arXiv:2603.07242v1 [math.FA] 7 Mar 2026

局部凸空间输出神经网络的通用逼近定理

Sachin Saini

摘要

本文建立了一种浅层神经网络的通用逼近定理（UAT）。该网络的输入属于拓扑向量空间（TVS），输出取值于豪斯多夫局部凸拓扑向量空间。网络由作用于输入的连续线性泛函上的标量激活函数构造，输出系数则属于目标空间。研究表明，对于由定义半范数所诱导的一致收敛拓扑，从输入空间的紧子集到目标空间的连续映射所构成的函数空间中，这类神经网络是稠密的。该结果推广了拓扑向量空间上的现有标量值逼近定理，并将巴拿赫值和希尔伯特值逼近作为特例涵盖在内。文中还给出了若干推论和示例，展示了其在函数空间之间算子逼近方面的应用。

关键词：通用逼近、向量值神经网络、局部凸空间、拓扑向量空间。
*2020*年数学主题分类：41A30, 41A65, 46A20, 46A22, 68T05。

1. 引言

神经网络（NNs）的函数逼近能力，已成为现代逼近理论和机器学习中的一个基础课题。除了在回归和分类任务中的实证成功之外，神经网络还提供了灵活的参数族，能够表示高度复杂的函数关系。从数学角度看，神经网络的表达能力可由通用逼近性质的各种形式刻画。

In 在经典设定中，人们考虑作用于

有限维欧几里得空间上的浅层前馈神经网络。给定一个输入向量 $s \in \mathbb{R}^d$，具有单隐藏层的网络所生成的函数形式如下

$$s \longmapsto \sum_{j=1}^{m} a_j \, \eta(\langle w_j, s \rangle - b_j),$$

其中 $w_j \in \mathbb{R}^d$, a_j, $b_j \in \mathbb{R}$ 和 $\eta : \mathbb{R} \to \mathbb{R}$ 是一个固定的激活函数。Cybenko、Funahashi、Hornik、Pinkus、Leshno 等人提出的基础性结果表明，在对 η施加温和的非退化条件时，对于每个紧集 $E \subset \mathbb{R}^d$ [1–5]，此类网络在 $C(E)$ 中稠密。因此， E 上的任意连续实值函数都可以由浅层神经网络一致逼近。

这些结果背后的一个关键观察是，映射 $s \mapsto \langle w, \ s \rangle$ 只是 $\mathbb{R}^d$上的一个连续线性泛函。这一视角自然促使人们考虑以更一般空间上的连续线性泛函取代有限维内积。这种方法使神经网络架构能够定义在抽象拓扑向量空间上，从而将逼近理论拓展到欧几里得框架之外。

受这一思想启发，若干研究者考察了作用于无限维空间上的神经网络。在巴拿赫空间的背景下，研究者使用脊函数技术建立了标量值逼近 [6, 7]的稠密性结果。神经网络也被用于逼近函数空间之间的非线性算子 [8]。近来，DeepONet 等算子学习架构 [9]重新激发了人们对无限维空间之间映射的逼近结果的兴趣。随后，研究者在希尔伯特空间和巴拿赫空间背景下开展了定量分析 [10, 11]。相关进展还包括有限维代数上向量值和超复值神经网络的通用逼近定理 [12]。

Ismailov [13], 提出了一种特别灵活的框架，其中考虑了定义在具有足够丰富的连续对偶的拓扑向量空间 S 上的前馈神经网络 S^*。在这一框架中，每个隐藏神经元都会在输入 $s \in S$处计算一个泛函 $\ell \in S^*$，随后应用标量激活函数 η $(\ell(s) - \theta)$。在对 S（例如满足 Hahn–Banach延拓性质（HBEP））以及激活函数 η作出适当假设的条件下，这些标量值网络在紧子集 $E \subset S$上构成 $C(E)$ 的稠密子集。该结果在统一的泛函分析框架内整合了有限维和无限维输入理论。然而，

[13] 中的逼近结果仅限于标量值映射，即函数 $F\colon S \to \mathbb{R}$。因此，将通用逼近定理推广到输出取值于一般拓扑向量空间的神经网络，仍然是一个重要的开放方向。

