通过线性测量和标量非线性，在 H^n 的紧子集上对连续泛函进行通用逼近

Andrey Krylov KRYL@CS.MSU.RU
Maksim Penkin PENKIN97@GMAIL.COM
俄罗斯莫斯科国立罗蒙诺索夫大学

arXiv:2602.03290v1 [cs.LG] 3 Feb 2026

摘要

我们研究乘积希尔伯特空间的紧子集上连续泛函的通用逼近。我们证明，任何此类泛函都可以通过如下模型实现一致逼近：模型首先对输入进行有限个连续线性测量，然后通过连续标量非线性组合这些测量。我们还将这一逼近原理推广到取值于巴拿赫空间的映射，从而得到有限秩逼近。这些结果为算子学习和成像中常见的“测量、应用标量非线性，然后组合”设计模式提供了紧集上的理论依据。

关键词： 通用逼近；连续映射；算子值逼近；巴拿赫空间；紧集；线性测量；标量非线性；有限秩逼近；Stone– Weierstrass 定理；神经算子；反问题、人工智能

1. 引言

许多现代学习任务本质上都涉及泛函：给定输入 $x = (x_1, \ldots\ldots, x_n)$ 属于乘积希尔伯特空间 H^n，我们希望在紧集 $f\colon K \to \mathbb{R}$ 上对连续映射 $K \subset H^n$进行建模。这种观点在算子学习和科学机器学习中十分自然，因为其中的数据是函数、场或信号，而非定长向量 Kovachki 等人(2021）；Li 等人(2021）；Lu 等人(2021）。

一种实用的设计模式是将感知与学习分开。首先，我们对每个函数值输入进行有限次线性测量（传感器读数、基系数、滤波响应、投影测量）。其次，将这些数值输入非线性模型（通常是神经网络）以生成输出。这种先测量、再进行非线性组合的结构，是许多算子学习架构的基础 Lu 等人(2021); Li 等人 (2021); Kovachki 等人 (2021)；经典成像流程也采用类似方式，将线性滤波／投影与逐点非线性运算相结合 Mallat¨(1999); Ronneberger 等人 (2015); Adler 和 Oktem (2018).

本文给出了一个紧集定理，从通用逼近的角度论证了这类架构的合理性。对于 H^n 的紧子集（其中 H 是希尔伯特空间），我们证明，任意连续泛函 $f\colon K \to \mathbb{R}$ 都可以由以下形式的有限和一致逼近

$$\sum_{j=1}^{r} \zeta_j \Big(\sum_{i=1}^{n} \varphi_{ji}(x_i) \Big),$$

其中，$\varphi_{ji} \in H^*$ 是连续线性测量，$\zeta_j\colon \mathbb{R} \to \mathbb{R}$ 是连续标量非线性函数。用算子学习的术语来说，$\{\varphi_{ji}\}$ 定义了测量（编码）阶段，而 $\{\zeta_j\}$定义了非线性解码器/混合器。

© A. Krylov 和 M. Penkin。

在附录 A 中，我们将同一紧集逼近原理扩展到算子值输出 $f\colon\ K \to Y$适用于一般巴拿赫空间 Y，并在附录 B–C 中简要讨论其在计算机视觉和人工智能中的应用。

除了陈述逼近结果外，我们还力求给出便于工程人员阅读的证明：借助紧致性，我们可以通过正交投影将问题化简为有限维代理，随后即可应用经典逼近结果。这一视角说明了为什么有限个测量便足以应对现实中的（紧）数据情形，也通过同一个两阶段框架将物理信息神经网络、隐式神经表示与成像／逆问题联系起来。

2. 主要结果

定理 1（紧集上的逼近）设 $K \subset H^n$ 为紧集，且 $f \in C(K)$。那么，对任意 $\varepsilon > 0$，存在一个形如

$$g(x_1, \dots, x_n) = \sum_{j=1}^{r} \zeta_j \Big(\sum_{i=1}^{n} \varphi_{ji}(x_i) \Big),$$

其中 $\varphi_{ji} \in H^*$ 和 $\zeta_j\colon\ \mathbb{R} \to \mathbb{R}$ 连续，使得

$$\sup_{(x_1,\dots,x_n)\in K} |f(x_1, \dots, x_n) - g(x_1, \dots, x_n)| < \varepsilon.$$

