Globally injective and bijective neural operators

Takashi Furuya

in collaboration with
Michael Puthawala
Matti Lassas
Maarten V. de Hoop

Motivation

Neural Operators learn mappings between infinite dimensional function spaces. Their analytical properties, including injectivity and bijectivity, are poorly understood.

In this work we extend prior work for finite-dimensional networks to the infinite-dimensional setting. Our work enables applications for

- Generative models in infinite-dimensional function space
- PDE-based inverse problems

We show that injective neural operators are universal approximators (Theorem 2) and, under appropriate assumptions, may be inverted by neural operators (Theorem 3).

Neural operator

- $D \subset \mathbb{R}^d$, Lipschitz bounded domain
- $L^2(D;\mathbb{R}^h) = L^2(D)^h$, L^2 space of $\mathbb{R}^h$-value function on D

Definition 1 (Neural operators [Kovachki et al., 2021])

We define a neural operator $G : L^2(D)^{d_{in}} \to L^2(D)^{d_{out}}$ by

$$G := T_{L+1} \circ \mathcal{L}_L \circ \cdots \mathcal{L}_1 \circ T_0,$$

$$\mathcal{L}_\ell : L^2(D)^{d_\ell} \to L^2(D)^{d_\ell+1}, \quad (\mathcal{L}_\ell v)(x) := \sigma(W_\ell(x)v(x) + K_\ell v(x) + b_\ell(x)),$$

- $\sigma : \mathbb{R} \to \mathbb{R}$, non-linear activation operating element-wise
- $W_\ell \in C(\overline{D}; \mathbb{R}^{d_{\ell+1} \times d_\ell})$, pointwise matrix multiplications,
- $K_\ell : L^2(D)^{d_\ell} \to L^2(D)^{d_\ell+1}$, linear integral operators,
- $b_\ell \in L^2(D)^{d_{\ell+1}}$, bias functions
- $T_0 : L^2(D)^{d_{in}} \to L^2(D)^{d_1}$, lifting operator
- $T_{L+1} : L^2(D)^{d_{L+1}} \to L^2(D)^{d_{out}}$, projection operator

Class of neural operators

We define

$$\begin{aligned}
&\mathrm{NO}_L(\sigma; D, d_{in}, d_{out}) := \Big\{ G : L^2(D)^{d_{in}} \to L^2(D)^{d_{out}} \Big| \\
&G = K_{L+1} \circ (K_L + b_L) \circ \sigma \cdots \circ (K_2 + b_2) \circ \sigma \circ (K_1 + b_1) \circ (K_0 + b_0), \\
&K_\ell : f \mapsto \int_D k_\ell(\cdot, y) f(y) dy \Big|_D, \ k_\ell \in L^2(D \times D; \mathbb{R}^{d_{\ell+1} \times d_\ell}), \\
&b_\ell \in L^2(D; \mathbb{R}^{d_{\ell+1}}), \ d_\ell \in \mathbb{N}, \ d_0 = d_{in}, \ d_{L+2} = d_{out}, \ \ell = 0, ..., L+2 \Big\},
\end{aligned}$$

and

$$\mathrm{NO}_L^{inj}(\sigma; D, d_{in}, d_{out}) := \{ G \in \mathrm{NO}_L(\sigma; D, d_{in}, d_{out}) : G \text{ is injective} \}.$$

Universal approximation theorem

Theorem 2

Let $G^+ : L^2(D)^{d_{in}} \to L^2(D)^{d_{out}}$ *be continuous such that for all* $R > 0$ *there is* $M > 0$ *so that*

$$\left\|G^+(a)\right\|_{L^2(D)^{d_{out}}} \leq M, \ \forall a \in L^2(D)^{d_{in}}, \ \|a\|_{L^2(D)^{d_{in}}} \leq R,$$

We assume that either $\sigma = \mathrm{Leaky\ ReLU}$ *or* $\sigma = \mathrm{ReLU}$. *Then, for any compact set* $K \subset L^2(D)^{d_{in}}$, $\epsilon \in (0,1)$, *there exists* $L \in \mathbb{N}$ *and* $G \in \mathrm{NO}_L^{inj}(\sigma; D, d_{in}, d_{out})$ *such that*

$$\sup_{a \in K} \left\|G^+(a) - G(a)\right\|_{L^2(D)^{d_{out}}} \leq \epsilon.$$

We don't have any dimensionality restrictions. In the case of Euclidean spaces $\mathbb{R}^d$, [Puthawala et al., 2022] requires that $2d_{in} + 1 \leq d_{out}$ before all continuous functions $G^+ : \mathbb{R}^{d_{in}} \to \mathbb{R}^{d_{out}}$ can be uniformly approximated in compact sets by injective neural networks.

Non-linear neural operator

We consider layers of the form

$$(\mathcal{L}_\ell v)(x) = \sigma(W_\ell(x)v(x) + K_\ell(v)(x)),\ x \in D,$$

where K_ℓ is non-linear integral operators

$$K_\ell(u)(x) = \int_D k_\ell(x, y, u(x), u(y))u(y)dy,$$

- Generalization of the attention mechanism in transformers [Kovachki et al., 2021]
$$k(x, y, v(x), v(y)) \equiv \mathrm{softmax} \circ < Av(x), Bv(y) >,$$
- Improve performance of integral autoencoders [Ong et al., 2022]

Construction of the inverse

As simple case, $n = 1$, $D \subset \mathbb{R}$ is a bounded interval. Consider a map $F : L^2(D) \to L^2(D)$ defined by

$$F(u)(x) = W(x)u(x) + \int_D k(x, y, u(y))u(y)dy, \ u \in L^2(D),$$

where $W \in C^1(\overline{D}; \mathbb{R})$ satisfies $0 < c_1 \leq W(x) \leq c_2$, $k \in C^3(\overline{D} \times \overline{D} \times \mathbb{R}; \mathbb{R})$ and

$$\|W\|_{C^1(\overline{D})} \leq c_0, \quad \|k\|_{C^3(\overline{D} \times \overline{D} \times \mathbb{R})} \leq c_0,$$

and for all $u_0 \in H^1(D)$, the Fréchet derivative

$$DF[u_0] : H^1(D) \to H^1(D) \text{ is injective.}$$

Theorem 3

Assume that $F : H^1(D) \to H^1(D)$ is bijective. Let $\mathcal{Y} \subset \overline{B}_{C^{1,\alpha}(\overline{D})}(0, R)$where $\alpha > 0$. The inverse of $F : H^1(D) \to H^1(D)$ in $\mathcal{Y}$ can be written as a limit of neural operators having distributional kernels.

References I

Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces. *arXiv preprint arXiv:2108.08481*, 2021.

Yong Zheng Ong, Zuowei Shen, and Haizhao Yang. Iae-net: Integral autoencoders for discretization-invariant learning. *arXiv preprint arXiv:2203.05142*, 2022.

Michael Puthawala, Konik Kothari, Matti Lassas, Ivan Dokmanić, and Maarten de Hoop. Globally injective relu networks. *Journal of Machine Learning Research*, 23(105):1–55, 2022.

Thank you ! !