然而，当代分析与科学计算中的许多应用，都需要对取值于无穷维函数空间的映射进行逼近。典型例子包括微分方程的解算子、参数到状态映射、函数到函数的回归问题，以及输出为函数或分布的模型。在此类情形中，人们希望逼近如下形式的映射

$$\mathcal{T} : S \longrightarrow T,$$

其中 T 并非仅仅是 $\mathbb{R}$，而是豪斯多夫局部凸拓扑向量空间。此类空间中的收敛由半范数族而非单一范数决定，这带来了额外的分析挑战。

本文的目标是为输入属于拓扑向量空间、输出位于局部凸拓扑向量空间的神经网络建立通用逼近定理。我们所考虑的架构保留标量激活函数 $\eta : \mathbb{R} \to \mathbb{R}$，同时允许输出系数在 T中取值。由此得到的函数类具有如下形式

$$s \longmapsto \sum_{j=1}^{m} \eta\big(\ell_j(s) - \theta_j\big)\, v_j, \qquad \ell_j \in S^*,\ v_j \in T.$$

此类表示可解释为图1所示的有限秩非线性算子逼近。该神经网络由函数 η 激活的单隐藏层构成。

我们的主要结果表明，在对 S、 T和 η的适当假设下，该类对每个紧子集 $E \subset S$在 $C(E; T)$中稠密，其中稠密性是相对于由 T的定义半范数所诱导的一致收敛而言的。特别地，当 $T = \mathbb{R}$ 时，我们恢复了 [13], 的标量值定理；而若 T 是巴拿赫空间，则半范数拓扑退化为标准的一致范数拓扑。因此，巴拿赫值情形表现为更一般的局部凸逼近理论的一个特例。

本文所得结果为作用于无限维空间之间的浅层神经网络模型提供了泛函分析基础，并将通用逼近定理推广到一大类局部凸空间。

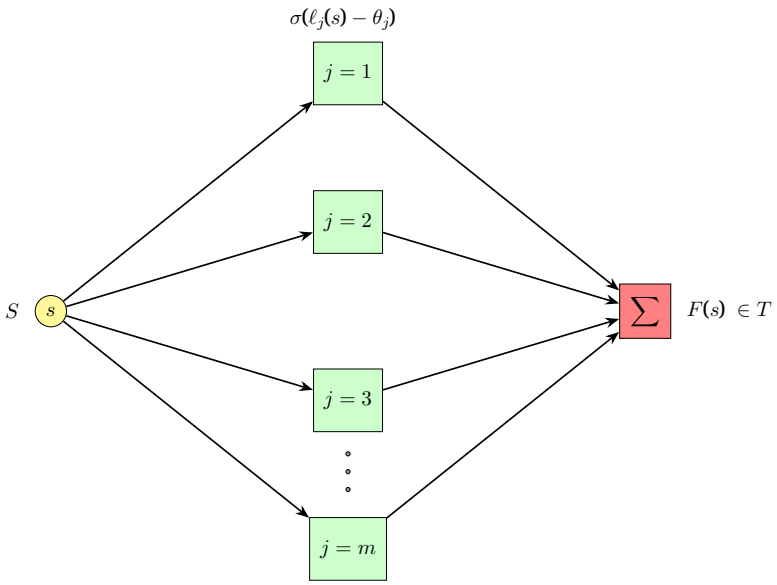

图1：上述定义的神经网络的架构。

2. 主要结果

本节将表述并证明一条通用逼近定理，适用于输入属于拓扑向量空间、输出取值于局部凸拓扑向量空间的神经网络。该结果将此前针对定义在拓扑向量空间上的神经网络所建立的标量逼近框架扩展为向量值情形。

在本节中，设 S 为实拓扑向量空间，T 为豪斯多夫局部凸拓扑向量空间。对于紧集 $E \subset S$，我们用 $C(E; T)$ 表示从 E 到 T 的连续映射空间，并赋予由 T 的连续半范数族 ρ 诱导的一致收敛拓扑。

我们考虑如下形式的非线性逼近函数

$$s \longmapsto \sum_{j=1}^{m} \eta\big(\ell_j(s) - \theta_j\big)\, v_j, \qquad \ell_j \in S^*,\ \theta_j \in \mathbb{R},\ v_j \in T,$$

