arXiv:2602.14847v1 [math.MG] 16 Feb 2026

凸体仿射最优位置的 John 型分解

Florian Grundbacher 和 Tomasz Kobos

2026年2月17日

摘要

凸几何中的许多经典问题都可以表述为在某些包含条件下的优化问题。其中可以说理解得最为透彻的例子，是包含于其他凸体之中的凸体的体积最大化；在此情形下，John 分解描述了——并在欧几里得情形下完全刻画了——最优位置。然而，对于许多其他此类问题，尚无已知的一般最优性条件。为解决这一问题，我们推广了 O. B. Ader 的方法，得到一种作为仿射最优包含链之必要条件的 John 型分解，即链 $rL_1 + c \subseteq K \subseteq RL_2 + d$，其针对凸体 $K, L_1, L_2 \subseteq \mathbb{R}^n$、平移向量 $c, d \in \mathbb{R}^n$ 以及实数 $r, R > 0$，使得比值 $\frac{R}{r}$ 不能通过线性变换 K 而减小。当涉及椭球时，我们再次得到最优性的充分性，并说明各种问题的最优性条件如何由我们的结果推出。我们的主要应用涉及 Banach–Mazur 距离：我们给出一般情形下的必要最优性条件，以及欧几里得情形下的完整刻画。最后，我们由这些最优性条件推导出若干推论，它们与到欧几里得球的 Banach–Mazur 距离有关。

1 引言与结果

$\mathbb{R}^n$ 中凸体几何的各个方面，都与用其他凸体对给定凸体进行最佳逼近的问题有关。这些问题可以针对不同类别的凸体来考察，也可以针对逼近所要优化的各种截然不同的性质来考察。最著名的例子之一，是用体积最大内接椭球逼近凸体。在这种情况下，所优化的凸体类别由椭球构成，任务是在满足包含约束的条件下使体积最大。本文的主要目标是提供一种通用工具，用于分析凸性领域各种逼近问题的最优解，并特别关注其在 Banach–Mazur 距离中的应用。

2020年 数学主题分类。主要分类 52A40；次要分类 46B07、46B20、52A21。关键词与短语。Banach–Mazur 距离、仿射最优逼近、Ader 分解、John 分解、均值椭球。

我们旨在开发的工具类型，主要受到 John 提出的以下著名结果的启发 [12]。该结果涉及上述以椭球进行体积极值逼近的问题，并根据与欧几里得单位球 $\mathbb{B}^n$ 的边界$\mathrm{bd}(K)$ 的凸体 $K \subseteq \mathbb{R}^n$ 的接触点，刻画了何时该单位球是最优解（即具有非空内部的紧凸集）。记 $\langle \cdot, \cdot \rangle$ 表示上的标准内积 $\mathbb{R}^n$。

定理 1.1 (John椭球定理)。 设 $K \subseteq \mathbb{R}^n$ 为满足 $\mathbb{B}^n \subseteq K$的凸体。那么，当且仅当存在接触点 $u^1, \ldots, u^N \in \mathrm{bd}(K) \cap \mathrm{bd}(\mathbb{B}^n)$ 以及权重 $\lambda_1, \ldots, \lambda_N > 0$，使得对任意 $x \in \mathbb{R}^n$， $\mathbb{B}^n$ 是包含于 K 中的唯一最大体积椭球，

$$\sum_{i=1}^{N} \lambda_i \langle x, u^i \rangle u^i = x \quad \text{and} \quad \sum_{i=1}^{N} \lambda_i u^i = 0.$$

在这种情况下， $\sum_{i=1}^{N} \lambda_i = n$ ，并且存在一种选择使得 $N \leq {}^{n(n+3)}/_{2}$。

等式 $x = \sum_{i=1}^{N} \lambda_i \langle x, u^i \rangle u^i$ 通常称为 *John* 分解。它是唯一需要的条件，也就是说，当凸体 $\sum_{i=1}^{N} \lambda_i u^i = 0$ 是关于原点对称时， K是，即 $-K := \{-x : \ x \in K\} = K$。

John椭球定理有许多深远的推论。可能其最为人熟知的推论是，任意凸体 K 都包含于其体积最大内接椭球的某个平移副本中 $\mathcal{E}$ ，当该椭球被放大至 $\rho = n$ 相对于集合 $\rho\mathcal{E} := \{\rho x : \ x \in \mathcal{E}\}$时。该放大因子在 $\rho = \sqrt{n}$ 当 K 关于中心对称时可以减小到平移 $K + t := \{x + t : \ x \in K\}$ ，且 $t \in \mathbb{R}^n$ 关于原点对称。值得注意的是，尽管John椭球定理刻画了体积极值逼近问题的最优解，但它还为另一种不同类型的逼近问题——通过位似副本进行的逼近——提供了强有力的信息。

后一类逼近构成了凸几何中的另一个经典研究领域，通常借助*Banach–Mazur* 距离这一概念来表述。对于两个凸体 K， $L \subseteq \mathbb{R}^n$ 其定义为

$$d_{BM}(K, L) := \inf\{\rho \geq 0 : K + u \subseteq T(L + v) \subseteq \rho(K + u)\}, \tag{1}$$

其中，下确界取遍所有作用于 $\mathbb{R}^n$ 上的可逆线性算子 T 以及所有平移向量 u， $v \in \mathbb{R}^n$。用这一术语来说，前面提到的 John椭球定理的推论表明，对于任意凸体 $K \subseteq \mathbb{R}^n$，有 $d_{BM}(K, \mathbb{B}^n) \leq n$；而对于中心对称的 K，有 $d_{BM}(K, \mathbb{B}^n) \leq \sqrt{n}$。

从历史上看，Banach–Mazur 距离主要是在泛函分析背景下的赋范空间中得到最广泛的研究。对于赋范空间 $X = (\mathbb{R}^n, \ \|\cdot\|_X)$、$Y = (\mathbb{R}^n, \ \|\cdot\|_Y)$，它们的 Banach–Mazur 距离是比较范数时能够达到的最小常数 $\rho \geq 0$

$$\|x\|_X \leq \|x\|_{Y'} \leq \rho \|x\|_X \quad \text{for all } x \in \mathbb{R}^n,$$

其中 $Y' = (\mathbb{R}^n, \|\cdot\|_{Y'})$ 是与 Y等距同构的任意赋范空间。换言之，Banach–Mazur 距离刻画了两个赋范空间与等距同构的接近程度。凸体之间 Banach–Mazur 距离的定义（1）是对这一概念的推广

。事实上，两个赋范空间之间的 Banach–Mazur 距离，等于其关于原点对称的单位球之间的 Banach–Mazur 距离；在这种情况下，式（1）中的平移向量 u, v 可以省略（即 $u = v = 0$）。因此，本文将始终使用凸体这一表述。关于 Banach–Mazur 距离的概述、它对泛函分析中诸多经典问题的重要性，以及它与 Banach 空间几何其他方面的联系，读者可参阅 Tomczak-Jaegermann 的经典专著 [23]。

尽管 Banach–Mazur 距离在理论上十分重要，但关于任意一对凸体之间距离的最优构型，我们几乎知之甚少。相反，许多关于 Banach–Mazur 距离的结果，都是通过研究其他问题中的最优构型间接获得的；上文关于凸体到欧几里得球的距离估计，以及近期针对一般凸体之间可能达到的最大 Banach–Mazur 距离所建立的近最优界（见 [5]），便是例证。虽然这种方法在渐近情形下往往足够，但通常无法为确定具体例子中的精确 Banach–Mazur 距离提供太多帮助。即便对于简单而熟悉的凸体，例如标准立方体和交叉多胞体，这一点依然成立；它们之间的精确距离也是直到最近才得以确定的（见 [14]）。正如前文所述，我们旨在进一步理解 Banach–Mazur 距离本身的最优构型，同时提供一种适用范围更广、也可用于其他问题的工具。

我们的研究始于 Ader 的一篇论文 [1]，发表于1938年，其重要性似乎在很大程度上未受到关注。或许有些出人意料的是，Ader 对于给定的中心对称凸体，完整刻画了与之达到 Banach–Mazur 距离的椭球 $\mathbb{R}^3$（下文的条件 (ii)），时间约早于 John的奠基性工作 [12]10 年。Ader 的刻画基于 Behrend 的一项更早的结果 [4],，后者以另一种较不易直接推广的方式处理了 $\mathbb{R}^2$ 的情形。根据历史记载 [15],，John 的工作受到其学生 Ader 的启发，但后来人们忽略了 Ader 本人并未研究体积极值椭球。本文作者在[10] 中指出，以现代视角重新审视 Ader 的结果，可以得到一种类似 John椭球定理的刻画，描述任意维数下与中心对称凸体达到 Banach–Mazur 距离的椭球。我们将在下文陈述该定理。记 $\mathcal{M}_s^n$ 为 $^{n}{}_{(n+1)\,2}$ 维实线性空间中的对称 $n \times n$ 矩阵。

定理 1.2. 设 $K \subseteq \mathbb{R}^n$ 为一个关于原点对称的凸体，且 $R \geq r > 0$ 为满足 $r\mathbb{B}^n \subseteq K \subseteq R\mathbb{B}^n$的实数。那么以下条件等价：

(i) $d_{BM}(K, \mathbb{B}^n) = \frac{R}{r}$.

(ii) 对任意矩阵 $A \in \mathcal{M}_s^n$，存在一个内接触点 $y \in \mathrm{bd}(K) \cap \mathrm{bd}(R\mathbb{B}^n)$ 以及一个外接触点 $z \in \mathrm{bd}(K) \cap \mathrm{bd}(r\mathbb{B}^n)$，使得

$$\left\langle \frac{y}{r}, A\left(\frac{y}{r}\right) \right\rangle \leq \left\langle \frac{z}{R}, A\left(\frac{z}{R}\right) \right\rangle.$$

(iii) 存在整数 N、$M \geq 1$，内接触点 y^1、…、$y^N \in \mathrm{bd}(K) \cap \mathrm{bd}(R\mathbb{B}^n)$，外接触点 z^1、…、$z^M \in \mathrm{bd}(K) \cap \mathrm{bd}(r\mathbb{B}^n)$，以及权重 λ_1、…、λ_N、μ_1、…、$\mu_M > 0$，使得对于任意 $x \in \mathbb{R}^n$,

$$\sum_{i=1}^{N} \lambda_i \langle x, y^i \rangle y^i = \sum_{j=1}^{M} \mu_j \langle x, z^j \rangle z^j.$$

在这种情况下， $r^2 \sum_{i=1}^{N} \lambda_i = R^2 \sum_{j=1}^{M} \mu_j$ ，并且存在一种选择使得 $N + M \leq {}^{n}{}_{(n+1)}{}_{2} + 1$。

为纪念 Ader 的工作，我们将类似于 (iii) 的分解称为*Ader 分解*。它也可以用矩阵语言等价地表示为

$$\sum_{i=1}^{N} \lambda_i y^i (y^i)^T = \sum_{j=1}^{M} \mu_i z^j (z^j)^T.$$

关于证明，我们请读者参阅 [10]，其中详细介绍了与这一刻画相关的历史，并给出了一些推论。这包括上界 $d_{BM}(K, \mathbb{B}^n) \leq \sqrt{n}$ [10, 推论 2.2]的另一种证明、其在 $n \leq 3$ [10, 定理 2.4]中的等号情形，以及一项归功于 Maurey、关于距离椭球唯一性的结果 [10, 定理 2.6]。后两项结果虽为专家所知，但在 [10]之前从未发表过其证明。我们的次要目标是基于上述定理的推广，将这些结果扩展到非对称情形。

鉴于对称情形下 Ader 分解与 John分解之间的直接对应关系，从关于体积极值逼近的结果中寻找可能的推广思路，似乎很自然。对于后者，已有多项结果将 John条件的必要性加以推广，其中最强的结果由 Gordon、Litvak、Meyer 和 Pajor 在 [9, 定理 3.5] 中给出。为陈述该定理并贯穿全文，凸体 $K \subseteq L$ 与 $y \in$ 的接触对 (y, a) 由一个接触点 $y \in \text{bd}(K) \cap \text{bd}(L)$ 和一个非零的 $a \in \mathbb{R}^n$ 构成，满足 $\langle a, y \rangle \geq \langle a, x \rangle$ ，对任意 $x \in L$ 均成立。因此，若 (y, a) 是一个接触对，那么对任意 $\rho > 0$， $(y, \rho a)$ 也同样是接触对。

定理 1.3 （一般情形下的 John 分解）。设 K、 $L \subseteq \mathbb{R}^n$ 为凸体，使得 $K \subseteq L$ 且 K 在包含于 L 的所有仿射变换中具有最大体积。则存在 K 与 L 的接触对 （y^1、 a^1）..... （y^N、 a^N） ，以及权重 λ_1、.....、 $\lambda_N > 0$ ，使得对于任意 $x \in \mathbb{R}^n$,

$$\sum_{i=1}^{N} \lambda_i \langle x, a^i \rangle y^i = x \quad and \quad \sum_{i=1}^{N} \lambda_i a^i = 0.$$

在此情形下， $\sum_{i=1}^{N} \lambda_i \langle y^i、\ a^i \rangle = n$ ，并且存在一种选择使得 $N \leq n\ (n+1)$ 。

如果允许对凸体 K 和 L 作适当平移，则可以在上述条件中加入 $\langle y^i,\ a^i \rangle = 1$ 对所有 $i = 1,\ .....,\ N$ 以及 $\sum_{i=1}^{N} \lambda_i y^i = 0$ （见 [9, 定理 3.8]） 。与 John椭球定理不同，这些条件一般不再保证最优性（见 [9, 例 5.7]） ，而只是必要条件。尽管如此，它们在证实 Grünbaum 关于所谓*Grünbaum 距离*最大可能值的一个古老猜想中起到了关键作用。Grünbaum 距离是 Banach–Mazur 距离的一种变体，允许负位似像；对于凸体 $K,\ L \subseteq \mathbb{R}^n$ 其定义为

$$d_G(K, L) = \inf\{|\rho| : K + u \subseteq T(L + v) \subseteq \rho(K + u)\},$$

其中，下确界仍取遍所有作用于 T 上的 $\mathbb{R}^n$ 可逆线性算子 u和 $v \in \mathbb{R}^n$平移向量。注意，当至少一个凸体中心对称时，Grünbaum 距离与 Banach–Mazur 距离一致。利用 [9, 定理 3.8]，Gordon、Litvak、Meyer 和 Pajor 证明了，对于所有凸体 $d_G(K, L) \leq n$ 均成立K。$L \subseteq \mathbb{R}^n$人们猜想，只有当 K或 L 为单纯形 [11], 时，这个不等式才取等号，不过目前仅有一些特殊情形得到验证 [11, 13]。这又是一个例子，说明针对实际问题的最优性条件，或许更适合用来获得清晰而完整的解答。

我们通过大幅推广定理 1.2中条件的必要性，达成本文的主要目标——为各种逼近问题（包括迄今讨论过的问题）提供最优性条件：将其推广至处于仿射最优位置的任意三个凸体。特别是，我们不要求任何凸体是欧几里得球，甚至不要求它们是中心对称的。更准确地说，对于凸体 K、L_1、$L_2 \subseteq \mathbb{R}^n$，向量 c、$d \in \mathbb{R}^n$，以及实数 r、$R > 0$，若包含链

$$rL_1 + c \subseteq K \subseteq RL_2 + d$$

称为 仿射最优，如果对于任意线性算子 T 作用于 $\mathbb{R}^n$，向量 c'、$d' \in \mathbb{R}^n$，以及实数 r'、$R' > 0$，满足 $r'L_1 + c' \subseteq T(K) \subseteq R'L_2 + d'$ 时，我们都有

$$\frac{R'}{r'} \geq \frac{R}{r}.$$

因此，当 $L_1 = L_2$时，仿射最优位置就是 Banach–Mazur 距离位置。虽然这类逼近不那么常见，但在各种情境中确实会自然出现。除了前面提到的 Grünbaum 距离（其中 $L_2 = -L_1$），其他著名例子还包括 Auerbach 引理（其中 L_1 而 L_2 分别是 $\mathbb{R}^n$中的标准交叉多胞体和标准立方体）以及 Dvoretzky–Rogers 引理（对应于 $L_1 = \mathbb{B}^n$ 和 L_2 均为立方体）。结果如下，其中我们用 $\mathrm{int}(X)$ 表示集合 $X \subseteq \mathbb{R}^n$的内部。

定理 1.4。设 $K, L_1, L_2 \subseteq \mathbb{R}^n$ 为凸体，且 $0 \in \mathrm{int}(L_1) \cap \mathrm{int}(L_2)$，设 $c, d \in \mathbb{R}^n$为向量，r, $R > 0$ 为实数，使得包含链 $rL_1 + c \subseteq K \subseteq RL_2 + d$为仿射最优。则满足以下条件：

(i) 对任意矩阵 $A \in \mathbb{R}^{n\times n}$ 及任意向量 $v, w \in \mathbb{R}^n$，存在 K 和 $rL_1 + c$ 的一个接触对(y, a)，以及 K 和 $RL_2 + d$ 的一个接触对(z, b)，使得

$$\langle a, Ay + v\rangle \leq \langle b, Az + w\rangle \quad and \quad \langle a, y - c\rangle = \langle b, z - d\rangle = 1.$$

(ii) 存在整数 $N, M \geq 1$、K和 $rL_1 + c$的接触对$(y^1, a^1) \ldots (y^N, a^N)$、$K$和 $RL_2 + d$的接触对$(z^1, b^1) \ldots, (z^M, b^M)$，以及权重 $\lambda_1, \ldots, \lambda_N, \mu_1, \ldots, \mu_M > 0$，使得对于任意 $x \in \mathbb{R}^n$，

$$\sum_{i=1}^{N} \lambda_i \langle x, a^i\rangle y^i = \sum_{j=1}^{M} \mu_j \langle x, b^j\rangle z^j \quad and \quad \sum_{i=1}^{N} \lambda_i a^i = \sum_{j=1}^{M} \mu_j b^j = 0.$$

在这种情况下，$\sum_{i=1}^{N} \lambda_i \langle y^i, a^i\rangle = \sum_{j=1}^{M} \mu_j \langle z^j, b^j\rangle$，并且存在一种选择，使得 $N+M \leq (n+1)^2$。

我们不知道是否还能额外施加某种涉及接触点 y^i，z^i 的加权和的条件，如同 [9, 定理 3.8] 中那样。此外，立方体与交叉多胞体在 $\mathbb{R}^n$ 中处于标准位置的例子表明，一般而言，不能指望该分解足以保证仿射最优性，即便纳入此类附加条件也是如此（在此情形下，由于关于原点的中心对称性，这些条件会自动满足）。尽管如此，虽然 Ader 分解的存在仅仅是一个必要条件，它确实为通常难以处理的仿射最优位置提供了一些启示。证明见第2节。其基于与定理1.2相同的基本思想，并结合一个简单的凸性结果，该结果推广了超平面分离定理，用以处理平移向量。

在第 3 节中，我们展示定理 1.4 在凸性领域中的若干逼近问题上的应用，这些问题并非都直接与仿射最优位置相关。由于 L_1 和L_2 彼此可能毫无关联，我们可以适当地选择它们，从而将问题化归为其他类型的问题，并得到针对这些问题的特定最优性条件。这些问题包括满足某些正则性条件时类似于定理 1.3 中的体积极值逼近、位似下的最优包含 [8]，以及仿射变换下直径与内切半径之比的最小化（见 [7, 第 5]节）。

随后，我们回到欧几里得球的情形。再次参照体积极值逼近的相关结果，我们会预期定理1.4 中的条件在此情形下足以保证最优性。我们通过推广定理 1.2验证了这一点，并给出了欧几里得球的 Banach–Mazur 距离位置的完整刻画。

定理 1.5。设 $K \subseteq \mathbb{R}^n$ 是一个凸体，c、$d \in \mathbb{R}^n$ 是向量，且 $R \geq r > 0$ 是实数，并满足 $r\mathbb{B}^n + c \subseteq K \subseteq R\mathbb{B}^n + d$。那么以下条件等价：

(i) $d_{BM}(K, \mathbb{B}^n) = \frac{R}{r}$.

