Bayesian ML: Project topics

Rémi Bardenet

December 14, 2020

1 Nature of the project

Groups can consist of either of 1 or 2 persons. Each group must pick a paper in Section 4, on a first-come-first-served basis. For the paper your group has chosen, you should:

1. explain the main theoretical, computational and/or empirical points of the paper,
2. apply it to real data of your choice when applicable,
3. be creative and add something insightful that is not in the original paper: this can be a theoretical point, an illustrative experiment, etc. **State clearly in your introduction what your creative addition is**.

The whole point is to show that you can read a scientific paper with a critical mind.

2 Assignment of papers

As a first step, **before 16 December**, please fill the spreadsheet at

https://lite.framacalc.org/9kke-paper-assignment-m2-lille

with the title of your paper and the group members.

3 Format of the deliverable

You should use Python for the programming part. Please send, by **11 January**,

- one report as a pdf (≤ 6 pages, with at most 3 pages dedicated to item 1 in Section 1) in the NeurIPS template,
- the link to a GitHub or GitLab repository containing your code and a detailed readme file with instructions to (compile/install and) run the code.

to remi.bardenet@gmail.com. There will be no deadline extension.

4 Proposed papers

All PDFs are here: https://nextcloud.univ-lille.fr/index.php/s/c8pKGidzJQzqwyi.

You can work alone or in a group of two. You can pick any paper from either the A and B list below. I will add 3 bonus points (out of 20) for people working alone, and the same amount of bonus points for choosing a paper from the B list.

To be clear, it is now OK to be a group of *two* and work on a paper in the A list. While you will not have the bonus points, this should allow you to reduce the time invested in the project and get an excellent grade if your work deserves it.

Shorter papers

[A1] C. M. Bishop. Bayesian PCA. In *Advances in neural information processing systems*, pages 382–388, 1999.

[A2] A. Corduneanu and C. M. Bishop. Variational Bayesian model selection for mixture distributions. In *Artificial intelligence and Statistics*, volume 2001, pages 27–34, 2001.

[A3] D. P. Kingma and M. Welling. Auto-encoding variational Bayes. In *International Conference on Learning Representations*, 2014.

[A4] M. Rabinovich, E. Angelino, and M. I. Jordan. Variational consensus Monte Carlo. In *Advances in Neural Information Processing Systems*, pages 1207–1215, 2015.

[A5] C. E. Rasmussen and Z. Ghahramani. Bayesian Monte Carlo. In *Advances in neural information processing systems*, pages 505–512, 2003.

[A6] A. Terenin, D. Simpson, and D. Draper. Asynchronous Gibbs sampling. In *International Conference on Artificial Intelligence and Statistics*, pages 144–154. PMLR, 2020.

[A7] M. Welling and Y. W. Teh. Bayesian learning via stochastic gradient Langevin dynamics. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2011.

Longer or more difficult papers

[B1] R. Bardenet, A. Doucet, and C. Holmes. Towards scaling up MCMC: an adaptive subsampling approach. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2014.

[B2] P. Germain, F. Bach, A. Lacoste, and S. Lacoste-Julien. Pac-Bayesian theory meets Bayesian inference. *Advances in Neural Information Processing Systems*, 29:1884–1892, 2016.

[B3] P. Grünwald, T. Van Ommen, et al. Inconsistency of Bayesian inference for misspecified linear models, and a proposal for repairing it. *Bayesian Analysis*, 12(4):1069–1103, 2017.

[B4] A. Kucukelbir, D. Tran, R. Ranganath, A. Gelman, and D. M. Blei. Automatic differentiation variational inference. *The Journal of Machine Learning Research*, 18(1):430–474, 2017.

[B5] T. Leonard and J. Hsu. Bayesian inference for a covariance matrix. *The Annals of Statistics*, 20(4):1669–1696, 1992.