其中 S^* S 的连续对偶，且 η： $\mathbb{R} \to \mathbb{R}$是一个固定的激活函数。

定理 2.1 (向量值通用逼近定理**)**。 假设 S 具有 ***HBEP***，且 $E \subset S$ 是紧的。设 η： $\mathbb{R} \to \mathbb{R}$ 连续，且在任何非空开区间上都不是多项式。

定义

$$\mathcal{A}^{\eta}_{S,T} := \operatorname{span}\left\{ s \mapsto \eta\big(\ell(s) - \theta\big)\, v : \ell \in S^*,\ \theta \in \mathbb{R},\ v \in T \right\}.$$

那么，$\mathcal{A}^{\eta}_{S,\,T}$在 C（E； T）中关于由 T 的连续半范数诱导的 E 上的一致收敛是稠密的。等价地，对于每个 F $\in C$（E； T）、 T 上的每个连续半范数 ρ，以及满足以下条件的 $\forall\, \varepsilon > 0$、 $\exists\, G \in \mathcal{A}^{\eta}_{S,\,T}$

$$\sup_{s \in E} \rho\big(F(s) - G(s)\big) < \varepsilon.$$

神经网络表示的良定义性：我们说明本文所考虑的神经网络表达式是良定义的。设 $s \in S$、$\ell \in S^*$、 $\theta \in \mathbb{R}$ 和 $v \in T$。

1. 由于 $\ell \in S^*$，它是 S 上的连续线性泛函。因此

$$\ell : S \to \mathbb{R},$$

因此 $\ell(s) \in \mathbb{R}$。

2. 由于 $\theta \in \mathbb{R}$，量 $\ell(s) - \theta$ 也属于 $\mathbb{R}$。
3. 激活函数 $\eta : \mathbb{R} \to \mathbb{R}$ 定义良好，因此

$$\eta(\ell(s) - \theta) \in \mathbb{R}.$$

4. 由于 T 是拓扑向量空间，标量乘法定义为一个映射

$$\mathbb{R} \times T \to T, \qquad (a, v) \mapsto av.$$

因此，

$$\eta(\ell(s) - \theta)\, v \in T.$$

5. 由于 T 是向量空间， T 中元素的有限和仍属于T。因此，形如

$$\sum_{j=1}^{m} \eta(\ell_j(s) - \theta_j)\, v_j$$

定义良好，且取值于 T。

注 2.2. 当 $T = \mathbb{R}$ 时，定理 *2.1* 便化为 *[13]* 中确立的标量值通用逼近定理。如果 T 是一个巴拿赫空间，则由 T 的半范数所诱导的一致收敛拓扑与通常的一致范数拓扑相同。

我们首先证明，在 T 的半范数拓扑下，向量值有限秩映射在 $C(E;\, T)$ 中稠密。

引理 2.3。 设 E 是一个紧豪斯多夫空间， T 是一个豪斯多夫局部凸拓扑向量空间。那么，以下形式映射的集合

$$s \longmapsto \sum_{j=1}^{m} \psi_j(s)\, v_j, \qquad \psi_j \in C(E),\ v_j \in T,$$

在由 T 的连续半范数所诱导的 E 上一致收敛的意义下，在 $C(E;\, T)$ 中稠密。

证明。 取 $F \in C(E;\, T)$、 T 上的一个连续半范数 ρ，以及 $\varepsilon > 0$。

由于 E 是紧的且 F 连续，像集 $F(E)$ 在 T 中是紧的。在豪斯多夫局部凸空间中，紧集关于每个连续半范数都是全有界的。因此，存在元素 $v_1, \ldots, v_m \in T$，使得

$$F(E) \subset \bigcup_{j=1}^{m} \{t \in T : \rho(t - v_j) < \varepsilon\}.$$

对每个 j，定义

$$V_j := \{s \in E : \rho(F(s) - v_j) < \varepsilon\}.$$

那么， $\{V_j\}_{j=1}^{m}$ 构成了 E的一个开覆盖。由于 E 是一个紧豪斯多夫空间，因此它是仿紧的。所以存在一个从属于开覆盖$\{V_j\}_{j=1}^{m}$的单位分解 $\{\psi_j\}_{j=1}^{m} \subset C(E)$，即，

$$0 \le \psi_j \le 1, \qquad \operatorname{supp}(\psi_j) \subset V_j, \qquad \sum_{j=1}^{m} \psi_j(s) = 1 \quad \text{for all } s \in E.$$

