Statistique bayésienne: Project topics

Arnak Dalalyan, Nicolas Chopin, Rémi Bardenet

November 21, 2017

1 Nature of the project

Students should form groups of three, each group undertaking one project. Groups of two students must be authorized by us. We suggest in Section 4 a few scientific papers that can each lead to a project, but you can choose another paper, subject to our approval.

For the paper your group will have chosen, you should: (1) explain the theoretical, computational and/or empirical methods, (2) emphasize the main points of the paper, and (3) apply it to real data (that you will find). 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.

2 Assignment of papers

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

https://goo.gl/forms/XM8Qvo1qOIjJnvQu1

to rank each paper of Section 4 by preference. Note that the row numbers in the Google form correspond to the numbers in Section 4. We ask that you fill in the form **before November 30th 23:59**. By that time, you will have had an outline of the last three courses, so that you can make your choices with enough information. Then, we will solve the assignment problem to find an optimal matching of papers and groups, and will make the result known to all as soon as possible. We will allow up to two groups per paper, but in that case we expect of course the deliverables to be significantly different.

3 Format of the deliverable

You can use either Python or R for the programming part. Please have each group send

- one report as a pdf (≤ 15 pages, with reasonable fonts and margins),

- one zipped folder containing your code and a detailed readme file with instructions to (compile/install and) run the code.

to all three teachers[1] **no later than January 30th.** There will be no deadline extension.

4 Proposed topics

[1] Anirban Bhattacharya, Debdeep Pati, Natesh S. Pillai, and David B. Dunson. Dirichletlaplace priors for optimal shrinkage. *Journal of the American Statistical Association*, 110(512):1479–1490, 2015.

[2] François-Xavier Briol, Chris Oates, Mark Girolami, and Michael A Osborne. Frank-wolfe bayesian quadrature: Probabilistic integration with theoretical guarantees. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 1162–1170. Curran Associates, Inc., 2015.

[3] Eric Brochu, Vlad M Cora, and Nando De Freitas. A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*, 2010.

[4] F. Caron and Y. W. Teh. Bayesian nonparametric models for ranked data. In *Advances in Neural Information Processing Systems*, pages 1520–1528, 2012.

[5] George Casella and Elí as Moreno. Objective Bayesian variable selection. *J. Amer. Statist. Assoc.*, 101(473):157–167, 2006.

[6] Hugh A. Chipman, Edward I. George, and Robert E. McCulloch. BART: Bayesian additive regression trees. *The Annals of Applied Statistics*, 4(1):266–298, mar 2010.

[7] W. Chu and Z. Ghahramani. Preference learning with gaussian processes. In *Proceedings of the 22nd international conference on Machine learning*, pages 137–144. ACM, 2005.

[8] A. S. Dalalyan. Further and stronger analogy between sampling and optimization: Langevin Monte Carlo and gradient descent. *ArXiv e-prints*, April 2017.

[9] A. S. Dalalyan, E. Grappin, and Q. Paris. On the Exponentially Weighted Aggregate with the Laplace Prior. *ArXiv e-prints*, November 2016.

[10] John S. Denker and Yann LeCun. Transforming neural-net output levels to probability distributions. In *Advances in Neural Information Processing Systems 3, [NIPS Conference, Denver, Colorado, USA, November 26-29, 1990]*, pages 853–859, 1990.

[1] if the above link is broken, this means: Arnak.Dalalyan@ensae.fr, nicolas.chopin@ensae.fr, and remi.bardenet@gmail.com

[11] B. Efron. *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing and Prediction (Chapter 1)*.

[12] Paul Fearnhead and Dennis Prangle. Constructing summary statistics for approximate bayesian computation: semi-automatic approximate bayesian computation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(3):419–474, 2012.

[13] Sylvia Frühwirth-Schnatter. Markov chain monte carlo estimation of classical and dynamic switching and mixture models. *Journal of the American Statistical Association*, 96(453):194–209, 2001.

[14] Y. Gal and Z. Ghahramani. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059, 2016.

[15] Clara Grazian and Christian P. Robert. *Jeffreys' Priors for Mixture Estimation*, pages 37–48. Springer International Publishing, Cham, 2015.

[16] Katherine A. Heller and Zoubin Ghahramani. A nonparametric bayesian approach to modeling overlapping clusters. In *Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics, AISTATS 2007, San Juan, Puerto Rico, March 21-24, 2007*, pages 187–194, 2007.

[17] Ferenc Huszár and David Duvenaud. Optimally-weighted herding is Bayesian quadrature. In *Proceedings of the conference on Uncertainty in Artificial Intelligence (UAI)*, 2012.

[18] Iain Murray, David MacKay, and Ryan P Adams. The gaussian process density sampler. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems 21*, pages 9–16. Curran Associates, Inc., 2009.

[19] Omiros Papaspiliopoulos and Gareth O Roberts. Retrospective markov chain monte carlo methods for dirichlet process hierarchical models. *Biometrika*, 95(1):169–186, 2008.

[20] Trevor Park and George Casella. The bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686, 2008.

[21] Nicholas G Polson, James G Scott, and Jesse Windle. The bayesian bridge. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(4):713–733, 2014.

[22] Judith Rousseau and Kerrie Mengersen. Asymptotic behaviour of the posterior distribution in overfitted mixture models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(5):689–710, 2011.

[23] D. M. Roy and Y. W. Teh. The Mondrian process. In *Advances in neural information processing systems*, pages 1377–1384, 2009.

[24] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical Bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems*, pages 2951–2959, 2012.

[25] 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.

[26] Yee W Teh, Michael I Jordan, Matthew J Beal, and David M Blei. Sharing clusters among related groups: Hierarchical Dirichlet processes. In *Advances in neural information processing systems*, pages 1385–1392, 2005.

[27] Yee Whye Teh. A hierarchical bayesian language model based on pitman-yor processes. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, pages 985–992. Association for Computational Linguistics, 2006.

[28] R. Thibaux and M. I. Jordan. Hierarchical beta processes and the Indian buffet process. In *International conference on artificial intelligence and statistics*, pages 564–571, 2007.

[29] Michalis K Titsias. Variational learning of inducing variables in sparse gaussian processes. In *International Conference on Artificial Intelligence and Statistics*, pages 567–574, 2009.

[30] Min Wang. Mixtures of g -priors for analysis of variance models with a diverging number of parameters. *Bayesian Anal.*, 12(2):511–532, 06 2017.