

Training and Generalization Errors for Underparameterized Neural Networks

Publié le 10 juillet 2024 – 2024 American Control Conference

Auteurs : Daniel Martin Xavier, Ludovic Chamoin, Laurent Fribourg

It has been theoretically explained, through the notion of Neural Tangent Kernel, why the training error of overparameterized networks converges linearly to 0. In this work, we focus on the case of small (or underparameterized) networks. An advantage of small networks is that they are faster to train while retaining sufficient precision to perform useful tasks in many applications. Our main theoretical contribution is to prove that the training error of small networks converges linearly to a (non-null) constant, of which we give a precise estimate. We verify this result on a neural network of 10 neurons simulating a Model Predictive Controller. We also observe that an upper bound of the generalization error follows a double-peak curve as the number of training data increases.