定义

$$G(s) := \sum_{j=1}^{m} \psi_j(s)\, v_j.$$

则 $G \in C(E;\, T)$。此外，对每个 $s \in E$,

$$\begin{aligned}\rho\big(F(s) - G(s)\big) &= \rho\left(\sum_{j=1}^{m} \psi_j(s)\big(F(s) - v_j\big)\right) \\ &\leq \sum_{j=1}^{m} \psi_j(s)\, \rho\big(F(s) - v_j\big) < \varepsilon,\end{aligned}$$

由于 $\psi_j(s) \neq 0$ 蕴含 $s \in V_j$。对 E 取上确界，得到

$$\sup_{s\in E} \rho\big(F(s) - G(s)\big) < \varepsilon.$$

□

备注 2.4. 引理 *2.3* 表明，$C(E;\ T)$ 在半范数拓扑下由有限个标量值连续函数与 T 中的系数构成的有限线性组合生成。若 T 是巴拿赫空间，则半范数拓扑退化为通常的一致范数拓扑。

现在，我们回顾拓扑向量空间背景下的标量值逼近结果，它将作为证明主定理的关键要素。

引理 2.5. 设 S 是具有 *HBEP* 的拓扑向量空间，且 $E \subset S$ 是紧的。假设 $\eta :\ \mathbb{R} \to \mathbb{R}$ 连续，并且在任何开区间上都不是多项式。

定义

$$\mathcal{A}_S^{\eta} := \operatorname{span}\left\{s \longmapsto \eta\big(\ell(s) - \theta\big) : \ell \in S^*,\ \theta \in \mathbb{R}\right\}.$$

则 $\mathcal{A}_S^{\eta}$ 在 $C(E)$中稠密。

证明。这一结论可由定义在拓扑向量空间上的神经网络的标量通用逼近定理推出；见 [13]。 □

定理 *2.1* 的证明。令 $F \in C(E;\, T)$，令 ρ为T上的连续半范，并固定 $\varepsilon > 0$。

根据引理 2.3，存在元素 $v_1, \ldots, v_m \in T$以及函数 $\phi_1, \ldots, \phi_m \in C(E)$，使得

$$\sup_{s\in E} \rho\left(F(s) - \sum_{j=1}^{m} \phi_j(s)\, v_j\right) < \frac{\varepsilon}{2}.$$

定义

$$C := \max_{1\le j\le m} \rho(v_j).$$

如果 $C = 0$，那么对于所有 j， $\rho(v_j) = 0$，因此

$$\rho\left(F(s) - \sum_{j=1}^{m} \phi_j(s) v_j \right) = \rho(F(s))$$

对于所有 $s \in E$。上述估计意味着

$$\sup_{s\in E} \rho(F(s)) < \frac{\varepsilon}{2},$$

并取 $G \equiv 0$，即可完成此情形下的证明。

现在假设 $C > 0$。根据引理 2.5，对于每个 $j = 1, \dots, m$，存在 $\psi_j \in \mathcal{A}_S^\eta$ 使得

$$\sup_{s\in E} |\phi_j(s) - \psi_j(s)| < \frac{\varepsilon}{2mC}.$$

定义

$$G(s) := \sum_{j=1}^{m} \psi_j(s)\, v_j.$$

根据构造，$G \in \mathcal{A}_{S,\,T}^\eta$。

对每个 $s \in E$，我们估计

$$\begin{aligned}
\rho\big(F(s) - G(s)\big) &\le \rho\left(F(s) - \sum_{j=1}^{m} \phi_j(s)\, v_j \right) + \rho\left(\sum_{j=1}^{m} (\phi_j(s) - \psi_j(s))\, v_j \right) \\
&\le \frac{\varepsilon}{2} + \sum_{j=1}^{m} |\phi_j(s) - \psi_j(s)|\, \rho(v_j) \\
&< \frac{\varepsilon}{2} + \sum_{j=1}^{m} \frac{\varepsilon}{2mC}\, \rho(v_j) \le \varepsilon.
\end{aligned}$$

