CRUNCH Seminars at Brown, Division of Applied Mathematics

Friday - July 23, 2021

Partition of Unity Networks for Deterministic and Probabilistic Regression

Mamikon Gulian, Sandia National Laboratories

In regression tasks, deep neural networks encode both a basis adapted to data and a representation of the data in that basis. Thus, training a DNN to solve a partial differential equations using a physics-informed loss is analogous to using an optimizer to construct a mesh, approximation space, and a representation within the approximation space all at once. We know from decades of experience in numerical simulations that this process is highly nonlinear, complex, and often requires specialized expertise. While DNNs have shown remarkable success using off-the-shelf architectures and optimizers in certain problems, it is not surprising that for such a Herculean task, DNNs frequently demonstrate a lack of convergence. This is despite theoretical work showing that optimal approximations within a specific architecture converge algebraically as the depth of the network increases.

Motivated by this, we propose partition of unity networks (POU-Nets) which incorporate partitions of unity and polynomial approximants directly into the architecture. Classification architectures of the type used to learn probability measures are used to build a mesh-free partition of space, while polynomial spaces with learnable coefficients are associated to each partition. The resulting hp-element-like approximation allows use of a fast least-squares optimizer, and the resulting architecture size need not scale exponentially with spatial dimension, breaking the curse of dimensionality. We then enrich POU-Nets with a Gaussian mixture noise model to obtain a probabilistic generalization amenable to gradient-based minimization of a maximum likelihood loss. The resulting architecture provides spatial representations of both noiseless and noisy data as Gaussian mixtures with closed form expressions for variance that provide estimates of local error. The partitions are amenable to a hierarchical refinement strategy that significantly improves the localization of the regression, and the framework scales more favorably to large data sets as compared to Gaussian process regression and allows for spatially varying uncertainty, leveraging the expressive power of deep neural networks while bypassing expensive training associated with other probabilistic deep learning methods. Compared to standard POU-Nets, the framework demonstrates hp-convergence without the use of regularizers to tune the localization of partitions.