
ELIZABETH COLLINS-WOODFIN, McGill University

High dimensional limit of streaming SGD for generalized linear models

We provide a characterization of the high dimensional limit of one-pass, single batch stochastic gradient descent (SGD) in the case where the number of samples scales proportionally with the problem dimension. We characterize the limiting process in terms of its convergence to a high-dimensional stochastic differential equation, referred to as the homogenized SGD. Our proofs assume Gaussian data but allow for a very general covariance structure. Our set-up covers a range of optimization problems including linear regression, logistic regression, and some simple neural nets. For each of these models, the convergence of SGD to homogenized SGD enables us to derive a close approximation of the statistical risk (with explicit and vanishing error bounds) as the solution to a Volterra integral equation. In a separate paper, we perform similar analysis without the Gaussian assumption in the case of SGD for linear regression. (Based on joint work with C. Paquette, E. Paquette, I. Seroussi).