对 E 取上确界可得

$$\sup_{s\in E} \rho\big(F(s) - G(s)\big) < \varepsilon,$$

这就证明了所需的逼近。 □

前述定理证明了在半范数拓扑下，具有向量值系数的浅层神经网络在 $C(E; T)$ 中的稠密性。当 T 是巴拿赫空间时，定义半范数族可简化为范数，因此该结果成为关于巴拿赫范数的一致逼近定理。

注 2.6 (对偶表述)。设 T 是一个豪斯多夫局部凸拓扑向量空间，其连续对偶为 T'。假设网 $\{F_\alpha\} \subset C(E;\ T)$ 在 E 上关于 T 的半范数拓扑一致收敛于 F。那么对于每个 $t' \in T'$,

$$\sup_{s\in E} \left|t'(F_\alpha(s) - F(s))\right| \longrightarrow 0.$$

此外，如果 T 的拓扑由其对偶生成，则逆命题也成立。

3. 推论

本节将推导主定理的若干推论，并举例说明这些理论结果。

神经网络 (F_α) 在 $C(E;\ T)$ 中收敛于 F，当且仅当

$$\sup_{s\in E} \rho\big(F_\alpha(s) - F(s)\big) \longrightarrow 0 \quad \text{for every continuous seminorm } \rho \text{ on } T.$$

每当涉及序列空间或 L^p 空间时，我们都假设在标准假设下成立通常的对偶关系（例如，$(L^p)^* = L^{p'}$ 对于$1 < p < \infty$）。

推论 3.1 （希尔伯特空间值逼近）。设 S 是具有 *HBEP* 的拓扑向量空间，且 $E \subset S$ 是紧的。假设 T 是希尔伯特空间，并且 $\eta : \mathbb{R} \to \mathbb{R}$ 在某个开区间上连续且非多项式。

则逼近类

$$\mathcal{A}^{\eta}_{S,T} = \operatorname{span}\left\{s \mapsto \eta\big(\ell(s) - \theta\big)\, v : \ell \in S^*,\ \theta \in \mathbb{R},\ v \in T\right\}$$

在 $C(E;\ T)$ 中关于 E 上的一致范数拓扑是稠密的。在这种情况下，半范数拓扑退化为由希尔伯特空间结构所诱导的范数拓扑。

推论 3.2 （函数到函数逼近）。 设 (Ω_1, μ_1) 和 (Ω_2, μ_2)为测度空间，其中 $1 < p < \infty$ 且 $1 \le q < \infty$。令 $S = L^p(\Omega_1)$ 和$T = L^q(\Omega_2)$。假设 $E \subset L^p(\Omega_1)$ 在 L^p 拓扑下是紧的，并令 η满足定理 *2.1* 的假设。那么，具有以下形式的映射

$$f \longmapsto \sum_{j=1}^{m} \eta\left(\int_{\Omega_1} f(s)\, \varphi_j(s)\, d\mu_1(s) - \theta_j \right) g_j, \qquad \varphi_j \in L^{p'}(\Omega_1),\ g_j \in L^q(\Omega_2),$$

在 $C(E; L^q(\Omega_2))$ 中关于 E 上的一致收敛稠密。

推论 3.3 (序列到序列逼近)。 设 $1 < p < \infty$ 和$1 \le q < \infty$，并考虑 $S = \ell^p$, $T = \ell^q$。若 $E \subset \ell^p$ 是紧的，且 η满足定理 *2.1* 的假设，则具有以下形式的映射

$$s \longmapsto \sum_{j=1}^{m} \eta\left(\sum_{n=1}^{\infty} a_{j,n} s_n - \theta_j \right) v_j, \qquad (a_{j,n})_{n\ge 1} \in \ell^{p'},\ v_j \in \ell^q,$$

