Bai, Ray; Ghosh, Malay
On the beta prime prior for scale parameters in high-dimensional Bayesian regression models. (English) Zbl 1470.62103
Stat. Sin. 31, No. 2, 843-865 (2021).

Summary: We study a high-dimensional Bayesian linear regression model in which the scale parameter follows a general beta prime distribution. Under the assumption of sparsity, we show that an appropriate selection of the hyperparameters in the beta prime prior leads to the (near) minimax posterior contraction rate when $p \gg n$. For finite samples, we propose a data-adaptive method for estimating the hyperparameters based on the marginal maximum likelihood (MML). This enables our prior to adapt to both sparse and dense settings and, under our proposed empirical Bayes procedure, the MML estimates are never at risk of collapsing to zero. We derive an efficient Monte Carlo expectation-maximization (EM) and variational EM algorithm for our model, which are available in the R package NormalBetaPrime. Simulations and an analysis of a gene expression data set illustrate our model’s self-adaptivity to varying levels of sparsity and signal strengths.

MSC:
62J05 Linear regression; mixed models
62F15 Bayesian inference
62P10 Applications of statistics to biology and medical sciences; meta analysis

Keywords:
beta prime density; empirical Bayes; high-dimensional data; posterior contraction; scale mixtures of normal distributions

Software:
glmnet; NormalBetaPrime; EBayesThresh

Full Text: DOI arXiv

References:
[1] Armagan, A., Clyde, M. and Dunson, D. B. (2011). Generalized beta mixtures of Gaussians. Advances in Neural Information Processing Systems 24, 523-531.
[2] Armagan, A., Clyde, M. and Dunson, D. B. (2013). Generalized double pareto shrinkage. Statistica Sinica 23, 119-143. - Zbl 1259.62061
[3] Bai, R. and Ghosh, M. (2019). Large-scale multiple hypothesis testing with the normal-beta prime prior. Statistics 53, 1210-1233. - Zbl 1435.62267
[4] Bhattacharya, A., Chakraborty, A. and Mallick, B. K. (2016). Fast sampling with Gaussian scale mixture priors in high-dimensional regression. Biometrika 103, 985-991.
[5] Bhattacharya, A., Pati, D., Pillai, N. S. and Dunson, D. B. (2015). Dirichlet-Laplace priors for optimal shrinkage. Journal of the American Statistical Association 110, 1479-1490. - Zbl 1373.62368
[6] Carvalho, C. M., Polson, N. G. and Scott, J. G. (2009). Handling sparsity via the horseshoe. In Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics, PMLR 5, 73-80.
[7] Carvalho, C. M., Polson, N. G. and Scott, J. G. (2010). The horseshoe estimator for sparse signals. Biometrika 97, 465-480. - Zbl 1341.62212
[8] Casella, G. (2001). Empirical Bayes gibbs sampling. Biostatistics 2, 485-500. - Zbl 1097.62505
[9] Castillo, I., Schmidt-Hieber, J. and van der Vaart, A. (2015). Bayesian linear regression with sparse priors. The Annals of Statistics 43, 1986-2018.
[10] Datta, J. and Ghosh, J. K. (2013). Asymptotic properties of Bayes risk for the horseshoe prior. Bayesian Analysis 8, 111-132. - Zbl 1329.62122
[11] Dobriban, E. and Fan, J. (2016). Regularity properties for sparse regression. Communications in Mathematics and Statistics 4, 1-19. - Zbl 1341.62212
[12] Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. Journal of the
Griffin, J. E. and Brown, P. J. (2010). Inference with normal-gamma prior distributions in regression problems. Bayesian
Ghosal, S., Ghosh, J. K. and van der Vaart, A. W. (2000). Convergence rates of posterior distributions. The Annals of
Griffin, J. E. and Brown, P. J. (2010). Inference with normal-gamma prior distributions in regression problems. Bayesian
Hahn, P. R. and Carvalho, C. M. (2015). Decoupling shrinkage and selection in Bayesian linear models: A posterior summary
Johnson, V. E. and Rossell, D. (2012). Bayesian model selection in high-dimensional settings. Journal of the American
Scott, J. G. and Berger, J. O. (2010). Bayes and empirical-Bayes multiplicity adjustment in the variable-selection problem.
Rousseau, J. and Szabó, B. (2017). Asymptotic behavior of the empirical Bayes posteriors associated to maximum marginal
van der Pas, S., Szabó, B. and van der Vaart, A. (2017). Adaptive posterior contraction rates for the horseshoe. Electronic
Johnson, V. E. and Russell, D. (2012). Bayesian model selection in high-dimensional settings. Journal of the American