证明 设 $X := H^n$，并将 X 视为赋范为以下范数的希尔伯特空间

$$\|(x_1, \dots, x_n)\|^2 := \sum_{i=1}^{n} \|x_i\|_H^2.$$

步骤 1（一致连续性）。由于 K 是紧的，且 $f \in C$（K），函数 f 在 K 上是一致连续的（海涅–康托尔定理）Rudin (1976)。取 $\delta > 0$ 使得对于任意$x,\ x' \in K$，条件 $\|x - x'\| < \delta$ 蕴含

$$|f(x) - f(x')| < \varepsilon/3.$$

步骤 2（化归到有限维情形）。由于 K 是紧的，因此可以从任意开覆盖中提取一个有限子覆盖Munkres (2000)；特别地，存在一个有限的 $\delta/3$覆盖：存在点 $x^{(1)}, \dots, x^{(m)} \in K$ 使得

$$K \subset \bigcup_{k=1}^{m} B\big(x^{(k)}, \delta/3\big).$$

令 V:= 张成$\{x^{(1)}, \dots, x^{(m)}\} \subset X$（因此维数为 $V \le m$），并令 $P\colon\ X \to V$为正交投影算子。

我们声称

$$\sup_{x\in K} \|x - Px\| < \delta/3.$$

事实上，对于每个 $x \in K$，选取 k 使得 $\|x - x^{(k)}\| < \delta/3$。于是 $x^{(k)} \in V$，因此从 x 到 V 的距离满足

$$\operatorname{dist}(x, V) \le \|x - x^{(k)}\| < \delta/3.$$

但根据正交投影的定义性质， $\|x - Px\| = \text{dist}(x,\ V)$，从而得到所需的估计。

步骤 3（将其转化为关于 $P(K)$ 的函数）。考虑紧集 $K' := P(K) \subset V$.

由构造可知， V =张成空间$\{x^{(1)}, \dots, x^{(m)}\}$ 我们有 $Px^{(k)} = x^{(k)}$。此外，步骤 2 表明，对于任意 $x \in K$，都有 $\|x - Px\| < \delta/3$。因此，由于 $K \subset \bigcup_{k=1}^{m} B\big(x^{(k)},\ \delta/3\big)$，对于每个 $x \in K$，都存在 k，使得

$$\|Px - x^{(k)}\| \le \|Px - x\| + \|x - x^{(k)}\| < \delta/3 + \delta/3 = 2\delta/3.$$

因此，

$$K' \subset \bigcup_{k=1}^{m} B_V\big(x^{(k)}, 2\delta/3\big),$$

其中 $B_V(\cdot,\cdot)$ 表示有限维空间 V 中的一个球。

选取一个连续单位分解 $\{\psi_k\}_{k=1}^{m}$ 于 K'上，其从属于覆盖 $\{B_V(x^{(k)},\ 2\delta/3)\}_{k=1}^{m}$（度量紧空间上的单位分解的存在性；参见 Munkres (2000)），即 $\psi_k \ge 0$，$\sum_k \psi_k \equiv 1$ 于 K'上，且 支集 $\psi_k \subset B_V(x^{(k)},\ 2\delta/3)$。定义连续函数 $F \in C(K')$ 如下

$$F(y) := \sum_{k=1}^{m} f\big(x^{(k)}\big)\, \psi_k(y), \qquad y \in K'.$$

估计 f 在 K上由 $F \circ P$ 给出的逼近误差。令 $x \in K$ 和 $y = Px \in K'$。如果$\psi_k(y) \neq 0$，那么 $y \in B_V(x^{(k)},\ 2\delta/3)$，因此

$$\|x - x^{(k)}\| \le \|x - y\| + \|y - x^{(k)}\| < \delta/3 + 2\delta/3 = \delta,$$

根据 δ的选取，我们得到 $|f(x) - f(x^{(k)})| < \varepsilon/3$。于是

$$|f(x) - F(Px)| = \Big|\sum_{k=1}^{m} \psi_k(Px)\,\big(f(x) - f(x^{(k)})\big)\Big| \le \sum_{k=1}^{m} \psi_k(Px)\,|f(x) - f(x^{(k)})| < \varepsilon/3.$$