在 $C(E; \ell^q)$ 中关于 E 上的一致收敛稠密。

推论 3.4 (矩阵输入)。 设 $S = \mathbb{R}^{n\times m}$，并设 T 是豪斯多夫局部凸拓扑向量空间。若 $E \subset \mathbb{R}^{n\times m}$ 是紧的，且 η 满足定理 *2.1* 的条件，则具有以下形式的映射

$$Z \longmapsto \sum_{j=1}^{r} \eta\big(\operatorname{tr}(W_j^\top Z) - \theta_j \big)\, v_j, \qquad W_j \in \mathbb{R}^{n\times m},\ v_j \in T,$$

在 $C(E; T)$ 中关于半范数拓扑稠密。

推论 3.5 (有限秩算子表示)。 在定理 *2.1* 的假设下，每个 $F \in C(E; T)$ 都可由以下形式的映射在E 上作一致逼近

$$s \longmapsto \sum_{j=1}^{m} \eta\big(\phi_j(s) - \theta_j \big)\, v_j, \qquad \phi_j \in S^*,\ v_j \in T.$$

因此，在这种情形下，神经网络逼近元可以解释为有限秩算子。

现在，我们通过指出若干定理 2.1 适用的重要局部凸拓扑向量空间类别来作结。这些例子展示了这一抽象框架的灵活性，以及它对无限维情形下算子学习的相关性。

例 3.6 (光滑函数空间)。 设 $\Omega \subset \mathbb{R}^d$ 为开集，并令 $T = C^\infty(\Omega)$。赋予 $C^\infty(\Omega)$ 由以下半范数生成的通常弗雷歇拓扑

$$\rho_{E,\alpha}(f) := \sup_{s\in E} |\partial^\alpha f(s)|,$$

其中 $E \subset \Omega$ 是紧集， α 是一个多重指标。赋予该拓扑后， $C^\infty(\Omega)$是豪斯多夫局部凸拓扑向量空间。

根据定理 *2.1*，系数取自 $C^\infty(\Omega)$ 的神经网络逼近元在所有紧集 $E_0 \subset S$ 上都在 $C(E_0; C^\infty(\Omega))$ 中稠密，其中收敛是相对于每个半范数 $\rho_{E,\alpha}$来衡量的。这一设定自然适用于微分方程光滑解算子的逼近。

例 3.7 (施瓦茨空间)。 令 $T = \mathcal{S}(\mathbb{R}^d)$ 表示由速降光滑函数构成的施瓦茨空间。其标准弗雷歇拓扑由以下半范数生成:

$$\rho_{\alpha,\beta}(f) := \sup_{s\in\mathbb{R}^d} |x^\alpha \partial^\beta f(s)|,$$

其中 α 和 β 是多重指标。施瓦茨空间是核弗雷歇空间，因此也是豪斯多夫局部凸拓扑向量空间。

因此，定理 *2.1* 的稠密性结果可推广至 $C(E; \mathcal{S}(\mathbb{R}^d))$，其中的逼近是指相对于所有定义半范数同时成立。此类框架在信号分析、时频表示以及涉及快速衰减核的算子模型中尤为重要。

例 3.8 （分布空间）。 设 $\Omega \subset \mathbb{R}^d$ 为开集，并考虑 $T = \mathcal{D}'(\Omega)$，即 Ω 上的分布空间，配备与测试函数空间 $\mathcal{D}(\Omega)$ 的有界子集上的一致收敛相对应的强对偶拓扑。在此拓扑下， $\mathcal{D}'(\Omega)$ 是豪斯多夫局部凸空间。

若 $F\colon\ E \to \mathcal{D}'(\Omega)$ 关于强拓扑连续，且其值域包含于 $\mathcal{D}'(\Omega)$的某个有界子集中，则定理 *2.1* 保证在 E 上关于每个定义半范数的一致逼近。这一设定非常适合偏微分方程弱解或分布解的逼近。

