \subset D$均范数收敛，则称 x 是 f 的一个 x^*-小直径点。
- 若对每个 $x^* \in \partial f(x)$，点 x都是 f的一个 x^*-小直径点，则称 x 是 f的一个小直径点。
- f 具有小直径性质（SDP），若每个 $y \in D$ 都是关于 f 的小直径点。

我们回顾以下结果，该结果将 X 中关于 f 的小直径点与 $X \times \mathbb{R}$ 中 epi(f) 的强暴露点联系起来。

引理 2.2（[3]）。设 D 是 X的凸子集， $f\colon D \to \mathbb{R}$ 是连续凸函数， $x \in X$，且 $x^* \in \partial f$（x）。那么，当且仅当（x， f（x））是 epi（f）上的强暴露点，且由（x^*， -1）强暴露时， x 才是 f 的 x^*-小直径点。

由此可推得，若 $f\colon X \to \mathbb{R}$ 具有 SDP，则对每个 $x \in X$，当且仅当（x， f（x））在 epi（f）中由（x^*， -1）强暴露时，（x， f（x））才由（x^*， -1）支撑。我们将证明一个关于 SDP 的初等事实；下一节将用到这一结果（参见 [11, 定理 4.2] 的证明）。

命题 2.3。设 f， $g\colon X \to \mathbb{R}$ 是连续凸函数。若 f具有 SDP，则 $f+g$ 具有 SDP。

证明。设 $x_0 \in X$ 和 $x^* \in \partial(f+g)(x_0)$。我们将证明 x_0 是 $f+g$的一个 x^*-小直径点。事实上，由于 $f+g$ 是一个凸函数，且$x^* \in \partial(f+g)(x_0)$，因此我们有

$$x^*(y) - (f+g)(y) \leqslant x^*(x_0) - (f+g)(x_0),$$

对任意 $y \in X$均成立。现在，设 $\{x_n\} \subset X$ 为一个序列，使得 $x^*(x_n) - (f+g)(x_n)$收敛到 $x^*(x_0) - (f+g)(x_0)$。根据 [10, 定理 3.23]，存在 $x_1^* \in \partial f(x_0)$ 和 $x_2^* \in \partial g(x_0)$，使得 $x^* = x_1^* + x_2^*$。固定 $\varepsilon > 0$，则最终有

$$\begin{aligned}
\varepsilon &> x^*(x_0) - (f+g)(x_0) - x^*(x_n) + (f+g)(x_n) \\
&= x_1^*(x_0) - f(x_0) - x_1^*(x_n) + f(x_n) + x_2^*(x_0) - g(x_0) - x_2^*(x_n) + g(x_n) \\
&\geqslant x_1^*(x_0) - f(x_0) - x_1^*(x_n) + f(x_n) \geqslant 0.
\end{aligned}$$

由于 f 具有 SDP，结合前述不等式链以及 $\varepsilon > 0$的任意性，可知 $\{x_n\}$ 收敛到 x_0。因此， $f+g$ 具有 SDP。□

下一个结果表明，在适当的假设下，如何将赋范空间 X 的范数中强暴露点提升到函数 $f\colon X \to \mathbb{R}$ 的图上。它为构造具有 SDP 的函数提供了一个很好的工具。

命题2.4. 设 $(X, \|\cdot\|)$ 是赋范空间， $x_0 \in S_X$，并假设 $x_0 \in B_X$ 被 $x^* \in S_{X^*}$ 强暴露。设 $h\colon [0,\infty) \to \mathbb{R}$ 是连续凸严格递增函数，并设 $\lambda \in \mathbb{R}^+$ 和 $\lambda^* \in \mathbb{R}$，假设 λ 是 λ^*-小直径点，属于 h。那么：

(i) $h(\|\cdot\|)$ 是 X 上的连续凸函数；(ii) λx_0 是 $h(\|\cdot\|)$ 的 $\lambda^* x^*$-小直径点；
(iii) X 的原点是 $h(\|\cdot\|)$ 的 0-小直径点，其中 $0 \in X^*$ 表示零泛函。

证明。第一点众所周知，且易于证明。我们来证明 (ii)。显然，由于 h 严格递增， λ^* 为正。因此，对于每个 $x \in X$，我们有

$$\lambda^* x^*(x) - h(\|x\|) \leqslant \lambda^*(x^*(x) - \|x\|) + \lambda^*\lambda - h(\lambda) \leqslant \lambda^* x^*(\lambda x_0) - h(\|\lambda x_0\|), \quad (1)$$

其中第一个不等式成立，是因为 $\lambda^* \in \partial h(\lambda)$。现在，我们来考虑一个序列 $\{x_n\} \subset X$，并假设

$$\lambda^* x^*(x_n) - h(\|x_n\|) \to \lambda^* x^*(\lambda x_0) - h(\|\lambda x_0\|).$$

由于对于 $n \in \mathbb{N}$,

$$\lambda^* x^*(x_n) - h(\|x_n\|) \leqslant \lambda^* \|x_n\| - h(\|x_n\|) \leqslant \lambda^* \lambda - h(\lambda),$$

我们有 $\|x_n\| \to \lambda$，特别是最终有 $\|x_n\| > 0$。此外，由 (1) 可得： $\lambda^*(x^*(x_n) - \|x_n\|) \to 0$，因此 $x^*\left(\frac{x_n}{\|x_n\|}\right) \to 1$。由于 $x_0 \in B_X$ 被 $x^* \in S_{X^*}$ 强暴露，我们有 $\frac{x_n}{\|x_n\|} \to x_0$，因此 $x_n \to \lambda x_0$。(ii) 的证明完成。

为了证明 (iii)，不失一般性，我们可以假设 $h(0) = 0$。然后，只需注意到，由 h 的反函数的连续性可知，若 $h(\|x_n\|) \to h(\|0\|) = 0$，则 $\|x_n\| \to 0$。□

若函数 $f\colon X \to \mathbb{R}$ 在凸集 $D \subset X$ 上满足严格凸性，则称其为严格凸函数；也就是说，对任意 x、$y \in D$，以下关系成立： $f(\frac{x+y}{2}) < \frac{1}{2}(f(x) + f(y))$。现在我们回顾一致凸函数这一更强的概念。