(ii) 对任意矩阵 $A \in \mathcal{M}_s^n$ 及任意向量 v、$w \in \mathbb{R}^n$，都存在一个内接触点$y \in \mathrm{bd}(K) \cap \mathrm{bd}(r\mathbb{B}^n + c)$ 以及一个外接触点 $z \in \mathrm{bd}(K) \cap \mathrm{bd}(R\mathbb{B}^n + d)$，使得

$$\left\langle \frac{y-c}{r}, A\left(\frac{y-c}{r}\right) + v \right\rangle \leq \left\langle \frac{z-d}{R}, A\left(\frac{z-d}{R}\right) + w \right\rangle.$$

(iii) 存在整数 N、$M \geq 1$，内接触点 y^1、$\ldots y^N \in \mathrm{bd}(K) \cap \mathrm{bd}(r\mathbb{B}^n + c)$，外接触点 z^1、……、$z^M \mathrm{bd}(K) \cap \mathrm{bd}(R\mathbb{B}^n + d)$，以及权重 λ_1、……、λ_N、μ_1、……、$\mu_M > 0$，使得对任意 $x \in \mathbb{R}^n$，

$$\sum_{i=1}^{N} \lambda_i \langle x, y^i - c\rangle (y^i - c) = \sum_{j=1}^{M} \mu_j \langle x, z^j - d\rangle (z^j - d) \quad and \quad \sum_{i=1}^{N} \lambda_i (y^i - c) = \sum_{j=1}^{M} \mu_j (z^j - d) = 0.$$

在这种情况下，$r^2 \sum_{i=1}^{N} \lambda_i = R^2 \sum_{j=1}^{M} \mu_j$ 并且存在一种选择，使得 $N + M \leq {}^{n(n+5)}\!/_{2} + 1$。

事实上，我们证明了一个更强的版本，其中只有中间的凸体是欧几里得球。与定理1.4类似，在以下结果中，内侧和外侧的凸体可以彼此不同。

定理 1.6。设 L_1、$L_2 \subseteq \mathbb{R}^n$ 为凸体，c、$d \in \mathbb{R}^n$ 为向量，r、$R > 0$ 为实数，且满足 $rL_1 + c \subseteq \mathbb{B}^n \subseteq RL_2 + d$。则以下条件等价：

(i) 包含链 $rL_1 + c \subseteq \mathbb{B}^n \subseteq RL_2 + d$ 是仿射最优的。

(ii) 对于任意矩阵 $A \in \mathcal{M}_s^n$ 和任意向量 $v, w \in \mathbb{R}^n$，存在一个内接触点$y \in \mathrm{bd}(\mathbb{B}^n) \cap \mathrm{bd}(rL_1 + c)$ 以及一个外接触点 $z \in \mathrm{bd}(\mathbb{B}^n) \cap \mathrm{bd}(RL_2 + d)$，使得

$$\langle y, Ay + v\rangle \leq \langle z, Az + w\rangle.$$

(iii) 存在整数 $N, M \geq 1$、内接触点 $y^1, \ldots y^N \in \mathrm{bd}(\mathbb{B}^n) \cap \mathrm{bd}(rL_1 + c)$、外接触点 $z^1, \ldots, z^M \mathrm{bd}(\mathbb{B}^n) \cap \mathrm{bd}(RL_2 + d)$，以及权重 $\lambda_1, \ldots, \lambda_N, \mu_1, \ldots, \mu_M > 0$，使得对于任意 $x \in \mathbb{R}^n$，

$$\sum_{i=1}^{N} \lambda_i \langle x, y^i\rangle y^i = \sum_{j=1}^{M} \mu_j \langle x, z^j\rangle z^j \quad and \quad \sum_{i=1}^{N} \lambda_i y^i = \sum_{j=1}^{M} \mu_j z^j = 0.$$

在这种情况下，$\sum_{i=1}^{N} \lambda_i = \sum_{j=1}^{M} \mu_j$ 并且存在一种选择使得 $N + M \leq \frac{n(n+5)}{2} + 1$。

定理 1.2 和定理 1.6 的证明之间的主要区别在于各自对最优性条件之充分性的证明。必要性是沿类似思路建立的（并且基本上由定理 1.4 得出），而在一般情形下，充分性要困难得多。如 [10, 节 2], 中所示，当假定$r\mathbb{B}^n \subseteq K \subseteq R\mathbb{B}^n$ 对于原点对称凸体 K 并非仿射最优时，很容易构造一个显式矩阵 A，使得定理 1.2 中的非分离条件 (ii) 不成立。然而，在定理 1.6 更为一般的设定中，类似的方法并不能立即得出同样的结论。此处的主要障碍在于额外的平移向量，而这些向量在对称情形中可以省略。我们借助非线性优化中的一种一般方法，通过证明底层问题的“凸性”来规避这一问题。具体而言，我们证明：若 K、$L \subseteq \mathbb{R}^n$ 是凸体，E_0 与 E_1 是以原点为中心的椭球，c_0、c_1、d_0、d_1 是向量，且 r_0、r_1、R_0、$R_1 > 0$ 是实数，使得

$$r_0 K - c_0 \subseteq E_0 \subseteq R_0 L - d_0 \quad \text{and} \quad r_1 K - c_1 \subseteq E_1 \subseteq R_1 L - d_1,$$

那么，对于任意的 $\lambda \in (0,1)$，存在一个 E_λ E_0 和 E_1 的（几何）均值椭球，以及向量 c_λ, $d_\lambda \in \mathbb{R}^n$，使得

$$r_0^{1-\lambda} r_1^{\lambda} K - c_\lambda \subseteq E_\lambda \subseteq R_0^{1-\lambda} R_1^{\lambda} L - d_\lambda.$$

基于此，我们可以证明，在定理 1.6 的情形下，若某个椭球 E 的位似比优于 $\mathbb{B}^n$，那么将 $\mathbb{B}^n$ 略微适当地朝 E 扰动，也能得到更优的位似比；再利用极限论证，便可确实推出这与定理中的条件 (ii) 矛盾。均值椭球结果以及定理 1.5 和 1.6 的证明分别见第 4 节和第5节。

我们认为，上述均值椭球结果本身就具有独立的研究价值。在第1.5节，我们转而讨论定理6的一些应用，这也凸显了该结果的重要性：我们的证明往往依赖这两个结果的相互作用。第一个应用是对不等式 $d_{BM}(K,\ \mathbb{B}^n) \leq n$ 适用于任意凸体 $K \subseteq \mathbb{R}^n$的另一种证明。与使用 John 椭球定理的标准证明相比，我们的方法能更直接地刻画等号情形，即等号成立当且仅当 K 是单纯形。作为定理 1.5的另一项应用，我们在定理6.3中将 Maurey 关于唯一距离椭球的结果推广到非对称情形，并以投影代替截面。

2 仿射最优位置下 Ader 分解的存在性

本节旨在证明定理 1.4。如前所述，这需要将标准超平面分离定理推广，以刻画两个凸体何时在给定线性子空间中的某一点相交，这可以通过某种非分离条件来表达。之所以需要这一推广，是因为定理 1.4 (ii) 中的平衡条件要求外法向量的加权和等于零。如果条件仅要求这些和彼此相等，那么定理 1.4 (ii) 就可以等价地表述为矩阵空间 $\mathbb{R}^{n\times(n+1)}$ 中的若干凸体相交，而这可以用通常的分离结果来处理。

为了得到定理 1.4 (ii) 中极小分解所含接触对数量的预期上界，我们还要验证 Kirchberger 定理针对所需分离类型的一个推广。事实上，这一推广可直接由 Kirchberger 定理推出；这一点从下文的证明中便显而易见，该证明的灵感来自 [22, 定理 1.3.11]中 Kirchberger 定理的证明。我们用 π_U 表示从 $\mathbb{R}^n$ 到线性子空间 $U \subseteq \mathbb{R}^n$的 $U^\perp$ 正交投影，并用 $U^\perp$ 表示 $U^\perp$ 的正交补。对于集合 $X \subseteq \mathbb{R}^n$，我们用 $\mathrm{conv}(X)$ 表示它的凸包，用 $\mathrm{lin}(X)$ 表示它的线性张成。

引理 2.1. 设 K、$L \subseteq \mathbb{R}^n$ 为紧集，$U \subseteq \mathbb{R}^n$ 为线性子空间。则以下条件等价：

(i) $\mathrm{conv}(K) \cap \mathrm{conv}(L) \cap U = \emptyset$。

(ii) 存在 $a \in U$和 $v, w \in U^\perp$，使得 $\langle a+v, x\rangle < \langle a+w, y\rangle$ 对所有 $x \in K$和$y \in L$ 均成立。*(iii)* 对于满足 $|K'| + |L'| \leq 2(n+1) - \dim(U)$ 的所有子集 $K' \subseteq K$ 和 $L' \subseteq L$，存在$a \in U$和 $v, w \in U^\perp$，使得 $\langle a+v, x\rangle < \langle a+w, y\rangle$ 对所有 $x \in K'$和 $y \in L'$ 均成立。

证明。首先，我们来证明 (ii) 推出 (i)。假设对于某些 $a \in U$ 和 $v,\ w \in U^\perp$，(ii) 成立。根据凸性，(ii) 中的不等式对于所有 $x \in \mathrm{conv}(K)$ 和 $y \in \mathrm{conv}(L)$ 都成立。因此，对于 $x \in \mathrm{conv}(K) \cap U$ 和 $y \in \mathrm{conv}(L) \cap U$，我们有

$$\langle a, x\rangle = \langle a+v, x\rangle < \langle a+w, y\rangle = \langle a, y\rangle.$$

这表明 $x \neq y$，因此，$\mathrm{conv}(K) \cap \mathrm{conv}(L) \cap U = \emptyset$。

接下来，我们证明 (i) 推出 (ii)。(ii) 中所需的条件可以改写为

$$\langle a, \pi_U(x)\rangle + \langle v, \pi_{U^\perp}(x)\rangle = \langle a+v, x\rangle < \langle a+w, y\rangle = \langle a, \pi_U(y)\rangle + \langle w, \pi_{U^\perp}(y)\rangle \qquad (2)$$

对所有 $x \in K$, $y \in L$ 以及某些 $a \in U$, $v, w \in U^\perp$。考虑线性向量空间 $V = U \times U^\perp \times U^\perp$和紧集 $K_0 := \{(\pi_U(x)\ \pi_{U^\perp}(x), 0) : x \in K\} \subseteq V$，以及 $L_0 := \{(\pi_U(y)\ 0, \pi_{U^\perp}(y)) : y \in L\} \subseteq V$。假设 (i) 蕴含 $\mathrm{conv}(K_0) \cap \mathrm{conv}(L_0) = \emptyset$。事实上，否则就会存在某些 $x \in \mathrm{conv}(K)$ 和 $y \in \mathrm{conv}(L)$，使得 $\pi_U(x) = \pi_U(y)$ 且 $\pi_{U^\perp}(x) = \pi_{U^\perp}(y) = 0$。然而，这意味着 $x = y \in U$，与 (i) 矛盾。因此，在 V 中应用标准超平面分离定理，可得到向量 $(a, v, w) \in V$，它严格分离 $\mathrm{conv}(K_0)$ 和 $\mathrm{conv}(L_0)$。不难验证，该向量满足 (2)。

最后，(ii) 和 (iii) 的等价性直接由 Kirchberger 定理得出。该定理在此处表明：K_0 与 L_0 能够被严格分离，当且仅当 K_0与 L_0的任意子集（其总基数至多为 $\dim(V) + 2 = 2n - \dim(U) + 2$）都能够被严格分离。再次利用（2），容易看出，后一个陈述等价于 (iii)。 □

在下面的注中，我们将对上述引理给出几何解释。我们还将说明，一般而言，条件 (iii) 所要求的点数上界无法改进。

注 2.2。引理 *2.1* 的一种更几何化的解释是，条件$\mathrm{conv}(K) \cap \mathrm{conv}(L) \cap U = \emptyset$ 等价于存在两个闭半空间 H_K，$H_L \subseteq \mathbb{R}^n$满足

$$K \subseteq H_K, \quad L \subseteq H_L \quad \textit{and} \quad H_K \cap H_L \cap U = \emptyset.$$

此外，该引理立即给出以下 ***Helly*** 型结果：$\mathrm{conv}(K) \cap \mathrm{conv}(L) \cap U = \emptyset$成立，当且仅当对于所有满足 $|K'| + |L'| \leq 2(n+1) - \dim(U)$的子集 $K' \subseteq K$ 和 $L' \subseteq L$，$\mathrm{conv}(K') \cap \mathrm{conv}(L') \cap U = \emptyset$ 均成立。若将U 视为仿射子空间而非线性子空间，上述两条观察仍然成立。

此外，我们指出，引理中的数值$2(n+1) - \dim(U)$是最佳可能的。事实上，设 $d \in \{0, \ldots, n\}$，令 $e^1, \ldots, e^n \in \mathbb{R}^n$ 取自 $\mathbb{R}^n$ 的一组基，并选取线性子空间 $U = \mathrm{lin}\{e^i : i = 1, \ldots, d\}$，使得 $d = \dim(U)$。我们将集合 $K, L \subseteq \mathbb{R}^n$ 定义为

$$K = \left\{ e^1, \ldots, e^n, -\frac{1}{2n}\sum_{i=1}^n e^i \right\} \quad \textit{and} \quad L = \left\{ -e^{d+1}, \ldots, -e^n, \frac{1}{2n}\sum_{i=1}^n e^i \right\}.$$

容易验证，$u := \frac{1}{3n-d}\sum_{i=1}^d e^i$ 是 $\mathit{conv}(L) \cap U$ 中的唯一点，它由一个唯一的凸组合生成，且该凸组合对 L 中所有的 $n + 1 - d$点都赋予正权重。类似地，可以验证 $u \in \mathit{conv}(K) \cap U$，同样存在一个唯一的凸组合，对 K 中所有的 $n + 1$ 点都赋予正权重。因此，点 u 是交集 $\mathit{conv}(K) \cap \mathit{conv}(L) \cap U$ 中的唯一点，其中来自 K 和 L 的凸组合所含点的总数等于 $2(n+1) - d$。因此，若 $K' \subseteq K$ 和$L' \subseteq L$ 满足 $|K'| + |L'| < 2(n+1) - d$，则 $\mathit{conv}(K') \cap \mathit{conv}(L') \cap U = \emptyset$。根据引理 *2.1*，对于任意这样的 K' 和 L'，所需的向量 a, v, w 均存在；而对于 K 和 L 则不存在，因为 $\mathit{conv}(K) \cap \mathit{conv}(L) \cap U \neq \emptyset$。

有了引理 2.1作为工具，我们现在可以着手实现本节的目标，即证明定理 1.4。其结构与 1.2 中定理的证明 [10], 相似，也与 [17, 定理 2.1]的证明有一些相似之处。我们证明方法中最大的不同在于，本情形可能涉及平移，而这在 [10] 和 [17]中并不存在。此外，执行证明时需要格外谨慎，以免不得不施加任何正则性假设；如果直接照搬 [17] 中的方法（见 [10, 注 2.1]），就有必要施加这些假设。

在给出证明之前，我们先介绍一些贯穿全文的记号。对于满足 $K \subseteq \mathbb{R}^n$ 且 $0 \in \operatorname{int}(K)$的凸体，其极体是凸体

$$K^\circ = \{a \in \mathbb{R}^n : \langle x, a\rangle \leq 1 \text{ for all } x \in K\}.$$

