Stochastic Newton with Arbitrary Sampling

Igor Melnikov

MIPT

Empirical Risk Minimization problem

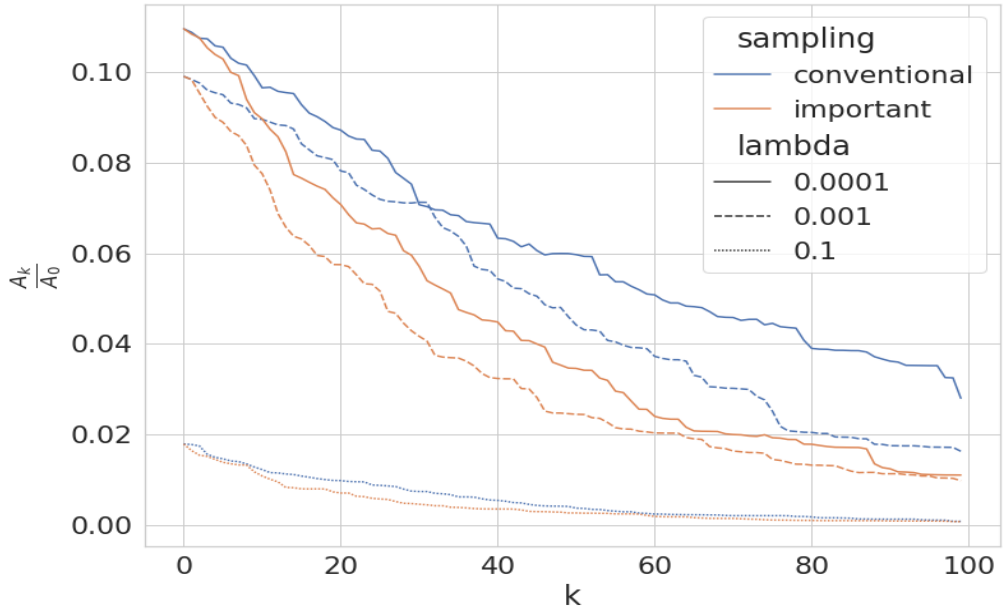

Figure 1: Convergence of NS algorithms

$f_i(x) = \log\left(1 + \exp(-b_i a_i^\top x)\right) + \frac{\lambda}{2}\|x\|$

Conventional sampling: $p(i) = \frac{1}{n}$

Important sampling: $p(i) = \frac{L_i}{\sum_{j=1}^{n} L_j}$

Convergence rate increased in 1.70 times