定义 2.5. 设 D 是赋范空间 X 的凸子集， $x_0 \in D$ 和 $f\colon D \to \mathbb{R}$。则称

- f 在 x_0 处一致凸，是指对于每个 $\varepsilon > 0$，都存在 $\delta > 0$，使得对于满足 $\|x_0 - y\| \geqslant \varepsilon$ 的每个 $y \in D$，都有
$$f\left(\frac{x_0 + y}{2}\right) \leqslant \frac{1}{2} f(x_0) + \frac{1}{2} f(y) - \delta.$$
- f 在 D 上一致凸，是指对于每个 $\varepsilon > 0$，都存在 $\delta > 0$，使得对于满足 $\|x - y\| \geqslant \varepsilon$ 的每个 x、$y \in D$，都有
$$f\left(\frac{x + y}{2}\right) \leqslant \frac{1}{2} f(x) + \frac{1}{2} f(y) - \delta.$$

注 2.6. 以下蕴含关系成立。

(i) 若函数 $f\colon X \to \mathbb{R}$ 在凸子集 $D \subset X$ 上一致凸，则对于每个 $x \in D$，它在 x 处一致凸。(ii) 若 $f\colon X \to \mathbb{R}$ 对每个 $x \in X$ 都在 x 处一致凸，则 f 具有 SDP。(iii) 若 $f\colon X \to \mathbb{R}$ 具有 SDP，则它严格凸。

证明。显然，（i）中包含的蕴含关系成立，下面证明（ii）。反设 f 不具有 SDP，则存在 $x \in X$、 $x^* \in \partial f(x)$、 $\varepsilon > 0$和一个序列 $\{x_n\} \subset X$，使得当 $n \to +\infty$时，$x^*(x_n) - f(x_n)$趋于 $x^*(x) - f(x)$，但只要 $n \in \mathbb{N}$，就有 $\|x_n - x\| \geqslant \varepsilon$。由于 f 在 x中一致凸，存在$\delta > 0$ 使得

$$f\left(\frac{x + x_n}{2}\right) \leqslant \frac{1}{2} f(x) + \frac{1}{2} f(x_n) - \delta,$$

对每个 $n \in \mathbb{N}$均成立。由此可得

$$2\delta + x^*(x) - f(x) + x^*(x_n) - f(x_n) \leqslant 2x^*\left(\frac{x+x_n}{2}\right) - 2f\left(\frac{x+x_n}{2}\right).$$

因此，当 n 足够大时，我们得到

$$\delta + 2[x^*(x) - f(x)] \leqslant 2\left[x^*\left(\frac{x+x_n}{2}\right) - f\left(\frac{x+x_n}{2}\right)\right].$$

最后一个不等式与点 $(x, f(x)) \in \mathrm{epi}(f)$ 由泛函 $(x^*, -1)$ 支撑这一事实相矛盾。

最后，让我们证明 (iii)。如果 $f\colon X \to \mathbb{R}$ 不是严格凸的，则其图像包含一条长度为 $\ell > 0$ 的线段，因此可以找到点 $x \in X$ 和 $x^* \in X^*$，使得 $\mathrm{diam}(S(y^*, \delta, \mathrm{epi}(f))) \geqslant \ell$，其中 $y^* = (x^*, -1)$。因此，f 不具有 SDP。□

一般而言，前述备注中包含的蕴含关系均不可逆（这一点对于 (i) 和 (ii) 很容易看出，而下面的例 3.7 给出了一个不具有 SDP 的严格凸函数示例）。另一方面，下一个命题（其证明留给读者）表明，在紧集上一致凸性与严格凸性是一致的。

命题 2.7. 设 K 是赋范空间 X 的紧凸子集。设 $f\colon K \to \mathbb{R}$ 是连续且严格凸的函数。则 f 在 K 上一致凸。

2.2. 凸多目标优化问题。多目标优化是优化理论的一个研究主题，在经济学、运筹学和工程学中有许多应用。这里我们仅简要介绍解的概念和线性标量化过程。关于这一主题的完整综述，请读者参阅如今已成为经典的专著 [9] 和 [8]。此外，我们也引用 [6], ，尽管该书仅讨论有限维情形。虽然本小节提及的结果很容易证明，但在上述著作中都能找到这些结果的证明，而且通常是在更一般的框架下给出的。

设 $f_1, \ldots, f_N$ 是定义在赋范线性空间上的 N 个实值函数 X.

定义 2.8. 当不存在满足以下条件的 $x \in X$ 时，点 $x_0 \in X$ 是 $f_1, \ldots, f_N$ 的帕累托最优点：对于每个 $k = 1, \ldots, N$，$f_k(x) \leqslant f_k(x_0)$，且对于某个 $i \in \{1, \ldots, N\}$，$f_i(x) < f_i(x_0)$。

多目标优化问题的一种较弱解概念是弱帕累托最优点。

定义 2.9. 当不存在满足以下条件的 $x \in X$ 时，点 $x_0 \in X$ 是 $f_1, \ldots, f_N$ 的弱帕累托最优点：对于每个 $k = 1, \ldots, N$，$f_k(x) < f_k(x_0)$。

当然，每个帕累托最优点也是弱帕累托最优点，但通常反之并不成立。然而，在适当的凸性假设下，这两个概念是一致的。具体而言，有如下结果。

命题 2.10。设 f_1，...，f_N 为严格凸函数，则对于 f_1，...，f_N 的每个弱帕累托最优点，也是 f_1，...，f_N 的帕累托最优点。

多目标优化问题旨在寻找目标函数 f_1、...、f_N 的（弱）帕累托最优点。显然，一般来说，帕累托最优点并不唯一。求解多目标优化问题的一种常见方法，是将其化为一族参数化的标量优化问题（这一过程称为多目标优化问题的标量化）。文献中已经发展出许多不同的标量化方法（参见上文引用的专著及其中的参考文献）。其中，最常用的方法之一是线性标量化或加权和标量化。记 Σ_N 为（$N-1$）维单纯形，即 $\Sigma_N = \{t \in [0,1]^N : \sum_{i=1}^N t_i = 1\}$；对于 $t =$（t_1，...，t_N）$\in \Sigma_N$，令 u_t（x）$= \sum_{i=1}^N t_i f_i$（x）。我们有以下充分最优性条件。

