Quantitative Approximation for Neural Operators in Nonlinear Parabolic Equations

Takashi Furuya

Doshisha University

Joint work with

Koichi Taniguchi (Shizuoka University)
Satoshi Okuda(Rikkyo University)

Neural operators

Neural operators have gained significant attention in deep learning as an extension of traditional neural networks between infinite dimensional function spaces.

A key application is in constructing surrogate models of solvers for PDEs by learning their solution operators.

Neural operators serve as universal approximators for general continuous operators. This indicates that these learning methods possess the capabilities to approximate a wide range of operators.

Quantitative approximation

However, operator learning for general operators suffers from "the curse of parametric complexity", where the number of learnable parameters exponentially grows as the desired approximation accuracy increases [Lanthaler and Stuart, 2023].

A common approach to mitigating the curse of parametric complexity is to restrict general operators to the solution operators of PDEs.

- [Kovachki et al., 2021, Lanthaler, 2023] ; Darcy and Navier-Stokes equations
- [Lanthaler and Stuart, 2023] ; Hamilton-Jacobi equations
- [Deng et al., 2022] ; Advection-diffusion equations
- [Chen et al., 2023, Lanthaler et al., 2022, Marcati and Schwab, 2023] ; Elliptic, parabolic, and hyperbolic equations

Here, we present the quantitative approximation for Nonlinear Parabolic equation:

Nonlinear Parabolic equation

We consider the nonlinear parabolic equation

$$\begin{cases} \partial_t u + \mathcal{L}u = F(u) & \text{in } (0,T) \times D, \\ u(0) = u_0 & \text{in } D, \end{cases}$$

where

- $\mathcal{L}$ is a certain operator (e.g. $\mathcal{L} = -\Delta$ is the Laplacian)
- $F \in C^1(\mathbb{R};\mathbb{R})$ satisfies $F(0) = 0$ and

$$|F(z_1) - F(z_2)| \leq C_F \max_{i=1,2} |z_i|^{p-1} |z_1 - z_2|,$$

  for any $z_1, z_2 \in \mathbb{R}$ and for some $p > 1$ and $C_F > 0$.
- $r, s \in [p, \infty]$ and $\frac{\nu}{s} + \frac{1}{r} < \frac{1}{p-1}$.

We denote $\Gamma^+ : L^\infty(D;\mathbb{R}) \to L^r(0,T;L^s(D;\mathbb{R}))$ by

$$\Gamma^+ : u_0 \mapsto u.$$

Nonlinear Parabolic equation

Definition 1

We define a neural operator $\Gamma : L^\infty(D;\mathbb{R}) \to L^r(0,T;L^s(D;\mathbb{R}))$ by $\Gamma : v_0 \mapsto v_{L+1}$ where v_{L+1} is give by the following steps:

$$v_{\ell+1}(t,x) = \sigma\left(W^{(\ell)} v_\ell(t,x) + (K_N^{(\ell)} v_\ell)(t,x) + b_N^{(\ell)}(t,x)\right),$$
$$v_{L+1}(t,x) = W^{(L)} v_L(t,x) + (K_N^{(L)} v_L)(t,x) + b_N^{(L)}(t,x),$$

where $\sigma : \mathbb{R} \to \mathbb{R}$ is a non-linear activation function, and $W^{(\ell)} \in \mathbb{R}^{d_{\ell+1} \times d_\ell}$ and $K_N^{(\ell)} : L^r(0,T;L^s(D;\mathbb{R}^{d_\ell})) \to L^r(0,T;L^s(D;\mathbb{R}^{d_{\ell+1}}))$ and $b_N^{(\ell)} \in L^r(0,T;L^s(D;\mathbb{R}^{d_{\ell+1}}))$ defined by

$$(K_N^{(\ell)} v)(t,x) := \sum_{m,n \in \Lambda_N} K_{m,n}^{(\ell)} \langle \psi_m, v \rangle \varphi_n(t,x) \quad \text{with } K_{m,n}^{(\ell)} \in \mathbb{R}^{d_{\ell+1} \times d_\ell},$$

$$b_N^{(\ell)}(t,x) := \sum_{n \in \Lambda_N} b_n^{(\ell)} \varphi_n(t,x) \quad \text{with } b_n^{(\ell)} \in \mathbb{R}^{d_{\ell+1}},$$

Nonlinear Parabolic equation

We have denoted by

- $\langle f, g \rangle = \int_0^T \int_D fg dx dt$
- $\varphi := \{\varphi_n\}_{n\in\Lambda}$ and $\psi := \{\psi_m\}_{m\in\Lambda}$ are families of functions in $L^r(0,T;L^s(D))$ and $L^{r'}(0,T;L^{s'}(D))$, respectively.
- Λ is an index set that is either finite or countably infinite and Λ_N a subset of Λ with its cardinality $|\Lambda_N| = N \in \mathbb{N}$ and the monotonicity $\Lambda_N \subset \Lambda_{N'}$ for any $N \leq N'$.

