CRUNCH Seminars at Brown, Division of Applied Mathematics

Friday - September 22, 2023

On the training and generalization of deep operator networks.

Yeonjong Shin, North Carolina State University

In this talk, we will present a novel training method for deep operator networks (DeepONets). DeepONets are constructed by two sub-networks, namely the branch and trunk networks. Typically, the two sub-networks are trained simultaneously, which amounts to solving a complex optimization problem in a high dimensional space. In addition, the nonconvex and nonlinear nature makes training very challenging. To tackle such a challenge, we propose a two-step training method that trains the trunk network first and then sequentially trains the branch network. The core mechanism is motivated by the divide-and-conquer paradigm and is the decomposition of the entire complex training task into two subtasks with reduced complexity. Therein the Gram-Schmidt orthonormalization process is introduced which significantly improves stability and generalization ability. On the theoretical side, we establish a generalization error estimate in terms of the number of training data, the width of DeepONets, and the number of input and output sensors. Numerical examples are presented to demonstrate the effectiveness of the two-step training method, including Darcy flow in heterogeneous porous media.