~命题 2.11。设 $x_0 \in X$。如果存在 $t \in \Sigma_N$，使得对于每个 $x \in X$都有 $u_{\tilde{t}}(x_0) \leqslant u_{\tilde{t}}(x)$，那么 x_0 是关于 f_1、...、f_N 的弱帕累托最优点。

如果对目标函数添加凸性假设，则上述充分条件也成为必要条件。

命题 2.12。设 $x_0 \in X$ 和 f_1、...、f_N 为凸函数。如果 x_0 是关于 f_1、...、f_N 的~弱帕累托最优点，那么存在 $t \in \Sigma_N$，使得对于每个 $x \in X$都有$u_{\tilde{t}}(x_0) \leqslant u_{\tilde{t}}(x)$。

在本文后续内容中，我们始终考虑严格凸的目标函数 f_1、...、f_N。因此，结合命题 2.10、2.11 和 2.12，我们得到以下备注；该备注在我们的设定中发挥着关键作用。

注记 2.13. 设有一个多目标优化问题，其目标函数 $f_1, \ldots, f_N$ 均为严格凸函数。那么，对于 $f_1, \ldots, f_N$，点 $x_0 \in X$ 是帕累托最优点，当且仅当存在某个 $t \in \Sigma_N$，使得 x_0 是函数 $u_t(x)$ 的最小点。

3. Continuity of 极小值点 map for a convex combinations of functions

本节旨在探讨一些条件，以确保连续严格凸函数 $f_1, \ldots, f_N$ 的凸组合的最小点连续依赖于该凸组合的系数。为了保证最小值存在，我们将讨论范围限定在自反空间内。

定义 3.1. 设 X 是一个自反巴拿赫空间。若函数 $f\colon X \to \mathbb{R}$ 满足以下条件，则称其为 U-函数：
(a) f 是一个在有界集上有界的连续严格凸函数。(b) f 是强制的（即，对于每个 $C \in \mathbb{R}$，集合 $\{x \in X\colon f(x) \leqslant C\}$ 都是有界的）。

注意到，U-函数有唯一的极小点。事实上，设 f 是一个U-函数，并取 $C \in \mathbb{R}$ 使得集合 $S_C = \{x \in X\colon f(x) \leqslant C\}$非空。由于 f 连续且凸， S_C 是闭凸集，因此

它是 X的弱闭子集，因此 S_C 是弱紧的。由于 f 是凸函数，它是弱下半连续的，因此在S_C中存在一个最小点。由 f的严格凸性可得其唯一性。

设 $f_1, \ldots, f_N$ 是一个有限的 U-函数族。对于

$$t = (t_1, \ldots, t_N) \in \Sigma_N,$$

我们回顾一下，$u_t\colon X \to \mathbb{R}$ 是由 $u_t = \sum_{i=1}^N t_i f_i$ 定义的函数。由于对于每个 $t \in \Sigma_N$，u_t 都是一个 U-函数，我们可以考虑由下式定义的映射 $\varphi\colon \Sigma_N \to X$

$$\varphi(t) := \operatorname{argmin}(u_t).$$

需要强调的是，映射 φ 隐式地依赖于函数 $f_1, \ldots, f_N$ 的选取。

注 3.2. 对于上述 U-函数族 $f_1, \ldots, f_N$，考虑集合 $\mathcal{M} = \varphi(\Sigma_N)$，它包含函数 $u_t (t \in \Sigma_N)$ 的所有极小点。按照 [2, 第 28], 页中的方法，不难看出集合 $\mathcal{M}$ 是有界的。

注 3.3（U-函数的多目标优化）。当我们考虑目标函数为 U-函数的凸多目标优化问题时，由注记 2.13 可知，集合 $\mathcal{M} = \varphi(\Sigma_N)$ 是函数 $f_1, \ldots, f_N$ 的所有帕累托最优点构成的集合。

如果 X 是有限维的，那么对于任意选取的 U-函数 f_1、…、f_N，通过标准的紧致性论证，可以证明函数 $\varphi\colon \Sigma_N \to X$ 是连续的（这里，Σ_N 赋予标准乘积拓扑）。需要指出的是，在无限维情形下，这通常并不成立：下面的例 3.7 给出了两个 U-函数 f_1、f_2，使得相应的函数 φ 不连续。本节的主要目的是给出适当的几何假设，以确保无限维情形下函数 φ 的连续性（见下文定理 3.5）。

我们先证明映射 φ 在 Σ_N 的给定点处的连续性。下文中，对于正实数 α 和 U-函数 f，我们用 $S_\alpha(f)$ 表示 f 在水平值 $(\min f + \alpha)$ 处的严格次水平集，即，

$$S_\alpha(f) = \{x \in X : f(x) < \min(f) + \alpha\}.$$

引理 3.4. 设 $f_1, \ldots, f_N$ 是一族有限的 U-函数，且 $\hat{}\, t \in \Sigma_N$。假设当 α 趋于零时，$\operatorname{diam}(S_\alpha(u_{\hat{t}}))$ 也趋于零。那么映射 $\varphi\colon \Sigma_N \to X$ 在 t 处是 $\hat{}$ 连续的。

证明。记 $x_0 = \varphi(\hat{t})$ =极小值点($u_{\hat{t}}$)。我们断言，如果序列 $\{v^n\} \subset \Sigma_N$ 收敛于 $\hat{}t$，那么函数序列 $\{u_{v^n}\}$ 在有界集上一致地收敛于$u_{\hat{t}}$ 。事实上，令 $D \subset X$ 为一个有界集，并记 $x \in D$,

$\hat{t}=(\hat{t}_1,\dots,\hat{t}_N)$ 和 $v^n=(v_1^n,\dots,v_N^n)$，则

$$\begin{aligned}|u_{\hat{t}}(x)-u_{v^n}(x)| &= \left|\sum_{i=1}^N \hat{t}_i f_i(x)-\sum_{i=1}^N v_i^n f_i(x)\right| \\ &\leqslant \max_i \sup_{y\in D}|f_i(y)|\sum_{j=1}^N|\hat{t}_j-v_j^n|.\end{aligned}$$

