

Friday - March 4, 2022

Over-parameterization in manifold based surrogates and deep neural operators

Somdatta Goswami, Brown University

Constructing accurate and generalizable approximators (surrogate models) for complex physico-chemical processes exhibiting highly non-smooth dynamics is challenging. The main question is what type of surrogate models we should construct and should these models be under-parametrized, deterministic or over-parametrized; the latter refers to the number of model parameters compared to available data. In this work, we propose new developments and perform comparisons for two promising approaches: manifold-based polynomial chaos expansion (m-PCE) and the deep neural operator (DeepONet), and we examine the effect of over-parametrization on generalization. While m-PCE enables the construction of a mapping by first identifying low-dimensional embeddings of the input functions, parameters, and quantities of interest (QoIs), a neural operator learns the nonlinear mapping via the use of deep neural networks. Both models are capable of overcoming the curse-of-dimensionality in the input space. We demonstrate the performance of these methods in terms of generalization accuracy for solving the 2D time-dependent Brusselator reaction-diffusion system with multiple sources, modeling an autocatalytic chemical reaction between two species. To tackle this problem, we first propose an extension of the m-PCE method by constructing a mapping between latent spaces formed by two separate embeddings that represent the input functions and QoIs. To further enhance the accuracy of DeepONet, we introduce weight self-adaptivity in the loss function. We demonstrate that the performance of m-PCE and DeepONet is comparable for cases of relatively smooth input-output mappings, however, when highly non-smooth dynamics is considered, DeepONet shows higher approximation accuracy. We also find that for m-PCE, modest over-parametrization leads to better generalization, both within and outside of distribution (OOD), whereas aggressive over-parametrization leads to over-fitting. In contrast, an even highly over-parametrized DeepONet leads to better generalization for both smooth and non-smooth dynamics. Furthermore, we compare the performance of the above models with another recently proposed operator learning model, namely the Fourier Neural Operator (FNO), and show that its over-parametrization also leads to better generalization. Taking together, our studies show that m-PCE can provide very good accuracy at very low training cost, whereas a highly over-parametrized DeepONet can provide better accuracy and robustness to noise but at higher training cost. In both methods, the inference cost is negligible.