

Large-scale Convex Optimization for Machine Learning

Francis Bach

INRIA - SIERRA project-team
Laboratoire d'Informatique de l'Ecole Normale Supérieure, Paris

Abstract: Many machine learning and signal processing problems are traditionally cast as convex optimization problems. A common difficulty in solving these problems is the size of the data, where there are many observations ("large n ") and each of these is large ("large p "). In this setting, online algorithms which pass over the data only once, are usually preferred over batch algorithms, which require multiple passes over the data. In this talk, I will present several recent results, showing that in the ideal infinite-data setting, online learning algorithms based on stochastic approximation should be preferred, but that in the practical finite-data setting, an appropriate combination of batch and online algorithms leads to unexpected behaviors, such as a linear convergence rate with an iteration cost similar to stochastic gradient descent. (joint work with Nicolas Le Roux, Eric Moulines and Mark Schmidt)
