Nonparametric Bayes

Review. Review of parametric Bayes, posterior calculation for simple parametric models, Bayesian testing, Bayes factors, Bayesian GLM, data augmentation, Bayesian variable selection.

Motivation of Nonparametric Bayes. Necessity to go beyond parametric models by providing motivating examples from epidemiology [Dun10] and machine learning, what is meant by "nonparametrics" in Bayesian paradigm, how to define distribution for an unknown distribution [GR03].

Bridge between parametrics and nonparametrics. Density estimation via histograms, Dirichlet priors, properties and computation, logistic-normal priors for histogram smoothing, applications to regression with unknown residual densities.

Construction of nonparametric priors. Ferguson's denition of the Dirichlet process (DP) prior [Fer73, Fer74] Blackwell and MacQueen's method [BM73] to introduce a Dirichlet prior, properties of the Dirichlet prior, calculation of the posterior distribution (conjugacy), and support [Fer73, Fer74, BM73, GR03], constructive definition of the Dirichlet prior [Set91], product partition models, Chinese restaurant process [Ald85, Pit95] and Indian buffet process [GG05].

Finite mixture models: Finite mixtures of Gaussians, estimation via the EM algorithm and Gibbs, prior choice, posterior computation, label switching [Ste00], model-based clustering [FR02], density estimation, classification, applications in semiparametric hierarchical modeling.

Polya Trees and DP mixtures. Nonparametric priors which select absolutely continuous distributions, Polya Trees [GR03, MSW92, Lav92, Lav94], DP mixture of Gaussians [EW95, MEW96], hierarchical Dirichlet processes [TJBB04], applications to density estimation and clustering for grouped and ungrouped data.

Computation for DP mixture of Gaussians. Collapsed Gibbs sampler [EW95, Mac98, MM98, Nea00, JN04], blocked Gibbs sampler [IJ01], slice sampler [Wal07], retrospective sampler [PR08], combining slice and retrospective sampler [Pap08].

Incorporating constraints. Priors for densities with quantile constraints [GK03], shape constraints [CCH+07], and stochastic ordering constraints [GK01, DP08b], applications to flexible residual density modeling and dose response modeling.

Gaussian processes (GP). Construction, properties, applications in regression and classification, adaptation of hyperparameters, kernel and basis function methods, approximation methods for large datasets [Ras06].

Density regression: predictor-dependent mixture models, dependent DP [Mac00], kernel mixtures of DPs, kernel stick-breaking process [DP08a], probit stick-breaking process [RD11, CD09], properties, applications.

Functional data analysis: priors for dependent collections of random functions/curves [RDG09, Dun10], nested Dirichlet process [RDG08], functional clustering [Bro08], classification from functional predictors [KTV06], posterior computation, inferences.

High-dimensional applications: mixtures of factor models [MPB03], non-linear embeddings, fast computation via variational methods [GB$^+$99], applications to gene expression studies, dictionary learning and application to image denoising [ZCR$^+$09].

Introduction to asymptotic theory: weak and strong posterior consistency, Schwartz theorem, applications to showing consistency in density estimation [GR03].

REFERENCES

[Ald85]   David Aldous. Exchangeability and related topics. *École d'Été de Probabilités de Saint-Flour XIII1983*, pages 1–198, 1985.

[BM73]   David Blackwell and James B MacQueen. Ferguson distributions via pólya urn schemes. *The annals of statistics*, pages 353–355, 1973.

[Bro08]   Duncan P Brown. Efficient functional clustering of protein sequences using the dirichlet process. *Bioinformatics*, 24(16):1765–1771, 2008.

[CCH$^+$07]   I-Shou Chang, Li-Chu Chien, Chao A Hsiung, Chi-Chung Wen, Yuh-Jenn Wu, et al. Shape restricted regression with random bernstein polynomials. In *Complex Datasets and Inverse Problems*, pages 187–202. Institute of Mathematical Statistics, 2007.

[CD09]   Yeonseung Chung and David B Dunson. Nonparametric bayes conditional distribution modeling with variable selection. *Journal of the American Statistical Association*, 104(488), 2009.

[DP08a]   David B Dunson and Ju-Hyun Park. Kernel stick-breaking processes. *Biometrika*, 95(2):307–323, 2008.

[DP08b]   David B Dunson and Shyamal D Peddada. Bayesian nonparametric inference on stochastic ordering. *Biometrika*, 95(4):859–874, 2008.

[Dun10]   David B Dunson. Nonparametric bayes applications to biostatistics. *Bayesian nonparametrics*, 28:223–273, 2010.

[EW95]   Michael D Escobar and Mike West. Bayesian density estimation and inference using mixtures. *Journal of the american statistical association*, 90(430):577–588, 1995.

[Fer73]   Thomas S Ferguson. A bayesian analysis of some nonparametric problems. *The annals of statistics*, pages 209–230, 1973.

[Fer74]   Thomas S Ferguson. Prior distributions on spaces of probability measures. *The annals of statistics*, pages 615–629, 1974.

[FR02]   Chris Fraley and Adrian E Raftery. Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, 97(458):611–631, 2002.