规范函数（或称闵可夫斯基泛函） K 定义为

$$\|x\|_K = \min\{\rho \geq 0 : x \in \rho K\} \quad \text{for all } x \in K.$$

对于矩阵 A, $B \in \mathbb{R}^{n\times m}$，它们的弗罗贝尼乌斯内积通过迹定义为

$$\langle A, B\rangle_F = \operatorname{Tr}(A^T B).$$

我们用 I_n 表示 $\mathbb{R}^{n\times n}$ 中的单位矩阵，并将矩阵 $A \in \mathbb{R}^{n\times n}$ 添加另一列 $x \in \mathbb{R}^n$ 后的扩展记作 $(A|x) \in \mathbb{R}^{n\times(n+1)}$。

定理 1.4 的证明。首先注意到，根据假设，0 位于 $K - c$ 的内部。任取 $A \in \mathbb{R}^{n\times n}$ 和 v, $w \in \mathbb{R}^n$。对于 $\varepsilon > 0$，我们定义线性算子 $T_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$，使得 $T_\varepsilon(x) = (I_n + \varepsilon A)x$。当 $\varepsilon > 0$ 充分小时，由诺伊曼级数可知 T_ε 可逆，并且有 $c - \varepsilon v \in \operatorname{int}(T_\varepsilon(K))$。现在，令 $r_\varepsilon > 0$ 为最大值，$R_\varepsilon > 0$ 为最小值，使得

$$r_\varepsilon L_1 + c - \varepsilon v \subseteq T_\varepsilon(K) \subseteq R_\varepsilon L_2 + d - \varepsilon w.$$

于是存在点 y^ε, $z^\varepsilon \in \operatorname{bd}(K)$ 的选取，使得 $T_\varepsilon(y^\varepsilon) + \varepsilon v - c \in \operatorname{bd}(r_\varepsilon L_1)$ 且$T_\varepsilon(z^\varepsilon) + \varepsilon w - d \in \operatorname{bd}(R_\varepsilon L_2)$。由 $rL_1 + c \subseteq K \subseteq RL_2 + d$ 的仿射最优性，我们有

$$\frac{R}{r} \leq \frac{R_\varepsilon}{r_\varepsilon} = \frac{\|T_\varepsilon(z^\varepsilon) + \varepsilon w - d\|_{L_2}}{\|T_\varepsilon(y^\varepsilon) + \varepsilon v - c\|_{L_1}}.$$

整理可得

$$\frac{\|T_\varepsilon(y^\varepsilon) + \varepsilon v - c\|_{L_1}}{r} \leq \frac{\|T_\varepsilon(z^\varepsilon) + \varepsilon w - d\|_{L_2}}{R}. \tag{3}$$

接下来，我们选取某些 $a^\varepsilon \in \operatorname{bd}((K - c)^\circ)$ 和 $b^\varepsilon \in \operatorname{bd}(\frac{1}{R}L_2^\circ)$，使得

$$\langle a^\varepsilon, y^\varepsilon - c\rangle = \|y^\varepsilon - c\|_{K-c} = 1 \quad \text{and} \quad \langle Rb^\varepsilon, T_\varepsilon(z^\varepsilon) + \varepsilon w - d\rangle = \|T_\varepsilon(z^\varepsilon) + \varepsilon w - d\|_{L_2} = R_\varepsilon.$$

则 $ra^\varepsilon \in r(K - c)^\circ \subseteq L_1^\circ$ 与 $y^\varepsilon - c \in \operatorname{bd}(K - c)$蕴含

$$\begin{aligned}\|T_\varepsilon(y^\varepsilon) + \varepsilon v - c\|_{L_1} &\geq \langle ra^\varepsilon, T_\varepsilon(y^\varepsilon) + \varepsilon v - c\rangle = r\langle a^\varepsilon, y^\varepsilon - c\rangle + r\varepsilon\langle a^\varepsilon, Ay^\varepsilon + v\rangle \\ &= r\|y^\varepsilon - c\|_{K-c} + r\varepsilon\langle a^\varepsilon, Ay^\varepsilon + v\rangle = r(1 + \varepsilon\langle a^\varepsilon, Ay^\varepsilon + v\rangle).\end{aligned}$$

类似地， $Rb^\varepsilon \in \mathrm{bd}(L_2^\circ)$与 $z^\varepsilon - d \in K - d \subseteq RL_2$ 得到

$$\begin{aligned}\|T_\varepsilon(z^\varepsilon) + \varepsilon w - d\|_{L_2} &= \langle Rb^\varepsilon, T_\varepsilon(z^\varepsilon) + \varepsilon w - d\rangle = \langle Rb^\varepsilon, z^\varepsilon - d\rangle + R\varepsilon\langle b^\varepsilon, Az^\varepsilon + w\rangle \\ &\le \|z^\varepsilon - d\|_{L_2} + R\varepsilon\langle b^\varepsilon, Az^\varepsilon + w\rangle \le R(1 + \varepsilon\langle b^\varepsilon, Az^\varepsilon + w\rangle).\end{aligned}$$

将这些估计代入（3），得到

$$1 + \varepsilon\langle a^\varepsilon, Ay^\varepsilon + v\rangle \le 1 + \varepsilon\langle b^\varepsilon, Az^\varepsilon + w\rangle,$$

化简后得

$$\langle a^\varepsilon, Ay^\varepsilon + v\rangle \le \langle b^\varepsilon, Az^\varepsilon + w\rangle. \tag{4}$$

现在，由于 y^ε， $z^\varepsilon \in \mathrm{bd}(K)$、 $a^\varepsilon \in \mathrm{bd}((K-c)^\circ)$ 以及 $b^\varepsilon \in \mathrm{bd}(\frac{1}{R}L_2^\circ)$ 对所有 $\varepsilon > 0$ 均成立，且这三个边界都是紧致的，因此存在一个序列，使得 $\varepsilon \to 0^+$，且 $y^\varepsilon \to y$、$z^\varepsilon \to z$、 $a^\varepsilon \to a$ 和 $b^\varepsilon \to b$ 对某个 y， $z \in \mathrm{bd}(K)$、 $a \in \mathrm{bd}((K-c)^\circ)$ 以及 $b \in \mathrm{bd}(\frac{1}{R}L_2^\circ)$ 成立。我们验证向量 y， z， a， b 满足 (i) 中要求的所有性质。首先，我们注意到

$$\frac{y-c}{r} = \lim_{\varepsilon\to 0^+} \frac{T_\varepsilon(y^\varepsilon) + \varepsilon v - c}{r_\varepsilon} \in \mathrm{bd}(L_1) \quad \text{and} \quad \frac{z-d}{R} = \lim_{\varepsilon\to 0^+} \frac{T_\varepsilon(z^\varepsilon) + \varepsilon w - d}{R_\varepsilon} \in \mathrm{bd}(L_2).$$

因此， y 是 K 和 $rL_1 + c$ 的公共边界点，而 z 是 K 和 $RL_2 + d$ 的公共边界点。其次，我们注意到

$$\langle a, y - c\rangle = \lim_{\varepsilon\to 0^+} \langle a^\varepsilon, y^\varepsilon - c\rangle = \lim_{\varepsilon\to 0^+} \|y^\varepsilon - c\|_{K-c} = 1$$

and

$$\langle b, z - d\rangle = \lim_{\varepsilon\to 0^+} \langle b^\varepsilon, T_\varepsilon(z^\varepsilon) + \varepsilon w - d\rangle = \lim_{\varepsilon\to 0^+} \frac{\|T_\varepsilon(z^\varepsilon) + \varepsilon w - d\|_{L_2}}{R} = \lim_{\varepsilon\to 0^+} \frac{R_\varepsilon}{R} = 1.$$

在（4）中取极限，得到

$$\langle a, Ay + v\rangle \le \langle b, Az + w\rangle.$$

这已经验证了 (i) 中所需的条件。事实上，矩阵 A 与向量 v、w 是任意选取的，而选取对 (y, a) 意味着 $\langle a, x - c\rangle \le \langle a, y - c\rangle = 1$，从而对所有 $x \in K \cup rL_1$ 有 $\langle a, x\rangle \le \langle a, y\rangle$。由于对 (z, b) 也有类似事实成立，因此 (y, a) 与 (z, b) 分别是 K 与 $rL_1 + c$ 以及 K 与 $RL_2 + d$ 的接触对，正如所要求的那样。

为了证明 (ii)，我们首先注意到，也可以对矩阵 $A' = A^T$ 与向量 $v' = v - A^T c$、$w' = w - A^T d$ 应用 (i)，从而如上得到接触对 (y, a) 与 (z, b)，且

$$\langle a, A^T(y - c) + v\rangle = \langle a, A'y + v'\rangle \le \langle b, A'z + w'\rangle = \langle b, A^T(z - d) + w\rangle$$

以及 $\langle a,\ y - c\rangle = \langle b,\ z - d\rangle = 1$。此外，我们有

$$\begin{aligned}\langle a, A^T(y - c) + v\rangle &= \mathrm{Tr}(a^T A^T(y - c)) + \langle a, v\rangle = \mathrm{Tr}(A^T(y - c)a^T) + \langle a, v\rangle \\ &= \langle A, (y - c)a^T\rangle_F + \langle a, v\rangle = \langle (A|v), ((y - c)a^T|a)\rangle_F\end{aligned}$$

并且类似地 $\langle b, A^T(z - d)\ + w\rangle = \langle (A|w), ((z - d)b^T|b)\rangle_F$。因此，

$$\langle (A|v), ((y - c)a^T|a)\rangle_F \le \langle (A|w), ((z - d)b^T|b)\rangle_F. \tag{5}$$

如果现在定义 $n\times(n+1)$ 矩阵的集合 $\mathcal{Y}$ 和 $\mathcal{Z}$ 为

$$\mathcal{Y}=\left\{((y-c)a^T|a)\in\mathbb{R}^{n\times(n+1)}:y\in\mathrm{bd}(K)\cap\mathrm{bd}(rL_1+c),a\in\mathrm{bd}((K-c)^\circ),\langle a,y-c\rangle=1\right\}$$

and

$$\mathcal{Z}=\left\{((z-d)b^T|b)\in\mathbb{R}^{n\times(n+1)}:z\in\mathrm{bd}(K)\cap\mathrm{bd}(RL_2+d),b\in\mathrm{bd}\left(\frac{1}{R}L_2^\circ\right),\langle b,z-d\rangle=1\right\},$$

那么它们两者都是紧致的。由于对 A、v、w 的任意选取，在适当选取 y、z、a、b时(5)均成立，因此若我们选取线性子空间 $\mathcal{U}:=\{(A|0)\in\mathbb{R}^{n\times(n+1)}:A\in\mathbb{R}^{n\times n}\}$，则引理2.1中关于 $\mathcal{Y}$和 $\mathcal{Z}$的第二个条件不成立。因此，存在某些整数 N、$M\geq 1$，矩阵$((y^1-c)(a^1)^T|a^1)\ldots,((y^N-c)(a^N)^T|a^N)\in\mathcal{Y}$，$((z^1-d)(b^1)^T|b^1)\ldots,((z^M-d)(b^M)^T|b^M)\in\mathcal{Z}$，以及权重 $\lambda_1,\ldots,\lambda_N,\mu_1,\ldots,\mu_M>0$，且满足$\sum_{i=1}^N\lambda_i=\sum_{j=1}^M\mu_j=1$，使得

$$\sum_{i=1}^N\lambda_i((y^i-c)(a^i)^T|a^i)=\sum_{j=1}^M\mu_j((z^j-c)(b^j)^T|b^j)\in\mathrm{conv}(\mathcal{Y})\cap\mathrm{conv}(\mathcal{Z})\cap\mathcal{U}.$$

这等价于

$$\sum_{i=1}^N\lambda_i(y^i-c)(a^i)^T=\sum_{j=1}^M\mu_j(z^j-d)(b^j)^T\quad\text{and}\quad\sum_{i=1}^N\lambda_ia^i=\sum_{j=1}^M\mu_jb^j=0.$$

为得到定理中分解的精确形式，我们计算

$$\begin{aligned}\sum_{i=1}^N\lambda_iy^i(a^i)^T&=\sum_{i=1}^N\lambda_iy^i(a^i)^T-c\left(\sum_{i=1}^N\lambda_ia^i\right)^T=\sum_{i=1}^N\lambda_i(y^i-c)(a^i)^T=\sum_{j=1}^M\mu_j(z^j-d)(b^j)^T\\&=\sum_{j=1}^M\mu_jz^j(b^j)^T-d\left(\sum_{j=1}^M\mu_jb^j\right)^T=\sum_{j=1}^M\mu_jz^j(b^j)^T.\end{aligned}$$

与之前类似，我们可以论证，(y^i,a^i) 是 K 和 rL_1+c 的接触对，其中 $i=1,\ldots,N$；而 (z^j,b^j) 是 K 和 RL_2+d 的接触对，其中 $j=1,\ldots,M$。

最后剩下的部分是等式 $\sum_{i=1}^N\lambda_i\langle y^i,a^i\rangle=\sum_{j=1}^M\mu_j\langle z^j,b^j\rangle$，以及对于适当选择的接触对，上界 $N+M\leq(n+1)^2$。前者可直接通过比较分解中矩阵的迹得出。对于后者，对任意 $((y-c)a^T|a)\in\mathcal{Y}$，我们有

$$\langle(I_n|0),((y-c)a^T|a)\rangle_F=\mathrm{Tr}(I_n^T(y-c)a^T)=\mathrm{Tr}(a^TI_n(y-c))=\langle a,y-c\rangle=1$$

并且对于任意 $((z-d)b^T|b)\in\mathcal{Z}$，我们有

$$\langle(I_n|0),((z-d)b^T|b)\rangle_F=\mathrm{Tr}(I_n^T(z-d)b^T)=\mathrm{Tr}(b^TI_n(z-d))=\langle b,z-d\rangle=1.$$

因此，$\mathcal{Y}$ 和 $\mathcal{Z}$ 是仿射子空间的子集

$$\mathcal{H}:=\{A\in\mathbb{R}^{n\times(n+1)}:\langle(I_n|0),A\rangle_F=1\}.$$

由于 conv($\mathcal{Y}$) 和 conv($\mathcal{Z}$) 在某个属于 $\mathcal{U}' := \mathcal{H} \cap \mathcal{U}$ 的点处相交，将注 2.2 应用于引理 2.1，并相对于仿射空间 $\mathcal{H}$ 考虑，可知已有某些子集 $\mathcal{Y}' \subseteq \mathcal{Y}$和 $\mathcal{Z}' \subseteq \mathcal{Z}$，满足 $|\mathcal{Y}'| + |\mathcal{Z}'| \leq 2(\dim(\mathcal{H}) + 1) - \dim(\mathcal{U}') = 2n(n+1) - (n^2 - 1) = (n+1)^2$satisfy conv($\mathcal{Y}'$)∩conv($\mathcal{Z}'$)∩$\mathcal{U}' \neq \emptyset$。通过与上述类似的计算，可将该交集中的一个点转化为一个与定理中类似的分解，其中 $N+M \leq (n+1)^2$。证明至此完成。 □

我们以一个简单的观察结束本节：Ader 分解在仿射变换下的存在性及其性质。该观察阐明了通过对包含链应用仿射变换，可以对 Ader 分解中出现的矩阵作出哪些修改。请注意，以下引理还表明，在定理1.5和1.6中，我们可以用任意椭球替换欧几里得球，而不失完整的等价性。

引理 2.3。 设 K、L_1、$L_2 \subseteq \mathbb{R}^n$ 为凸体， c、$d \in \mathbb{R}^n$ 为向量， r、$R > 0$ 为实数，且满足 $rL_1 + c \subseteq K \subseteq RL_2 + d$。此外，设 $B \in \mathbb{R}^{n \times n}$ 为可逆矩阵， $v \in \mathbb{R}^n$ 为向量，并设 T：$\mathbb{R}^n \to \mathbb{R}^n$，其中 $T(x) = Bx + v$ 为仿射算子。那么，对于包含链 $rL_1 + c \subseteq K \subseteq RL_2 + d$, 存在一个分解中矩阵为 $A \in \mathbb{R}^{n \times n}$ 的 *Ader* 分解，当且仅当对于包含链 $T(rL_1 + c) \subseteq T(K) \subseteq T(RL_2 + d)$，存在一个分解中矩阵为 BAB^{-1} 的 *Ader* 分解。*</output>*

证明。 我们只需证明：若 $rL_1+c \subseteq K \subseteq RL_2+d$存在一个矩阵为 A的 Ader 分解，则 $T(rL_1+c) \subseteq T(K) \subseteq T(RL_2+d)$存在一个矩阵为 BAB^{-1}的 Ader 分解；反向结论可通过应用 T^{-1}以相同方式得到。现在，设 $(y^1, a^1) \ldots, (y^N, a^N)$ 是 K和 rL_1+c的接触对，$(z^1, b^1) \ldots (z^M, b^M)$ 是 K 和 $RL_1 + d$ 的接触对，并且权重 $\lambda_1, \ldots, \lambda_N, \mu_1, \ldots, \mu_M > 0$ 满足

$$\sum_{i=1}^{N} \lambda_i y^i (a^i)^T = \sum_{j=1}^{M} \mu_j z^j (b^j)^T = A \quad \text{and} \quad \sum_{i=1}^{N} \lambda_i a^i = \sum_{j=1}^{M} M \mu_j b^j = 0.$$