因此，

$$\sup_{x \in K} |f(x) - F(Px)| < \varepsilon/3.$$

步骤 4（在有限维紧集上逼近）。通过标准正交基将 V 与 $\mathbb{R}^d$ (d =维数 V) 识别。

考虑由以下函数生成的子代数 $\mathcal{A} \subset C(K')$（在 $\mathbb{C}$上）

$$y \longmapsto e^{i\langle t, y\rangle}, \qquad t \in \mathbb{R}^d.$$

也就是说， $\mathcal{A}$ 中的元素是有限个线性组合

$$y \longmapsto \sum_{j=1}^{r} c_j\, e^{i\langle t_j, y\rangle}.$$

显然， $\mathcal{A}$包含常数函数，且在复共轭运算下封闭。此外， $\mathcal{A}$能分离紧集 K'中的点：若 $y \neq y'$，令 $u := y - y' \neq 0$ 并选取 $t := u$，则有$\langle t,\ y\rangle - \langle t,\ y'\rangle = \|u\|^2 \neq 0$，因此 $e^{i\langle t,\ y\rangle} \neq e^{i\langle t,\ y'\rangle}$。

根据（复）Stone–Weierstrass 定理Rudin (1987)，代数 $\mathcal{A}$在 $C(K')$中稠密。因此，存在 $T \in \mathcal{A}$使得

$$\sup_{y \in K'} |F(y) - T(y)| < \varepsilon/3.$$

由于 F是实值的，将 T替换为

$$\widetilde{T} := \operatorname{Re} T = \tfrac{1}{2}\big(T + \overline{T}\big) \in \mathcal{A}.$$

于是，对所有 $y \in K'$，我们有 $|F(y) - \widetilde{T}(y)| \le |F(y) - T(y)|$，因此 $\sup_{y \in K'} |F(y) - \widetilde{T}(y)| < \varepsilon/3$。将 $\widetilde{T}$重新命名为 T，便得到一个实值函数，它可以写成正弦函数和余弦函数的有限和：

$$T(y) = \sum_{j=1}^{r} \alpha_j \cos\langle t_j, y\rangle + \sum_{j=1}^{r} \beta_j \sin\langle t_j, y\rangle.$$

每一项的形式为 $\zeta(\langle t_j, y\rangle)$，其中 ζ 连续： $\mathbb{R} \to \mathbb{R}$ （即， $\zeta(s) = \alpha_j \cos s$ 或 $\zeta(s) = \beta_j \sin s$），因此

$$T(y) = \sum_{j=1}^{r} \zeta_j\big(\langle t_j, y\rangle\big).$$

步骤 5（回到 $K \subset H^n$）。定义 $g(x) := T(Px)$ 对于 $x \in K$。于是

$$\sup_{x \in K} |f(x) - g(x)| \le \sup_{x \in K} |f(x) - F(Px)| + \sup_{x \in K} |F(Px) - T(Px)| < \varepsilon/3 + \varepsilon/3 = 2\varepsilon/3 < \varepsilon.$$

剩下的工作是将 g 化为所需形式。注意，任意连续线性泛函 $L \in (H^n)^*$都具有如下形式 Kreyszig(1978)

$$L(x_1, \dots, x_n) = \sum_{i=1}^{n} \varphi_i(x_i)$$

对于某个 $\varphi_i \in H^*$ （即 $(H^n)^* \cong (H^*)^n$）。特别地，泛函 $L_j(x) := \langle t_j,\ Px\rangle$ 在 $X = H^n$ 上是线性且连续的；因此存在 φ_{j1}，……， $\varphi_{jn} \in H^*$ 使得

$$\langle t_j, P(x_1, \dots, x_n)\rangle = \sum_{i=1}^{n} \varphi_{ji}(x_i).$$

因此，

$$g(x_1, \dots, x_n) = \sum_{j=1}^{r} \zeta_j\Big(\sum_{i=1}^{n} \varphi_{ji}(x_i)\Big),$$

■

3. 定量推论（度量熵视角）

尽管斯通-魏尔施特拉斯定理这一步是定性的，但此前的紧致性约化已经给出了显式的有限维结构。我们用覆盖数（度量熵）来表述一个简单的推论Kolmogorov 和 Tikhomirov （1959）。

