arXiv:2304.03212v2 [math.NA] 6 Aug 2024

关于通过体采样逼近向量值函数

Daniel Kressner* Tingting Ni† Andr´e Uschmajew‡

摘要

给定希尔伯特空间 $\mathcal{H}$ 和有限测度空间 Ω，通过 k维子空间 $\mathcal{U} \subset \mathcal{H}$ 对向量值函数 $f : \Omega \to \mathcal{H}$ 的逼近，在降维技术中发挥着重要作用，例如用于求解参数相关偏微分方程的约化基方法。对于勒贝格–博赫纳空间 $L^2(\Omega;\mathcal{H})$ 中的函数，最佳子空间逼近误差$d_k^{(2)}$ 由 f 的奇异值刻画。然而，出于实际原因，$\mathcal{U}$ 通常被限制为由 f 的点样本张成。我们证明，这一限制对可达到的误差影响很小；始终存在 k 个样本，使得所得误差不大于 $\sqrt{k+1}\cdot d_k^{(2)}$。我们的工作扩展了 Binev 等人（*SIAM J. Math. Anal.*，43(3):1457–1472，2011）关于上确界范数逼近的现有结果，以及 Deshpande 等人（*Theory Comput.*，2:225–247，2006）关于矩阵列子集选择的现有结果。

1 引言

设$(\Omega, \mathcal{A}, \mu)$为一个有限测度空间，且 $\mu(\Omega) > 0$。设 $\mathcal{H}$ 为一个带有内积 $\langle\cdot,\cdot\rangle$ 及其诱导范数 $\|\cdot\|$ 的实希尔伯特空间。在本文中，我们研究向量值函数 $f : \Omega \to \mathcal{H}$ 通过投影到有限维子空间$\mathcal{U} \subset \mathcal{H}$所得到的逼近。对于 $1 \le p \le \infty$，可能的误差度量则是逼近数

$$d_k^{(p)}(f) = \inf\left\{\|f-g\|_{L^p(\Omega;\mathcal{H})} : \mathcal{U} \text{ is a } k\text{-dimensional subspace of } \mathcal{H} \text{ and } g \in L^p(\Omega;\mathcal{U})\right\},$$

其中$\|\cdot\|_{L^p(\Omega;\mathcal{H})}$ 定义了 Lebesgue-Bochner 空间 $L^p(\Omega;\mathcal{H})$ 上的范数。

数值 $d_k^{(p)}(f)$ 衡量了用一个将（Ω 中几乎所有点）映射到某个 k 维子空间 $\mathcal{U} \subseteq \mathcal{H}$ 的函数 g，对 f 在 $L^p(\Omega;\mathcal{H})$ 中进行逼近所能达到的最佳效果。显然，对于给定的 $\mathcal{U}$，g 的最佳选择是几乎处处取 $g(y) = P_{\mathcal{U}} f(y)$，其中 $P_{\mathcal{U}}$ 是到 $\mathcal{U}$ 上的 $\mathcal{H}$-正交投影。采用这一记号，我们有

$$d_k^{(p)}(f) = \inf\left\{\|f - P_{\mathcal{U}} f\|_{L^p(\Omega;\mathcal{H})} : \mathcal{U} \text{ is a } k\text{-dimensional subspace of } \mathcal{H}\right\}. \tag{1}$$

*ANCHP, EPFL数学研究所, CH-1015 洛桑, 瑞士
†SYCAMORE实验室，瑞士洛桑 EPFL，CH-1015
‡数学研究所与高级分析与预测科学中心，奥格斯堡大学，86159 奥格斯堡，德国

在特殊情形 $p=\infty$ 下，该量

$$d_k^{(\infty)}(f)=\inf\left\{\operatorname*{ess\,sup}_{y\in\Omega}\|f(y)-P_{\mathcal{U}}f(y)\|:\mathcal{U}\text{ is a }k\text{-dimensional subspace of }\mathcal{H}\right\}$$

是 [14, 16] f 在 $\mathcal{H}$中的“本质”像的 Kolmogorov 宽度。对于一般的 $p<\infty$，量 $d_k^{(p)}(f)$ 称为平均 Kolmogorov 宽度，甚至可以推广到非希尔伯特空间的情形，参见例如 [15, 第 2 节2] 及其中的参考文献。对于 $\mathcal{H}$ 为 Sobolev 空间且 f 为高斯随机变量的情形，已有若干关于 $d_k^{(p)}(f)$ 渐近阶的结果；综述参见 [19] 。

在许多情况下，人们希望甚至预期，上述宽度会随着子空间维数 k 的增加而快速衰减。于是，目标是构造一个实际的k维子空间 $\mathcal{U}$，使其误差接近 $d_k^{(p)}(f)$。一种常用方法是在 k个精心选取的点 $y_1,\ldots,y_k\in\Omega$ 处对 f进行采样，并将 $\mathcal{U}$定义为$f(y_1),\ldots,f(y_k)$的张成空间。例如，用于求解参数依赖偏微分方程的约化基方法 [11, 17]通常采用贪心策略，依次选取参数点 $y_1,\ldots,y_k$。类似地，正交匹配追踪 [18]和经验插值方法 [2]也可视为贪心策略。近年来，深度神经网络已被证明能够有效地根据（含噪）样本近似 f；例如，参见 [1]。