对于 $i = 1, \ldots, N$和 $x \in K \cup (rL_1 + c)$，根据定义可得

$$\begin{aligned} \langle T(x), (B^{-1})^T a^i \rangle &= \langle B^{-1}(Bx + v), a^i \rangle = \langle x, a^i \rangle + \langle B^{-1} v, a^i \rangle \\ &\leq \langle y^i, a^i \rangle + \langle B^{-1} v, a^i \rangle = \langle T(y^i), (B^{-1})^T a^i \rangle. \end{aligned}$$

因此，$(T(y^i), (B^{-1})^T a^i)$ 是 $T(K)$ 和 $T(rL_1 + c)$ 的接触对。同样地，对于$j = 1$，……， M，可得 $(T(z^j), (B^{-1})^T z^j)$ 是 $T(K)$ 和 $T(RL_2 + d)$ 的接触对。最后，

$$\sum_{i=1}^{N} \lambda_i ((B^{-1})^T a^i) = (B^{-1})^T \left(\sum_{i=1}^{N} \lambda_i a^i \right) = 0 = (B^{-1})^T \left(\sum_{j=1}^{M} \mu_j b^j \right) = \sum_{j=1}^{M} \mu_j ((B^{-1})^T b^j)$$

and

$$\begin{aligned} \sum_{i=1}^{N} \lambda_i (T(y^i))((B^{-1})^T a^i)^T &= B \left(\sum_{i=1}^{N} \lambda_i y^i (a^i)^T \right) B^{-1} + v \left(\sum_{i=1}^{N} \lambda_i a^i \right)^T B^{-1} = BAB^{-1} \\ &= B \left(\sum_{j=1}^{M} \mu_j z^j (b^j)^T \right) B^{-1} + v \left(\sum_{j=1}^{M} \mu_j b^j \right)^T B^{-1} = \sum_{j=1}^{M} \mu_j (T(z^j))((B^{-1})^T b^j)^T \end{aligned}$$

为带有矩阵 BAB^{-1} 的 $T(rL_1+c) \subseteq T(K) \subseteq T(RL_2+d)$ 生成 Ader 分解。 □

3 分解型结果的应用

在继续讨论涉及椭球时 Ader 分解对于最优性的充分性之前，我们先在本节介绍定理 1.4 的一些应用。我们的目标是说明，如何利用该定理中内外逼近凸体彼此独立这一自由度，为那些本质上与仿射最优包含链甚至仿射变换无关的问题获得最优性条件。

这类应用中的第一个，是对以下广为人知的结果给出一种新的证明。该结果为位似下的最优包含给出了一个必要条件。需要指出的是，这些条件也已知是充分的（例如，参见 [8, 定理 2.3]），不过本文不讨论这一点。

推论3.1。设 $K \subseteq \mathbb{R}^n$ 是一个紧集，$C \subseteq \mathbb{R}^n$ 是一个凸体。假设 $v \in \mathbb{R}^n$ 是一个向量，$R \geq 0$ 是一个实数，且 $K \subseteq RC + v$，并且对于任意实数 $0 \leq R' < R$ 和向量 $v' \in \mathbb{R}^n$，K 都不包含于 $R'C + v'$。那么，存在 K 和 $RC + v$ 的公共支撑超平面的外法向量a^1、……、$a^N \in \mathbb{R}^n \setminus \{0\}$，使得 $0 \in$凸包$\{a^1, \ldots, a^N\}$。

证明。如有必要，调整 v，我们不妨假设 $0 \in \text{int}(C)$。若 $R = 0$，则 K 和 $RC+v$ 均为单点集，因此可以任取一个非零向量作为 a^1，并令$a^2 = -a^1$，以满足推论中的条件。若 $R > 0$，则取某个半径为正、包含于 int(C) 的原点中心欧氏球 B，并将凸体 $K' \subseteq \mathbb{R}^n$ 定义为 $K' = \text{conv}(K \cup (RB + v))$。显然，$v \in \text{int}(K')$。

我们声称包含链 $K' - v \subseteq K' - v \subseteq RC$ 是仿射最优的。事实上，若 $T \in \mathbb{R}^{n\times n}$ 是可逆线性算子，c，$d \in \mathbb{R}^n$ 是向量，且 R'，$r' > 0$ 是实数，使得

$$r'(K' - v) + c \subseteq T(K' - v) \subseteq R'C + d,$$

那么还有 $K \subseteq \frac{R'}{r'}C + t$ 对于 $t = \frac{1}{r'}(d - c) + v$。由推论中的假设可知 $\frac{R'}{r'} \geq R$，因此所考虑的包含链确实是仿射最优的。由于$0 \in \text{int}(K' - v) \cap \text{int}(C)$，我们可以应用定理 1.4，特别是可以得到一些接触对 $(y^1, a^1) \ldots, (y^N, a^N)$，它们属于 K'和 $RC + v$，并满足

$$\sum_{i=1}^{N} \lambda_i a^i = 0$$

对于某些权重 λ_1、……、$\lambda_N > 0$。由于 $RB + v \subseteq$ int($RC + v$)，显然所有接触点 y^i 都必须属于 K。因此，由等式 $\langle x$、$a^i\rangle = \langle y^i$、$a^i\rangle$所确定的超平面，对于所有 $i = 1$、……、N，同时支撑 K 和 $RC + v$，证明完毕。 □

如引言所述，历史上，人们常借助定理中的 John 分解等工具来估计 Banach–Mazur 距离 1.3，以研究体积极值逼近。下面，我们对定理 1.4 的第二个应用表明，只要适当选取包含链中的凸体，体积极值逼近便可作为一种特殊情形纳入仿射最优包含链的框架，从而至少在满足正则性假设时得出定理 1.3 。从这个意义上说，我们的方法可以视为对这一经典理论的推广，它逆转了通常的研究视角——通常人们利用体积极值位置来研究其他逼近问题。

推论3.2。 设 $K \subseteq L \subseteq \mathbb{R}^n$ 为凸体，且 $K \subseteq L$ 在包含于 L 的所有仿射变换中具有最大体积。另设 K 光滑。那么存在 $N \leq n^2+n$ 个接触对$(y^1, a^1) \ldots, (y^N, a^N)$，它们属于 K 和 L；此外还存在权重 $\lambda_1, \ldots, \lambda_N > 0$，使得对于任意 $x \in \mathbb{R}^n$,

$$\sum_{i=1}^{N} \lambda_i \langle x, a^i \rangle y^i = x \quad and \quad \sum_{i=1}^{N} \lambda_i a^i = 0.$$

证明。 设 $S \subseteq K$ 是包含于 K中的体积最大单纯形。我们断言，包含链 $S \subseteq K \subseteq L$ 是仿射最优的。事实上，设 $T\colon \mathbb{R}^n \to \mathbb{R}^n$ 是一个可逆线性算子，且 v， $w \in \mathbb{R}^n$ 是向量，使得

$$rS + v \subseteq T(K) \subseteq RL + w$$

对于某个 r， $R > 0$。由于 $\frac{1}{R}(T(K) - w) \subseteq L$，由 K在L 中的体积最大性可得

$$\operatorname{vol}(K) \geq \operatorname{vol}\left(\frac{1}{R}(T(K) - w)\right) = \frac{|\det(T)|}{R^n} \operatorname{vol}(K),$$

即， $|\det(T)| \leq R^n$。类似地，由于 $S \subseteq K$被选为体积最大，且$T^{-1}(rS + v) \subseteq K$，我们有

$$\operatorname{vol}(S) \geq \operatorname{vol}\left(T^{-1}(rS + v)\right) = \frac{r^n}{|\det(T)|} \operatorname{vol}(S),$$

因此 $|\det(T)| \geq r^n$。由此可得 $R \geq r$，进而得到$S \subseteq K \subseteq L$的仿射最优性。

现在，设 $T\colon \mathbb{R}^n \to \mathbb{R}^n$ 是一个仿射算子，它将 S 映射为内接于 $\mathbb{B}^n$的正则单纯形，即 $T(S) = \operatorname{conv}\{u^1, \ldots, u^{n+1}\}$，其中 $u^i \in \operatorname{bd}(\mathbb{B}^n)$，且对于所有$i, j = 1, \ldots, n+1$, $i \neq j$，都有 $\langle u^i, u^j \rangle = -\frac{1}{n}$。显然，包含链 $T(S) \subseteq T(K) \subseteq T(L)$仍是仿射最优的，并且 $T(S)$是包含于 $T(K)$中的体积最大单纯形。特别地，对于每个 $i = 1, \ldots, n+1$，都存在一个垂直于 u^i 的超平面，同时支撑$T(S)$和 $T(K)$于 u^i；否则，在 $T(K)$内便会存在一个体积更大的单纯形。根据我们的假设， $T(K)$是光滑的，因此对于 $T(K)$，这是 u^i处唯一的支撑超平面。此外， u^i 是 $T(S)$与 $T(K)$唯一的公共边界点。事实上，否则便会得到一个超平面 H ，它在某个$x \in \operatorname{bd}(T(S)) \setminus \{u^1, \ldots, u^{n+1}\}$处支撑 $T(S)$和 $T(K)$，而且必然也会在某个顶点 u^i处支撑 $T(S)$。特别地， H 也会在 u^i处支撑 $T(K)$，但根据 $x \in H \cap \operatorname{int}(\mathbb{B}^n)$，它不可能垂直于 u^i，这与 $T(K)$在 u^i处的光滑性矛盾。总之， $T(S)$与 $T(K)$的任一接触对均具有形式$(u^i, \rho u^i)$，其中某个 $i = 1, \ldots, n+1$ 和 $\rho > 0$。

现在，将定理 1.4 应用于 $T(S) \subseteq T(K) \subseteq T(L)$所得的 Ader 分解可写为

$$\sum_{j=1}^{n+1} \mu_j u^j (u^j)^T = \sum_{i=1}^{N} \lambda_i y^i (a^i)^T \quad \text{and} \quad \sum_{j=1}^{n+1} \mu_j u^j = \sum_{i=1}^{N} \lambda_i a^i = 0 \tag{6}$$

对于 $T(K)$ 和 $T(L)$ 的接触对 $(y^1,\ a^1)\ldots,\ (y^N,\ a^N)$，以及权重 $\lambda_1,\ \ldots,\ \lambda_N > 0$和 $\mu_1,\ \ldots,\ \mu_{n+1} \geq 0$，其中 $\sum_{j=1}^{n+1} \mu_j = n$。 u^i的仿射无关性表明，它们中任意一个等于 0 的线性组合，其系数都相等，因此 $\mu_1 = \ldots = \mu_{n+1} = \frac{n}{n+1}$。此外，对于任意固定的 $i = 1,\ \ldots,\ n+1$，我们有

$$\sum_{j=1}^{n+1} \langle u^i, u^j \rangle u^j = u^i - \frac{1}{n} \sum_{j \neq i} u^j = u^i + \frac{1}{n} u^i = \frac{n+1}{n} u^i.$$

因此，对于所有 $\sum_{j=1}^{n+1} \mu_j \langle x,\ u^j \rangle u^j = x$ 等式 $x = u^i,\ \ i = 1,\ \ldots\ldots,\ \ n+1$均成立，因而对所有 $x \in \mathbb{R}^n$也成立。因此，(6) 得出

$$\sum_{i=1}^{N} \mu_i \langle x, a^i \rangle y^i = x$$

对每个 $x \in \mathbb{R}^n$都成立。总之，对于 $T(S) \subseteq T(K) \subseteq T(L)$，存在一个包含单位矩阵的 Ader 分解。引理2.3表明，通过 $S \subseteq K \subseteq L$ 回到 T^{-1}后，这一结论仍然成立，从而得到所声称的分解。最后，定理 1.4给出某种选择，使得 $N + n + 1 \leq (n+1)^2$,，也就是说， $N \leq n^2 + n$。 □

注 3.3. 上述推论中的正则性假设可以放宽为存在一个光滑凸体或严格凸体 C，满足 $K \subseteq C \subseteq L$；采用类似但稍复杂的方法即可证明。此处省略细节，因为定理*1.3* 已经在完全一般的情形下给出了该结果。

本节中定理 1.4 的第三个也是最后一个应用，涉及在仿射变换下最小化直径与内切半径之比。对于凸体 $K, L \subseteq \mathbb{R}^n$ （其中 L 关于原点对称）， K 相对于 L 的直径为 $D(K, L) = \max\{\|x - y\|_L : x, y \in K\}$，而内切半径为 $r(K, L) = \max\{\rho \geq 0 :\ \rho L + c \subseteq K,\ c \in \mathbb{R}^n\}$。由于 $\frac{D(K,L)}{r(K,L)}$ 一般可以任意大，我们的目标是先通过仿射变换 K 来最小化这一比值，再寻求其上界。近期已有研究探讨了这个问题，并为一般范数和欧几里得范数建立了紧确上界。此外，文中还指出了它与 Gordon、Litvak、Meyer 和 Pajor 于 [9] 提出的不等式 $d_G(K, L) \leq n$ 之间可能存在的联系，以及该不等式猜想的等号情形 [11]；当 L 是椭球或平行多面体时，该不等式可直接由 [7] 中的界推出。尽管如此， [7] 并未试图刻画 K 的最优变换，而是采用体积极值位置。为了或许有助于理解更一般情形下与界 $d_G(K, L) \leq n$ 的联系，下面我们将刻画直径与内切半径之比的最优位置。在证明的最后部分，我们将直接应用定理 1.6，该定理将在第 5 节中得到验证。

对于下述推论，我们称 (y, z, b) 是 相对于 K 而言的直径三元组，条件是 L满足 $y, z \in K$, $b \in L^\circ$，并且

$$\langle y - z, b \rangle = \|y - z\|_L = D(K, L).$$

我们将集合 X、 $Y \subseteq \mathbb{R}^n$ 的闵可夫斯基和记为 $X + Y = \{x + y :\ \ x \in X,\ \ y \in Y\}$。

推论 3.4. 设 K、 $L \subseteq \mathbb{R}^n$ 为凸体，且 L 关于原点对称。I f

$$\frac{D(K,L)}{r(K,L)} = \min_{\det(A)\neq 0} \frac{D(AK,L)}{r(AK,L)}, \tag{7}$$

那么，对于任意满足 $r(K,L)L + c \subseteq K$的向量 $c \in \mathbb{R}^n$，存在整数 $N, M \geq 1$，以及 K 和 $r(K,L)L + c$的接触对$(x^1, a^1) \ldots, (x^N, a^N)$，还有相对于 L的 K的直径三元组(y^1, z^1, b^1) $\ldots, (y^M, z^M, b^M)$，以及权重 $\lambda_1, \ldots, \lambda_N, \mu_1, \ldots, \mu_M > 0$，使得对于任意 $x \in \mathbb{R}^n$,

$$\sum_{i=1}^{N} \lambda_i \langle x, a^i \rangle x^i = \sum_{j=1}^{M} \mu_j \langle x, b^j \rangle (y^j - z^j) \quad and \quad \sum_{i=1}^{N} \lambda_i a^i = \sum_{j=1}^{M} \mu_j b^j = 0.$$

如果 L 是一个椭球，那么反过来结论也成立，也就是说，如果对于任意 $c \in \mathbb{R}^n$ 满足$r(K,$ $L)L + c \subseteq K$，都存在如上所述的接触对、直径三元组和权重，那么(7) 成立。

证明。注意到一个众所周知的事实： $D(K,$ $L)$是满足 $K-K \subseteq \rho L$的最小实数 $\rho \geq 0$。因此，对于任意满足 $r(K,$ $L)L + c \subseteq K$ 的 $c \in \mathbb{R}^n$，有

$$\frac{1}{D(K,L)}(K-K) \subseteq L \subseteq \frac{1}{r(K,L)}(K-c). \tag{8}$$

我们声 称这条包含链是仿射最优的，当且仅当（7）成立 .