[GB$^+$99]   Zoubin Ghahramani, Matthew J Beal, et al. Variational inference for bayesian mixtures of factor analysers. In *NIPS*, pages 449–455, 1999.

[GG05]   Thomas L Griffiths and Zoubin Ghahramani. Infinite latent feature models and the indian buffet process. In *NIPS*, volume 18, pages 475–482, 2005.

[GK01]   Alan E Gelfand and Athanasios Kottas. Nonparametric bayesian modeling for stochastic order. *Annals of the Institute of Statistical Mathematics*, 53(4):865–876, 2001.

[GK03]   Alan E Gelfand and Athanasios Kottas. Bayesian semiparametric regression for median residual life. *Scandinavian Journal of Statistics*, 30(4):651–665, 2003.

[GR03]   Jayanta K Ghosh and RV Ramamoorthi. *Bayesian nonparametrics*, volume 1. Springer, 2003.

[IJ01]   Hemant Ishwaran and Lancelot F James. Gibbs sampling methods for stick-breaking priors. *Journal of the American Statistical Association*, 96(453), 2001.

[JN04]   Sonia Jain and Radford M Neal. A split-merge markov chain monte carlo procedure for the dirichlet process mixture model. *Journal of Computational and Graphical Statistics*, 13(1), 2004.

[KTV06]   Sinae Kim, Mahlet G Tadesse, and Marina Vannucci. Variable selection in clustering via dirichlet process mixture models. *Biometrika*, 93(4):877–893, 2006.

[Lav92]   Michael Lavine. Some aspects of polya tree distributions for statistical modelling. *The Annals of Statistics*, pages 1222–1235, 1992.

[Lav94]   Michael Lavine. More aspects of polya tree distributions for statistical modelling. *The Annals of Statistics*, pages 1161–1176, 1994.

[Mac98]   Steven N MacEachern. Computational methods for mixture of dirichlet process models. In *Practical nonparametric and semiparametric Bayesian statistics*, pages 23–43. Springer, 1998.

[Mac00]   Steven N MacEachern. Dependent dirichlet processes. *Unpublished manuscript, Department of Statistics, The Ohio State University*, 2000.

[MEW96]   Peter Müller, Alaattin Erkanli, and Mike West. Bayesian curve fitting using multivariate normal mixtures. *Biometrika*, 83(1):67–79, 1996.

[MM98]   Steven N MacEachern and Peter Müller. Estimating mixture of dirichlet process models. *Journal of Computational and Graphical Statistics*, 7(2):223–238, 1998.

[MPB03]   Geoffrey J McLachlan, David Peel, and RW Bean. Modelling high-dimensional data by mixtures of factor analyzers. *Computational Statistics & Data Analysis*, 41(3):379–388, 2003.

[MSW92]   R Daniel Mauldin, William D Sudderth, and SC Williams. Polya trees and random distributions. *The Annals of Statistics*, pages 1203–1221, 1992.

[Nea00]   Radford M Neal. Markov chain sampling methods for dirichlet process mixture models. *Journal of computational and graphical statistics*, 9(2):249–265, 2000.

[Pap08]   Omiros Papaspiliopoulos. A note on posterior sampling from dirichlet mixture models. 2008.

[Pit95]   Jim Pitman. Exchangeable and partially exchangeable random partitions. *Probability theory and related fields*, 102(2):145–158, 1995.

[PR08]   Omiros Papaspiliopoulos and Gareth O Roberts. Retrospective markov chain monte carlo methods for dirichlet process hierarchical models. *Biometrika*, 95(1):169–186, 2008.

[Ras06]   Carl Edward Rasmussen. Gaussian processes for machine learning. 2006.

[RD11]   Abel Rodriguez and David B Dunson. Nonparametric bayesian models through probit stick-breaking processes. *Bayesian analysis (Online)*, 6(1), 2011.

[RDG08]   Abel Rodriguez, David B Dunson, and Alan E Gelfand. The nested dirichlet process. *Journal of the American Statistical Association*, 103(483), 2008.

[RDG09]   Abel Rodríguez, David B Dunson, and Alan E Gelfand. Bayesian nonparametric functional data analysis through density estimation. *Biometrika*, 96(1):149–162, 2009.

[Set91]   Jayaram Sethuraman. A constructive definition of dirichlet priors. Technical report, DTIC Document, 1991.

[Ste00]   Matthew Stephens. Dealing with label switching in mixture models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(4):795–809, 2000.

[TJBB04]   Yee Whye Teh, Michael I Jordan, Matthew J Beal, and David M Blei. Sharing clusters among related groups: Hierarchical dirichlet processes. In *NIPS*, 2004.

[Wal07]   Stephen G Walker. Sampling the dirichlet mixture model with slices. *Communications in Statistics–Simulation and Computation®*, 36(1):45–54, 2007.

[ZCR⁺09]   Mingyuan Zhou, Haojun Chen, Lu Ren, Guillermo Sapiro, Lawrence Carin, and John W Paisley. Non-parametric bayesian dictionary learning for sparse image representations. In *Advances in neural information processing systems*, pages 2295–2303, 2009.