We denote by

$$\mathcal{NO}^{L,H,\sigma}_{N,\varphi,\psi}(L^\infty(D;\mathbb{R}), L^r(0,T;L^s(D;\mathbb{R})))$$

the class of neural operators defined above, with the depth L, the number of neurons $H = \sum_{\ell=1}^{L} d_\ell$, the activation function σ, the rank N, and the families of functions φ, ψ.

Nonlinear Parabolic equation

Let $G(t,x,y)$ be the Green's function for the linear differential operator $\partial_t + \mathcal{L}$. Assume that

$$E_G(N) := \left\| \| G(t-\tau,x,y) - G_N(t-\tau,x,y) \|_{L^{r'}_\tau(0,T;L^{s'}_y)} \right\|_{L^r_t(0,T;L^s_x)} \to 0,$$

where

$$G_N(t-\tau,x,y) := \sum_{m,n\in\Lambda_N} c_{n,m}\psi_m(\tau,y)\varphi_n(t,x), \quad 0 \le \tau, t < T, \ x,y \in D$$

Nonlinear Parabolic equation

Theorem 2

In the above setting, for any $R > 0$, there exists a time $T > 0$ such that the following statement holds: For any $\epsilon \in (0, 1)$, there exist $L, H, N \in \mathbb{N}$, and $\Gamma \in \mathcal{NO}^{L,H,ReLU}_{N,\varphi,\psi}(L^\infty(D;\mathbb{R}), L^r(0,T;L^s(D;\mathbb{R})))$ such that

$$\sup_{u_0 \in B_{L^\infty}(0,R)} \|\Gamma^+(u_0) - \Gamma(u_0)\|_{L^r(0,T;L^s)} \leq \epsilon.$$

Moreover, L, H, and N satisfy

$$L \lesssim (\log(\epsilon^{-1}))^2, \quad H \lesssim \epsilon^{-1}(\log(\epsilon^{-1}))^2, \quad E_G(N) \lesssim \epsilon$$

Sketch of proof

$$\partial_t u + \mathcal{L}u = F(u) \quad \text{in } (0,T) \times D, \quad u(0) = u_0 \quad \text{in } D \iff$$

$$u = \Phi(u) := \int_D G(t,x,y) u_0(y)\, dy + \int_0^t \int_D G(t-\tau,x,y) F(u(\tau,y))\, dy d\tau$$

Showing that Φ is a contraction map, the above solution can be approximated by the Picard's iteration u_L where

$$u_{\ell+1} = \Phi(u_\ell).$$

Then, we approximate the map Φ with a layer of neural operator.

This proof technique of the Picard' iteration could be applied to a wide range of PDE for the application of neural operators.

References I

Ke Chen, Chunmei Wang, and Haizhao Yang. Deep operator learning lessens the curse of dimensionality for PDEs. *Transactions on Machine Learning Research*, 2023. ISSN 2835-8856.

Beichuan Deng, Yeonjong Shin, Lu Lu, Zhongqiang Zhang, and George Em Karniadakis. Approximation rates of deeponets for learning operators arising from advection–diffusion equations. *Neural Networks*, 153:411–426, 2022. ISSN 0893-6080.

Nikola Kovachki, Samuel Lanthaler, and Siddhartha Mishra. On universal approximation and error bounds for fourier neural operators. *Journal of Machine Learning Research*, 22:Art–No, 2021.

Samuel Lanthaler. Operator learning with PCA-Net: upper and lower complexity bounds. *Journal of Machine Learning Research*, 24(318): 1–67, 2023.

References II

Samuel Lanthaler and Andrew M Stuart. The curse of dimensionality in operator learning. *arXiv preprint arXiv:2306.15924*, 2023.

Samuel Lanthaler, Siddhartha Mishra, and George E Karniadakis. Error estimates for deeponets: A deep learning framework in infinite dimensions. *Transactions of Mathematics and Its Applications*, 6(1): tnac001, 2022.

Carlo Marcati and Christoph Schwab. Exponential convergence of deep operator networks for elliptic partial differential equations. *SIAM Journal on Numerical Analysis*, 61(3):1513–1545, 2023.