事实上，设 $A \in \mathbb{R}^{n\times n}$ 为可逆矩阵，并设 $r > 0$ 为极大且 $R > 0$ 为极小，使得存在向量 c', $d' \in \mathbb{R}^n$ 满足

$$r(K-K) + c' \subseteq AL \subseteq R(K-c) + d'.$$

$K - K$ 与 L 的关于原点的中心对称性表明，同样有 $r(K-K) \subseteq AL$，因此

$$r(A^{-1}K - A^{-1}K) \subseteq L \subseteq R(A^{-1}K - A^{-1}c) + A^{-1}d'.$$

现在， r的最大性与 R的最小性给出 $D(A^{-1}K,$ $L) = \frac{1}{r}$， $r(A^{-1}K,$ $L) = \frac{1}{R}$，以及

$$\frac{R}{r} = \frac{D(A^{-1}K,L)}{r(A^{-1}K,L)}.$$

若（8）是仿射最优的，则由定义有 $\frac{D(K,L)}{r(K,L)} \leq \frac{R}{r}$，且由 A 的任意性可得（7）。反之，若（8）不是仿射最优的，则可以选取 A、 r和 R，使得 $\frac{R}{r} < \frac{D(K,L)}{r(K,L)}$，从而（7）不成立。

验证了我们的断言后，由定理1.4 可知，(7)蕴含存在一个 Ader 分解，其依据是 L 与 $D\frac{1}{K(\cdot,L)}(K-K)$ 的接触对，以及 L 与 $r\frac{1}{K(\cdot,L)}(K-c)$ 的接触对。 L 与 $D\frac{1}{K(\cdot,L)}(K-K)$ 的接触对可以表示为相对于 L的 K 的直径三元组，因此所述分解的存在性得证。

若 L 是一个椭球，则定理1.6 表明，对于包含链（8），Ader 分解的存在足以保证其仿射最优性。根据上文所述的等价关系，此时可得（7）。

□

4 均值椭球定理

如引言所述，我们证明：当适当涉及椭球时，Ader 分解足以刻画包含链的仿射最优性。此证明依赖一项技术性结果，该结果允许我们在保持某些包含关系的同时，对椭球取（几何）均值。本节的目标大致是证明：若两个可能不同心的椭球包含于凸体 $K \subseteq \mathbb{R}^n$中（或包含凸体 K），那么这两个椭球的某种均值也包含于 K中（或也包含 $K \subseteq \mathbb{R}^n$）。下面我们分别处理这些内椭球和外椭球。我们的方法可视为对 [10, 引理 2.8]的推广；该引理仅讨论了以原点为中心的椭球。正如下文所示，对于内椭球，考虑一般位置的情形可以直接处理；但对于外椭球，这会带来显著的技术复杂性。

在以原点为中心的情形下，我们所用的（几何）均值椭球可以从凸体几何均值的一般理论中得到（例如参见 [18, 6]）。然而，由于针对我们的情形所需的分析比该理论所提供的更为细致，因此我们将在下文明确给出所有必要的定义。对于如下给出的椭球

$$E = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^{n} \frac{\langle x, v^i \rangle^2}{\alpha_i^2} \leq 1 \right\} \tag{9}$$

对于某个正交归一基 v^1、……、$v^n \in \mathbb{R}^n$和实数 α_1、……、$\alpha_n > 0$，权重为 $\lambda \in [0,1]$ 的 $\mathbb{B}^n$ 和 E 的均值椭球 E_λ定义为

$$E_\lambda = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^{n} \frac{\langle x, v^i \rangle^2}{\alpha_i^{2\lambda}} \leq 1 \right\}.$$

特别地，$E_0 = \mathbb{B}^n$ 是欧几里得球，且 $E_1 = E$。此外，我们通过以下方式将线性子空间 $\mathcal{V}_E$ 与 E 关联起来：

$$\mathcal{V}_E = \operatorname{lin}\{v^i : \alpha_i = 1, i = 1, \ldots, n\}. \tag{10}$$

显然，对于任意 $\lambda \in [0,1]$，交集 $E_\lambda \cap \mathcal{V}_E$ 都是 $\mathcal{V}_E$ 中的标准欧几里得球。

在下面的证明中，我们记 $h_K : \mathbb{R}^n \to R$ 为凸体$K \subseteq \mathbb{R}^n$的支撑函数。此外，我们反复使用加权几何平均—算术平均不等式，该不等式指出，对于 x，$y > 0$ 以及 $\lambda \in [0,1]$，有

$$x^{1-\lambda} y^\lambda \leq (1-\lambda)x + \lambda y.$$

当 $\lambda \in (0,\ 1)$ 时，等号成立当且仅当 $x = y$。特别地，若 $\lambda \in [0,1]$ 且 $\alpha > 0$，则分别令 $x := 1$ 和 $y := \alpha$ 代入上式可得

$$\alpha^\lambda \leq 1 - \lambda + \lambda\alpha, \tag{11}$$

而分别令 $x := \alpha^\lambda$ 和 $y := \alpha^{\lambda-1}$ 代入可得

$$1 - (1-\lambda)\alpha^\lambda \leq \lambda\alpha^{\lambda-1}. \tag{12}$$

当且仅当 $\lambda \in \{0,1\}$ 或 $\alpha = 1$时，两种情形下等号成立。

引理 4.1. 设向量 $v^1, \ldots, v^n \in \mathbb{R}^n$ 构成一组正交归一基，设 $c_0, c_1 \in \mathbb{R}^n$ 为向量，设 $\alpha_1, \ldots, \alpha_n > 0$ 和 $\lambda \in [0,1]$ 为实数，并设 $E \subseteq \mathbb{R}^n$ 为如 (9) 所示的椭球。则

$$E_\lambda + c_\lambda \subseteq (1-\lambda)(E_0 + c_0) + \lambda(E_1 + c_1),$$

其中 $c_\lambda = (1-\lambda)c_0 + \lambda c_1$。若 $\lambda \in (0,1)$，则还可以得到

$$(E_\lambda + c_\lambda) \cap \mathrm{bd}((1-\lambda)(E_0 + c_0) + \lambda(E_1 + c_1)) = \mathrm{bd}(E_\lambda + c_\lambda) \cap (\mathcal{V}_E + c_\lambda).$$

证明。要证明所述集合包含关系，只需证明 $E_\lambda \subseteq (1-\lambda)E_0 + \lambda E_1$。这等价于不等式

$$\begin{aligned}\left(\sum_{i=1}^n \alpha_i^{2\lambda}\langle a, v^i\rangle^2\right)^{\frac{1}{2}} &= h_{E_\lambda}(a) \le (1-\lambda)h_{E_0}(a) + \lambda h_{E_1}(a) \\ &= (1-\lambda)\left(\sum_{i=1}^n \langle a, v^i\rangle^2\right)^{\frac{1}{2}} + \lambda\left(\sum_{i=1}^n \alpha_i^2\langle a, v^i\rangle^2\right)^{\frac{1}{2}}\end{aligned}$$

对所有 $a \in \mathbb{R}^n$均成立。进而，这直接由式（11）和闵可夫斯基不等式推出，二者共同给出

$$\begin{aligned}\left(\sum_{i=1}^n \alpha_i^{2\lambda}\langle a, v^i\rangle^2\right)^{\frac{1}{2}} &\le \left(\sum_{i=1}^n \left((1-\lambda)\langle a, v^i\rangle + \lambda\alpha_i\langle a, v^i\rangle\right)^2\right)^{\frac{1}{2}} \\ &\le (1-\lambda)\left(\sum_{i=1}^n \langle a, v^i\rangle^2\right)^{\frac{1}{2}} + \lambda\left(\sum_{i=1}^n \alpha_i^2\langle a, v^i\rangle^2\right)^{\frac{1}{2}}.\end{aligned}$$

若 $\lambda \in (0,1)$，则上述估计取等价于 $\alpha_i = 1$ 对所有$i = 1$，……，n均成立，且 $\langle a, v^i\rangle \neq 0$，这是由式（11）的等号情形得到的。因此，$a$是一个支撑 E_λ和 $(1-\lambda)E_0 + \lambda E_1$ 的超平面的外法向量，且该超平面在公共边界点处支撑二者，当且仅当 a 属于子空间 $\mathcal{V}_E$。这立即推出关于边界的结论，即使再平移向量 c_λ之后也成立。 □

在下面的证明中应用上述引理时，我们总是有某个凸体 K 同时包含 $E_0 + c_0$ 和 $E_1 + c_1$（这里假设这两个椭球之一是欧几里得单位球，但其中心位置可以任意）。由 K 的凸性可知，对任意 $\lambda \in \in (0,\ 1)$，

$$E_\lambda + c_\lambda \subseteq (1-\lambda)(E_0 + c_0) + \lambda(E_1 + c_1) \subseteq \mathrm{conv}((E_0 + c_0) \cup (E_1 + c_1)) \subseteq K.$$

引理已经提供了一些关于 $E_\lambda + c_\lambda$和 K 的公共边界点的信息，因为这些点也必须是 $(1-\lambda)(E_0 + c_0) + \lambda(E_1 + c_1)$ 的边界点。事实上，我们还可以推断出更多信息。在这样的公共边界点处，共同支撑超平面的外法向量 a 必须满足

$$\begin{aligned}(1-\lambda)h_{E_0+c_0}(a) + \lambda h_{E_1+c_1}(a) &= h_{(1-\lambda)(E_0+c_0)+\lambda(E_1+c_1)}(a) \\ &= h_{\mathrm{conv}((E_0+c_0)\cup(E_1+c_1))}(a) = \max\{h_{E_0+c_0}(a), h_{E_1+c_1}(a)\}.\end{aligned}$$

若 $\lambda \in (0,1)$，则只有当 $h_{E_0+c_0}(a) = h_{E_1+c_1}(a)$ 时才可能成立。由于我们已知 a 属于线性子空间 $\mathcal{V}_E$，等式 $h_{E_0+c_0}(a) = h_{E_1+c_1}(a)$ 可简化为$\langle a, c_0\rangle = \langle a, c_1\rangle$。因此，向量 a 可能来源的空间进一步缩小为

$$\mathcal{V}_E \cap (\operatorname{lin}\{c_0 - c_1\})^{\perp}. \tag{13}$$

接下来，我们继续讨论一个与上述引理类似的结果，不过对象是外椭球。在这种情况下，取适当的均值在技术上更为复杂，因为包含关系

$$E_\lambda + c_\lambda \supseteq (E_0 + c_0) \cap (E_1 + c_1)$$

可能对线段 $[c_0, c_1]$ 上的所有 c_λ 取值都不成立，更不用说前一条引理中特定的取值了。$\mathbb{R}^2$ 中会出现这种情况的一个例子是 $v^1 = (1, 0)$、$v^2 = (0, 1)$、$\alpha_1 = 2$、$\alpha_2 = \frac{1}{4}$、$c_0 = (0, 0)$、$c_1 = (7\,5, 4\,5)$ 和 $\lambda = \frac{1}{2}$(见图1)。可以验证，$x = (-\frac{3}{5}, 4\,5) \in (E_0 \dotplus c_0) \cap (E_1 + c_1)$，然而对任意 $c_\lambda \in [c_0, c_1]$，$x \notin E_\lambda \dotplus \bar{c}_\lambda$。

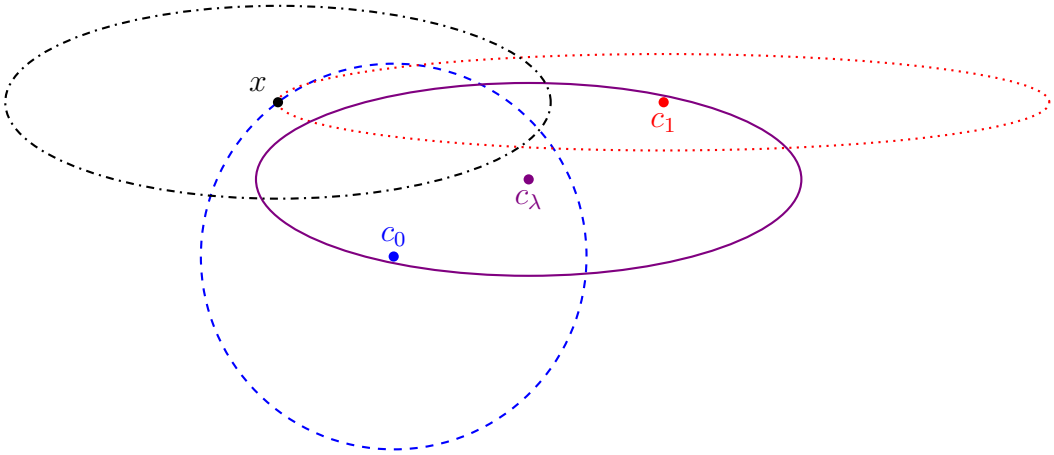

图1：一个示例，说明“自然候选”外均值椭球可能不满足所需的包含关系：$E_0 + c_0$（蓝色，虚线），$E_1 + c_1$（红色，点线），$E_\lambda + c_\lambda$（紫色，实线），$E_\lambda + x$（黑色，点划线）。

如图1 所示，主要困难在于找到均值椭球的适当平移。我们将看到，这仍然可以明确地做到，但公式比前一个结果中的公式更复杂。为此，我们需要以下技术引理。它是为一般情形下的外均值椭球找到适当平移的关键。

引理 4.2. 设 $\alpha > 0$、$\lambda \in [0, 1]$ 和 $\mu \in [0, 1]$ 是实数，且 $1 - (1 - \lambda)\alpha^\lambda \le \mu \le \lambda\alpha^{\lambda-1}$。那么对于任意的 x、$y \in \mathbb{R}$，

$$\frac{((1-\mu)x + \mu y)^2}{\alpha^{2\lambda}} \le (1-\lambda)x^2 + \lambda\frac{y^2}{\alpha^2}.$$

当且仅当 $\lambda \in \{0, 1\}$，或 $\alpha = 1$ 和 $x = y$，或 $x = y = 0$时，等号成立。

证明。 注意，对于固定的 $x, y \in \mathbb{R}$，所述不等式的左端作为 μ 的函数是凸函数。由于定义在紧区间上的凸函数会在某个端点处取得最大值，因此只需考虑 $\mu \in \{0, 1, 1-(1-\lambda)\alpha^{\lambda}, \lambda\alpha^{\lambda-1}\}$ 的情形。此外，所述不等式可改写为

$$\left((1-\lambda)\alpha^{2\lambda} - (1-\mu)^2\right) x^2 - 2\mu(1-\mu)xy + \left(\lambda\alpha^{2\lambda-2} - \mu^2\right) y^2 \geq 0. \tag{14}$$

x^2 和 y^2 的系数均非负，因为

$$(1-\mu)^2 \leq ((1-\lambda)\alpha^{\lambda})^2 \leq (1-\lambda)\alpha^{2\lambda} \quad \text{and} \quad \mu^2 \leq (\lambda\alpha^{\lambda-1})^2 \leq \lambda\alpha^{2\lambda-2} \tag{15}$$

根据关于 μ的假设，特别是，对14， $x = 0$ 或 $y = 0$，不等式成立。因此，我们可以除以 $y^2 > 0$ ，得到等价不等式

$$\left((1-\lambda)\alpha^{2\lambda} - (1-\mu)^2\right) t^2 - 2\mu(1-\mu)t + \left(\lambda\alpha^{2\lambda-2} - \mu^2\right) \geq 0 \tag{16}$$

对于 $t = -\frac{x}{y} \in \mathbb{R}$。由于首项系数非负，这等价于上文关于 t 的二次函数的判别式非正，即

$$\left((1-\lambda)\alpha^{2\lambda} - (1-\mu)^2\right)\left(\lambda\alpha^{2\lambda-2} - \mu^2\right) \geq \mu^2(1-\mu)^2.$$

如前所述，左端的因子均非负，因此若 $\mu = 0$ 或 $\mu = 1$，则无需证明。下面考虑 $\mu = \lambda\alpha^{\lambda-1}$的情形。只有当 $\lambda\alpha^{\lambda-1} \leq 1$时才会出现这种情况，否则，**1** 就是限制 μ的区间交集的上端点。所需的不等式现在可改写为

$$\left((1-\lambda)\alpha^{2\lambda} - \left(1-\lambda\alpha^{\lambda-1}\right)^2\right)\left(\lambda\alpha^{2\lambda-2} - \lambda^2\alpha^{2\lambda-2}\right) \geq \lambda^2\alpha^{2\lambda-2}\left(1-\lambda\alpha^{\lambda-1}\right)^2.$$

若 $\lambda = 0$，两边均等于 **0**。否则，两边同时除以 $\lambda\alpha^{2\lambda-2} > 0$ ，可简化为

$$\left((1-\lambda)\alpha^{2\lambda} - \left(1-\lambda\alpha^{\lambda-1}\right)^2\right)(1-\lambda) \geq \lambda\left(1-\lambda\alpha^{\lambda-1}\right)^2,$$

然后合并同类项，得到

$$(1-\lambda)^2\alpha^{2\lambda} \geq \left(1-\lambda\alpha^{\lambda-1}\right)^2,$$

最后再取平方根，得到

$$(1-\lambda)\alpha^{\lambda} \geq 1-\lambda\alpha^{\lambda-1}.$$

这不过就是不等式（12），由此证明了 $\mu = \lambda\alpha^{\lambda-1}$的情形。 此外， $\mu = 1 -$ （$1-\lambda$）α^{λ} 的情形可立即由上述情形应用于 $\lambda' := 1-\lambda$ 和$\alpha' := \alpha^{-1}$ 即可得出（注意，在此替换下，边界条件也彼此对应）。

接下来还需验证等号情形的刻画。容易验证，在所声称的三种情形中等号均成立，其中我们利用了由关于 μ的假设可知 $\lambda \in \{0, 1\}$ 蕴含 $\mu = \lambda$ 。余下只需证明在所有其他情形中不等式是严格的。由左端作为 μ的函数的凸性，我们只需对 μ的四个边界值验证严格性。

首先，注意到对于 $\lambda \in (0,\ 1)$，（15）中的两个第二不等式都是严格的。因此，（14）对于 $x = 0$ 与 $y \neq 0$，或 $x \neq 0$ 与 $y = 0$是严格的。此后我们可以假设$x,\ y$ 均非零，否则我们将落入所声称的第三种等号情形。

若 $\alpha = 1$，则（12）中的等式给出 $\mu = \lambda = \lambda\alpha^{\lambda-1}$，且由（ t ）得到的关于 16）的二次函数的判别式等于 0。因此，它有唯一实根。容易验证，这个唯一实根为 $t = 1$，对应于 $x = y$。换言之，（14）对于 $x \neq y$是严格的。

若 $\alpha \neq 1$，则（12）是严格的。根据上述推理，这意味着由（ t ）得到的关于 16）的二次函数，在 μ的所有四个边界值处判别式均为负，因此始终严格为正。由此可得，（14）对于非零 x、y均为严格不等式。证明完毕。 □

上述技术引理已经证明，现在我们可以处理外椭球的均值了。

引理 4.3。 设向量 v^1、……、$v^n \in \mathbb{R}^n$ 来自一个正交归一基，设 d_0、$d_1 \in \mathbb{R}^n$ 为向量，设 α_1、……、$\alpha_n > 0$ 和 $\lambda \in [0,1]$ 为实数，并设 $E \subseteq \mathbb{R}^n$ 为（(9)）中的椭球。对于 $i = 1$、……、n，令 $\mu_i \in [0,1]$ 满足 $1-(1-\lambda)\alpha_i^{\lambda} \leq \mu_i \leq \lambda\alpha_i^{\lambda-1}$。那么

$$(E_0 + d_0) \cap (E_1 + d_1) \subseteq E_\lambda + d_\lambda,$$

其中 $d_\lambda = \sum_{i=1}^n ((1-\mu_i)\langle d_0, v^i\rangle + \mu_i\langle d_1, v^i\rangle)v^i$。如果 $\lambda \in (0,\ 1)$，则此外

$$(E_0+d_0)\cap(E_1+d_1)\cap \mathrm{bd}(E_\lambda+d_\lambda) = \begin{cases} \mathrm{bd}(E_0+d_0)\cap \mathrm{bd}(E_1+d_1)\cap(\mathcal{V}_E+d_\lambda), & \text{if } d_0 = d_1, \\ \emptyset, & \text{if } d_0 \neq d_1. \end{cases}$$