Johnstone, I. M. and Silverman, B. W. (2004). Needles and straw in haystacks: Empirical Bayes estimates of possibly sparse
Leday, G. G. R., de Gunst, M. C. M., Kpogbezan, G. B., van der Vaart, A. W., van Wieringen, W. N. and van de Wiel, M. A. (2017). Gene network reconstruction using global-local shrinkage priors. The Annals of Applied Statistics 11, 41-68. · Zbl 1366.62227
Martin, R., Mass, R. and Walker, S. G. (2017). Empirical Bayes posterior concentration in sparse high-dimensional linear models. Bernoulli 23, 1822-1847. · Zbl 1450.62085
Narisetti, N. N. and He, X. (2014). Bayesian variable selection with shrinking and diffusing priors. The Annals of Statistics 42, 789-817. · Zbl 1302.62158
Park, T. and Casella, G. (2008). The Bayesian Lasso. Journal of the American Statistical Association 103, 681-686. · Zbl 1366.62292
Pérez, M.-E., Pericchi, L. R. and Ramírez, I. C. (2017). The scaled beta2 distribution as a robust prior for scales. Bayesian Analysis 12, 615-637. · Zbl 1384.62040
Polson, N. G. and Scott, J. G. (2010). Shrink globally, act locally: Sparse Bayesian regularization and prediction. Bayesian Statistics 9, 501-538.
Polson, N. G. and Scott, J. G. (2012). On the half-cauchy prior for a global scale parameter. Bayesian Analysis 7, 887-902. · Zbl 1330.62128
Raskutti, G., Wainwright, M. J. and Yu, B. (2011). Minimax rates of estimation for high-dimensional linear regression over q-balls. IEEE Transactions on Information Theory 57, 6976-6994. · Zbl 1343.62012
Rocková, V. and George, E. I. (2018). The spike-and-slab Lasso. Journal of the American Statistical Association 113, 431-444. · Zbl 1398.62186
Rossell, D. and Telesca, D. (2017). Nonlocal priors for high-dimensional estimation. Journal of the American Statistical Association 112, 254-265.
Rousseau, J. and Szabó, B. (2017). Asymptotic behavior of the empirical Bayes posteriors associated to maximum marginal likelihood estimator. The Annals of Statistics 45, 833-865. · Zbl 1371.62048
Scheetz, T. E., Kim, K.-Y. A., Swiderski, R. E., Philp, A. R., Braun, T. A., K ultson, K. L., Dorrance, A. M., DiBona, G. F., Huang, J., Casavant, T. L., Sheffield, V. C. and Stone, E. M. (2006). Regulation of gene expression in the mammalian eye and its relevance to eye disease. In Proceedings of the National Academy of Sciences 103, 14429-14434.
Scott, J. G. and Berger, J. O. (2010). Bayes and empirical-Bayes multiplicity adjustment in the variable-selection problem. The Annals of Statistics 38, 258-2619. · Zbl 1200.62020
Shin, M., Bhattacharya, A. and Johnson, V. E. (2018). Scalable Bayesian variable selection using nonlocal prior densities in ultrahigh-dimensional settings. Statistica Sinica 28, 1053-1078. · Zbl 1390.62125
Song, Q. and Liang, F. (2017). Nearly optimal Bayesian shrinkage for high dimensional regression. ArXiv Preprint arXiv:1712.08964.
Sparks, D. K., Khare, K. and Ghoosh, M. (2015). Necessary and sufficient conditions for high-dimensional posterior consistency under g-priors. Bayesian Analysis 10, 627-664. · Zbl 1335.62066
Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (2002). Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64, 583-639. · Zbl 1067.62010
Tiao, G. C. and Tan, W. Y. (1965). Bayesian analysis of random-effect models in the analysis of variance. I. Posterior distribution of variance-components. Biometrika 52, 37-54. · Zbl 0144.42204
Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 58, 267-288. · Zbl 0850.62358
van der Pas, S., Salomond, J.-B. and Schmidt-Hieber, J. (2016). Conditions for posterior con-traction in the sparse normal means problem. Electronic Journal of Statistics 10, 976-1000. · Zbl 1343.62012
van der Pas, S., Szabó, B. and van der Vaart, A. (2017). Adaptive posterior contraction rates for the horseshoe. Electronic Journal of Statistics 11, 3196-3225. · Zbl 1373.62140
[42] Yang, Y., Wainwright, M. J. and Jordan, M. I. (2016). On the computational complexity of high-dimensional Bayesian variable selection. The Annals of Statistics 44, 2497-2532. - Zbl 1359.62088

[43] Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. The Annals of Statistics 38, 894-942. - Zbl 1183.62120

[44] Zou, H. (2006). The adaptive Lasso and its oracle properties. Journal of the American Statistical Association 101, 1418-1429. - Zbl 1171.62326

[45] Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 67, 301-320. Ray Bai - Zbl 1069.62054

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.