A general approximation lower bound in L^p norm, with applications to feed-forward neural networks

El Mehdi Achour
RWTH Aachen University, Germany
achour@mathc.rwth-aachen.de

We study the fundamental limits to the expressive power of neural networks. Given two sets F, G of real-valued functions, we first prove a general lower bound on how well functions in F can be approximated in $L^p(\mu)$ norm by functions in G, for any $p \geq 1$ and any probability measure μ. The lower bound depends on the packing number of F, the range of F, and the fat-shattering dimension of G. We then instantiate this bound to the case where G corresponds to a piecewise-polynomial feed-forward neural network, and describe in details the application to two sets F: Hölder balls and multivariate monotonic functions. Beside matching (known or new) upper bounds up to log factors, our lower bounds shed some light on the similarities or differences between approximation in L^p norm or in sup norm, solving an open question by DeVore et al. (2021). Our proof strategy differs from the sup norm case and uses a key probability result of Mendelson (2002).

Joint work with Armand Foucault (Institut de Mathématiques de Toulouse), Sébastien Gerchinovitz (Institut de Recherche Technologique Saint Exupéry) and François Malgouyres (Institut de Mathématiques de Toulouse).