对 $x\in D$ 取上确界，并令 n 趋于无穷大，即得所述结论。

现在假设函数 $\varphi\colon \Sigma_N\to X$ 在 $\hat{t}$ 处不连续。那么存在一个收敛到 $\hat{t}$ 的序列 $\{t_n\}\subset\Sigma_N$，以及一个正实数 $\xi>0$，使得 $\|x_n-x_0\|>\xi$；其中，对每个 $n\in\mathbb{N}$，$x_n:=\varphi(t_n)$。由于 $\operatorname{diam}(S_\alpha(u_{\hat{t}}))$ 趋于零，存在 $\alpha>0$ 使得 $\operatorname{diam}(S_\alpha(u_{\hat{t}}))<\xi/2$。由注 3.2，集合 $\mathcal{M}$ 有界。由于 $\{u_{t_n}\}$ 在 $\mathcal{M}$ 上一致收敛到 $u_{\hat{t}}$，存在 $n\in\mathbb{N}$ 使得 $|u_{t_n}(x_0)-u_{\hat{t}}(x_0)|<\alpha/3$，并且 $|u_{t_n}(x_n)-u_{\hat{t}}(x_n)|<\alpha/3$。由于 $x_0\in S_\alpha(u_{\hat{t}})$ 且 $\|x_n-x_0\|>\xi$，我们有 $u_{\hat{t}}(x_n)\geqslant\alpha+u_{\hat{t}}(x_0)$。因此，我们得到

$$u_{t_n}(x_n)=u_{t_n}(x_n)-u_{\hat{t}}(x_n)+u_{\hat{t}}(x_n)>2\alpha/3+u_{\hat{t}}(x_0)>\alpha/3+u_{\hat{t}}(x_0)>u_{t_n}(x_0).$$

这与 $x_n=\varphi(t_n)=$ 为 u_{t_n}的极小值点这一事实矛盾。 □

定理 3.5. 设 $f_1,\dots,f_N$是一个由 U-函数组成的有限族。以下断言成立

(i) 若 f_i对某个 $i\in\{1,\dots,N\}$具有 SDP，则映射 $\varphi\colon \Sigma_N\to X$在 $\{t=(t_1,\dots,t_n)\in\Sigma_N\colon t_i\neq 0\}$上连续。(ii) 若 f_i对每个 $i=1,\dots,N$都具有 SDP，则映射 $\varphi\colon \Sigma_N\to X$在 Σ_N上连续。

证明. 我们证明 (i)，(ii) 可直接得出。设 $t=(t_1,\dots,t_n)\in\Sigma_N$满足 $t_i\neq 0$，并证明 φ 在 t处连续。由命题 2.3，函数 u_t 是满足 SDP 的 U- 函数。若 $x_t=\varphi(t)$，由引理 2.2，$(x_t,u_t(x_t))$ 被 $(0,-1)\in X^*\times\mathbb{R}$强暴露。因此，当 α 趋于零时，$\operatorname{diam}(S_\alpha(u_t))$也趋于零。由引理 3.4 即得所述结论。□

注3.6. 指出，若连续凸函数 $f\colon X\to\mathbb{R}$具有 SDP，则根据 [11, 命题 3.13], ，它是强制的。因此，在这一附加假设下，定义 3.1 中的条件（b）是多余的。

以下示例表明，映射 $\varphi\colon \Sigma_N\to X$ 并非总是连续的，即使 $f_1,\dots,f_N$都是 U -函数。因此，该示例指出了 SDP 在定理 3.5 中发挥的关键作用。

例 3.7. 设 $X=\ell_2$, $h_1,h_2\colon X\to\mathbb{R}$ 和 $g\colon \mathbb{R}\to\mathbb{R}$ 定义如下:

$$h_1(x)=\sum_{n=2}^{+\infty}\frac{(x_n-x_1)^2}{2^n},, \qquad x=(x_1,x_2,x_3,\dots)\in\ell_2,$$

$$h_2(x)=\sum_{n=2}^{+\infty}\frac{(x_n+x_1-1)^2}{2^n}, \qquad x=(x_1,x_2,x_3,\dots)\in\ell_2,$$

$$g(y)=\begin{cases}(y-2)^2 & \text{if } y\geqslant 2,\\ 0 & \text{if } 0\leqslant y<2.\end{cases}$$

然后定义

$$f_1(x)=h_1(x)+g(\|x\|),\quad f_2(x)=h_2(x)+g(\|x\|)\qquad (x\in\ell_2).$$

那么， f_1 和 f_2 是定义在 ℓ_2 上的实值 U-函数，使得：
(1) f_1 在 $x=0$处取得其最小值；(2) 对每个 $t\in(0,1)$，函数 $u_t:=(1-t)f_1+tf_2$ 在 $2U_X$ 之外取得其最小值（即 $\|\varphi(t)\|\geqslant 2$）。

特别地，映射 $\varphi\colon\ \Sigma_2\to X$ 在$(t_1,t_2)=(1,0)\in\Sigma_2$处不连续。

证明。我们先证明 f_1 和 f_2 是 U-函数。由于对每个 $x=(x_1,x_2,x_3,\dots)\in\ell_2$，我们有

$$h_2\left((x_1,x_2,x_3,\dots)\right)=h_1\left((1-x_1,x_2,x_3,\dots)\right),$$

只需对 f_1证明即可。

f_1 的连续性显然，而定义 3.1 中的条件（b）由函数 g的定义可得。我们断言 h_1 在有界集上有界。事实上，若 D 是一个有界集且 $x\in D$，则有

$$|h_1(x)|=\left|\sum_{n=2}^{+\infty}\frac{(x_n-x_1)^2}{2^n}\right|\leqslant\sup_{n\in\mathbb{N}}|x_n-x_1|^2\sum_{n=2}^{+\infty}\frac{1}{2^n},$$

由此，结论得证。根据我们的结论和构造， f_1 在有界集上有界。接下来只需证明 f_1是严格凸的。由于 $g(\|\cdot\|)$是凸函数，只需证明 h_1 是严格凸的。设 x、 $y\in\ell_2$ 互不相同，$\lambda\in(0,1)$，下面证明

$$\theta_1:=h_1(\lambda x+(1-\lambda)y)-[\lambda h_1(x)+(1-\lambda)h_1(y)]<0.$$

为此，考虑函数 k_1： $\ell_2\times\ell_2\to\mathbb{R}$，其定义如下：

$$k_1(x,y)=\sum_{n=2}^{+\infty}\frac{(x_n-x_1)(y_n-y_1)}{2^n},\quad x=(x_1,x_2,\dots),y=(y_1,y_2,\dots)\in\ell_2,$$

并注意到

(i) $h_1(x)>0$ if and only if $x\neq 0$;
(ii) $h_1(x\pm y)=h_1(x)+h_1(y)\pm 2k_1(x,y)$.