对此类基于样本方法的现有收敛性分析 [2, 4, 9, 18] 通常会证明误差界在定性上接近 (1)，但其中涉及的前置因子可能非常大，且往往随 k呈指数增长。例如，[4, 中关于 $p=\infty$ 的第 2] 节结果表明，贪心选择 k 个点所导致的误差几乎有界于 $2^{k+1}(k+1)d_k^{(\infty)}(f)$；[3, 中的定理 4.4]将此界改进为 $2^{k+1}/\sqrt{3}\cdot d_k^{(\infty)}(f)$。[3] 中的结果还表明，贪心算法能够恢复 Kolmogorov 宽度的代数衰减或指数衰减。另一方面，在使用贪心选择时，因子 2^k 通常无法避免，这引出了一个问题：使用样本是否必然会导致较大的前置因子，还是存在能给出更有利界 (1) 的样本选择？同样，对于 $p=\infty$ 这一情形，[3, 中的定理 4.1] 证明，总存在样本点 $y_1,\ldots,y_k$，使得前置因子确实可以降至 $k+1$。该结果的证明是非构造性的，并使用了最大体积论证这一逼近理论中的经典工具。最大体积原理也是 [10] 中一个结果的基础：该结果通过交叉插值，在逐元素最大范数下对矩阵进行秩为 k 的逼近。然而，有证据 [8]表明，这类最大体积论证无法推广到情形 $p=2$，即矩阵的 Frobenius 范数逼近；不过，可以改用随机采样论证。

主要结果 在本文中，我们讨论 $p=2$ 的情形，并建立一个前置因子 $\sqrt{k+1}$.

定理 1. 设 $f\in L^2(\Omega;\mathcal{H})$。那么存在一个可测集 $\hat{\boldsymbol{\Omega}}\subseteq\Omega\times\cdots\times\Omega=\Omega^k$，其乘积测度为正 $\mu^{\otimes k}$，使得对于 $\mu^{\otimes k}$-几乎处处的 $\mathbf{y}=(y_1,\ldots,y_k)\in\hat{\boldsymbol{\Omega}}$，有

$$d_k^{(2)}(f)\le\|f-\Pi_{\mathbf{y}}f\|_{L^2(\Omega;\mathcal{U})}\le\sqrt{k+1}\cdot d_k^{(2)}(f),\tag{2}$$

其中，$\Pi_{\mathbf{y}}f$表示将 f正交投影到由 $f(y_1)\ldots,f(y_k)$张成的子空间上的 $\mathcal{H}$-正交投影。

这里我们不假设 f 连续。一般而言， $L^2(\Omega;\mathcal{H})$中的元素 f 是函数等价类，它们仅在 μ-几乎处处相等。定理的表述方式使其适用于整个等价类，也就是说，对于每个代表元，集合$\mathbf{\Omega}$ 都相同。注意，对于两个代表元，实际满足 (2) 的$\mathbf{y}\ \in\mathbf{\Omega}$ 集合可能不同，但差异仅在于一个乘积测度为零的集合。本注的其余部分将证明定理 1。

有限维情形 定理 1 的有限维类比由 Deshpande 等人提出 [8]，内容如下：给定矩阵 $A=[a_1,\dots,a_n]\in\mathbb{R}^{m\times n}$和$k<m\le n$，存在 k个列索引 $j_1,\dots,j_k\in\{1,\dots,n\}$，使得

$$(\sigma_{k+1}^2+\cdots+\sigma_m^2)^{1/2}\le\|A-P_{j_1,\dots,j_k}A\|_F\le\sqrt{k+1}\cdot(\sigma_{k+1}^2+\cdots+\sigma_m^2)^{1/2},\qquad(3)$$

其中， $\|\cdot\|_F$表示 Frobenius 范数， $\sigma_1\ge\cdots\ge\sigma_m\ge0$是 A的奇异值，$P_{j_1,\dots,j_k}$表示到 span$\{a_{j_1},\dots,a_{j_k}\}$上的正交投影。将 $\Omega=\{1,\dots,n\}$上的均匀测度代入定理 1，取欧几里得空间 $\mathcal{H}=\mathbb{R}^m$，并令 $f(i)=a_i$（其中 $i=1,\dots,n$），即可得到这一结果。文献 [8] 中的证明采用概率方法，表明当列的抽样概率与其诱导体积成正比时，式 (3) 在期望意义下成立。

[8] 中的一个例子表明，(3)中的前置因子 $\sqrt{k+1}$一般而言无法改进。在后续工作中，Deshpande和 Rademacher [7] 推导出了一种更高效的算法，用于计算满足(3)的列子集选择；关于其数值实现的进一步改进，参见 [5] 。对于固定的 $j_1,\dots,j_k$，对 A哪怕极其微小的改变也可能导致显著更大的误差$\|A-P_{j_1,\dots,j_k}A\|_F$ [6]。这种缺乏连续性的情况使得很难将 [5, 7] 中的算法与无限维情形的离散化结合起来。目前，定理1是一个基于概率论证的存在性结果，而设计一种高效的采样过程，使其产生的误差（近似地）由 $\sqrt{k+1}\cdot d_k^{(2)}(f)$界定，仍然是一个开放问题。

