

Langevin Monte Carlo as convex optimization in the space of measures

Szymon Majewski

The central problem considered in this thesis is approximate computation of integrals with respect to log-concave measures and approximate sampling from such measures using Monte Carlo algorithms based on discretizations of the Langevin equation. The main results of this thesis rest on observation, that sampling from a log-concave target distribution using Unadjusted Langevin Algorithm can be meaningfully interpreted as convex optimization in the space of probability measures endowed with the Wasserstein metric. Thanks to this observation, a series of results for ULA can be obtained using a proof technique analogous to proof techniques used in classical analyses of first-order convex optimization algorithms. In particular, in this thesis there are presented nonasymptotic bounds on the convergence of the Markov Chain to the target measure, as well as nonasymptotic bounds on the bias of Monte Carlo estimators of integrals with respect to this measure. An important advantage of the approach presented in this thesis is that it can be effortlessly extended to modifications of ULA that use stochastic gradients or use subgradient and proximal operators to sample from nonsmooth distributions. In this dissertation, two such modifications of ULA, which use stochastic first-order updates and can be used when the target measure is non-smooth, are presented and analyzed. In addition, using the approach presented in this thesis a number of bounds on the discrepancy between the target Gibbs measure and stationary measure of Unadjusted Langevin Algorithm with small constant stepsize are obtained.

Algorytmy Monte Carlo oparte o równanie Langevina jako optymalizacja wypukła w przestrzeni miar

Szymon Majewski

Niniejsza rozprawa jest poświęcona problemowi przybliżonego obliczania całek względem rozkładów log-wklęsłych, oraz przybliżonego generowania zmiennych losowych z rozkładów log-wklęsłych przy użyciu algorytmów Monte Carlo opartych o dyskretyzację równania Langevina. Główne wyniki przedstawione w rozprawie są oparte o obserwację, że gdy docelowa miara jest log-wklęsła, Unadjusted Langevin Algorithm (ULA) można zinterpretować jako algorytm optymalizacji wypukłej w przestrzeni miar probabilistycznych z metryką Wassersteina. Korzystając z tej obserwacji, można uzyskać szereg rezultatów dla ULA przy użyciu rozumowań analogicznych do rozumowań z klasycznych analiz algorytmów pierwszego rzędu w optymalizacji wypukłej. W szczególności, w rozprawie zaprezentowane zostały nieasymptotyczne ograniczenia na zbieżność generowanego łańcucha Markova do miary docelowej oraz nieasymptotyczne ograniczenia na obciążenie estymatorów Monte Carlo całek względem tejże miary. Ważną zaletą prezentowanego w rozprawie podejścia jest to, że można je z łatwością zastosować również do wyprowadzenia analogicznych wyników dla modyfikacji ULA wykorzystujących stochastyczne gradienty, stochastyczne subgradienty i operatory proksymalne. W rozprawie zostały przedstawione i przeanalizowane dwie takie modyfikacje ULA, które mogą zostać wykorzystane gdy gęstość docelowej miary nie jest różniczkowalna lub gdy krok gradientowy ULA chcemy zastąpić stochastycznym krokiem gradientowym. Ponadto, w pracy zaprezentowane są wyniki dotyczące ograniczeń na entropię względną oraz odległość Wassersteina pomiędzy miarą docelową a miarą stacjonarną łańcucha Markova generowanego przez ULA ze stałą długością kroku.