因此，利用 (i) 和 (ii)，我们得到

$$\begin{aligned}\theta_1&=\sum_{n=2}^{\infty}\frac{(\lambda x_n-\lambda x_1+(1-\lambda)y_n-(1-\lambda)y_1)^2}{2^n}-\lambda h_1(x)-(1-\lambda)h_1(y)\\&=\lambda^2h_1(x)+(1-\lambda)^2h_1(y)+2\lambda(1-\lambda)k_1(x,y)-\lambda h_1(x)-(1-\lambda)h_1(y)\\&=\lambda(\lambda-1)h_1(x)-(1-\lambda)(\lambda)h_1(y)+2\lambda(1-\lambda)k_1(x,y)\\&=-\lambda(1-\lambda)h_1(x-y)<0.\end{aligned}$$

由此证明了 f_1, f_2 是 U -函数这一事实。

现在，条件 (1) 显然成立，下面证明 (2)。固定 $t \in (0,1)$，我们断言函数 $\widetilde{u}_t = (1-t)h_1 + th_2$ 不具有最小点。通过标准计算容易看出，定义为 $x_t^* \in \ell_2$ 的泛函

$$\left(\sum_{n=2}^{+\infty} \frac{(x_1 - x_n)(1-t) + (x_n + x_1 - 1)t}{2^{n-1}}, \dots, \frac{(1-t)(x_n - x_1) + t(x_n + x_1 - 1)}{2^{n-1}}, \dots \right)$$

是函数 $\widetilde{u}_t$ 在点 $x = (x_1, x_2, \dots)$ 处的导数。根据 [10, 命题 1.26]，$\widetilde{u}_t$ 在 x 处有最小值，当且仅当 $x_t^* = 0$。反过来假设$x_t^* = 0$，那么对于每个 $n \geqslant 2$，以下条件都必须成立

$$\frac{(1-t)(x_n - x_1) + t(x_n + x_1 - 1)}{2^{n-1}} = 0.$$

由此得到 $x_n = x_1 - 2tx_1 + t$。将其代入第一个坐标，我们得到 t

$$\sum_{n=2}^{+\infty} \frac{(1-t)(x_1 - x_1 + 2tx_1 - t) + t(x_1 - 2tx_1 + t + x_1 - 1)}{2^{n-1}},$$

当且仅当级数中的每一项都等于零时，该式才为零。因此可得

$$\begin{aligned} &(1-t)(2tx_1 - t) + (2x_1 - 2tx_1 + t - 1) = 0 \\ &4tx_1 - 4t^2 x_1 + 2t^2 - 2t = 0 \\ &x_1 = \frac{1}{2}. \end{aligned}$$

这意味着对于每个 $n \in \mathbb{N}$，都有 $x_n = \frac{1}{2}$。由于 $x \in \ell_2$，我们得出矛盾，断言得证。

现在，反过来假设 $\|\varphi(t)\| < 2$。那么，由于 $u_t|_{2U_X} = \widetilde{u}_t|_{2U_X}$，我们得到 $\widetilde{u}_t$ 在 $\varphi(t)$ 处取得其最小值，这与我们的断言矛盾。证明至此完成。□

我们以指出以下结论来结束本节：借助定理 3.5，我们可以获得关于以 f_1、.....、f_N 为目标函数的多目标优化问题的帕累托最优点集的拓扑性质的一些信息。这类结果在多目标优化领域已得到广泛研究（例如参见 [9, 第 6] 章和 [6, 第 3.4] 节）。这里我们仅引用一个由注 3.3 和定理 3.5 直接推出的推论。

推论 3.8。设 f_1，.....，f_N 是 U-函数的有限族。若对于每个 $i = 1$，.....，N，f_i 都具有 SDP，则 f_1，.....，f_N 的帕累托最优点集是 X 的一个紧致连通子集。

4. Approximation of the solutions for a convex multiobjective optimization problem

本节旨在借鉴论文 [2] 的思路，为凸多目标优化问题的解建立一个逼近结果。我们首先回顾 [2] 中证明的主要定理。

定理 4.1. 设 X 是一个一致凸巴拿赫空间。设 $f_1,\dots,f_N$ 是一个有限族 U-函数，且还满足以下条件：
(a) 对每个 $i\in\{1,\dots,N\}$，f_i 在 X 的每一点均为盖托可微。(b) 对每个凸有界集 $D\subset X$，存在常数 C_D 和 $q\geqslant 2$，使得对于每个 x_1，$x_2\in D$

$$f_i\left(\frac{x_1+x_2}{2}\right)\leqslant\frac{1}{2}(f_i(x_1)+f_i(x_2))-C_D\|x_1-x_2\|^q,$$

对每个 $i\in\{1,\dots,N\}$ 均成立。

那么，存在 K、$\gamma>0$，使得若 $t\in\Sigma_N$、$x_t=\varphi(t)$，且 $\{F_k\}$ 是定义在同一概率空间 (Ω,μ) 上、取值为 f_1、……、f_N，相应概率为 t_1、……、t_N 的独立随机变量序列，则对每个 $n\in\mathbb{N}$

$$\int_\Omega\left\|\operatorname{argmin}\left(\frac{1}{n}\sum_{k=1}^n F_k(\omega)\right)-x_t\right\|d\mu\leqslant\frac{K}{n^\gamma}. \tag{2}$$