请注意，即使 $d_0 = d_1$，$(E_0 + d_0) \cap (E_1 + d_1) \cap \mathrm{bd}(E_\lambda + d_\lambda)$ 仍可能为空。

证明。 设 $x \in \mathbb{R}^n$。则

$$\begin{aligned} \sum_{i=1}^n \frac{\langle x - d_\lambda, v^i\rangle^2}{\alpha_i^{2\lambda}} &= \sum_{i=1}^n \frac{\langle x - \sum_{j=1}^n ((1-\mu_j)\langle d_0, v^j\rangle + \mu_j\langle d_1, v^j\rangle)v^j, v^i\rangle^2}{\alpha_i^{2\lambda}} \\ &= \sum_{i=1}^n \frac{((1-\mu_i)\langle x - d_0, v^i\rangle + \mu_i\langle x - d_1, v^i\rangle)^2}{\alpha_i^{2\lambda}}. \end{aligned}$$

对每个被加项应用引理 4.2可得

$$\sum_{i=1}^n \frac{((1-\mu_i)\langle x - d_0, v^i\rangle + \mu_i\langle x - d_1, v^i\rangle)^2}{\alpha_i^{2\lambda}} \leq \sum_{i=1}^n (1-\lambda)\langle x - d_0, v^i\rangle^2 + \lambda\frac{\langle x - d_1, v^i\rangle^2}{\alpha_i^2}.$$

如果现在假设 $x \in (E_0 + d_0) \cap (E_1 + d_1)$，则

$$\begin{aligned} \sum_{i=1}^n (1-\lambda)\langle x - d_0, v^i\rangle^2 + \lambda\frac{\langle x - d_1, v^i\rangle^2}{\alpha_i^2} &= (1-\lambda)\sum_{i=1}^n \langle x - d_0, v^i\rangle^2 + \lambda\sum_{i=1}^n \frac{\langle x - d_1, v^i\rangle^2}{\alpha_i^2} \\ &\leq (1-\lambda) + \lambda = 1. \end{aligned} \tag{17}$$

综上，我们得出 $(E_0 + d_0) \cap (E_1 + d_1) \subseteq E_\lambda + d_\lambda$。

现在，假设 $\lambda \in (0,1)$，且 $x \in (E_0 + d_0) \cap (E_1 + d_1) \cap \mathrm{bd}(E_\lambda + d_\lambda)$。那么式 (17) 中必须取等号，这立即表明 $x \in \mathrm{bd}(E_0 + d_0) \cap \mathrm{bd}(E_1 + d_1)$。接下来，对于每个 $i = 1, \ldots, n$，应用引理 4.2 时也必须取等号。由于 $\lambda \in (0,1)$，这意味着对于 $\alpha_i = 1$，有 $\langle x - d_0, v^i \rangle = \langle x - d_1, v^i \rangle$；对于 $\alpha_i \neq 1$，有 $\langle x - d_0, v^i \rangle = \langle x - d_1, v^i \rangle = 0$。在这两种情况下，我们也立即得到 $\langle d_0, v^i \rangle = \langle d_1, v^i \rangle$。向量 $v^1, \ldots, v^n$ 构成了 $\mathbb{R}^n$ 的正交归一基，因此可推出 $d_0 = d_1 = d_\lambda$。此外，我们有 $x - d_\lambda = \sum_{i=1}^n \langle x - d_\lambda, v^i \rangle v^i$，其中系数 $\langle x - d_\lambda, v^i \rangle = \langle x - d_0, v^i \rangle$ 在 $\alpha_i \neq 1$ 时为零。由此可得 $x \in \mathcal{V}_E + d_\lambda$。

最后，我们需要证明，当 $d_0 = d_1$ 时，$x \in \mathrm{bd}(E_0 + d_0) \cap \mathrm{bd}(E_1 + d_1) \cap (\mathcal{V}_E + d_\lambda)$ 是 $E_\lambda + d_\lambda$ 的边界点。这意味着要证明对于每个 $i = 1, \ldots\ldots, n$，引理 4.2 中等号成立；对于式 (17)，等号显然成立。对于前者，注意到根据 $d_0 = d_1$，对所有 $i = 1, \ldots\ldots, n$，都有 $\langle x - d_0, v^i \rangle = \langle x - d_1, v^i \rangle$。若 $\alpha_i \neq 1$，则 $x - d_0 = x - d_\lambda \in \mathcal{V}_E$ 还表明 $\langle x - d_0, v^i \rangle = 0$。因此，对于每个 $i = 1, \ldots\ldots, n$，引理 4.2 的某个等号情形均适用，证明完毕。□

在处理完内均值椭球和外均值椭球后，我们现在可以用本节的主要结果作结。该结果考虑用椭球同时逼近两个可能不同的凸体，并汇总了本节中确立的所有事实。需要指出的是，将该结果应用于一般椭球，只需进行适当的归一化。

定理 *4.4.* 设 $v^1, \ldots, v^n \in \mathbb{R}^n$ 是正交归一基中的向量，$c_0, c_1, d_0, d_1 \in \mathbb{R}^n$ 是向量，$\alpha_1, \ldots, \alpha_n, r_0, r_1, R_0, R_1 > 0$ 和 $\lambda \in [0,1]$ 是实数，且 $E \subseteq \mathbb{R}^n$ 是式 *(9)* 中的椭球。对于 $i = 1, \ldots, n$，令 $\mu_i \in [0,1]$ 满足 $1 - (1-\lambda)(\frac{R_1}{R_0}\alpha_i)^\lambda \leq \mu_i \leq \lambda(\frac{R_1}{R_0}\alpha_i)^{\lambda - 1}$。若 $K, L \subseteq \mathbb{R}^n$ 是满足下列条件的凸体

$$(r_0 E_0 + c_0) \cup (r_1 E_1 + c_1) \subseteq K \quad \textit{and} \quad L \subseteq (R_0 E_0 + d_0) \cap (R_1 E_1 + d_1),$$

then

$$r_\lambda E_\lambda + c_\lambda \subseteq K \quad \textit{and} \quad L \subseteq R_\lambda E_\lambda + d_\lambda,$$

其中 $c_\lambda = (1-\lambda)c_0 + \lambda c_1$，$d_\lambda = \sum_{i=1}^n ((1-\mu_i)\langle d_0, v^i \rangle + \mu_i \langle d_1, v^i \rangle) v^i$，$r_\lambda = r_0^{1-\lambda} r_1^\lambda$ 和 $R_\lambda = R_0^{1-\lambda} R_1^\lambda$。若 $\lambda \in (0,1)$，则还满足

(i) $r_\lambda E_\lambda + c_\lambda$ 和 K 的任一公共边界点都包含在仿射子空间 $(\mathrm{lin}\{v^i : \alpha_i = \frac{r_0}{r_1}, i = 1, \ldots, n\} \cap (\mathrm{lin}\{c_0 - c_1\})^\perp) + c_\lambda$ 中，并且

(ii) $R_\lambda E_\lambda + d_\lambda$ 和 L 的任一公共边界点都包含在仿射子空间 $\mathrm{lin}\{v^i : \alpha_i = \frac{R_0}{R_1}, i = 1, \ldots, n\} + d_\lambda$ 中，并且也是 $R_0 E_0 + d_0$ 和 $R_1 E_1 + d_1$ 的边界点。若存在这样的点，则 $d_0 = d_1$。

证明。我们从内椭球开始。为此，我们对所有 $\lambda \in [0,1]$ 定义椭球

$$F_\lambda = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n \frac{\langle x, v^i \rangle^2}{(\frac{r_1}{r_0}\alpha_i)^{2\lambda}} \leq 1 \right\}.$$

于是 $F_\lambda = (\frac{r_1}{r_0})^\lambda E_\lambda$，因为这些椭球具有相同的主轴，且 F_λ 与 E_λ 的半轴长相差一个因子 $(\frac{r_1}{r_0})^\lambda$。因此，引理 4.1 得出

$$K \supseteq \operatorname{conv}((r_0E_0 + c_0) \cup (r_1E_1 + c_1)) = r_0 \operatorname{conv}\left(\left(F_0 + \frac{1}{r_0}c_0\right) \cup \left(F_1 + \frac{1}{r_0}c_1\right)\right)$$
$$\supseteq r_0\left(F_\lambda + \frac{1-\lambda}{r_0}c_0 + \frac{\lambda}{r_0}c_1\right) = r_0^{1-\lambda}r_1^\lambda E_\lambda + (1-\lambda)c_0 + \lambda c_1 = r_\lambda E_\lambda + c_\lambda.$$

此外，(13) 之前的讨论表明，在第一组和最后一组的接触点处，共同支撑超平面的外法向量属于 $(\mathcal{V}_{F_1} \cap (\operatorname{lin}\{\frac{1}{r_0}c_0 - \frac{1}{r_0}c_1\})^\perp)$。这意味着该公共边界点本身必须位于 $r_\lambda E_\lambda + c_\lambda$ 中与此子空间平行的中心截面内。(i) 中的断言得证。

下面讨论关于外接椭球的断言。对于 $\lambda \in [0,1]$，我们定义椭球 F'_λ，如上所述 F_λ，但用 R_0 和 R_1 分别替换 r_0 和 r_1。类似地，我们有 $F'_\lambda = (\frac{R_1}{R_0})^\lambda E_\lambda$。因此，应用引理 4.3 可得

$$L \subseteq (R_0E_0 + d_0) \cap (R_1E_1 + d_1) = R_0\left(\left(F'_0 + \frac{1}{R_0}d_0\right) \cap \left(F'_1 + \frac{1}{R_0}d_1\right)\right)$$
$$\subseteq R_0\left(F'_\lambda + \sum_{i=1}^n \left((1-\mu_i)\left\langle \frac{1}{R_0}d_0, v^i\right\rangle + \mu_i\left\langle \frac{1}{R_0}d_1, v^i\right\rangle\right)v^i\right)$$
$$= R_0^{1-\lambda}R_1^\lambda E_\lambda + \sum_{i=1}^n ((1-\mu_i)\langle d_0, v^i\rangle + \mu_i\langle d_1, v^i\rangle)v^i = R_\lambda E_\lambda + d_\lambda.$$

最后，引理 4.3还表明，第一组和最后一组的任意公共边界点也是 $R_0(F'_0 + \frac{1}{R_0}d_0)$ 和 $R_0(F'_1 + \frac{1}{R_0}d_1)$ 的边界点，位于子空间 $R_0(\mathcal{V}_{F'_1} + \sum_{i=1}^n ((1-\mu_i)\langle \frac{1}{R_0}d_0, v^i\rangle + \mu_i\langle \frac{1}{R_0}d_1, v^i\rangle)v^i)$ 中，并且其存在意味着 $\frac{1}{R_0}d_0 = \frac{1}{R_0}d_1$。这就化简为 (ii) 中的附加断言。 □

5 涉及椭球的仿射最优位置的刻画

在本节中，我们证明定理 1.5 与 1.6，它们将定理 1.2 推广到非对称情形。特别地，我们得到任意凸体 $K \subseteq \mathbb{R}^n$ 相对于欧几里得球的最优 Banach–Mazur 位置的完整刻画。

我们首先从定理1.6 出发，并随后以直接的方式导出定理 1.5 。非分离条件 (ii) 与分解形式 (iii) 对于仿射最优性的必要性，主要来自适配于欧氏情形的定理 1.4 。具体针对 (ii)，若直接移植该条件，则会得到一个略有不同的不等式

$$\left\langle \frac{y}{1-\langle y, c\rangle}, Ay + v\right\rangle \le \left\langle \frac{z}{1-\langle z, d\rangle}, Az + w\right\rangle.$$

这是因为，对于任意其他凸体，与 $\mathbb{B}^n$ 相接触的外法向量始终是公共边界点的正倍数，而定理 1.4 (i)规定了

外法向量的缩放。与定理1.6 (ii)所述版本的差异，主要在于非分离条件可以用若干不同但等价的形式表述；我们选择了最适合定理1.4 和1.6各自语境的形式。尽管如此，我们仍可将这些分解作为中间步骤，推导出所需的形式。要证明条件(ii)和(iii)对于仿射最优性也具有充分性，关键在于关于椭球的均值的定理4.4 。

定理 *1.6* 的证明。 在整个证明过程中，我们不妨假设 $0 \in \operatorname{int}(L_1) \cap \operatorname{int}(L_2)$，否则我们可以平移 L_1 和 L_2，同时相应地调整 c 和 d 。注意，这不会影响集合$\operatorname{bd}(\mathbb{B}^n) \cap \operatorname{bd}(rL_1 + c)$ 和$\operatorname{bd}(\mathbb{B}^n) \cap \operatorname{bd}(RL_2 + d)$。

现在，假设定理 1.6 的条件 (i) 成立。那么，我们可以取定理 1.4 (ii) 中给出的接触对和权重。由于支撑 $\mathbb{B}^n$ 且经过其边界点的超平面的外法向量，唯一可能是这些边界点的正倍数，因此，对于所有 $a^i = \alpha_i y^i$， $i = 1, \ldots, N$以及 $b^j = \beta_j z^j$对于所有 $j = 1, \ldots, M$，其中 $\alpha_1, \ldots, \alpha_N, \beta_1, \ldots, \beta_M > 0$ 均为实数。因此，将所有权重 λ_i 替换为 $\lambda_i \alpha_i > 0$，并将 μ_j 替换为 $\mu_j \beta_j > 0$，即可得到 (iii) 中所需的分解。

接下来，假设存在一个如 (iii) 所示的分解。此时所声称的等式 $\sum_{i=1}^{N} \lambda_i = \sum_{j=1}^{M} \mu_j$ 可直接通过比较该分解中矩阵的迹得出。此外，通过改进定理 1.4中类似界限的论证，可以得到界限 $N + M \leq {}^{n}{}_{(n+5)}{}_{2} + 1$ 。这些集合 ——

$$\mathcal{Y} = \{(yy^T | y) \in \mathbb{R}^{n \times (n+1)} : y \in \operatorname{bd}(\mathbb{B}^n) \cap \operatorname{bd}(rL_1 + c)\}$$

and

$$\mathcal{Z} = \{(zz^T | z) \in \mathbb{R}^{n \times (n+1)} : z \in \operatorname{bd}(\mathbb{B}^n) \cap \operatorname{bd}(RL_2 + d)\},$$

位于空间 $\mathcal{M}_s^n \times \mathbb{R}^n$ 中一个超平面内，该超平面由迹为 1 的矩阵部分给出。等式 $\sum_{i=1}^{N} \lambda_i = \sum_{j=1}^{M} \mu_j$ 表明，我们可以将所有权重按同一因子缩放，从而假设这两个和都等于 1，这意味着 $\mathcal{Y}$ 和 $\mathcal{Z}$的凸包交于一点，该点位于向量部分也为零的子空间中。由于 $\dim(\mathcal{M}_s^n) = \underline{{}^{n}{}_{(n+1)}{}_{2}}$，引理 2.2 中的注记 2.1 现在使我们可以将分解中的点总数减少到

$$N + M \leq 2\left(\left(\frac{n(n+1)}{2} - 1 + n\right) + 1\right) - \left(\frac{n(n+1)}{2} - 1\right) = \frac{n(n+5)}{2} + 1.$$

此外，对于任意矩阵 $A \in \mathcal{M}_s^n$ 和任意向量 v， $w \in \mathbb{R}^n$ ，我们得到 t

$$\min_{i=1,\ldots,N} \langle y^i, Ay^i + v \rangle \leq \sum_{i=1}^{N} \lambda_i \langle y^i, Ay^i + v \rangle = \sum_{j=1}^{M} \mu_j \langle z^j, Az^j + w \rangle \leq \max_{j=1,\ldots,M} \langle z^j, Az^j + w \rangle,$$

从而证明了 (ii)。

剩下要证明的是从 (ii) 推出 (i)。为简单起见，我们可以通过考虑 $L_1' := rL_1$ 和 $L_2' := RL_2$，假设 $r = R = 1$ 。因此，我们有

$$L_1 + c \subseteq \mathbb{B}^n \subseteq L_2 + d, \tag{18}$$

并假设用反证法证明该包含链不是仿射最优的。我们的目标是证明，对矩阵 $A \in \mathcal{M}_s^n$ 和向量 v，$w \in \mathbb{R}^n$，存在某些适当的取值，使得 (ii) 不成立；也就是说，所有接触点 $y \in \mathrm{bd}(\mathbb{B}^n) \cap \mathrm{bd}(L_1 + c)$ 和 $z \in \mathrm{bd}(\mathbb{B}^n) \cap \mathrm{bd}(L_2 + d)$ 都满足

$$\langle y, Ay + v\rangle > \langle z, Az + w\rangle.$$

若（18）不是仿射最优的，则存在一个以原点为中心的椭球 $E \subseteq \mathbb{R}^n$、向量 c'，$d' \in \mathbb{R}^n$ 以及一个实数 $\rho > 1$ 使得

$$\rho(L_1 + c') \subseteq E \subseteq L_2 + d'.$$

对 $\lambda \in [0, 1]$ 应用定理 4.4 可得

$$\rho^\lambda(L_1 + c_\lambda) \subseteq E_\lambda \subseteq L_2 + d_\lambda, \tag{19}$$

