

On the Expected Complexity of Maxout Networks

Hanna Tseran

Max-Planck-Institut für Mathematik in den Naturwissenschaften, 04103 Leipzig

Learning with neural networks relies on the complexity of the representable functions, but more importantly, the particular assignment of typical parameters to functions of different complexity. Taking the number of activation regions as a complexity measure, recent works have shown that the practical complexity of deep ReLU networks is often far from the theoretical maximum. In this work we show that this phenomenon also occurs in networks with maxout (multi-argument) activation functions and when considering the decision boundaries in classification tasks. We also show that the parameter space has a multitude of full-dimensional regions with widely different complexity, and obtain nontrivial lower bounds on the expected complexity. Finally, we investigate different parameter initialization procedures and show that they can increase the speed of convergence in training.