注 4.2。从多目标优化的角度来看，这一结果可作如下解读。首先，x_t 是与权重 $(t_1,\dots,t_N)=t\in\Sigma_N$ 相关的函数 f_1、……、f_N 的帕累托最优点；也就是说，x_t 是函数 u_t 的唯一极小点（见注 3.3）。其次，随机变量 $\frac{1}{n}\sum_{k=1}^n F_k(\omega)$ 的每个实现都是从 f_1、……、f_N 中选出的 n 个函数的凸组合，其中每一项的系数均为 $1/n$。因此，对于固定的 $\omega\in\Omega$，$\operatorname{argmin}(\frac{1}{n}\sum_{k=1}^n F_k(\omega))$ 是函数 $F_1(\omega)$、……、$F_n(\omega)$ 的帕累托最优点，其中 $F_j(\omega)\in\{f_1,\dots,f_N\}$。考虑到不等式 (2) 中各量的上述含义，我们得出结论：将 x_t 近似为某个特殊帕累托最优点（即所有权重相等时对应的点）所产生的误差的期望值，受到 K/n^γ 的控制。这里所说的是一个凸多目标优化问题，其 n 个目标函数选自 f_1、……、f_N；对于每个 $j=1$、……、N，所选的 n 个函数中，每个函数以概率 t_j 成为 f_j。

定理 4.1 所给出的估计中最有趣的特点是，多目标问题的每个解（该问题具有 f_1，……，f_N 个目标函数）都可以通过考察从原多目标问题的目标函数中选取的 n 个函数的凸组合（系数相等）的最小点，以良好的方式得到逼近。此外，在原有 N 个目标函数中选取 n 个函数的大多数方式都能给出良好的逼近（正如 [2], 中所指出的，另请参阅本文引言）。事实上，平均而言，我们可以对误差给出良好的估计，该误差仅取决于所选函数的数量 n 以及它们的几何性质。最后值得指出的是，该估计不依赖于 $t\in\Sigma_N$；特别是，它不取决于帕累托最优点 x_t。

本节旨在介绍上述结果的一个新版本，其假设是（这显著弱化了上述定理中的条件 (b)）：所涉及的函数 $f_1,\dots,f_N$ 均为盖托可微的

且满足 SDP 的 U-函数。另一方面，我们证明，若去掉 SDP 假设，我们的新版本定理 4.1 将不再成立，见例 4.7。

注意到，定理 4.1 中的条件（b）意味着每个 f_i 在有界集上都是一致凸的。另一方面，在 [2] 中声称，可以用假设“每个 f_i 在有界集上都是一致凸的”来替代条件（b）。由于 [2], 中没有包含这一结论的证明，为方便读者，我们决定在稿件中加入相应的标准修改，补充一些预备结果；这些结果在 [2] 中用于证明定理 4.1。下面的技术引理表明，对于有限族一致凸函数，存在一个具有若干附加性质的公共凸性模。证明受到 [13]的启发。

引理 4.3. 设 $f_1, \ldots, f_N$是一族定义在巴拿赫空间 X的凸子集 D 上的实值连续函数。假设每个 f_i 在 D上一致凸。那么存在一个函数 $\delta\colon \mathbb{R}^+ \to \mathbb{R}^+$，使得

(i) 对每个 x、$y \in D$ 和 $i \in \{1, \ldots, N\}$

$$f_i(\frac{x+y}{2}) \leqslant \frac{1}{2}(f_i(x) + f_i(y)) - \delta(\|x-y\|)$$

成立。(ii) 函数 $t \mapsto {}^{\delta(t)}\!/_{t}$ 在 $\mathbb{R}^+$ 上严格递增。

(iii) $\lim_{t\to 0} {}^{\delta(t)}\!/_{t} = 0$。

证明。令 $i \in \{1, \ldots, N\}$。由于 f_i 在 D 上一致凸，结合 [13, 备注 2.6] 和 [12, 命题 3.5.1]，存在一个映射 $\delta_i\colon \mathbb{R}^+ \to \mathbb{R}^+$，满足：(a) 对每个 x，$y \in D$

$$f_i(\frac{x+y}{2}) \leqslant \frac{1}{2}(f_i(x) + f_i(y)) - \delta_i(\|x-y\|)$$

成立；

(b) 函数 $t \mapsto {}^{\delta_i(t)}\!/_{t^2}$ 在 $\mathbb{R}^+$ 上单调不减。

令 $s > t > 0$，则有

$$\frac{\delta_i(t)}{t^2} \leqslant \frac{\delta_i(s)}{s^2},$$

因此，可以得出

$$0 < \frac{\delta_i(t)}{t} \leqslant t\frac{\delta_i(s)}{s^2},$$

当 t 趋于零时，它也趋于零。同理，令 $s > t > 0$，则有

$$\frac{\delta_i(t)}{t} \leqslant t\frac{\delta_i(s)}{s^2} < \frac{\delta_i(s)}{s}.$$

因此，函数 ${}^{\delta_i(t)}\!/_{t}$ 严格递增。最后，定义

$$\delta(t) = \min\{\delta_1(t), \ldots, \delta_N(t)\},$$

我们得到所需的函数。 □

现在，我们可以陈述本节的主要结果了。

定理 4.4. 设 X 是一个一致凸巴拿赫空间。设 $f_1,\dots,f_N$ 是一组有限的 U-函数，并且还满足以下条件：
(a) 对每个 $i\in\{1,\dots,N\}$， f_i 在 X 的每一点都盖托可微。(b) 对每个 $i\in\{1,\dots,N\}$， f_i 都具有 SDP。

那么存在一个趋于 0 的正数序列 $\{\varepsilon_n\}$，使得若 $t=(t_1,\dots,t_N)\in\Sigma_N$，且 $\{F_k\}$ 是定义在同一概率空间 (Ω,μ) 上的一列独立随机变量，它们分别以概率 $t_1,\dots,t_N$ 取值 $f_1,\dots,f_N$，则对每个 $n\in\mathbb{N}$

$$\int_\Omega \|\operatorname{argmin}(\frac{1}{n}\sum_{k=1}^{n}F_k(\omega))-\varphi(t)\|d\mu\leqslant\varepsilon_n.$$

正如我们在注 4.2 中已经指出的，定理 4.4 也可解释为多目标优化问题的一个逼近结果。除了相较于定理 4.1 的假设更弱之外，值得指出的是，误差平均上界 ε_n 也可以显式计算。事实上，从定理 4.4 的证明中我们将看到

$$\varepsilon_n=\operatorname{diam}(\mathcal{M}_1)n^{\beta}+\eta(n^{\beta}),$$

其中，$\mathcal{M}_1$ 是 f_1的帕累托最优点集合的闭凸包，……，f_N；β 是一个负实常数，η 是一个当自变量趋于 0 时趋于 0，且取决于目标函数几何性质的函数。

