

Reza Gheissari: High-dimensional limit theorems for stochastic gradient descent

Stochastic gradient descent (SGD) is the go-to method for large-scale optimization problems in modern data science. Often, these settings are data constrained, so the sample size and the dimension of the parameter space have to scale together. In this "high-dimensional" scaling, we study pathwise limits of the SGD. Namely, we show limit theorems for trajectories of reasonable summary statistics (finite-dimensional functions of the SGD) as the dimension goes to infinity and the step size simultaneously goes to zero. The limits that arise can be complicated ODE's or SDE's, depending on the initialization, step-size and space rescaling. We present these general results, and then discuss their implications for some concrete tasks including classification of XOR Gaussian mixtures via two layer networks. Based on joint work with Gerard Ben Arous and Aukosh Jagannath.