Kullback Leibler property of kernel mixture priors in Bayesian density estimation

Yuefeng Wu
Subhashis Ghosal

Abstract

Positivity of the prior probability of Kullback-Leibler neighborhood around the true density, commonly known as the Kullback-Leibler property, plays a fundamental role in posterior consistency. A popular prior for Bayesian estimation is given by a Dirichlet mixture, where the kernels are chosen depending on the sample space and the class of densities to be estimated. The Kullback-Leibler property of the Dirichlet mixture prior has been shown for some special kernels like the normal density or Bernstein polynomial, under appropriate conditions. In this paper, we obtain easily verifiable sufficient conditions, under which a prior obtained by mixing a general kernel possesses the Kullback-Leibler property. We study a wide variety of kernel used in practice, including the normal, t, histogram, Weibull, gamma densities and so on, and show that the Kullback-Leibler property holds if some easily verifiable conditions are satisfied at the true density. This gives a catalog of conditions required for the Kullback-Leibler property, which can be readily used in applications.

AMS 2000 subject classifications: Primary 62G07, 62G20.

Keywords: Bayesian density estimation, Dirichlet process, kernel mixture, Kullback-Leibler property, posterior consistency.

Full Text: PDF

References

[1]   Antoniak, C. (1974). Mixtures of Dirichlet processes with application to Bayesian non–parametric problems. Ann. Statist. 2 1152–1174. MR0365969

[2]   Arfken, G. (1985) Digamma and Polygamma Functions in Mathematical Methods for Physicists. 3rd ed. Academic Press, Orlando.

[3]   Bouezmarni, T. and Scaillet, O. (2003). Consistency of asymmetric kernel density estimators and smoothed histograms with application to income data. DP0306, Institut de Statistique. Université Catholique de Louvain.

[4]   Chen, S. (2000). Probability density function estimation using gamma kernels. Ann. Inst. Statist. Math. 52 No. 3, 471–480. MR1794247

[5]   Diaconis, P. and Ylvisaker, D. (1985). Quantifying prior opinion. In Bayesian Statistics 2 (J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, Eds.) 133–156. North-Holland, Amsterdam. MR0862488

[6]   Escobar, M. and West, M. (1995). Bayesian density estimation and inference using mixtures. J. Amer. Statist. Assoc. 90 577–588. MR1340510

[7]   Escobar, M. and West, M. (1998). Computing nonparametric hierarchical models. In Practical Nonparametric and Semiparametric Bayesian Statistics. Lecture Notes in Statistics, 133 1–22. Springer, New York. MR1630073

[8]   Feller, W. (1957). An Introduction to Probability Theory and Its Application, Vol I. & II. John Wiley & Sons, Inc. MR0088081

[9]   Ferguson, T. S. (1983). Bayesian density estimation by mixtures of normal distributions. In Recent Advances in Statistics (M. Rizvi, J. Rustagi, and D. Siegmund, Eds.) 287–302. Academic Press, New York. MR0736538

[10]   Ghosal, S. (2001). Convergence rates for density estimation with Bernstein polynomials. Ann. Statist. 29 1264–1280. MR1873330

[11]   Ghosal, S., Ghosh, J. K. and Ramamoorthi, R. V. (1999). Posterior consistency of Dirichlet mixtures in density estimation. Ann. Statist. 27 143–158. MR1701105

[12]   Ghosal, S., Ghosh, J. K. and Ramamoorthi, R. V. (1999). Consistent semiparametric Bayesian inference about a location parameter. J. Statist. Plann. Inference 77 181–193. MR1687955

[13]   Ghosal, S. and van der Vaart, A. W. (2001). Entropies and rates of convergence for maximum likelihood and Bayes estimation for mixtures of normal densities. Ann. Statist. 29 1233–1263. MR1873329

[14]   Ghosal, S. and van der Vaart, A. W. (2007). Posterior convergence rates of Dirichlet mixtures at smooth densities. Ann. Statist. 35 697–723. MR2336864

[15]   Ghosal, S. and van der Vaart, A. W. (2009) Theory of Nonparametric Bayesian inference. Cambridge University Press (to appear). MR2003184

[16]   Ghosh, J. K. and Ramamoortrhi, R. V. (2003). Bayesian Nonparametrics. Springer-Verlag, New York. MR1992245

[17]   Ghosh, S. K. and Ghosal, S. (2006). Semiparametric accelerated failure time models for censored data. In Bayesian Statistics and its Applications (S. K. Upadhyay et al., eds.) 213–229 Anamaya Publishers, New Delhi.

[18]   Hason, T. (2006). Modeling censored lifetime data using a mixture of gammas baseline. Bayesian Analysis, 1 575–594. MR2221289

[19]   Kruijer, W. and van der Vaart. (2005). Posterior convergence rates for Dirichlet mixtures of beta densities. Preprint.

[20]   Lavine, M. (1992). Some aspects of Polya tree distributions for statistical modeling. Ann. Statist. 20 1222–1235. MR1186248

[21]   Lo, A. Y. (1984). On a class of Bayessian nonparametric estimates I: density estimates. Ann. Statist. 1 38–53. MR0733519

[22]   Lorentz, G. (1953). Bernstein Polynomials. University of Toronto Press, Toronto. MR0057370

[23]   Petrone, S. (1999). Random Bernstein polynomials. Scand. J. Statist. 26 373–393. MR1712051

[24]   Petrone, S. (1999). Bayesian density estimation using Bernstein polynomials. Canad. J. Statist. 27 105–126. MR1703623

[25]   Petrone, S. and Veronese, P. (2007). Feller operators and mixture priors in Bayesian nonparametrics. Preprint.

[26]   Petrone, S. and Wasserman, L. (2002). Consistency of Bernstein polynomial posteriors. J. Roy. Statist. Soc., Ser. B 64 79–100. MR1881846

[27]   Royden, H. L. (1988). Real Analysis. Macmillan, New York; Collier MacMillan, London. MR1013117

[28]   Schwartz, L. (1965). On Bayes procedures. Z. Wahrsch. Verw. Gebiete 4 10–26. MR0184378

[29]   Tokdar, S. (2006). Posterior consistency of Dirichlet location-scale mixture of normals in density estimation and regression. Sankhya: The Indian Journal of Statistics. 67 90–110. MR2301566

[30]   West, M. (1992). Modeling with mixtures. In Bayesian Statistcs 4 (J. M. Bernardo, J. O. Berger, A. P. David, and A. F. M. Smith, Eds.) 503–524. Oxford Univ. Press. MR1380294

[31]   West, M., Müller, P. and Escobar, M. (1994). Hierarchical priors and mixture models, with applications in regressions and density estimation. In Aspects of Uncertainty: A Tribute to D. V. Lindley (P. R. Freeman and A. F. M. Smith, Eds.) 363–386. Wiley, Chichester. MR1309702