推论（通过网构造有限维代理）。固定 $\varepsilon > 0$，并令 $\delta > 0$ 按照步骤 1 选取。令 $\mathcal{N}(K,\ \delta/3)$ 表示在 H^n 上的乘积范数下，K 的 $\delta/3$-网的最小大小。则存在一个子空间 $V \subset H^n$，满足

$$\dim V \le \mathcal{N}(K, \delta/3)$$

以及一个正交投影算子 $P\colon\ H^n \to V$，使得

$$\sup_{x\in K} \|x - Px\| < \delta/3$$

因此

$$\sup_{x\in K} |f(x) - f(Px)| < \varepsilon/3.$$

特别地，若能对紧集 $P(K) \subset V \cong \mathbb{R}^{\dim V}$ 上的步骤 4 作出任何定量强化，则可立即得到关于 K 的定量逼近保证。

4. 应用

我们给出一组典型示例，主要用于说明连续线性测量与标量非线性的结构。

4.1. L^2 上的积分型泛函

Let $H = L^2(0,1)$ 和 $n = 2$。考虑一个紧集 $K \subset H^2$ 和一个连续函数 a 1

$$f(u,v) = \int_0^1 \Phi\big(u(x), v(x)\big)\, dx,$$

其中 $\Phi\colon\ \mathbb{R}^2 \to \mathbb{R}$ 连续。

那么对于任意 $\varepsilon > 0$，都存在 $r \in \mathbb{N}$、元素 $h_{j1}, h_{j2} \in L^2(0,1)$ 以及连续函数 ζ_j: $\mathbb{R} \to \mathbb{R}$，使得

$$\sup_{(u,v)\in K} \left| f(u,v) - \sum_{j=1}^{r} \zeta_j\Big(\langle u, h_{j1}\rangle + \langle v, h_{j2}\rangle\Big) \right| < \varepsilon,$$

其中 $\langle u,\ \ h\rangle := \int_0^1 u(x)h(x)\ dx$。这可以理解为对输入进行有限测量 +和标量非线性后处理。

4.2. 断层成像及其他线性逆问题

在断层成像（计算机断层成像、正电子发射断层成像等）中，数据是对未知图像进行的有限个（近似）线性测量 u，例如离散化拉东变换的采样Natterer (2001）。因此，学习式重建方法可视为从测量到重建对象的非线性映射

¨对象 Adler 和 Oktem (2018)；相关观点也见于压缩感知领域Cand`es 等 (2006)。该定理给出一个存在性结论：对于可容许图像的紧集，任意关于测量对象的连续标量泛函，都可以通过测量 + 和标量非线性进行一致逼近。

4.3. 与 Kolmogorov–Arnold网络（柯尔莫哥洛夫–阿诺德网络）的关系

该公式

$$g(x_1,\dots,x_n)=\sum_{j=1}^{r}\zeta_j\Big(\sum_{i=1}^{n}\varphi_{ji}(x_i)\Big)$$

可以理解为一个带有单个隐藏层的两层网络：从输入 x_i 到隐藏神经元 j 的边上有一个一维函数 $x_i\mapsto\varphi_{ji}(x_i)$；随后，隐藏神经元应用一个一维非线性函数 ζ_j，之后对 j求和。

从现代柯尔莫哥洛夫–阿诺德网络的角度来看 Liu 等人 (2024)，其中可学习的一维函数位于边上，该陈述意味着，这样的模型类在 $C(K)$ 中稠密，前提是输入边函数选为线性且连续的（对输入信号进行测量），而非线性函数ζ_j 则仍可为任意连续函数。

其他关联（简述）。 相同的结构模式也出现在变换器中 Vaswani 等人(2017）、算子学习领域 Kovachki 等人 (2021)； Li 等人 (2021)； Lu 等人 (2021)、物理信息神经网络 Raissi 等人 (2019)， 以及隐式神经表示 Sitzmann 等人 (2020)； Mildenhall 等人 (2020)。我们在此简要讨论这些关联，因为本文关注的是在 H^n的紧子集上的逼近定理。