虽然我们证明定理1的主要策略遵循 [8], ，但为了处理无限维情形，仍需进行若干非平凡的修改。

博赫纳积分与可测函数 在继续之前，让我们回顾一下有关博赫纳积分的基本定义。函数 $f:\Omega\to\mathcal{H}$ 被称为强 μ- 可测，如果它是简单函数 $f_\ell:\Omega\to\mathcal{H}$ 的 μ- 几乎处处逐点强极限。如果此外满足$\int_\Omega\|f-f_\ell\|\mathrm{d}\mu\to0$ 对于 $\ell\to\infty$，则称其为博赫纳可积函数。在这种情况下， f 的博赫纳积分 $\int_\Omega f\,\mathrm{d}\mu$ 定义为简单函数 f_ℓ 的积分的极限，并且该极限不依赖于序列 f_ℓ的选取。由于实值函数 $y\mapsto\|f(y)-f_\ell(y)\|$ 本身是强 μ-可测的，因此该定义中的积分 $\int_\Omega\|f-f_\ell\|$ 是有定义的。特别地，如果 μ 是完备测度，则被积函数是$\mathcal{A}$-可测函数；否则，它们几乎处处等于某个 $\mathcal{A}$-可测函数 [13, 命题 1.1.16]。注意，在本文其他地方由强 μ $\mathcal{H}$-值函数复合实值函数时，也将隐含地采用类似的推理，而不再另行说明。

希尔伯特空间 $L^2(\Omega;\mathcal{H})$ 由博赫纳可积函数 f的等价类组成，这些函数满足 $\int_\Omega\|f\|^2\,\mathrm{d}\mu<\infty$。虽然通常不会区分函数及其等价类，但我们经常会针对具体代表元进行逐点论证

。关于博赫纳积分和空间$L^p(\Omega;\mathcal{H})$的更多详情，请参阅 [13]。

2 施密特分解

希尔伯特空间 $L^2(\Omega;\ \mathcal{H})$ 与从 $L^2(\Omega)$ 到 $\mathcal{H}$ 的希尔伯特–施密特算子空间 $HS(L^2(\Omega);\mathcal{H})$ 等距同构。该等距同构通过将 $f \in L^2(\Omega;\mathcal{H})$ 与有界积分算子联系起来实现

$$T_f : L^2(\Omega) \to \mathcal{H}, \qquad v \mapsto T_f v = \int_\Omega f v \,\mathrm{d}\mu.$$

为说明这一点，设 $\{u_i\colon\ i \in I\}$ 和 $\{v_j\colon\ j \in J\}$ 分别是 $\mathcal{H}$ 和 $L^2(\Omega)$ 的标准正交基，其中 J 是可数的。然后，利用 Bochner 积分的性质，可以例行验证 $\{u_i v_j\colon\ i \in I,\ j \in J\}$ 是 $L^2(\Omega;\mathcal{H})$ 的标准正交基。与 $u_i v_j$ 相关联的积分算子是秩一算子 $u_i\langle v_j,\ \cdot\rangle_{L^2(\Omega)}$。这些秩一算子构成了 $HS(L^2(\Omega);\mathcal{H})$的标准正交基；例如参见 [12, 定理 4.4。5]。这表明 $f \mapsto T_f$ 是 $L^2(\Omega;\ \mathcal{H})$ 与 $HS(L^2(\Omega);\mathcal{H})$ 之间的等距同构，并导出了 f 的以下众所周知的分解。

定理 2（施密特分解）。设 $f \in L^2$（Ω；$\mathcal{H}$）。存在至多可数的正交归一系统 $\{u_i \in \mathcal{H}\colon\ i = 1、2、.....、r\}$，$\{v_i \in L^2$（$\Omega$ μ）：$\ i = 1、2、.....、r\}$（满足$r \in \mathbb{N} \cup \{+\infty\}$）以及奇异值 $\sigma_1 \geq \sigma_2 \geq \ldots$，其中 $\sigma_i > 0$ 对于 $i = 1、2、.....、r$成立，使得

$$f = \sum_{i=1}^{r} \sigma_i u_i v_i,$$

级数在 L^2（Ω；$\mathcal{H}$）中收敛。并且

$$f(y) \in \overline{\operatorname{span}}\{u_i : i = 1, 2, \ldots, r\} \quad \mu\text{-}a.e.$$

and

$$f(y) = \sum_{i=1}^{r} \sigma_i u_i v_i(y) \quad \mu\text{-}a.e., \tag{4}$$

级数在 $\mathcal{H}$中收敛。

证明。由 $T = T_f$ 的奇异值分解，存在 σ_i、u_i 和 v_i，具有上述性质，且满足 $T = \sum_{i=1}^{r} \sigma_i u_i \langle v_i,\ \cdot\rangle_{L^2(\Omega)}$，其中级数在希尔伯特–施密特范数下收敛；例如参见 [12, 定理 4.3。2]。由上文讨论的等距同构，即可得到 f 在 L^2（Ω；$\mathcal{H}$）中的所述级数表示。

为了证明定理的第二部分，我们首先注意到，T 的值域是可分希尔伯特空间 $\mathcal{H}_0 = \overline{\operatorname{span}\{u_i\colon\ i = 1, 2, \ldots, r\}}$。因此，

$$\int_A f \,\mathrm{d}\mu = \int_\Omega f \chi_A \,\mathrm{d}\mu = T\chi_A \in \mathcal{H}_0 \quad \forall A \in \mathcal{A}.$$

由于该测度是有限的，根据 [13] 中的命题1.2.13，这意味着 $f(y) \in \mathcal{H}_0$对 μ-几乎处处的 $y \in \Omega$。由于 $\{u_i\}$ 是 $\mathcal{H}_0$的一组正交归一基，这又意味着

$$f(y) = \sum_{i=1}^{r} \langle f(y), u_i \rangle u_i \quad \mu\text{-a.e.} \tag{5}$$

