Bayesian ML: project topics

Rémi Bardenet and Julyan Arbel

January 29, 2024

1 Nature of the project

A group is made of three students. Each group is to pick one of the papers listed below, and each paper can be chosen only once. The assignment in on a first-come first-served basis. There will a small bonus for picking a paper identified as "long/difficult", but it will be possible to get the maximum grade with a paper identified as "short/easy". You can also choose to work on a paper of your choice, subject to our explicit approval; in that case, we trust you not to pick a paper that you are or have been working on in another class.

The whole point is to read your paper with a critical mind. For the paper your group will have chosen, you should: (1) explain the contents of the paper, (2) emphasize the strong and weak points of the paper, and (3) apply it to real data of your choice when applicable. Bonus points will be considered if you are creative and add something insightful that is not in the original paper: this can be a theoretical point, an illustrative experiment, etc. If you have such a creative contribution, state it explicitly in the introduction of your report.

2 Assignment of papers

As a first step, we ask each group to fill the spreadsheet at

https://lite.framacalc.org/euwu5cehat-a5rr

with the title of the paper, a link to it (if available), and the composition of the group. Please fill the form **before Wednesday 7 February**.

3 Format of the deliverable

Please have each group send

- one report as a pdf (≤ 5 pages) 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 both remi.bardenet@gmail.com and julyan.arbel@inria.fr **no later than March 14th**. We will then meet on March 21th for presentations (10+5 minutes per group).

4 Proposed papers

Short and "easy" papers

[Easy1] Baris Alparslan, Sinan Yıldırım, and Ilker Birbil. Differentially private distributed Bayesian linear regression with MCMC. In *International Conference on Machine Learning*, pages 627–641. PMLR, 2023.

[Easy2] Fadhel Ayed, Juho Lee, and Francois Caron. Beyond the Chinese restaurant and Pitman-Yor processes: Statistical models with double power-law behavior. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 395–404, 2019.

[Easy3] Francis Bach, Simon Lacoste-Julien, and Guillaume Obozinski. On the equivalence between herding and conditional gradient algorithms. In *Proceedings of the 29th International Coference on International Conference on Machine Learning*, pages 1355–1362, 2012.

[Easy4] Christopher Bishop. Bayesian PCA. *Advances in neural information processing systems*, 11, 1998.

[Easy5] B. Calderhead. A general construction for parallelizing Metropolis-Hastings algorithms. *Proceedings of the National Academy of Sciences*, 111(49):17408–17413, 2014.

[Easy6] Carlos M Carvalho, Nicholas G Polson, and James G Scott. Handling sparsity via the horseshoe. In *Artificial Intelligence and Statistics*, pages 73–80. PMLR, 2009.

[Easy7] Adam D Cobb, Stephen J Roberts, and Yarin Gal. Loss-calibrated approximate inference in Bayesian neural networks. *arXiv preprint arXiv:1805.03901*, 2018.

[Easy8] Francesco D'Angelo and Vincent Fortuin. Repulsive deep ensembles are Bayesian. *Advances in Neural Information Processing Systems*, 2021.

[Easy9] Erik Daxberger, Agustinus Kristiadi, Alexander Immer, Runa Eschenhagen, Matthias Bauer, and Philipp Hennig. Laplace redux - effortless Bayesian deep learning. In *Advances in Neural Information Processing Systems*, 2021.

[Easy10] Gintare Karolina Dziugaite and Daniel M Roy. Computing nonvacuous generalization bounds for deep (stochastic) neural networks with many more parameters than training data. *arXiv preprint arXiv:1703.11008*, 2017.

[Easy11] Pavel Izmailov, Dmitrii Podoprikhin, Timur Garipov, Dmitry Vetrov, and Andrew Gordon Wilson. Averaging weights leads to wider optima and better generalization. *Uncertainty in Artificial Intelligence (UAI)*, 2018.

[Easy12] Pavel Izmailov, Sharad Vikram, Matthew D Hoffman, and Andrew Gordon Gordon Wilson. What are Bayesian neural network posteriors really like? In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 4629–4640. PMLR, 18–24 Jul 2021.

[Easy13] Ghassen Jerfel, Serena Wang, Clara Wong-Fannjiang, Katherine A Heller, Yian Ma, and Michael I Jordan. Variational refinement for importance sampling using the forward Kullback-Leibler divergence. In *Uncertainty in Artificial Intelligence*, pages 1819–1829. PMLR, 2021.

[Easy14] Romain Lopez, Pierre Boyeau, Nir Yosef, Michael I Jordan, and Jeffrey Regier. Decision-making with auto-encoding variational Bayes. In *Advances in Neural Information Processing Systems*, 2020.