与巴拿赫空间情形的关系。如果输出空间 T 是巴拿赫空间，则局部凸结构由单个范数生成。在这种情况下，由半范数族确定的一致收敛拓扑

与 $C(E;\,T)$ 上通常的一致范数拓扑一致。因此，取值于巴拿赫空间的逼近理论是本文所建立的更一般局部凸框架的一个特例。

4. 算子逼近的应用

本节将说明，上文所确立的向量值通用逼近定理如何自然地应用于函数空间之间的非线性算子逼近。这类映射常见于微分方程、反问题以及现代算子学习框架的分析中。

4.1. 非线性算子的逼近

设 S 和 T 为拓扑向量空间，$E \subset S$ 为紧集。考虑一个非线性算子

$$\mathcal{F} : E \longrightarrow T.$$

若 $\mathcal{F}$ 连续，则定理 2.1 保证 $\mathcal{F}$ 可以在 E 上由以下形式的映射一致逼近

$$\mathcal{G}(s) = \sum_{j=1}^{m} \eta\big(\ell_j(s) - \theta_j\big)\, v_j, \qquad \ell_j \in S^*,\ v_j \in T.$$

因此，采用标量激活函数和向量值系数的浅层神经网络，为无限维空间之间的连续算子提供了通用逼近器。

4.2. 积分算子的逼近

分析和科学计算中一类特别重要的算子是积分算子。设 $\Omega \subset \mathbb{R}^d$ 是一个有界区域，并考虑算子

$$(\mathcal{F}f)(x) = \int_{\Omega} K(x,s)\, f(s)\, ds,$$

其中 $K:\ \Omega \times \Omega \to \mathbb{R}$ 是一个连续核。

设 $S = L^p(\Omega)$ 配备 $1 < p < \infty$，且 $T = L^q(\Omega)$ 配备 $1 \le q < \infty$。假设 $E \subset L^p(\Omega)$ 是紧的，且 $\mathcal{F} : E \to L^q(\Omega)$ 是连续的。

根据定理 2.1，对于每个 $\varepsilon > 0$ 以及 $L^q(\Omega)$ 上的每个连续半范 ρ，存在一个神经网络

$$\mathcal{G}(f) = \sum_{j=1}^{m} \eta\left(\int_{\Omega} f(s)\, \varphi_j(s)\, ds - \theta_j \right) g_j(x), \qquad \varphi_j \in L^{p'}(\Omega),\ g_j \in L^q(\Omega),$$

使得

$$\sup_{f \in E} \rho\big(\mathcal{F}(f) - \mathcal{G}(f)\big) < \varepsilon.$$

因此，上述形式的神经网络能够在 $L^p(\Omega)$ 的紧子集上一致逼近积分算子。

4.3. 与神经算子学习的联系
该表示

$$\mathcal{G}(f) = \sum_{j=1}^{m} \eta\big(\ell_j(f) - \theta_j\big)\, v_j$$

在结构上类似于科学机器学习中使用的现代神经算子架构。在这些模型中，函数 $\ell_j(f)$ 充当传感器或测量泛函，作用于输入函数 f；而向量 v_j 则表示输出空间中的基函数。

因此，定理 2.1 为浅层神经算子模型提供了严格的泛函分析基础，表明此类架构能够逼近无限维空间之间的连续算子。

4.4. 偏微分方程解算子的逼近数学物理和工程学中的许多问题都可以表述为算子方程，其中偏微分方程（PDE）的解取决于一个输入函数。在这类情形下，人们关注的是对相应解算子进行逼近。

设 $\Omega \subset \mathbb{R}^d$ 为一个有界区域，并考虑如下形式的偏微分方程：

$$\mathcal{L}u = f \quad \text{in } \Omega,$$

在适当的边界条件下，其中 $\mathcal{L}$ 是微分算子，f 表示输入项。

在标准假设下，该问题定义了一个解算子

$$\mathcal{S} : S \longrightarrow T, \qquad f \longmapsto u,$$

其中， S 是一个包含容许输入的函数空间（例如 $L^2(\Omega)$ 或 $H^{-1}(\Omega)$），而 T 是一个描述解的函数空间（例如 $H_0^1(\Omega)$ 或 $C(\Omega)$ ）。

如果解算子 $\mathcal{S}$ 在紧子集 $E \subset S$上连续，那么定理 2.1 表明， $\mathcal{S}$ 可以在 E 上由如下形式的神经网络一致逼近：

$$\mathcal{G}(f) = \sum_{j=1}^{m} \eta\big(\ell_j(f) - \theta_j\big)\, v_j, \qquad \ell_j \in S^*,\ v_j \in T.$$