文献评述。相关论述（关于由有限个连续线性泛函构成的柱函数的稠密性，以及通过线性形式的一维非线性函数之和进行表示）见于逼近理论和神经网络逼近文献；例如，参见 Pinkus (2015)；Cybenko (1989)；Hornik (1991)。

5. 结论

我们证明了一个关于紧集 $K\subset H^n$ 上连续泛函的通用逼近结果：每个 $f\in C(K)$ 都可以由有限个连续线性测量和标量非线性构成的表达式一致逼近，

$$g(x_1,\dots,x_n)=\sum_{j=1}^{r}\zeta_j\Big(\sum_{i=1}^{n}\varphi_{ji}(x_i)\Big),$$

其中 $\varphi_{ji}\in H^*$ 以及连续 $\zeta_j\colon\ \mathbb{R}\to\mathbb{R}$。这与本文主题直接对应：通过线性测量和标量非线性，在 H^n 的紧子集上实现通用逼近。

总体而言，证明遵循两个阶段的路径。首先，紧致性通过投影将问题约化为一个有限维代理；随后，有限维空间中的经典逼近（Stone–Weierstrass）产生脊函数型逼近函数。中间的有限维代理步骤凸显了其与覆盖数/度量熵之间的明确联系Kolmogorov 和 Tikhomirov（1959）。

仍有多个方向有待研究。一个自然的下一步是在对 f 施加额外正则性假设、对 K 施加几何假设的条件下，推导定量界（ε 中的速率），并研究关于测量 φ_{ji} 的结构约束（例如局部性、卷积结构、带宽限制）如何影响逼近与可学习性。

在附录 A 中，我们通过有限秩展开记录了该定理的巴拿赫空间值（算子输出）扩展；在附录 B–C 中，我们概述了由此得到的线性测量 + 标量非线性 + 线性合成模板如何与成像和人工智能中的常见模型相契合。

参考文献

¨Jonas Adler 和 Ozan Oktem。学习型原始-对偶重建。*IEEE*医学影像汇刊，37：1322–1332，2018。Emmanuel J. Cand`es、Justin Romberg 和 Terence Tao。鲁棒不确定性原理：从高度不完整的频率信息中精确信号重建。*IEEE*信息论汇刊，52：489–509，2006。George Cybenko。通过 S 形函数的叠加进行逼近。控制、信号与系统数学，2：303–314，1989。Kurt Hornik。多层前馈网络的逼近能力。《神经网络》，4：251–257，1991。A. N. Kolmogorov 和 V. M. Tikhomirov。函数空间中集合的 ε-熵与 ε-容量。《数学科学进展》，14：3–86，1959。Nikola Kovachki、Zongyi Li、Burigede Liu、Kamyar Azizzadenesheli、Kaushik Bhattacharya、Andrew Stuart 和 Anima Anandkumar。神经算子：学习函数空间之间的映射。arXiv:2108.08481，2021。Erwin Kreyszig。《应用泛函分析导论》。约翰·威立父子出版公司，纽约，1978。Zongyi Li、Nikola Kovachki、Kamyar Azizzadenesheli、Burigede Liu、Kaushik Bhattacharya、Andrew Stuart 和 Anima Anandkumar。用于参数化偏微分方程的傅里叶神经算子。发表于国际学习表征会议（*ICLR*），2021。Ziming Liu 等。柯尔莫哥洛夫–阿诺德网络。arXiv:2404.19756，2024。Lu Lu、Pengzhan Jin 和 George Em Karniadakis。基于算子的通用逼近定理，通过 DeepONet 学习非线性算子。《自然·机器智能》，3：218–229，2021。St´ephane Mallat。《信号处理小波之旅》。学术出版社，第2版，1999。Ben Mildenhall、Pratul P. Srinivasan、Matthew Tancik、Jonathan T. Barron、Ravi Ramamoorthi 和 Ren Ng。NeRF：将场景表示为用于视图合成的神经辐射场。发表于欧洲计算机视觉会议（*ECCV*），2020。James R. Munkres。《拓扑学》。普伦蒂斯·霍尔出版社，萨德尔河上游，新泽西州，第2版，2000年。

