CRUNCH Seminars at Brown, Division of Applied Mathematics

Friday - June 11, 2021

Can stable and accurate neural networks be computed? On the barriers of deep learning and Smale's 18th problem

Matthew Colbrook, University of Cambridge

Deep learning (DL) has had unprecedented success and is now entering scientific computing with full force. However, DL suffers from a universal phenomenon: instability, despite universal approximating properties that often guarantee the existence of stable neural networks (NNs). We show the following paradox. There are basic well-conditioned problems in scientific computing where one can prove the existence of NNs with great approximation qualities, however, there does not exist any algorithm, even randomised, that can train (or compute) such a NN. Indeed, for any positive integers $K > 2$ and L, there are cases where simultaneously: (a) no randomised algorithm can compute a NN correct to K digits with probability greater than $1/2$, (b) there exists a deterministic algorithm that computes a NN with $K-1$ correct digits, but any such (even randomised) algorithm needs arbitrarily many training data, (c) there exists a deterministic algorithm that computes a NN with $K-2$ correct digits using no more than L training samples. These results provide basic foundations for Smale's 18th problem and imply a potentially vast, and crucial, classification theory describing conditions under which (stable) NNs with a given accuracy can be computed by an algorithm. We begin this theory by initiating a unified theory for compressed sensing and DL, leading to sufficient conditions for the existence of algorithms that compute stable NNs in inverse problems. We introduce Fast Iterative REstarted NETworks (FIRENETs), which we prove and numerically verify are stable. Moreover, we prove that only $\mathcal{O}(|\log(\epsilon)|)$ layers are needed for an ϵ accurate solution to the inverse problem (exponential convergence), and that the inner dimensions in the layers do not exceed the dimension of the inverse problem. Thus, FIRENETs are computationally very efficient