为了证明我们的主要结果，我们回顾 [2]中两个技术性引理的表述。值得指出的是，以下结果中出现的常数 p 恰好是巴拿赫空间 X^* 的型（关于型和余型的定义，参见例如 [1, 第 137]页）。

引理 4.5. 设 X 是一致凸巴拿赫空间，A 是 X^*的一个有界子集。那么，存在 $p\in(1,2]$ 和 $C>0$，使得

$$\int_\Omega \|\frac{1}{n}\sum_{k=1}^{n}G_k-x^*\|d\mu\leqslant Cn^{\frac{1}{p}-1}$$

只要 $x^*\in X^*$，且 $\{G_k\}$ 是定义在同一概率空间（Ω μ）上的一列独立随机变量，其取值属于 A，并满足

$$\int_\Omega G_k d\mu=x^*,\qquad\qquad k\in\mathbb{N}.$$

下一个结果以略有不同的表述形式见于 [2]。为完整起见，我们在此给出证明。

引理 4.6. 设 $f_1,\dots,f_N$ 是定义在 X 上的一族实值 U-函数，且 $D\subset X$ 是一个凸子集。设 $t\in\Sigma_N$, $x_t=\varphi(t)$，并假设存在一个函数 $\delta\colon\mathbb{R}^+\to\mathbb{R}^+$，使得对每个 $x,y\in D$ 和 $i\in\{1,\dots,N\}$

$$f_i\left(\frac{x+y}{2}\right)\leqslant\frac{1}{2}(f_i(x)+f_i(y))-\delta(\|x-y\|)$$

成立。因此，对每个 $z \in D$，有

$$2\delta(\|x_t - z\|) < u_t(z) - u_t(x_t).$$

证明。只需注意到，对每个 $z \in D$，我们都有

$$u_t(x_t) < u_t\left(\frac{x_t + z}{2}\right) < \frac{1}{2}(u_t(x_t) + u_t(z)) - \delta(\|x_t - z\|),$$

由此即可得到该结论。 □

定理 4.4 的证明。考虑由函数 $\{f_1, \ldots, f_N\}$ 的凸组合的所有最小点构成的集合 $\mathcal{M}$，即 $\mathcal{M} = \varphi(\Sigma_N)$。由于 Σ_N 是一个紧集，且根据定理 3.5，函数 φ 连续，因此 $\mathcal{M}$ 是一个紧集。因此，$\mathcal{M}_1 \equiv \operatorname{conv}(\mathcal{M})$，即 $\mathcal{M}$ 的凸包的闭包，也为紧集（例如参见 [5, 定理 V.2.6]）。于是，根据命题 2.7，限制在集合 $\mathcal{M}_1$ 上的函数 $f_1, \ldots, f_N$ 都是一致凸的。因此，根据引理 4.3，存在一个函数 $\delta\colon \mathbb{R}^+ \to \mathbb{R}^+$，满足引理 4.3 中的条件 (i)-(iii)。若 $k \in \mathbb{N}, t \in \Sigma_N$，且 $x_t = \varphi(t)$，则在 (Ω, μ) 上定义一个取值于 X^* 的随机变量 G_k^t 如下：若 $\omega \in \Omega$ 且 $F_k(\omega) = f_i$，令 $G_k^t(\omega) = f_i'(x_t)$。则，$\{G_k^t\}_k$ 是一列取值于 X^* 的独立随机变量，其以概率 $t_1, \ldots, t_N$ 取值 $f_1'(x_t) \ldots, f_N'(x_t)$。此外，由于函数 $f_1, \ldots, f_N$ 在有界集上有界，容易看出 $\sup_{x \in \mathcal{M}_1} \max_i \|f_i'(x)\|$ 是有限的，因而存在一个有界集 $A \subset X^*$，使得当 $k \in \mathbb{N}$ 且 $t \in \Sigma_N$ 时，G_k^t 取值于 A。由于对于每个 $k \in \mathbb{N}$ 和 $t \in \Sigma_N$，都有 $\int_\Omega G_k^t d\mu = \sum_{i=1}^N t_i f_i'(x_t) = 0$，引理 4.5 保证存在常数 $p \in (1, 2]$ 和 $C > 0$，使得

$$\int_\Omega \|\frac{1}{n}\sum_{k=1}^n G_k^t\| d\mu \leqslant C n^{\frac{1}{p}-1},$$

每当 $t \in \Sigma_N$。因此

$$\mu(\Omega_1) \leqslant n^{\frac{1-p}{2p}},$$

其中 $\Omega_1 = \{\omega \in \Omega\colon \ \|\frac{1}{n}\sum_{k=1}^n G_k^t(\omega)\| \geqslant C n^{\frac{1-p}{2p}}\}$。令 $\omega \in \Omega \backslash \Omega_1$ 和 $n \in \mathbb{N}$，则对于每个 $k \leqslant n$，$G_k^t(\omega) = f_{i_k}'(x_t)$，其中某个 $i_k \in \{1, \ldots, N\}$。令 $x =$极小值点$(\frac{1}{n}\sum_{k=1}^n f_{i_k})$ 并

$$\psi(r) = \frac{1}{n}\sum_{k=1}^n f_{i_k}(x + r(x_t - x)).$$

映射 ψ' 在单位区间 $[0, 1]$ 上递增，并在 $r = 0$ 处取得最小值。根据引理 4.6，我们有

$$\begin{aligned} 2\delta(\|x - x_t\|) &\leqslant \psi(1) - \psi(0) = \int_0^1 \psi'(r) dr \leqslant |\psi'(1)| \\ &\leqslant \left\| \frac{1}{n}\sum_{k=1}^n f_{i_k}'(x_t) \right\| \|x - x_t\| \leqslant C n^{\frac{1-p}{2p}} \|x - x_t\|. \end{aligned}$$

定义 $h(r) = \delta(r)/r$ 和 $\eta(r) = h^{-1}(r)$，我们得到

$$\frac{\delta(\|x - x_t\|)}{\|x - x_t\|} \leqslant \frac{Cn^{\frac{1-p}{2p}}}{2},$$