Frank Natterer。计算机断层成像数学。SIAM，2001年。

Allan Pinkus. 脊函数. 剑桥大学出版社，剑桥，2015年。

Maziar Raissi、Paris Perdikaris 和 George Em Karniadakis。物理信息神经网络：一种用于求解涉及非线性偏微分方程的正问题和反问题的深度学习框架。计算物理学杂志，378:686–707，2019年。

Olaf Ronneberger、Philipp Fischer 和 Thomas Brox。U-Net：用于生物医学图像分割的卷积网络。载于国际医学图像计算与计算机辅助干预会议 *(MICCAI)*，2015年。

Walter Rudin。数学分析原理。麦格劳-希尔，纽约，第3版，1976年。

Walter Rudin。实分析与复分析。麦格劳-希尔，纽约，第3版，1987年。

Vincent Sitzmann、Julien N. P. Martel、Alexander W. Bergman、David B. Lindell 与 Gordon Wetzstein。具有周期激活函数的隐式神经表示。载于神经信息处理系统进展（*NeurIPS*），2020。

Ashish Vaswani 等人。注意力就是你所需要的一切。载于 神经信息处理系统进展（*NeurIPS*），2017。

附录 A. 算子值扩展

设 Y为一个巴拿赫空间。

定理 2（有限秩逼近 Y值映射）设 $K \subset H^n$ 为紧，设 Y为一个巴拿赫空间，并设 $f\colon K \to Y$ 连续。则对每个 $\varepsilon > 0$，存在 $r \in \mathbb{N}$、向量 y_1，……，$y_r \in Y$、连续线性泛函 $\varphi_{ji} \in H^*$以及连续标量函数$\zeta_j\colon \mathbb{R} \to \mathbb{R}$，使得映射

$$g(x_1, \dots, x_n) := \sum_{j=1}^{r} y_j \, \zeta_j\Big(\sum_{i=1}^{n} \varphi_{ji}(x_i)\Big)$$

满足

$$\sup_{x \in K} \|f(x) - g(x)\|_Y < \varepsilon.$$

特别地，$g(K)$ 包含在 Y 的一个有限维子空间中。

证明 由于 K 是紧的，且 f 连续，因此像 $f(K) \subset Y$ 是紧的。固定 $\varepsilon > 0$。

步骤 1（值域的有限维逼近）。选取有限个点 $y^{(1)}, \dots, y^{(m)} \in f(K)$ 使得

$$f(K) \subset \bigcup_{k=1}^{m} B_Y\big(y^{(k)}, \varepsilon/3\big).$$

设 $W := $张成空间$\{y^{(1)}, \dots, y^{(m)}\} \subset Y$（故维数为 $W \le m$）。设 $q\colon W \to \mathbb{R}^d$为任意线性同构（其中 $d =$维数为 W），并设 $i\colon W \hookrightarrow Y$表示包含映射。

步骤 2（化为 $\mathbb{R}^d$-值映射）。考虑紧集 $f(K)$ 由球 $\{B_Y(y^{(k)},\varepsilon/3)\}_{k=1}^m$ 构成的开覆盖。在 $f(K)$ 上选取从属于该覆盖的连续单位分解 $\{\psi_k\}_{k=1}^m$。定义连续映射 $P:\ f(K)\to W$ 如下

$$\widetilde{P}(z):=\sum_{k=1}^{m} y^{(k)}\,\psi_k(z).$$

若 $\psi_k(z)\neq 0$，则 $z\in B_Y(y^{(k)},\ \varepsilon/3)$，因此 $\|z-y^{(k)}\|_Y<\varepsilon/3$。所以，对所有 $z\in f(K)$，

$$\|z-\widetilde{P}(z)\|_Y=\Big\|\sum_{k=1}^{m}\psi_k(z)\,(z-y^{(k)})\Big\|_Y\le\sum_{k=1}^{m}\psi_k(z)\,\|z-y^{(k)}\|_Y<\varepsilon/3.$$

现在定义 $F:\ K\to\mathbb{R}^d$，其定义为 $F(x):=q\big(\widetilde{P}(f(x))\big)$。那么，对所有$x\in K$，$\|f(x)-i\big(\widetilde{P}(f(x))\big)\|_Y<\varepsilon/3$。

