Robust and accelerated randomized algorithms for scientific machine learning

Rachel Ward, Professor, Mathematics

Abstract: In the past several decades, randomized algorithms have gained popularity and established legitimacy in machine learning and scientific computing for their ability to operate with ease at problem scales involving massive amounts of high-dimensional data fit with high-dimensional models. The resulting accurate and stable randomized models can be used as numerical approximations to highly nonlinear systems in order to reduce the cost associated with high-dimensional simulations, predict timedependent outcomes from limited experimental or computed data, and gain insight into complicated datasets. Nevertheless, the scalability offered by randomization often comes at the price of new *hyperparameters* which must be tuned by hand in practice. While hyperparameter sensitivities can be tolerated in *static* machine learning problems such as image or text classification where one has an abundance of carefully curated data from homogeneous sources, we need more robust algorithms for practical implementations in real-world scientific machine learning tasks where data is scarce, the environment is dynamic, and decisions must be made in real time. The main goal of this proposal is to enhance and state-of-art randomized algorithms in machine learning such as s tochastic gradient descent and random Fourier features by carefully incorporating techniques from scientific computing such as momentum and importance sampling in order to reduce hyperparameter sensitivity while maintaining scalability.