由此可得

$$\|x - x_t\| \leqslant \eta\left(\frac{Cn^{\frac{1-p}{2p}}}{2}\right).$$

因此，

$$\int_{\Omega} \|\operatorname{argmin}\left(\frac{1}{n}\sum_{k=1}^{n} F_k(\omega)\right) - x_t\| d\mu \leqslant \int_{\Omega_1} \|\operatorname{argmin}\left(\frac{1}{n}\sum_{k=1}^{n} F_k(\omega)\right) - x_t\| d\mu$$
$$+ \int_{\Omega\setminus\Omega_1} \|\operatorname{argmin}\left(\frac{1}{n}\sum_{k=1}^{n} F_k(\omega)\right) - x_t\| d\mu \leqslant \operatorname{diam}(\mathcal{M}_1) n^{\frac{1-p}{2p}} + \eta(\frac{Cn^{\frac{1-p}{2p}}}{2}).$$

若对于 $n \in \mathbb{N}$，我们定义

$$\varepsilon_n = \operatorname{diam}(\mathcal{M}_1) n^{\frac{1-p}{2p}} + \eta(\frac{Cn^{\frac{1-p}{2p}}}{2}),$$

证明得证。 □

例 4.7。令 f_1 和 f_2 如例 3.7 中所定义。显然， f_1 和 f_2在 X上都是盖托可微的。若 $t = 0$，则 $\operatorname{argmin}(u_t) = \{0\}$，其中 $u_t = (1-t)f_1 + tf_2$。另一方面，若 $t \in (0, 1]$，，则 $\|\operatorname{argmin}(u_t)\| \geqslant 1$。设 $t \in (0, 1)$。令 (Ω, μ) 为一个概率空间。考虑满足以下条件的随机变量 $F_k : \Omega \to \{f_1, f_2\}$： $\mu(\{\omega \in \Omega : F_k(\omega) = f_1\}) = 1 - t$且 $\mu(\{\omega \in \Omega : F_k(\omega) = f_2\}) = t$。令

$$\Omega_1 = \left\{\omega \in \Omega : \operatorname{argmin}(\frac{1}{n}\sum_{k=1}^{n} F_k(\omega)) = \{0\}\right\}$$
$$= \{\omega \in \Omega : F_k(\omega) = f_1 \,\forall\, k = 1, \ldots, n\}.$$

那么，我们注意到 $\mu(\Omega_1) = (1-t)^n$。这意味着，

$$\int_{\Omega} \|\operatorname{argmin}\left(\frac{1}{n}\sum_{k=1}^{n} F_k(\omega)\right) - x_t\| d\mu \geqslant \int_{\Omega_1} \|\operatorname{argmin}\left(\frac{1}{n}\sum_{k=1}^{n} F_k(\omega)\right) - x_t\| d\mu$$
$$\geqslant \mu(\Omega_1) = (1-t)^n.$$

因此，无法得到一个不依赖于 t 的估计。

注4.8。需要指出的是，存在满足 SDP 但不满足定理 4.1 中条件（b）的盖托可微 U-函数。为说明这一点，请注意， [4] 中给出了一个例子：在每个无限维可分巴拿赫空间（以及每个自反空间）中，都存在一个平均局部一致凸、盖托可微的范数 $|\cdot|$，但它不是局部一致严格凸的。对于自反巴拿赫空间，根据 [4], ，一个范数是平均局部一致凸的，当且仅当每个 $x \in S_X$都被每个支撑 x的泛函强暴露。因此，命题2.4

意味着函数 $|\cdot|^2$ 是一个满足 SDP 但不满足定理 4.1 中条件（b）的盖托可微 U-函数。

上述构造表明，存在这样的情形：定理 4.4 的假设得到满足，但定理 4.1 无法应用。

References

[1] F. Albiac 和 N. Kalton, *Topics in Banach space theory*, 数学研究生教材，施普林格，纽约，2006。[2] B. Beauzamy 和 P. Enflo, *Théorème de point fixe et d'approximation*, （法文）《Arkiv för Matematik》23（1985），19–34。[3] J.B. Collier, *A class of strong differentiability spaces*, 《美国数学会会刊》53（1975），420–422。[4] C.A. De Bernardi 和 J. Somaglia, *Rotund Gâteaux smooth norms which are not locally uniformly rotund*，预印本。

[5] N. Dunford 和 J.T. Schwartz, *Linear operators. Part I. General theory*, 1958 年原版重印本，John Wiley & Sons, Inc.，纽约，1988。[6] M. Ehrgott, *Multicriteria optimization*。第二版。施普林格出版社，柏林，2005。[7] M. Fabian、P. Habala、P. H´ajek、V. Montesinos 和 V. Zizler, *Banach Space Theory.The basis for linear and nonlinear analysis*, CMS数学丛书/Ouvrages de Math´ematiques de la SMC，施普林格，纽约，2011。[8] J. Jahn, *Vector optimization*。理论、应用与拓展。施普林格出版社，柏林，2004。[9] –D.T. Luc。 *Theory of vector optimization*。经济学与数学系统讲义，319 1989 。施普林格出版社，柏林，R.R. Phelps，, [10] *Convex functions monotone operators and differentiability*，数学讲义 1364，施普林格出版社，1993。[11] M. Volle 和 C. Z˘alinescu, *Strongly adequate functions on Banach sapces*, 《凸分析杂志》20（2013），655–668。[12] C. Z˘alinescu, *Convex analysis in general vector spaces*，世界科学出版社，2002。[13] C. Z˘alinescu, *On uniformly convex functions*, 《数学分析与应用杂志》95（1983），344–374。

Dipartimento di Matematica per le Scienze economiche, finanziarie ed attuariali, Università Cattolica del Sacro Cuore, Via Necchi 9, 20123 Milano, Italy*Email address*: carloalberto.debernardi@unicatt.it*Email address*: carloalberto.debernardi@gmail.com

Dipartimento di Matematica per le Scienze economiche, finanziarie ed attuariali, Università Cattolica del Sacro Cuore, Via Necchi 9, 20123 Milano, Italy*Email address*: enrico.miglierina@unicatt.it

Dipartimento di Scienze economiche e Aziendali, Università degli Studi di Pavia, Via San Felice 5, 27100 Pavia, Italy*Email address*: elena.molho@unipv.it

Politecnico di Milano, Dipartimento di Matematica, Piazza Leonardo da Vinci 32, 20133 Milano , Italy. *Email address*: jacopo.somaglia@polimi.it