对于每个 i，函数强 $y \mapsto \langle f(y)\ u_i \rangle$ 是 μ-可测的，因此几乎处处等于一个 μ-可测函数 $\mathcal{A}$ 命题 1.1. [13, 。于是，对于每个 16],

$$\begin{aligned} \int_A \langle f(y), u_i \rangle \, d\mu &= \Big\langle \int_A f(y)\, d\mu, u_i \Big\rangle \\ &= \langle T\chi_A, u_i \rangle = \sum_{j=1}^{r} \sigma_j \langle u_j, u_i \rangle \langle v_j, \chi_A \rangle_{L^2(\Omega)} = \int_A \sigma_i v_i(y)\, d\mu \end{aligned}$$

这里我们用到了 Bochner 积分可以与有界线性泛函交换这一性质；例如参见 [13,式 (1.2)]。这意味着

$$\langle f(y), u_i \rangle = \sigma_i v_i(y) \quad \mu\text{-a.e.},$$

结合 (5)，这就证明了 (4)。 □

奇异值由 f 唯一确定，且与其代表元无关。正奇异值的个数 $r \in \mathbb{N} \cup \{+\infty\}$ 称为 f 的秩，记作 rank(f)。由 Schmidt–Mirsky 定理 [12, 定理 4.4.7] 以及上文所述的同构可知，截断函数

$$f_k = \sum_{i=1}^{k} \sigma_i u_i v_i = P_{\mathcal{U}_k} f, \quad \mathcal{U}_k = \operatorname{span}\{u_1, \ldots, u_k\},$$

是 f 在秩至多为 k 的函数中关于 $L^2(\Omega;\mathcal{H})$-范数的最佳逼近。由于 f 在一个 k 维子空间上的任意其他投影的秩至多为 k，因此可得

$$d_k^{(2)}(f) = \|f - f_k\|_{L^2(\Omega;\mathcal{H})} = \big(\sigma_{k+1}^2 + \sigma_{k+2}^2 + \cdots\big)^{1/2}. \tag{6}$$

3 Expected volume

对于 $k \geq 1$，我们考虑乘积测度空间 $(\Omega^k,\ \mathcal{A}^{\otimes k},\ \boldsymbol{\mu})$，其中 $\mathcal{A}^{\otimes k}$ 表示乘积 σ-代数（包含笛卡尔积 $A_1 \times \cdots \times A_k$的集合 $A_1,\ \ldots\ldots,\ A_k \in \mathcal{A}$ 的最小 σ-代数），而 $\boldsymbol{\mu} := \mu^{\otimes k}$ 表示乘积测度（在$\mathcal{A}^{\otimes k}$上满足 $\boldsymbol{\mu}(A_1 \times \cdots \times A_k) = \mu(A_1) \cdots \mu(A_k)$ 的唯一测度）。令 $f \in L^2(\Omega;\ \mathcal{H})$。给定 k 个样本点 $\mathbf{y} = (y_1,\ \ldots\ldots,\ y_k)$，$f(y_1),\ \ldots\ldots,\ f(y_k) \in \mathcal{H}$ 的 Gram 矩阵为

$$G^{(k)}(\mathbf{y}) = \begin{bmatrix} \langle f(y_1), f(y_1) \rangle & \cdots & \langle f(y_1), f(y_k) \rangle \\ \vdots & & \vdots \\ \langle f(y_k), f(y_1) \rangle & \cdots & \langle f(y_k), f(y_k) \rangle \end{bmatrix} \in \mathbb{R}^{k \times k}.$$

其行列式

$$\det G^{(k)}(\mathbf{y}) = \sum_{\pi\in\mathcal{S}_k} \operatorname{sign}(\pi) \prod_{i=1}^{k} \langle f(y_i), f(y_{\pi(i)})\rangle,$$

其中， $\mathcal{S}_k$ 表示 $(1,\ldots,k)$ 的所有排列组成的集合，通常称为 $f(y_1)\ldots,f(y_k)$ 的体积。注意，就 f 的等价类而言，该量仅在 $\boldsymbol{\mu}$-几乎处处唯一确定。因此，我们将 $\det G^{(k)}(\mathbf{y})$视为 $\boldsymbol{\mu}$-几乎处处相等的 $\boldsymbol{\mu}$-可测实值函数的等价类。进而，可以定义 $\det G^{(k)}$的勒贝格积分，称为期望体积。下面的引理表明，与 [8, 引理 31]类似，该量可以根据 f 的奇异值计算得到。

引理 3. 采用上述引入的记号，有如下结论：

$$\int_{\Omega^k} \det G^{(k)}\,\mathrm{d}\boldsymbol{\mu} = \sum_{\substack{j_1,\ldots,j_k=1\\ j_1,\ldots,j_k \text{ mutually distinct}}}^{r} \sigma_{j_1}^2\cdots\sigma_{j_k}^2.$$

*Pr*哎。 考虑定理 2 中的施密特分解。对于每个有限的 $n \le r$， le t

$$f_n(y) = \sum_{i=1}^{n} \sigma_i u_i v_i(y).$$

由于 v_i 是 $\mathcal{A}$-可测的，相应的行列式

$$\det G_n^{(k)}(\mathbf{y}) = \sum_{\pi\in\mathcal{S}_k} \operatorname{sign}(\pi) \prod_{i=1}^{k} \langle f_n(y_i), f_n(y_{\pi(i)})\rangle = \sum_{\pi\in\mathcal{S}_k} \operatorname{sign}(\pi) \prod_{i=1}^{k}\sum_{j=1}^{n} \sigma_j^2 v_j(y_i) v_j(y_{\pi(i)})$$

