Échantillonnage de lois en grande dimension: diffusion de Langevin et régularisation de Moreau

Eric Moulines, École Polytechnique

Alain Durmus, Télécom ParisTech

Nicolas Brosse, École Polytechnique

L'échantillonnage de lois en grandes dimensions est devenu un enjeu fondamental dans de nombreuses applications. Cette question se pose naturellement dans les approches bayésiennes pour les problèmes inverses, l'inférence statistique bayésienne en grande dimension, l'aggrégation d'experts ou encore la sumulation des estimateurs en bayésien non paramétrique.
Nous nous intéressons dans cet exposé au cas où la loi à simuler est log-concave. Cette hypothèse peut paraître restrictive, mais elle est vérifiée dans de nombreux exemples en statistique bayésienne et en apprentissage. C'est en particulier le cas, dans le cadre bayésien, lorsque la log-vraisemblance et la loi a priori des paramètres sont log-concaves.
Lorsque la distribution est log-concave et différentiable, nous avons démontré l'intérêt d'utiliser des méthodes de simulation "sans rejet", basées sur la discrétisation de la diffusion de Langevin. L'idée de simuler une loi en discrétisant la diffusion de Langevin n'est pas nouvelle [1], [2] [3]. Nous avons obtenu des bornes de convergence explicites à horizon fini. Nous avons en particulier établi dans de telles situations que la vitesse de convergence vers l'équilibre (en distance de Wasserstein ou de variation totale) dépendait polynomialement de la dimension; [4].
Dans de nombreux cas toutefois, les lois a priori utilisées en statistique en grande dimension sont non-différentiables (c'est le cas par exemple de la loi a priori utilisée dans la version bayésienne du LASSO, de l'elastic net ou des normes nucléaires). Dans ce travail, nous introduisons une nouvelle méthode permettant d'utiliser les méthodes basées sur la discrétisation de Langevin pour de telles distributions. Notre approche s'appuie sur la régularisation de Moreau de la composante non-différentiable de la loi. Cette approche permet d'exploiter les résultats théoriques et numériques récents obtenus récemment (opérateurs proximaux). Nous présenterons dans cet exposé également des bornes de convergence et des résultats théoriques associés à cette méthode. Une attention toute particulière sera portée sur la dépendance de ces bornes en la dimension sous différentes hypothèses (régularité, concavité forte, etc...)
Finalement nous montrerons l'intérêt de cette nouvelle méthodologie pour la résolution un problème de déconvolution et de reconstruction tomographique en traitement d'images.

Références

[1] Grenander, U. . Tutorial in pattern theory. Division of Applied Mathematics, Brown University, Providence, 1983.

[2] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion. *Bernoulli*, 8(3):367–405, 2002.

[3] Roberts, G. O. and Tweedie, R. L. Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli*, 2(4):341–363, 1996

[4] Durmus, A. and Moulines, E. Non-asymptotic convergence analysis for the unadjusted langevin algorithm. Accepté pour publications, *Annals of Applied Probability*, 2017

Eric Moulines, Centre de Mathématiques Appliquées, Ecole Polytechnique, route de Saclay, Palaiseau
`eric.moulines@polytechnique.edu`