[Easy15] Sanae Lotfi, Pavel Izmailov, Gregory Benton, Micah Goldblum, and Andrew Gordon Wilson. Bayesian model selection, the marginal likelihood, and generalization. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*. PMLR, 2022.

[Easy16] Thomas Möllenhoff and Mohammad Emtiyaz Khan. SAM as an optimal relaxation of Bayes. In *ICLR*, 2023.

[Easy17] Y. W. Teh, D. Newman, and M. Welling. A collapsed variational Bayesian inference algorithm for latent Dirichlet allocation. In *Advances in neural information processing systems*, pages 1353–1360, 2007.

[Easy18] Michael Tipping and Christopher Bishop. Bayesian image super-resolution. *Advances in neural information processing systems*, 15, 2002.

[Easy19] Michalis Titsias. Variational learning of inducing variables in sparse gaussian processes. In *Artificial intelligence and statistics*, pages 567–574. PMLR, 2009.

Long or difficult papers

[Hard1] Filippo Ascolani, Antonio Lijoi, Giovanni Rebaudo, and Giacomo Zanella. Clustering consistency with Dirichlet process mixtures. *Biometrika*, 2022.

[Hard2] Francis Bach. On the equivalence between kernel quadrature rules and random feature expansions. *The Journal of Machine Learning Research*, 18(1):714–751, 2017.

[Hard3] A. Belhadji, R. Bardenet, and P. Chainais. Kernel interpolation with continuous volume sampling. In *International Conference on Machine Learning (ICML)*, 2020.

[Hard4] Gabriel Cardoso, Yazid Janati El Idrissi, Sylvain Le Corff, and Eric Moulines. Monte carlo guided diffusion for Bayesian linear inverse problems. *arXiv preprint arXiv:2308.07983*, 2023.

[Hard5] Randal Douc, Pierre E Jacob, Anthony Lee, and Dootika Vats. Solving the Poisson equation using coupled markov chains. *arXiv preprint arXiv:2206.05691*, 2022.

[Hard6] P. Germain, F. Bach, A. Lacoste, and S. Lacoste-Julien. PAC-Bayesian theory meets Bayesian inference. In *Advances in Neural Information Processing Systems*, pages 1884–1892, 2016.

[Hard7] P. Grünwald and T. Van Ommen. Inconsistency of Bayesian inference for misspecified linear models, and a proposal for repairing it. *Bayesian Analysis*, 12(4):1069–1103, 2017.

[Hard8] Yuqi Gu and David B Dunson. Bayesian pyramids: identifiable multilayer discrete latent structure models for discrete data. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 85(2):399–426, 2023.

[Hard9] Matthew D Hoffman and Andrew Gelman. The no-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15(1):1593–1623, 2014.

[Hard10] Pierre E Jacob, John O'Leary, and Yves F Atchadé. Unbiased Markov chain Monte Carlo methods with couplings. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82(3):543–600, 2020.

[Hard11] Mohammad Emtiyaz Khan and Havard Rue. The Bayesian learning rule. *Journal of Machine Learning Research*, 24(281):1–46, 2023.

[Hard12] Jeremias Knoblauch, Jack Jewson, and Theodoros Damoulas. An optimization-centric view on Bayes' rule: Reviewing and generalizing variational inference. *The Journal of Machine Learning Research*, 23(1):5789–5897, 2022.

[Hard13] A. Korba, A. Salim, M. Arbel, G. Luise, and A. Gretton. A non-asymptotic analysis for Stein variational gradient descent. *Advances in Neural Information Processing Systems*, 33:4672–4682, 2020.

[Hard14] Jeffrey W Miller and Matthew T Harrison. Inconsistency of Pitman-Yor process mixtures for the number of components. *The Journal of Machine Learning Research*, 15(1):3333–3370, 2014.

[Hard15] T. Papamarkou, J. Hinkle, M. T. Young, and D. Womble. Challenges in Markov chain Monte Carlo for Bayesian neural networks. *Statistical Science*, 37(3):425–442, 2022.

[Hard16] P. Rigollet and A. Tsybakov. Exponential screening and optimal rates of sparse estimation. *The Annals of Statistics*, 39(2):731–771, 2011.

[Hard17] Sanvesh Srivastava, Cheng Li, and David B Dunson. Scalable Bayes via barycenter in Wasserstein space. *The Journal of Machine Learning Research*, 19(1):312–346, 2018.

[Hard18] Achille Thin, Nikita Kotelevskii, Arnaud Doucet, Alain Durmus, Eric Moulines, and Maxim Panov. Monte carlo variational auto-encoders. In *International Conference on Machine Learning*, pages 10247–10257. PMLR, 2021.