Convergence rates for stochastic gradient descent algorithms in non-convex loss landscapes

Benjamin Fehrman

University of Oxford, Mathematical Institute, UK

In this talk, we will establish a rate of convergence for stochastic gradient descent algorithms in loss landscapes that are not necessarily locally convex or globally attracting. We will make only local assumptions on the objective function near its minimizers, and will rely on a quantitative use of mini-batches to control the loss of SGD iterates to non-attracted regions. In particular, the critical points are not assumed to be isolated, and the analysis can be used to treat the type of degeneracies observed practically in the optimization of certain neural networks.