Developing a Multi-fidelity Approach to Numerically Optimize Stochastic Variance Reduced Gradient Descent

Neymika Jain
Mentors: Elizabeth Qian and Pan Xu

Stochastic gradient descent (SGD) is a commonly used optimization method for regression problems but has a sublinear convergence rate due to variance introduced by random data selection. Stochastic variance reduced gradient descent (SVRG) achieves linear convergence for strongly convex functions by correcting SGD's update rule using the full gradient. This method of correction is similar to principles used in multifidelity Monte Carlo methods, which reduce Monte Carlo estimator variance. We empirically test different theoretically-supported choices of SVRG hyperparameters to achieve the optimal convergence results for a given computational budget. We then compare these results with SGD and other SVRG estimators with different hyperparameters for a strongly convex function using a linear loss function.