

Mean field inference for the Dirichlet process mixture model

Oliver Zobay, *University of Bristol*

Abstract

We present a systematic study of several recently proposed methods of mean field inference for the Dirichlet process mixture (DPM) model. These methods provide approximations to the posterior distribution and are derived using the truncated stick-breaking representation and related approaches. We investigate their use in density estimation and cluster allocation and compare to Monte-Carlo results. Further, more specific topics include the general mathematical structure of the mean field approximation, the handling of the truncation level, the effect of including a prior on the concentration parameter α of the DPM model, the relationship between the proposed variants of the mean field approximation, and the connection to maximum a-posteriori estimation of the DPM model.

AMS 2000 subject classifications: Primary 62E17; secondary 62G07.

Keywords: Bayesian nonparametrics, approximation methods, variational inference, density estimation.



Full Text: [PDF](#)

Zobay, Oliver, Mean field inference for the Dirichlet process mixture model, *Electronic Journal of Statistics*, 3, (2009), 507-545 (electronic). DOI: 10.1214/08-EJS339.

References

- Antoniak, C.E. (1974). Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *Ann. Statist.* 2, 1152–1174. [MR0365969](#)
- Bishop, C.M. (2006). *Pattern recognition and machine learning*. Springer, New York. [MR2247587](#)
- Blei, D.M. and Jordan, M.I. (2006). Variational inference for Dirichlet process mixtures. *Bayesian Anal.* 1, 121–143. [MR2227367](#)
- Blei, D.M., Ng, A.Y., and Jordan, M.I. (2003). Latent Dirichlet allocation. *J. Mach. Learning Res.* 3, 993–1022.
- Escobar, M.D. (1988). Estimating the means of several normal populations by nonparametric estimation of the distribution of the means. Unpublished Ph.D. dissertation, Yale University, Department of Statistics.
- Escobar, M.D. (1994). Estimating normal means with a Dirichlet process prior. *J. Amer. Statist. Assoc.* 89, 268–277. [MR1266299](#)
- Escobar, M.D. and West, M. (1995). Bayesian density estimation and inference using mixtures. *J. Amer. Statist. Assoc.* 90, 577–588. [MR1340510](#)
- Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* 1, 209–230. [MR0350949](#)
- Ishwaran, H. and James, L.F. (2001). Gibbs sampling methods for stick-breaking priors.

Ishwaran, H. and Zarepour, M. (2002). Exact and approximate sum representations for the Dirichlet process. *Canad. J. Statist.* 30, 269–283. [MR1926065](#)

Jaakkola, T.S. and Jordan, M.I. (1998). Improving the mean field approximation via the use of mixture distributions. In *Learning in Graphical Models*, ed. M.I. Jordan, MIT Press, Cambridge, MA, 163–174.

Kurihara, K., Welling, M., and Teh, Y.W. (2007). Collapsed variational Dirichlet process mixture models. In *Proceedings of IJCAI-07*, 2796–2801.

Kurihara, K., Welling, M., and Vlassis, N. (2007). Accelerated variational Dirichlet process mixtures. In *Advances in Neural Information Processing Systems*, Vol. 19, eds. B. Schölkopf, J.C. Platt and T. Hofmann, MIT Press, Cambridge, MA, 761–768. [MR2441316](#)

Lo, A.Y. (1984). On a class of Bayesian nonparametric estimates. I. Density estimates. *Ann. Statist.* 12, 351–357. [MR0733519](#)

MacEachern, S.N. (1994). Estimating normal means with a conjugate style Dirichlet process prior. *Comm. Statist. Simulation Comput.* 23, 727–741. [MR1293996](#)

MacEachern, S.N. and Müller, P. (1998). Estimating mixture of Dirichlet process models. *J. Comput. Graph. Statist.* 7, 223–238.

MacKay, D.J.C. (2003). *Information theory, inference and learning algorithms*. Cambridge University Press, New York. [MR2012999](#)

Mézard, M., Parisi, G., and Virasoro, M.A. (1987). *Spin glass theory and beyond*. Lecture Notes in Physics, Vol. 9. World Scientific Publishing, Teaneck, NJ. [MR1026102](#)

Mukherjee, I. and Blei, D.M. (2009). Relative performance guarantees for approximate inference in latent Dirichlet allocation. In *Advances in Neural Information Processing Systems 21*, ed. D. Koller, Y. Bengio, D. Schuurmans, L. Boutou, and A. Culotta, 1129–1136.

Neal, R.M. (2000). Markov chain sampling methods for Dirichlet process mixture models. *J. Comput. Graph. Statist.* 9, 249–265. [MR1823804](#)

Opper, M. and Saad, D. (eds.) (2001). *Advanced mean field methods: theory and practice*. Neural Information Processing Series. MIT Press, Cambridge, MA. [MR1863214](#)

Roeder, K. (1990). Density estimation with confidence sets exemplified by superclusters and voids in the galaxies. *J. Amer. Statist. Assoc.* 85, 617–624.

Sethuraman, J. (1994). A constructive definition of Dirichlet priors. *Statist. Sinica* 4, 2, 639–650. [MR1309433](#)

Teh, Y.W., Kurihara, K., and Welling, M. (2008). Collapsed variational inference for HDP. In *Advances in Neural Information Processing Systems 20*, ed. J.C. Platt, D. Koller, Y. Singer, and S. Roweis, Cambridge, MA: MIT Press, 1481–1488.

Wainwright, M.J. and Jordan, M.I. (2008). Graphical models, exponential families, and variational inference. *Found. Trends Mach. Learning* 1, 1–305.

Walker, S.G. (2007). Sampling the Dirichlet mixture model with slices. *Comm. Statist. Simulation Comput.* 36, 45–54. [MR2370888](#)

Wang, B. and Titterton, D.M. (2006). Convergence properties of a general algorithm for calculating variational Bayesian estimates for a normal mixture model. *Bayesian Anal.* 1, 625–649. [MR2221291](#)

Weiss, Y. (2001). Comparing the mean field method and belief propagation for approximate inference in MRFs. In *Advanced mean field methods: theory and practice*,

[Home](#) | [Current](#) | [Past volumes](#) | [About](#) | [Login](#) | [Notify](#) | [Contact](#) | [Search](#)

Electronic Journal of Statistics. ISSN: 1935-7524