Implicit Bias of Mirror Descent for Shallow Neural Networks in Univariate Regression

Guido Montufar

University of California, Los Angeles

We examine the implicit bias of mirror flow in univariate least squares error regression with wide and shallow neural networks. For a broad class of potential functions, we show that mirror flow exhibits lazy training and has the same implicit bias as ordinary gradient flow when the network width tends to infinity. For ReLU networks, we characterize this bias through a variational problem in function space. Our analysis includes prior results for ordinary gradient flow as a special case and lifts limitations which required either an intractable adjustment of the training data or networks with skip connections. We further introduce scaled potentials and show that for these, mirror flow still exhibits lazy training but is not in the kernel regime. For networks with absolute value activations, we show that mirror flow with scaled potentials induces a rich class of biases, which generally cannot be captured by an RKHS norm. A takeaway is that whereas the parameter initialization determines how strongly the curvature of the learned function is penalized at different locations of the input space, the scaled potential determines how the different magnitudes of the curvature are penalized. This is work with Shuang Liang.