是一个非负的 $\mathcal{A}^{\otimes k}$-可测函数，利用托内利定理可知，它满足 ,

$$\begin{aligned}
\int_{\Omega^k} \det G_n^{(k)}\, d\boldsymbol{\mu} &= \sum_{\pi\in\mathcal{S}_k} \operatorname{sign}(\pi) \int_{\Omega^k} \prod_{i=1}^{k}\sum_{j=1}^{n} \sigma_j^2 v_j(y_i) v_j(y_{\pi(i)})\,\mathrm{d}\boldsymbol{\mu} \\
&= \sum_{j_1=1}^{n}\cdots\sum_{j_k=1}^{n} \sigma_{j_1}^2\cdots\sigma_{j_k}^2 \sum_{\pi\in\mathcal{S}_k} \operatorname{sign}(\pi) \int_{\boldsymbol{\Omega}} \prod_{i=1}^{k} v_{j_i}(y_i) v_{j_{\pi(i)}}(y_i)\,\mathrm{d}\boldsymbol{\mu} \\
&= \sum_{j_1=1}^{n}\cdots\sum_{j_k=1}^{n} \sigma_{j_1}^2\cdots\sigma_{j_k}^2 \sum_{\pi\in\mathcal{S}_k} \operatorname{sign}(\pi) \prod_{i=1}^{k} \langle v_{j_i}, v_{j_{\pi(i)}}\rangle_{L^2(\Omega)} \\
&= \sum_{j_1=1}^{n}\cdots\sum_{j_k=1}^{n} \sigma_{j_1}^2\cdots\sigma_{j_k}^2 \sum_{\pi\in\mathcal{S}_k} \operatorname{sign}(\pi) \prod_{i=1}^{k} \delta_{j_i,j_{\pi(i)}} \\
&= \sum_{j_1=1}^{n}\cdots\sum_{j_k=1}^{n} \sigma_{j_1}^2\cdots\sigma_{j_k}^2 \det(P_{j_1,\ldots,j_k}),
\end{aligned}$$

对于矩阵 $[P_{j_1,\ldots,j_k}]_{\alpha,\beta} = \delta_{j_\alpha,j_\beta}$， 当 $\alpha,\beta = 1,\ldots,k$ 时。如果所有 $j_1,\ldots,j_d$ 都互不相同，则 $\det(P_{j_1,\ldots,j_k}) = 1$；否则 $\det(P_{j_1,\ldots,j_k}) = 0$，因为此时 $P_{j_1,\ldots,j_k}$ 包含相同的行。

因此，

$$\int_{\Omega^k} \det G_n^{(k)}\, d\boldsymbol{\mu} = \sum_{\substack{j_1,\dots,j_k=1 \\ j_1,\dots,j_k \text{ mutually distinct}}}^{n} \sigma_{j_1}^2 \cdots \sigma_{j_k}^2.$$

如果 r 是有限的，取 $n = r$即可证明该断言。否则我们取极限 $n \to \infty$并论证

$$\int_{\Omega^k} \det G^{(k)}\, d\boldsymbol{\mu} = \lim_{n\to\infty} \int_{\Omega^k} \det G_n^{(k)}\, d\boldsymbol{\mu}$$

根据控制收敛定理。这是可行的，因为 $\det G_n^{(k)}(\mathbf{y}) \to \det G^{(k)}(\mathbf{y})$ $\boldsymbol{\mu}$-几乎处处成立（这是 (4) 的推论），并且

$$\left|\det G_n^{(k)}(\mathbf{y})\right| \le \sum_{\pi\in\mathcal{S}_k} \left|\prod_{i=1}^{k} \langle f_n(y_i), f_n(y_{\pi(i)})\rangle\right| \le \sum_{\pi\in\mathcal{S}_k} \prod_{i=1}^{k} \|f_n(y_i)\|^2 \le \sum_{\pi\in\mathcal{S}_k} \prod_{i=1}^{k} \|f(y_i)\|^2.$$

Sinc在 $f \in L^2(\Omega;\mathcal{H})$ 中，右侧是 Ω^k 上的一个可积支配函数。 □

4 定理 1 的证明

假设 $\operatorname{rank}(f) \ge k$。那么我们可以定义

$$\varrho = \frac{\det G^{(k)}}{\int_{\Omega^k} \det G^{(k)}\, \mathrm{d}\boldsymbol{\mu}}.$$

根据引理 3，分母为正。我们有 $\varrho \ge 0$ （$\boldsymbol{\mu}$-几乎处处）以及$\int_{\Omega^k} \varrho \mathrm{d}\boldsymbol{\mu} = 1$。因此，我们可以将 ϱ 解释为乘积测度空间上的概率密度函数。现在我们旨在界定

$$\mathbb{E}_\varrho\big(\|f - \Pi_{\mathbf{y}} f\|^2_{L^2(\Omega;\mathcal{H})}\big) = \int_{\Omega^k} \|f - \Pi_{\mathbf{y}} f\|^2_{L^2(\Omega;\mathcal{H})} \varrho(\mathbf{y})\, \mathrm{d}\boldsymbol{\mu}(\mathbf{y}),$$

