Phase Transition of Landscape From Narrow to Wide Neural Networks

Ruoyu Sun

University of Illinois at Urbana-Champaign
E-mail: ruoyus@illinois.edu

Abstract: Wide neural networks are believed to have nice landscape, but what rigorous results can we prove, using just the condition of "wide"? We will show that: (i) From under-parameterized to over-parameterized networks, there is a phase transition from having sub-optimal basins to no sub-optimal basins. More specifically, for a dense set of activations and generic data, narrow networks have sub-optimal basins, while for all continuous activations, wide deep neural-nets have no sub-optimal basins. (ii) Over-parameterization alone cannot eliminate bad non-strict local minima, at least for a class of neurons. These results lead to a conjecture for lottery ticket hypothesis (Frankel and Carbin, ICLR'19): the lottery tickets in narrow networks are bad basins. These results are for un-regularized networks; time permitting, we will discuss results showing that with proper regularizers, even non-strict local minima can be eliminated.