特别地，当 $S = L^p(\Omega)$ 且 $1 < p < \infty$时，泛函 ℓ_j 可表示为

$$\ell_j(f) = \int_\Omega f(s)\, \varphi_j(s)\, ds, \qquad \varphi_j \in L^{p'}(\Omega).$$

因此，逼近算子具有如下形式：

$$\mathcal{G}(f) = \sum_{j=1}^{m} \eta\left(\int_\Omega f(s)\, \varphi_j(s)\, ds - \theta_j\right) g_j(x), \qquad g_j \in T.$$

因此，浅层神经算子架构能够在输入空间的紧子集上逼近偏微分方程的解算子。

此类逼近结果为现代算子学习框架提供了理论依据。在这些框架中，神经网络经过训练，将输入函数映射为相应的偏微分方程解。

5. 结论

本文为输入属于拓扑向量空间、输出取值于豪斯多夫局部凸拓扑向量空间的浅层神经网络建立了通用逼近定理。所提出的框架表明，具有标量激活函数和向量值系数的神经网络所生成的有限神经展开，在 $C(E;\, T)$ 中

紧子集 $E \subset S$ 上稠密，其中收敛是相对于目标空间 T 的半范数拓扑来理解的。

该结果将神经网络的标量值通用逼近定理推广到定义在拓扑向量空间上的情形，并自然地将巴拿赫空间值和希尔伯特空间值逼近纳入为特例。因此，向量值映射的经典一致逼近结果可视为更一般的局部凸逼近框架中的特例。由此，该理论为作用于无限维空间之间的浅层神经网络模型奠定了泛函分析基础。

我们还展示了这些理论结果在若干重要场景中的适用性。具体而言，该框架能够逼近函数空间之间的非线性算子，包括积分算子以及偏微分方程的解算子。这些例子表明，所提出的神经网络架构可作为算子学习和科学计算中各类映射的通用逼近器。

未来仍有若干研究方向值得探索。例如，可以在附加正则性假设下推导定量逼近率，研究在同一局部凸框架下向更深层神经网络架构的扩展，并探索涉及随机输入或算子值激活机制的变体。

利益冲突声明

作者声明，其不存在可能影响本文所报告工作的竞争性经济利益或个人关系。

参考文献

[1] G. Cybenko，通过S形函数的叠加进行逼近，《数学控制与信号系统》2 (4) (1989) 303–314。[2] K. I. Funahashi，神经网络对连续映射的近似实现，《神经网络》2 (3) (1989) 183–192。[3] K. Hornik，多层前馈网络的逼近能力，《神经网络》4 (2) (1991) 251–257。[4] A. Pinkus，神经网络中MLP模型的逼近理论，《数值学报》8 (1999) 143–195。

[5] M. Leshno、V. Y. Lin、A. Pinkus、S. Schocken，具有非多项式激活函数的多层前馈网络能够逼近任意函数，《神经网络》6 (6) (1993) 861–867。[6] X. Sun、E. W. Cheney，脊函数集的基础性，等式数学44 (2) (1992) 226–235。[7] W. Light，脊函数、S形函数与神经网络，逼近理论杂志 (1992) VIIi–44。[8] T. Chen、H. Chen，具有任意激活函数的神经网络对非线性算子的通用逼近及其在动力系统中的应用，IEEE 神经网络汇刊6 (4) (1995) 911–917。[9] L. Lu、P. Jin、G. Pang、Z. Zhang、G. E. Karniadakis，基于算子通用逼近定理，通过DeepONet学习非线性算子，《自然·机器智能》3 (3) (2021) 218–229。[10] S. Lanthaler、S. Mishra、G. E. Karniadakis，DeepONet的误差估计：无限维中的深度学习框架，《数学应用汇刊》6 (1) (2022) tnac001。[11] Y. Korolev，取值于巴拿赫空间的双层神经网络，SIAM数学分析杂志54 (6) (2022) 6358–6389。[12] M. E. Valle、W. L. Vital、G. Vieira，向量值与超复数值神经网络的通用逼近定理，《神经网络》180 (2024) 106632。[13] V. E. Ismailov，输入来自拓扑向量空间的神经网络通用逼近定理，《信息处理快报》 (2026) 106623。