其中 $\Pi_{\mathbf{y}} f\colon \Omega \to \mathcal{H}$是将 f正交投影到由 $\{f(y_1)\dots, f(y_k)\}$张成的子空间上的 $\mathcal{H}$-正交投影。注意，投影算子 $\Pi_{\mathbf{y}}$ 取决于所选的等价类代表元 f；事实上，同一等价类中两个不同代表元所张成的子空间 $f(y_1)\dots, f(y_k)$可能完全不同，但这种差异仅发生在乘积测度为零的集合上。下述定理的陈述与证明是 [8, 定理 1.3], 的对应版本，请注意这些细微之处。

定理 4. 给定 $f \in L^2(\Omega;\mathcal{H})$，考虑施密特分解 (4) 并假设$r = \operatorname{rank}(f) \ge k$。则有：

(i) 对每个固定的 $\mathbf{y} \in \Omega^k$，函数 $y' \mapsto f(y') - \Pi_{\mathbf{y}} f(y')$（其中 $\Pi_{\mathbf{y}} f$如上定义）确定了一个 *Bochner* 可积函数的等价类，且该等价类属于$L^2(\Omega;\mathcal{H})$。因此，非负函数 $\mathbf{y} \mapsto \|f - \Pi_{\mathbf{y}} f\|^2_{L^2(\Omega;\mathcal{H})}$ 是良定义的。

(ii) 几乎处处相等的函数类 $\mathbf{y} \mapsto \|f - \Pi_{\mathbf{y}} f\|^2_{L^2(\Omega;\mathcal{H})} \cdot \varrho(\mathbf{y})$包含一个 $\mathcal{A}^{\otimes k}$-可测的代表元。因此， $\mathbb{E}_\varrho\big(\|f - \Pi_{\mathbf{y}} f\|^2_{L^2(\Omega;\mathcal{H})}\big)$ 定义良好。

(iii) 成立：

$$\sum_{i=k+1}^{r} \sigma_i^2 \le \mathbb{E}_\varrho\big(\|f - \Pi_{\mathbf{y}} f\|^2_{L^2(\Omega;\mathcal{H})}\big) \le (k+1) \sum_{i=k+1}^{r} \sigma_i^2.$$

证明。 (i) 取 f 的一个固定代表元。于是，$\{f(y_1)\dots, f(y_k)\}$ 在 $\mathcal{H}$ 上构成一个有界线性算子，其中的正交投影算子$T = \Pi_{\mathbf{y}}$作用于 span。进而， Tf 是博赫纳可积的。此外，对于每个 $y' \in \Omega$，$\|f(y') - Tf(y')\| = \|(I-T)f(y')\| \le \|f(y')\|$。因此，$\int_\Omega \|f(y') - Tf(y')\|^2 \mathrm{d}\mu \le \int_\Omega \|f(y')\|^2 \mathrm{d}\mu < \infty$。显然，对于代表元 f 的不同选取，函数 $f - \Pi_{\mathbf{y}} f$ 在 $\boldsymbol{\mu}$-几乎处处相等。

(ii) 取 f 的一个代表元，使其处处都是简单函数 [13, Proposition 1.1.16] 的逐点强极限。相应的体积函数 $\mathbf{y} \mapsto \det G^{(k)}(\mathbf{y})$ 因而是 $\mathcal{A}^{\otimes k}$-可测的。同样，函数 $(\mathbf{y}\ y') \mapsto \det G^{(k+1)}(\mathbf{y}\ y')$ 是 $\mathcal{A}^{\otimes(k+1)}$-可测的。根据托内利定理，非负函数

$$g(\mathbf{y}) = \int_\Omega \det G^{(k+1)}(\mathbf{y}, y')\, \mathrm{d}\mu(y')$$

于是是 $\mathcal{A}^{\otimes k}$-可测的。由于代表元 f 已固定，投影算子 $\Pi_{\mathbf{y}}$ 对每个 $\mathbf{y} \in \Omega^k$ 都有明确的定义。我们声称，对于每个 $\mathbf{y}$，都有

$$g(\mathbf{y}) = \det G^{(k)}(\mathbf{y}) \int_\Omega \|f(y') - \Pi_{\mathbf{y}} f(y')\|^2\, \mathrm{d}\mu(y'). \tag{7}$$

除以 $\int_{\Omega^k} \det G^{(k)}\, \mathrm{d}\boldsymbol{\mu} > 0$ 后，这就推出了 (ii) 部分。为证明 (7)，我们首先注意到，如果 $f(y_1)\dots, f(y_k)$ 线性相关，那么 (7) 显然成立，因为这意味着 $\det G^{(k)}(\mathbf{y}) = 0$ 以及 $g(\mathbf{y}) = 0$（因为对于所有 $y' \in \Omega$，都有 $\det G^{(k+1)}(\mathbf{y}\ y') = 0$）。因此，我们可以假设 $f(y_1)\dots, f(y_k)$ 线性无关。于是 $G^{(k)}(\mathbf{y})$ 是可逆的，并且对所有 $y' \in \Omega$，都有

$$\begin{aligned} G^{(k+1)}(\mathbf{y}, y') &= \begin{bmatrix} G^{(k)}(\mathbf{y}) & w_{\mathbf{y},y'} \\ w_{\mathbf{y},y'}^T & \|f(y')\|^2 \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ w_{\mathbf{y},y'}^T G^{(k)}(\mathbf{y})^{-1} & 1 \end{bmatrix} \begin{bmatrix} G^{(k)}(\mathbf{y}) & w_{\mathbf{y},y'} \\ 0 & \|f(y')\|^2 - w_{\mathbf{y},y'}^T G^{(k)}(\mathbf{y})^{-1} w_{\mathbf{y},y'} \end{bmatrix}, \end{aligned}$$

