Geometric unification of central MCMC algorithms via rate distortion theory and factorizability of multivariate Markov chains

Michael Choi
Department of Statistics and Data Science, National University of Singapore
mchchoi@nus.edu.sg

Coauthor(s): Youjia Wang, Geoffrey Wolfer

Motivated by the notion of mutual information between two random variables and rate distortion theory, we propose and analyze a rate distortion optimization problem for Markov chains based upon information divergences between transition matrices. This framework offers a unified variational view on the optimality of a comprehensive suite of MCMC algorithms such as Metropolis-Hastings, Glauber dynamics, swapping algorithm and Feynman-Kac path models. Along the way, we put forward an alternative notion of distance to independence, or more generally factorizability, of a given Markov chain on a finite product state space, which is of independent interest. We prove a counterpart of the Shearer's lemma governing the distance to independence and entropy rate of Markov chains and their information projections. In addition, we develop mixing and hitting time comparison results between the original chain and its information projections, as well as a Sanov's theorem in this context. These considerations eventually lead us to investigate the geometry of multivariate Markov chains induced by information divergences.