Sergey Samsonov

Summary: Concentration inequalities for functionals of Markov chains

During my research into the theoretical properties of different machine learning algorithms, such as MCMC, stochastic approximation (Robbins-Monroe algorithm) and reinforcement learning, I noticed that this analysis is largely limited to non-asymptotic analysis of functionals of Markov chains. Although the behaviour of additive functionals is quite well-studied, practical algorithms require studying not only linear statistics, but more complicated objects, such as quadratic forms (or, more generally, U-statistics) and products of random Markovian matrices.

In the recent papers [1] and [2], we perform theoretical analysis of variance reduction schemes for MCMC and prove high-probability bounds on the excess variance. Particular applications of this result require concentration bounds for the quadratic forms. We are able to obtain Hanson-Wright-type inequality given that the corresponding Markov kernel P is W_2-uniformly geometrically ergodic and the measure $P(x, \cdot)$ satisfy L^2-transportation-information cost inequality. We also study concentration of quadratic forms of bounded functions under the assumption that the Markov kernel P is W_1-uniformly geometrically ergodic. In this setting we prove new Rosenthal-type bound on higher order moments of the quadratic forms.

Future research include studying L_p stability for the products of random Markovian matrices. Such results are not only of independent interest, but are one of the key ingredients for analyzing the recurrent schemes, such as the linear stochastic approximation. Another research direction is an alternative approach to variance reduction for additive functionals of Markov chains via the martingale representations. Its theoretical analysis also require bounds on L_p norms of Markovian matrices. The problem becomes more complicated since tracking additional dependencies, for example, on the disctretization step in the Langevin-based algorithms, is needed.

References

[1] D. Belomestny, L. Iosipoi, E. Moulines, A. Naumov, and S. Samsonov. Variance reduction for dependent sequences with applications to stochastic gradient mcmc. 2020.

[2] D. Belomestny, L. Iosipoi, E. Moulines, A. Naumov, and S. Samsonov. Variance reduction for Markov chains with application to MCMC. *Statistics and Computing*, 30(4):973–997, 2020.