对于向量

$$w_{\mathbf{y},y'} = [\langle f(y_1), f(y')\rangle, \dots, \langle f(y_k), f(y')\rangle]^T \in \mathbb{R}^k.$$

容易证明， $f(y')$ 在 $f(y_1)\dots, f(y_k)$ 张成的空间上的正交投影 $\Pi_{\mathbf{y}} f(y')$ 满足 $\|\Pi_{\mathbf{y}} f(y')\|^2 = w_{\mathbf{y},y'}^T G^{(k)}(\mathbf{y})^{-1} w_{\mathbf{y},y'}$。因此我们有

$$\det G^{(k+1)}(\mathbf{y}, y') = \det G^{(k)}(\mathbf{y}) \cdot \|f(y') - \Pi_{\mathbf{y}} f(y')\|^2,$$

从而得到（7）。

关于 (iii)。下界可由 (6) 立即得出。接下来证明上界。使用与 (ii) 中相同的 f 代表元，由 (7) 可得

$$\mathbb{E}_\varrho\big(\|f - \Pi_{\mathbf{y}} f\|^2_{L^2(\Omega;\mathcal{H})}\big) \cdot \int_{\Omega^k} \det G^{(k)}\, \mathrm{d}\boldsymbol{\mu} = \int_{\Omega^{k+1}} \det G^{(k+1)}\, \mathrm{d}\mu^{\otimes(k+1)}.$$

对两边应用引理 3，得到

$$\mathbb{E}_{\varrho}\big(\|f-\Pi_{\mathbf{y}}f\|_{L^2(\Omega;\mathcal{H})}^2\big)=\frac{\sum_{j_1,\dots,j_{k+1}}\sigma_{j_1}^2\cdots\sigma_{j_k}^2\sigma_{j_{k+1}}^2}{\sum_{j_1,\dots,j_k}\sigma_{j_1}^2\cdots\sigma_{j_k}^2},\tag{8}$$

其中，求和遍历从 1 到 r 的互不相同的指标。由于排列 $k+1$ 个互不相同的数共有 $(k+1)!$ 种方式，因此可得

$$\begin{aligned}\sum_{j_1,\dots,j_{k+1}}\sigma_{j_1}^2\cdots\sigma_{j_k}^2\sigma_{j_{k+1}}^2&=(k+1)!\sum_{j_1<\cdots<j_{k+1}}\sigma_{j_1}^2\cdots\sigma_{j_k}^2\sigma_{j_{k+1}}^2\\&=(k+1)!\sum_{j=k+1}^{r}\sigma_j^2\sum_{j_1<\cdots<j_k<j}\sigma_{j_1}^2\cdots\sigma_{j_k}^2\\&\le(k+1)!\sum_{j=k+1}^{r}\sigma_j^2\sum_{j_1<\cdots<j_k}\sigma_{j_1}^2\cdots\sigma_{j_k}^2\\&=(k+1)\sum_{j=k+1}^{r}\sigma_j^2\sum_{j_1,\dots,j_k}\sigma_{j_1}^2\cdots\sigma_{j_k}^2.\end{aligned}$$

将此不等式代入 (8)，即可完成证明。 □

现在可以陈述本节的主要结果；结合 (6)，该结果也证明了定理 1。

定理 5. 设 $f\in L^2(\Omega;\mathcal{H})$ 的秩为 $r\in\mathbb{N}\cup\{\infty\}$。存在一个乘积测度为正的可测集 $\hat{\mathbf{\Omega}}\subseteq\Omega^k$，使得对于等价类 $\hat{}$ 中的每个代表元 f，以及对几乎所有 样本元组 $\mathbf{y}\in\hat{\mathbf{\Omega}}$，均有

$$\|f-\Pi_{\mathbf{y}}f\|_{L^2(\Omega;\mathcal{H})}^2\le(k+1)\sum_{i=k+1}^{r}\sigma_i^2.$$

证明. 暂且假设 $r\ge k$。在这种情况下，证明归结为以下事实：随机变量不大于其期望值的概率为正。为完整起见，我们给出完整论证。如定理 4(ii) 的证明中所述，存在一个合适的代表元 f，使得两个函数 $\mathbf{y}\mapsto\varrho(\mathbf{y})$ 和 $\mathbf{y}\mapsto\|f-\Pi_{\mathbf{y}}f\|_{L^2(\Omega;\mathcal{H})}^2\cdot\varrho(\mathbf{y})$ 均为 $\mathcal{A}^{\otimes k}$-可测。因此，集合 $B=\{\mathbf{y}\in\Omega^k\colon\varrho(\mathbf{y})>0\}$ 以及

$$\begin{aligned}\hat{\mathbf{\Omega}}&=\left\{\mathbf{y}\in B\colon\|f-\Pi_{\mathbf{y}}f\|_{L^2(\Omega;\mathcal{H})}^2\le(k+1)\sum_{i=k+1}^{r}\sigma_i^2\right\}\\&=\left\{\mathbf{y}\in B\colon\|f-\Pi_{\mathbf{y}}f\|_{L^2(\Omega;\mathcal{H})}^2\cdot\varrho(\mathbf{y})-(k+1)\sum_{i=k+1}^{r}\sigma_i^2\cdot\varrho(\mathbf{y})\le0\right\}\end{aligned}$$

