A Bernstein-type Result for Stable Approximation with RNNs

Shida Wang
National University of Singapore, Singapore
e0622338@u.nus.edu

We prove an inverse approximation theorem for the approximation of nonlinear sequence-to-sequence relationships using RNNs. This is a so-called Bernstein-type result in approximation theory, which deduces properties of a target function under the assumption that it can be effectively approximated by a hypothesis space. In particular, we show that nonlinear sequence relationships, viewed as functional sequences, that can be stably approximated by RNNs with hardtanh/tanh activations must have an exponential decaying memory structure - a notion that can be made precise. This extends the previously identified curse of memory in linear RNNs into the general nonlinear setting and quantifies the essential limitations of the RNN architecture for learning sequential relationships with long-term memory. Our theoretical results are confirmed by numerical experiments.

Joint work with Qianxiao Li(National University of Singapore) and Zhong Li(MSRA).