其中，定义 c_λ 时的参数 μ_i 可在各自允许的范围内任意选取，且 $d_\lambda = (1-\lambda)\, d + \lambda d'$（注意，与定理4.4相比，此处字母 c和 d的角色互换了）。

接下来，取一个对称正定矩阵 $S_\lambda \in \mathcal{M}_s^n$，使得 $\|x\|_{E_\lambda}^2 = \langle x,\ S_\lambda x\rangle$对所有 $x \in \mathbb{R}^n$成立。于是，由（19）可得，对于任意 $x \in L_1 + c_\lambda$，

$$\langle x, S_\lambda x\rangle = \|x\|_{E_\lambda}^2 \le \rho^{-2\lambda}, \tag{20}$$

并且对于任意 $x \in \mathrm{bd}(L_2 + d_\lambda)$,

$$\langle x, S_\lambda x\rangle = \|x\|_{E_\lambda}^2 \ge 1. \tag{21}$$

取 $A_\lambda := I_n - S_\lambda \in \mathcal{M}_s^n$以及 $v_\lambda := -2S_\lambda (c_\lambda - c)$，（20）表明，对于 $y \in \mathrm{bd}(\mathbb{B}^n)$, $\cap\, \mathrm{bd}(L_1 + c)$有

$$\begin{aligned} 1 - \rho^{-2\lambda} &\le \langle y, y\rangle - \langle y - c + c_\lambda, S_\lambda(y - c + c_\lambda)\rangle \\ &= \langle y, y\rangle - \langle y, S_\lambda y\rangle - 2\langle y, S_\lambda(c_\lambda - c)\rangle - \langle c_\lambda - c, S_\lambda(c_\lambda - c)\rangle \\ &= \langle y, A_\lambda y + v_\lambda\rangle - \|c_\lambda - c\|_{E_\lambda}^2 \le \langle y, A_\lambda y + v_\lambda\rangle. \end{aligned}$$

类似地，s 置 $w_\lambda := -2S_\lambda(d_\lambda - d)$ 并利用（21）可得，对于 $z \in \mathrm{bd}(\mathbb{B}^n) \cap \mathrm{bd}(L_2 + d)$ 该 at

$$\begin{aligned} 0 &\ge \langle z, z\rangle - \langle z - d + d_\lambda, S_\lambda(z - d + d_\lambda)\rangle \\ &= \langle z, z\rangle - \langle z, S_\lambda z\rangle - 2\langle z, S_\lambda(d_\lambda - d)\rangle - \langle d_\lambda - d, S_\lambda(d_\lambda - d)\rangle \\ &= \langle z, A_\lambda z + w_\lambda\rangle - \|d_\lambda - d\|_{E_\lambda}^2 = \langle z, A_\lambda z + w_\lambda\rangle - \lambda^2\|d' - d\|_{E_\lambda}^2. \end{aligned}$$

由紧致性，存在某个 $C > 1$，与 λ 无关，且对所有$\lambda \in [0, 1]$满足 $\|d' - d\|_{E_\lambda}^2 \le C$。综合上述结果，因此可得

$$\langle y, A_\lambda y + v_\lambda\rangle \ge 1 - \rho^{-2\lambda} \ge 1 - \rho^{-2\lambda} - C\lambda^2 + \langle z, A_\lambda z + w_\lambda\rangle$$

对所有 $\lambda \in [0, 1]$，$y \in \mathrm{bd}(\mathbb{B}^n) \cap \mathrm{bd}(L_1 + c)$，以及 $z \in \mathrm{bd}(\mathbb{B}^n) \cap \mathrm{bd}(L_2 + d)$。最后，我们注意到映射 $\lambda \mapsto 1 - \rho^{-2\lambda}$是凹的。因此，选取 $\lambda \in \left(0, \frac{1-\rho^{-2}}{C}\right)$（由 $\rho > 1$可知这是可能的），可得

$$1 - \rho^{-2\lambda} \ge (1 - \rho^{-2})\lambda > C\lambda^2.$$

总之，我们看到，只要 $\lambda > 0$ 足够小，选取 A_λ、v_λ 和 w_λ 时，条件 (ii) 就不成立，这正是我们所要的。 □

定理 1.5 的证明。初始假设 $r\mathbb{B}^n + c \subseteq K \subseteq R\mathbb{B}^n + d$ 可改写为$\frac{1}{R}(K-d) \subseteq \mathbb{B}^n \subseteq \frac{1}{r}(K-c)$。此外，等式 $d_{BM}(K, \mathbb{B}^n) = \frac{R}{r}$等价于后一个包含链是仿射最优的。换言之，在 1.5 定理的情形下，定理 1.6 中的第一个条件 $L_1 = L_2 = K$等价于定理 1.5 中的第一个条件。我们将证明，在这种情形下，这些定理中相应的条件 (ii) 和 (iii) 也等价。

事实上，假设定理 1.6 中的 (ii) 成立，即对任意 $A \in \mathcal{M}_s^n$ 和 $v,\ w \in \mathbb{R}^n$ 存在接触点 $y \in \mathrm{bd}(\mathbb{B}^n) \cap \mathrm{bd}(\frac{1}{R}(K-d))$ 和 $z \in \mathrm{bd}(\mathbb{B}^n) \cap \mathrm{bd}(\frac{1}{r}(K-c))$ 使得

$$\langle y, Ay + v\rangle \leq \langle z, Az + w\rangle.$$

将此应用于 $A' = -A$、$v' = -w$ 和 $w' = -v$，可得 $y' \in \mathrm{bd}(\mathbb{B}^n) \cap \mathrm{bd}(\frac{1}{R}(K-d))$ 和 $z' \in \mathrm{bd}(\mathbb{B}^n) \cap \mathrm{bd}(\frac{1}{r}(K-c))$，其中

$$\langle z', Az' + v\rangle = -\langle z', A'z' + w'\rangle \leq -\langle y', A'y' + v'\rangle = \langle y', Ay' + w\rangle.$$

因此，$y = rz' + c \in \mathrm{bd}(K) \cap \mathrm{bd}(r\mathbb{B}^n + c)$ 和 $z = Ry' + d \in \mathrm{bd}(K) \cap \mathrm{bd}(R\mathbb{B}^n + d)$ 满足

$$\left\langle \frac{y-c}{r}, A\left(\frac{y-c}{r}\right) + v \right\rangle = \langle z', Az' + v\rangle \leq \langle y', Ay' + w\rangle \leq \left\langle \frac{z-d}{R}, A\left(\frac{z-d}{R}\right) + w \right\rangle.$$

这表明定理 1.5 中的 (ii) 同样成立。将上述所有步骤逆转，即可得到反向蕴含。

关于两个定理的条件 (iii)，同样的论证表明，对于形如 $\frac{1}{R}(K-d) \subseteq \mathbb{B}^n \subseteq \frac{1}{r}(K-c)$ 的包含关系，其分解条件 (iii)

$$\sum_{i=1}^{N} \lambda_i y^i (y^i)^T = \sum_{j=1}^{M} \mu_j z^j (z^j)^T \quad \text{and} \quad \sum_{i=1}^{N} \lambda_i y^i = \sum_{j=1}^{M} \mu_j z^j = 0$$

对应于形如

$$\sum_{i=1}^{N} \lambda_i (y^i - c)(y^i - c)^T = \sum_{j=1}^{M} \mu_j (z^j - d)(z^j - d)^T \quad \text{and} \quad \sum_{i=1}^{N} \lambda_i (y^i - c) = \sum_{j=1}^{M} \mu_j (z^j - d) = 0$$

的分解，对于包含关系 $r\mathbb{B}^n + c \subseteq K \subseteq R\mathbb{B}^n + d$。注意，如有必要，重新缩放因子总是可以吸收到权重中。证明至此完成。 □

最后，在本节结束时，我们指出，对于椭球 $E \subseteq \mathbb{R}^n$，包含链 $rE + c \subseteq K \subseteq RE + d$ 或 $rL_1 + c \subseteq E \subseteq RL_2 + d$ 是仿射最优的，当且仅当 E 是局部最优椭球（即，不存在一个与 $E' \subseteq \mathbb{R}^n$ 足够接近的椭球 E 能使比值小于 $\frac{R}{r}$）。直接证明这一点可借助定理 4.4。另一种方法是注意到，局部最优性足以证明定理 1.4，而定理 1.5 和 1.6则表明，由此得到的条件反过来也足以保证全局最优性。

6 Banach–Mazur 距离到欧几里得球的应用

本文最后一节将讨论定理 1.5 和 4.4 关于 Banach–Mazur 距离至欧几里得球的应用。首先，我们将说明如何从定理 1.5推出已知的欧几里得球距离的紧致一般界。关于中心对称凸体 $d_{BM}(K,\ \mathbb{B}^n) \leq \sqrt{n}$ 的不等式$K \subseteq \mathbb{R}^n$ 的另一种证明已在 [10, 推论 2.2] 中作为定理1.2的推论给出。类似地，根据非对称情形下距离位置的刻画，我们给出任意凸体 $d_{BM}(K,\ \mathbb{B}^n) \leq n$ 所满足的不等式 $K \subseteq \mathbb{R}^n$的另一种证明。

John椭球定理给出了 $d_{BM}(K,\mathbb{B}^n) \leq n$ 的经典证明，该证明归功于 John [12]。也可以根据 Gordon、Litvak、Meyer 和 Pajor 提出的更一般的不等式 $d_G(K,L) \leq n$ 推导出这一估计（参见 [9, 定理 5.1]）。据我们所知，我们基于定理 1.5 的方法是首个完全避免体积概念的方法。尽管我们的证明可能比采用 John椭球定理的经典证明稍显复杂，但其优点在于能够直接刻画出单纯形是唯一达到等号的凸体。标准方法则并非如此，通常需要在不等式本身的证明之外另行论证。Leichtweiss [16], 首先给出了这类证明，后来 Palmon [19], 也给出了另一个证明，但他似乎并不知道 Leichtweiss 先前的德文研究。Jiménez 和 Naszódi [11], 的一项结果也导出了另一种证明；他们分析了 [9] 中估计 $d_G(K,L) \leq n$ 的等号情形。最近，有人将 Jung 不等式与欧几里得直径内切半径比方面的结果结合起来，给出了一个证明 [7]。

推论 6.1. 设 $K \subseteq \mathbb{R}^n$ 是一个凸体。则

$$d_{BM}(K,\mathbb{B}^n) \leq n,$$

当且仅当 K 是一个单纯形时，等号成立。

证明。通过施加适当的仿射变换，不失一般性地设 $\mathbb{B}^n \subseteq K \subseteq R\mathbb{B}^n + d$，其中 $d \in \mathbb{R}^n$ 且 $R = d_{BM}(K,\mathbb{B}^n)$。由定理 1.5，存在整数 N、$M \geq 1$，内接触点 y^1、……、$y^N \in \mathrm{bd}(K) \cap \mathrm{bd}(\mathbb{B}^n)$，外接触点 z^1、……、$z^M \in \mathrm{bd}(K) \cap \mathrm{bd}(R\mathbb{B}^n + d)$，以及权重 λ_1、……、λ_N、μ_1、……、$\mu_M > 0$，使得

$$\sum_{i=1}^{N} \lambda_i y^i (y^i)^T = \sum_{j=1}^{M} \mu_i (z^j - d)(z^j - d)^T = A \quad \text{and} \quad \sum_{i=1}^{N} \lambda_i y^i = \sum_{j=1}^{M} \mu_i (z^j - d) = 0 \tag{22}$$

对于某个矩阵 $A \in \mathcal{M}_s^n$。此外，我们还有

$$\mathrm{Tr}(A) = \sum_{i=1}^{N} \lambda_i \langle y^i, y^i \rangle = \sum_{i=1}^{N} \lambda_i = R^2 \sum_{j=1}^{M} \mu_j > 0 \tag{23}$$

并且可以假设，对于 $i \neq j$，$y^i \neq y^j$；类似地，对于 $i \neq j$，$z^i \neq z^j$。

对于任意 $i=1$、……、N，y^i 是 K 和 $\mathbb{B}^n$ 的公共边界点，因此它也是支撑 K 于 y^i 处的超平面的外法向量。特别地，$z^j \in K$ 得出 $\langle y^i,\ z^j\rangle \le \langle y^i,\ y^i\rangle = 1$，其中 $j=1$、……、M。柯西–施瓦茨等式还表明 $\langle y^i,\ z^j - d\rangle \ge -R$，因此

$$
\begin{aligned}
0 &\le \sum_{i=1}^{N}\sum_{j=1}^{M} \lambda_i\mu_j(1-\langle y^i, z^j\rangle)(R+\langle y^i, z^j-d\rangle) \qquad (24)\\
&= R\sum_{i=1}^{N}\lambda_i\sum_{j=1}^{M}\mu_j + \left\langle \sum_{i=1}^{N}\lambda_i y^i, \sum_{j=1}^{M}\mu_j((1-R)z^j-d)\right\rangle - \sum_{i=1}^{N}\lambda_i\left\langle y^i, \sum_{j=1}^{M}\mu_j\langle y^i, z^j-d\rangle z^j\right\rangle.
\end{aligned}
$$

利用（22）和（23），可将其简化为

$$
\begin{aligned}
0 &\le \frac{\operatorname{Tr}(A)^2}{R} - \sum_{i=1}^{N}\lambda_i\langle y^i, Ay^i\rangle = \frac{\operatorname{Tr}(A)^2}{R} - \sum_{i=1}^{N}\lambda_i \operatorname{Tr}(y^i(y^i)^T A)\\
&= \frac{\operatorname{Tr}(A)^2}{R} - \operatorname{Tr}\left(\sum_{i=1}^{N}\lambda_i y^i(y^i)^T A\right) = \frac{\operatorname{Tr}(A)^2}{R} - \operatorname{Tr}(A^2).
\end{aligned}
$$

重新整理，并对 $(\mathcal{M}_s^n)^2$ 应用弗罗贝尼乌斯内积下的柯西–施瓦茨不等式，可得

$$
R \le \frac{\operatorname{Tr}(A)^2}{\operatorname{Tr}(A^2)} = \frac{\langle I_n, A\rangle_F^2}{\operatorname{Tr}(A^2)} \le \frac{\langle I_n, I_n\rangle_F\langle A, A\rangle_F}{\operatorname{Tr}(A^2)} = \frac{\operatorname{Tr}(I_n^2)\operatorname{Tr}(A^2)}{\operatorname{Tr}(A^2)} = n. \qquad (25)
$$

接下来只需证明，在等号情形下，K 必须是一个单纯形。

如果等号 $R=n$ 成立，那么特别地，在（25）中应用柯西–施瓦茨不等式时也取等号，这意味着半正定矩阵 A是单位矩阵的正倍数，即 $A = \frac{\operatorname{Tr}(A)}{n} I_n$。此外，（24）中也取等号，这意味着对所有 $i=1$、……、N以及 $j=1$、……、M，

$$
\langle y^i, z^j\rangle = 1 \quad \text{or} \quad \langle y^i, z^j - d\rangle = -R = -n. \qquad (26)
$$

在后一种情况成立时，柯西–施瓦茨不等式的等号情形表明 $z^j - d = -ny^i$。

现在，反设对于某个固定的 i，有 $\langle y^i,\ z^j\rangle = 1$ 对所有 j都成立。那么，由（22），

$$
0 = \left\langle y^i, \sum_{j=1}^{M}\mu_j(z^j-d)\right\rangle = \sum_{j=1}^{M}\mu_j\langle y^i, z^j-d\rangle = (1-\langle y^i, d\rangle)\sum_{j=1}^{M}\mu_j
$$

这将表明 $\langle y^i,\ d\rangle = 1$ 并因此 $\langle y^i,\ z^j - d\rangle = 0$ 对所有 j 都成立。然而，现在利用$A = \frac{\operatorname{Tr}(A)}{n} I_n$ 即可导出所需的矛盾

$$
0 \ne y^i = \frac{n}{\operatorname{Tr}(A)}Ay^i = \frac{n}{\operatorname{Tr}(A)}\sum_{j=1}^{n}\mu_j\langle y^i, z^j-d\rangle(z^j-d) = 0.
$$

因此，对于任意 i，存在 j 使得 $\langle y^i,\ z^j-d\rangle=-n$ 从而 $z^j-d=-ny^i$。类似地，对于固定的 j，由式（22）可得

$$0=\left\langle\sum_{i=1}^{N}\lambda_i y^i,z^j\right\rangle=\sum_{i=1}^{N}\lambda_i\langle y^i,z^j\rangle, \tag{27}$$

因此存在某个 i，使得 $\langle y^i,z^j\rangle\le 0$。在这种情况下，根据（26），必有 $\langle y^i,z^j-d\rangle=-n$ 和 $z^j-d=-ny^i$。由于我们假设 y^i 和 z^j 分别两两不同，综上可得 $N=M$；适当重新编号后，对所有 $i=1,\dots,N$，都有 $z^i-d=-ny^i$。这也意味着，当 $i\neq j$时，$-ny^i\neq -ny^j=z^j-d$，因此由（26）可知，对所有 $i\neq j$，都有$\langle y^i,z^j\rangle=1$。将这些内积代入（27），可得对任意 i，

$$0=-n\lambda_i+\lambda_i\langle y^i,d\rangle+\sum_{\substack{j=1\\ j\neq i}}^{N}\lambda_j=-n\lambda_i+\lambda_i\langle y^i,d\rangle+\operatorname{Tr}(A)-\lambda_i$$

因此

$$\lambda_i\langle y^i,d\rangle=\lambda_i(n+1)-\operatorname{Tr}(A).$$

对所有 i求和，由（22）可得

$$0=\left\langle\sum_{i=1}^{N}\lambda_i y^i,d\right\rangle=\sum_{i=1}^{N}\lambda_i\langle y^i,d\rangle=\sum_{i=1}^{N}\left(\lambda_i(n+1)-\operatorname{Tr}(A)\right)=(n+1)\operatorname{Tr}(A)-N\operatorname{Tr}(A),$$