均属于 $\mathcal{A}^{\otimes k}$。随后，通过标准论证可得，我们必须有 $\boldsymbol{\mu}(\hat{\mathbf{\Omega}})>0$，否则定理 4(iii) 将不成立。对于 f 的任何其他代表元，

$\|f - \Pi_{\mathbf{y}} f\|^2_{L^2}$ 的值$(\Omega;\mathcal{H})$ 对于 $\boldsymbol{\mu}$-几乎所有的 $\mathbf{y}$都保持不变。这就完成了 $r \geq k$ 情形下的证明。

在 rank(f) $= r < k$ 的情况下，我们可以将该定理应用于 r，而不是 k。然后

$$\|f - \Pi_{(y_1,\ldots,y_r)} f\|^2_{L^2(\Omega;\mathcal{H})} \leq (r+1) \sum_{i=r+1}^{r} \sigma_i^2 = 0$$

对于子集 $\tilde{\boldsymbol{\Omega}}\, \Omega^r$ 中的所有元组 $(y_1, \ldots, y_r)$，其具有正的 $\mu^{\otimes r}$-测度。则对于所有$\mathbf{y} \in \tilde{\boldsymbol{\Omega}} \times \Omega^{k-r}$，其具有正的 $\boldsymbol{\mu}$-测度，我们也有

$$\|f - \Pi_{\mathbf{y}} f\|^2_{L^2(\Omega;\mathcal{H})} \leq \|f - \Pi_{(y_1,\ldots,y_r)} f\|^2_{L^2(\Omega;\mathcal{H})} = 0$$

因为所投影到的子空间只会变大。 □

致谢 A.U. 的研究工作得到了德国研究联合会（DFG）——项目编号 506561557 的资助。作者感谢审稿人细致审阅并提出建设性意见。

参考文献

[1] B. Adcock、S. Brugiapaglia、N. Dexter 和 S.Morage。从有限数据中学习高维 Hilbert 值函数：深度神经网络非常有效。载于 J. Bruna、J. Hesthaven 和L. Zdeborova（编），第 *2* 届数学与科学机器学习会议论文集，机器学习研究论文集第 145 卷，1–36页。PMLR, 2022年。[2] M. Barrault、Y. Maday、N. C. Nguyen和A.T.Patera。“经验插值”方法：应用于偏微分方程的高效约化基离散化

巴黎科学院数学学报, 339(9):667–672, 2004.[3] P. Binev、A. Cohen、W.Dahmen、R. DeVore、G.Petrova和P. Wojtaszczyk。约化基方法中贪心算法的收敛速率。*SIAM*数学分析杂志, 43(3):1457–1472,2011.[4] A. Buffa、Y. Maday、A. T. Patera、C. Prud'homme 和 G. Turinici。参数化约化基方法中贪心算法的先验收敛性。*ESAIM* 数学建模与数值分析, 46(3):595–603, 2012.[5] A.Cortinovis 和 D.Kressner。通过列和行子集选择在 Frobenius范数下进行低秩近似。*SIAM* 矩阵分析与应用杂志, 41(4):1651–1673,2020.[6] Alice Cortinovis。个人通信，2023 年。[7] A.Deshpande和L. Rademacher。用于行/列子集选择的高效体积采样。载于*2010*年 *IEEE* 第 *51*届计算机科学基础年度研讨会——*FOCS2010*，第 329–338 页。2010 年。[8] A. Deshpande、L.Rademacher、S.Vempala 和 G. Wang。通过体积采样进行矩阵近似和射影聚类。计算理论, 2:225–247, 2006.[9] Z. Drmaˇc和S.Gugercin。离散经验插值方法的新选择算子——改进的先验误差界及扩展。*SIAM* 科学计算杂志, 38(2):A631– A648, 2016。

[10] S. A.Goreinov 和E. E.Tyrtyshnikov。《低秩矩阵逼近中的最大体积概念》。载于数学、计算机科学与工程中的结构化矩阵，*I*（科罗拉多州博尔德，*1999*），第280 卷当代数学，第47–51页。美国数学学会，普罗维登斯，罗得岛州，2001 年。[11] J. S. Hesthaven、G.Rozza 和 B. Stamm。《参数化偏微分方程的经认证约化基方法》。Springer，湛，2016 年。[12] T.Hsing和 R. Eubank。《函数型数据分析的理论基础：线性算子导论》。JohnWiley& Sons, Ltd.，奇切斯特，2015 年。[13] T. Hyt¨onen、J. vanNeerven、M.Veraar 和 L. Weis。《巴拿赫空间分析。第一卷。鞅与Littlewood-Paley理论》。Springer，湛，2016年。.[14] A. Kolmogoroff。《关于给定函数类中函数的最佳逼近》。数学年刊（第二辑），37(1)：107–110，1936 年。[15] Y. Malykhin。《宽度与刚性》，2022 年。arXiv:2205.03453。[16] A. Pinkus。 n逼近理论中的 -宽度。Springer-Verlag，柏林，1985 年。[17] A. Quarteroni、A. Manzoni 和F. Negri。《偏微分方程的约化基方法》。Springer，湛，2016 年。[18] V.Temlyakov。《贪婪逼近》。剑桥大学出版社，剑桥，2011 年。[19] H.Wang。《配备高斯测度的球上加权 Sobolev 空间的概率线性宽度与平均线性宽度》。逼近理论杂志，241：11–32，2019 年。