步骤 3（利用定理 1 逼近每个坐标）。写作 $F=(F_1,\ \ldots\ldots,\ F_d)$。将定理 1 应用于每个标量映射 $F_\ell:\ K\to\mathbb{R}$，并取精度 $\varepsilon/(3\,\|q^{-1}\|\,\|i\|\,d)$，得到函数

$$G_\ell(x)=\sum_{j=1}^{r} a_{\ell j}\,\zeta_j\Big(\sum_{i=1}^{n}\varphi_{ji}(x_i)\Big)$$

其中 $a_{\ell j}\in\mathbb{R}$ 满足 $\sup_{x\in K}|F_\ell(x)-G_\ell(x)|<\varepsilon/(3\,\|q^{-1}\|\,\|i\|\,d)$。令 $G:=(G_1,\ \ldots\ldots,\ G_d)$。

步骤 4（提升回 Y）。定义

$$g(x):=i\big(q^{-1}(G(x))\big)\in Y.$$

于是，g具有所需形式，其向量为 $y_j:=i\big(q^{-1}((a_{1j},\ldots,a_{dj}))\big)\in Y$。此外，对于所有$x\in K$，

$$\|i\big(\widetilde{P}(f(x))\big)-g(x)\|_Y\le\|i\|\,\|q^{-1}\|\,\|F(x)-G(x)\|_1\le\|i\|\,\|q^{-1}\|\sum_{\ell=1}^{d}|F_\ell(x)-G_\ell(x)|<\varepsilon/3.$$

与步骤 2 结合可得 $\sup_{x\in K}\|f(x)-g(x)\|_Y<\varepsilon$。 ■

附录 B. 定理 2 在计算机视觉与成像中的应用

作为一个简单示例，取 $H=L^2(\Omega)$，其中 $\Omega\subset\mathbb{R}^2$ 为图像域，并令 $K\subset H$ 为可容许图像的一个紧集（例如，归一化后的“自然图像”的紧模型）。令 $Y=L^2(\Omega)$ 表示一个输出像/场（去噪图像、重建结果、深度/光流场、分割logits 等），并令 $f:\ K\to Y$为一个连续的图像到图像映射。

定理 2 意味着，对于每个 $\varepsilon>0$，都存在有限个输出“模板” $y_1,\ \ldots,\ y_r\in Y$、有限个线性测量 $\varphi_j\in H^*$ (例如滤波器响应、傅里叶/拉东采样、小波系数)以及连续标量非线性 ζ_j，使得Natterer (2001)；Adler 和 Öktem (2018)；Mallat(1999)

$$\sup_{x\in K}\Big\|f(x)-\sum_{j=1}^{r} y_j\,\zeta_j\big(\varphi_j(x)\big)\Big\|_{L^2(\Omega)}<\varepsilon.$$

用成像领域的术语来说，这为如下处理流程提供了通用逼近论证（适用于紧数据域）：先计算输入图像的一组有限线性特征，再将有限个基像的线性组合合成为输出，其中系数为非线性（标量）系数。

附录 C. 定理 2 在人工智能中的应用

定理 2 也广泛适用于输入与输出自然地被建模为函数空间中元素的人工智能场景。例如，在算子学习中，人们旨在学习函数空间之间的映射（例如，将映射系数/强迫项映射为偏微分方程解），这些映射可视为连续映射 $f\colon\ K \to Y$ 在紧数据域上 Kovachki 等人 （2021）；Li 等人 （2021）；Lu 等人 （2021）。定理 2 给出了一种存在性保证，即此类映射可在 K 上由基于输入信号的有限多个线性测量与标量非线性构建的有限秩模型一致逼近。

更具体地说，若 $H = L^2(\Omega)$ 且 $K \subset H$ 是可容许输入（场、信号、上下文）的紧集，则任意连续的 $f\colon\ K \to Y$ 均可被逼近为

$$g(x) = \sum_{j=1}^{r} y_j \, \zeta_j\big(\varphi_j(x)\big),$$

这与常见的“编码–非线性–解码”模板相符：线性编码器 φ_j（特征）、标量非线性混合 ζ_j，以及由输出原子y_j所张成的线性解码器。这一观点有助于将一系列架构（例如，线性特征前端后接多层感知机、低秩解码器以及神经算子模型）统一为紧集上同一通用逼近原理的实例。