因此 $N=n+1$。对于 $i\neq j$，我们也得到

$$\langle y^i,y^j\rangle=\left\langle y^i,\frac{d-z^j}{n}\right\rangle=\frac{\langle y^i,d\rangle-1}{n}=1-\frac{\operatorname{Tr}(A)}{n\lambda_i}.$$

交换 i 和 j 的角色不会改变 $\langle y^i,\ y^j\rangle$，但会将右侧的 λ_i 替换为 λ_j。因此，所有 λ_i 都相等，且必等于 $\operatorname{Tr}(A)/(n+1)$。特别地，对所有 ~~$i\neq j$~~，都有 $\langle y^i,\ d\rangle=0$ 以及 $\langle y^i,\ y^j\rangle=-\frac{1}{n}$。由此可知，$y^i$ 是内接于 $\mathbb{B}^n$的正则单纯形 S 的顶点。此外，$A=\frac{\operatorname{Tr}(A)}{n}I_n$ 表明

$$d=\frac{n}{\operatorname{Tr}(A)}Ad=\sum_{i=1}^{N}\lambda_i\langle d,y^i\rangle y^i=0.$$

还需注意，K 包含 $-nS$ 的顶点 $-ny^i=z^i-d=z^i$，且其边界上有 $-nS$的面心 $y^1,\dots,y^{n+1}$。由此可知，$K=-nS$是一个单纯形。

众所周知，单纯形到欧几里得球的 Banach–Mazur 距离等于 n，我们也可以根据定理 1.5 (iii) 如下证明这一点：对于外切于 $\mathbb{B}^n$的正则单纯形 S，直接计算可知，其顶点位于 $\operatorname{bd}(n\mathbb{B}^n)$中。此外，存在一个基于 S 与 $\mathbb{B}^n$的公共边界点（即 S的面心）以及 S 与 $n\mathbb{B}^n$ 的公共边界点（即 S 的顶点）的 John 分解。因此，存在一个如定理 1.5 (iii) 所述的 Ader 分解，其中公共算子 A是单位矩阵的倍数。这立即验证了 $d_{BM}(S,\mathbb{B}^n)=n$。

□

我们的最后一个目标是证明若干关于给出凸体 Banach–Mazur 距离的椭球唯一性的结果。如果一对椭球 (E, F) 是 的距离椭球对 K，则称其为距离椭球对，当且仅当 E 与 F 位似，位似比为 $d_{BM}(K, \mathbb{B}^n)$，并且满足 $E \subseteq K \subseteq F$。注意，$E$ 和 F 的中心可以任意放置。

我们的核心结果是将 Maurey 椭球定理推广到 $\mathbb{R}^n$ 中的一般凸体。在对称情形下，该定理表明，对于原点对称凸体 $K \subseteq \mathbb{R}^n$，存在一个线性子空间 $U \subseteq \mathbb{R}^n$（可能是 $U = \mathbb{R}^n$），使得 $d_{BM}(K \cap U, \mathbb{B}^n \cap U) = d_{BM}(K, \mathbb{B}^n)$，并且对于 $K \cap U$，距离椭球对是唯一的。在本文及下文中，低维凸体的 Banach–Mazur 距离均指相对于它们的共同线性张成空间的距离。Maurey 的结果已在多篇不同的论文中被提及（见 [2, 3, 20]），但其首次发表的证明直到最近才出现在 [10, 定理 2.6] 中，该证明基于定理1.2。

下文将建立 Maurey 的结果的一个变体，该变体涵盖非对称情形，并使用投影而非截面。投影与截面之间众所周知的对偶性表明，这一新变体确实推广了原点对称情形。我们从定理 4.4的一个推论开始。该推论表明，任意凸体都包含一个特定点，该点始终是距离椭球对中外接椭球的中心。需要指出的是，这一点可能位于凸体的边界上。例如，若 $K \subseteq \mathbb{R}^3$ 是 $\mathbb{B}^3$ 与一个内切半径为 1 且与 $\mathbb{B}^3$ 相切、形心为 c 的正三角形的凸包，那么定理 1.5 表明，以 2 为半径、以 c 为中心的欧几里得球属于 K 的一个距离椭球对。

引理 6.2。 设 $K \subseteq \mathbb{R}^n$ 为一个凸体。则存在某个点 $x \in K$，使得每当 $(E、F)$ 是关于 K 的一对距离椭球时，点 x 就是 F 的中心。

证明。在不失一般性的情况下，我们可以假设 $(\mathbb{B}^n + c, R\mathbb{B}^n)$ 是关于 K 的一对距离椭球，其中 $c \in \mathbb{R}^n$ 和 $R \geq 1$ 分别满足相应条件。设 (E, F) 是另一对距离椭球。根据关于 $\lambda \in (0, 1)$ 的定理 4.4，存在一个以原点为中心的椭球 E_λ 以及某些向量 $c'、d' \in \mathbb{R}^n$，使得 $E_\lambda + c' \subseteq K \subseteq RE_\lambda + d'$，并且 $\mathrm{bd}(K) \cap \mathrm{bd}(RE_\lambda + d') \neq \emptyset$ 当且仅当 F 的中心为零。事实上，$R = d_{BM}(K, \mathbb{B}^n)$ 表明 $\mathrm{bd}(K) \cap \mathrm{bd}(RE_\lambda + d') \neq \emptyset$。因此，$x = 0$ 满足所需性质。 □

我们 现在可以证明 Maurey 椭球定理的非对称变体 .

定理 6.3. 设 $K \subseteq \mathbb{R}^n$ 是一个凸体。那么存在一个线性子空间 $U \subseteq \mathbb{R}^n$（可能是 $U = \mathbb{R}^n$）和一个线性投影 $\pi:\ \mathbb{R}^n \to U$，使得 $d_{BM}(\pi(K)\ \mathbb{B}^n \cap U) = d_{BM}(K, \mathbb{B}^n)$，且 $\pi(K)$ 的距离椭球对是唯一的。此外，若 K 的距离椭球对为一对欧几里得球，则可将 π 取为到 U 上的正交投影。

证明。我们对 n 进行归纳。不失一般性，设 $\mathbb{B}^n + c \subseteq K \subseteq R\mathbb{B}^n$，其中 $c \in \mathbb{R}^n$ 且 $R = d_{BM}(K, \mathbb{B}^n)$。如果这是 K 的唯一一对距离椭球（当 $n = 1$ 时总是如此），则无需证明，因为我们只需选取 $U = \mathbb{R}^n$。反之，假设对于某个以原点为中心的椭球 E 以及向量 $c'、d' \in \mathbb{R}^n$，$(E + c', RE + d')$ 是 K 的另一对距离椭球。引理 6.2 表明 $d' = 0$。现在，令 $\mathcal{V}_E$ 为 (10) 中定义的线性子空间，并令 $U' = \mathcal{V}_E \cap (\mathrm{lin}\{c - c'\})^\perp$。由于假设 $(E + c', RE)$ 与 $(\mathbb{B}^n + c, R\mathbb{B}^n)$ 不同，因此 $\mathcal{V}_E \neq \mathbb{R}^n$ 或 $c' \neq c$。

无论哪种情况， $U' \neq \mathbb{R}^n$。对 $\lambda \in$ 应用定理 4.4（0, 1），可得一个以原点为中心的均值椭球 $E_\lambda \subseteq \mathbb{R}^n$ 和一个向量 $t \in \mathbb{R}^n$，使得 $E_\lambda + t \subseteq K \subseteq RE_\lambda$，bd($K$) ∩bd($E_\lambda$+$t$) 中的任意点都属于 $U'+t$，而 bd(K)∩bd(RE_λ) 中的任意点都属于 $\mathcal{V}_E$。因此，在公共边界点处支撑 K 和 $E_\lambda + t$ 的超平面的外法向量属于 U'，而支撑 K 和 RE_λ 的超平面的外法向量属于 $\mathcal{V}_E$。

令 u s 选取一个 Ader 分解，使其满足 $E_\lambda + t \subseteq K \subseteq RE_\lambda$ 如定理1.4所示，即 .,

$$\sum_{i=1}^{N} \lambda_i y^i (a^i)^T = \sum_{j=1}^{M} \mu_j z^j (b^j)^T \quad \text{and} \quad \sum_{i=1}^{N} \lambda_i a^i = \sum_{j=1}^{M} \mu_j b^j = 0,$$

其中，(y^i, a^i) 是 K 与 $E_\lambda + t$ 的接触对，而 (z^j, b^j) 是 K 与 RE_λ 的接触对。如前所述，对于所有 $i = 1, \ldots, N$，都有 $y^i \in U' + t$ 和 $a^i \in U'$。我们断言，对于所有 $j = 1, \ldots, M$，也有 $z^j, b^j \in U'$。

根据子空间 $\mathcal{V}_E$的定义（10），交集 $E_\lambda \cap U'$ 是 U'中的标准欧几里得球，因此，支撑 E_λ 于 y^i 的超平面的外法向量只有 $y^i - t$的正倍数。因此，经过适当的重新缩放，我们可以假设$a^i = y^i - t$ 对于 $i = 1, \ldots, N$。同样，由于 $z^j \in \mathcal{V}_E$且 $E_\lambda \cap \mathcal{V}_E$是$\mathcal{V}_E$中的欧几里得球，我们可以假设，对于所有$b^j = z^j$， $j = 1, \ldots, M$。因此，上述分解可写为

$$\sum_{i=1}^{N} \lambda_i y^i (y^i - t)^T = \sum_{j=1}^{M} \mu_j z^j (z^j)^T \quad \text{and} \quad \sum_{i=1}^{N} \lambda_i (y^i - t) = \sum_{j=1}^{M} \mu_j z^j = 0. \tag{28}$$

现在，令 $x \in \mathbb{R}^n$ 为垂直于子空间 U'的任意向量。则对于所有 $i = 1$，……，N，都有 $\langle x, y^i - t \rangle = 0$，因此

$$0 = \sum_{i=1}^{N} \lambda_i \langle x, y^i - t \rangle \langle x, y^i \rangle = \sum_{j=1}^{M} \mu_j \langle x, z^j \rangle^2.$$

由此可得，对于所有 $j = 1$，……，M，都有 $\langle x, z^j \rangle = 0$，这意味着所有 z^j确实属于 U'，因为 x 是垂直于 U'的任意向量。

现在，使用正交投影 π'将 K正交投影到 U'上：$\mathbb{R}^n \to U'$。于是，对于所有 $i = 1$，……，N，都有 $\pi'(y^i - t) = y^i - t$，特别地，当 $t' = \pi'(t) - t$时，$\pi'(y^i) = y^i + t'$。此外，对于所有 $j = 1$，……，M，都有 $\pi'(z^j) = z^j$，并且

$$\pi'(E_\lambda) + \pi'(t) \subseteq \pi'(K) \subseteq R\pi'(E_\lambda).$$

我们进一步注意到，所有支撑 K 且外法向量属于 U' 的超平面也支撑 $\pi'(K)$。因此，$(y^i + t',\ y^i - t)$ 是 $\pi'(K)$ 与 $\pi'(E_\lambda) + \pi'(t)$ 的接触对，而 $(z^j,\ z^j)$ 是 $\pi'(K)$ 与 $R\pi'(E_\lambda)$ 的接触对；其中我们注意到 $\pi'(E_\lambda)$ $\mathbb{B}^n \cap U'$ $\sum_{i=1}^{N} \lambda_i\ y^i - t\ = 0$ 28 。此外，由 ( ) 中的 ( ) 可知

$$\sum_{i=1}^{N} \lambda_i (y^i - t)(y^i - t)^T = \sum_{i=1}^{N} \lambda_i y^i (y^i - t)^T = \sum_{j=1}^{M} \mu_j z^j (z^j)^T.$$

利用 $y^i - t = y^i + t' - \pi'$（$t$），由定理 1.5可得 d_{BM}（π'（K） $\mathbb{B}^n \cap U'$） $= R$。

根据 $U' \neq \mathbb{R}^n$，归纳假设给出一个线性子空间 $U'' \subseteq U'$，使得对于正交投影 π'': $U' \to U''$，我们有

$$d_{BM}(\pi''(\pi'(K)), \mathbb{B}^n \cap U'') = d_{BM}(\pi'(K), \mathbb{B}^n \cap U') = d_{BM}(K, \mathbb{B}^n)$$

并且 $\pi''(\pi'(K))$ 的距离椭球对是唯一的。令 $U = U''$ 和 $\pi = \pi'' \circ \pi'$即可完成证明。

□

我们以如下推论结束本文：关于与欧几里得球的距离接近最大值的凸体的上述定理。关于中心对称凸体的类似变体见 [10, 推论 2.11]。值得注意的是，对称平面情形下距离椭球对的唯一性早已得到确立 [4]并被反复研究（例如参见 [20, 21]），但在一般情形下，这似乎是一个新结果。

推论 6.4。设 $K \subseteq \mathbb{R}^n$ 是一个凸体。若 $d_{BM}(K, \mathbb{B}^n) > n - 1$，则其距离椭球对是唯一的。特别地，若 $n = 2$，则 K 的距离椭球对总是唯一的。

证明。取子空间 U 和投影 π : $\mathbb{R}^n \to U$ 如定理 6.3中所述。令$1 \leq k \leq n$ 为 U的维数。结合推论 6.1，我们得到

$$k \geq d_{BM}(\pi(K), \mathbb{B}^n \cap U) = d_{BM}(K, \mathbb{B}^n) > n - 1,$$

这意味着 $k = n$。由此可知， $U = \mathbb{R}^n$ ，且 π 是单位矩阵。因此，对于 K ，距离椭球对是唯一的。 □

致谢

这项研究合作由克拉科夫雅盖隆大学卓越倡议——研究型大学计划资助。

参考文献

[1] O. B. Ader r, 凸区域的一个仿射不变量, 《杜克数学杂志》4(2) (1938), 291–299.

[2] R. Anisca, A. Tcaciuc, N. Tomczak-Jaegermann, 与欧几里得空间距离达到极值的赋范空间结构, 《休斯敦数学杂志》31(1)(2005),267–283.

[3] A.Arias, T. Figiel, W. B.Johnson, G. Schechtman, 具有 2-可和性质的巴拿赫空间, 《美国数学会汇刊》347(10) (1995), 3835–3857.

[4] F. Behrend, 关于凸区域的若干仿射不变量, 《数学年刊》113(1937) 713–747（德文）。

[5] P. Bizeul, B. Klartag, 非对称凸体之间的距离：多对数范围内的最优界，arXiv: 2510.20511v2（2025年）。

[6] R. Brandenberg，F. Grundbacher，基于凸体几何均值之间关系的几何均值全面定义，arXiv:2312.17512v2（2024年）。[7]R. Brandenberg，F. Grundbacher，利用不对称度量收紧体积极值椭球的不等式，Mathematika 71(4)（2025年），e70051。[8] R. Brandenberg，S. König，位似变换下包含问题不存在与维数无关的核集，Discrete Comput. Geom. 49(1)（2013年），3–21。[9] Y. Gordon，A. E. Litvak，M. Meyer，A. Pajor，一般情形下的 John 分解及其应用，J. Differential Geom. 68(1)（2004年），99–119。[10] F. Grundbacher，T. Kobos，关于某些极值 Banach–Mazur 距离及距离椭球的 Ader 刻画，Mathematika 72(1)（2026年），e70062。[11] C. H. Jiménez，M. Naszódi，关于两个凸体之间的极值距离，Israel J. Math. 183（2011年），103–115。[12] F. John，带有不等式辅助条件的极值问题，载于《Studies and essays presented to R. Courant on his 60th Birthday, January 8, 1948》，Interscience Publishers, Inc.，纽约（1948年），187–204。[13] T. Kobos，平面上对称凸体与任意凸体之间的极值 Banach–Mazur 距离，Mathematika 66(1)（2020年），161–177。[14] T. Kobos，M. Varivoda，关于低维情形下的 Banach–Mazur 距离，Discrete Comput. Geom. 74(2)（2025年），399–427。[15]H. W. Kuhn，非线性规划：历史回顾，载于 R. W. Cottle 和 C. E. Lemke（编），《Nonlinear Programming》，SIAM-AMS 会议录第9卷，美国数学会，罗得岛州普罗维登斯（1976年），1–26。[16] K. Leichtweiss，《Über die affine Exzentrizität konvexer Körper》，Arch. Math. 10（1959年），187–199（德文）。[17] D. R. Lewis，由 Banach 理想范数定义的椭球，Mathematika 26(1)（1979年），18–29。[18] V. Milman，L. Rotem，凸体的加权几何平均，载于《Functional Analysis and Geometry: Selim Grigorievich Krein 百年诞辰纪念文集》，Contemp. Math. 第733卷，美国数学会，罗得岛州普罗维登斯（2019年），233–249。[19] O. Palmon，唯一一个与球的距离达到极值的凸体是单纯形，Israel J. Math. 80(3)（1992年），337–349。[20]D. Praetorius，关于距离椭球的评述与例子，Colloq. Math. 93(1)（2002年），41–53。[21] M. Romney，具有小膨胀率的度量曲面的拟共形参数化，Indiana Univ. Math. J. 68(3)（2019年），1003–1011。

[22] R. Schneider，《凸体：布伦–闵可夫斯基理论》，第 2 版增订版，《数学及其应用百科全书》**151**，剑桥大学出版社，剑桥（2014）。

[23] N.Tomczak-Jaegermann，《*Banach–Mazur* 距离与有限维算子理想》，皮特曼纯粹与应用数学专著和综述 **38**，朗文科技出版社，哈洛（1989）。