Building hyper Dirichlet processes for graphical models

Daniel S. Heinz, *Carnegie Mellon University*

Abstract
Graphical models are used to describe the conditional independence relations in multivariate data. They have been used for a variety of problems, including log-linear models (Liu and Massam, 2006), network analysis (Holland and Leinhardt, 1981; Strauss and Ikeda, 1990; Wasserman and Pattison, 1996; Pattison and Wasserman, 1999; Robins et al., 1999), graphical Gaussian models (Roverato and Whittaker, 1998; Giudici and Green, 1999; Marrelec and Benali, 2006), and genetics (Dobra et al., 2004). A distribution that satisfies the conditional independence structure of a graph is Markov. A graphical model is a family of distributions that is restricted to be Markov with respect to a certain graph. In a Bayesian problem, one may specify a prior over the graphicalmodel. Such a prior is called a hyper Markov law if the random marginals also satisfy the independence constraints. Previous work in this area includes (Dempster, 1972; Dawid and Lauritzen, 1993; Giudici and Green, 1999; Letac and Massam, 2007). We explore graphicalmodels based on a non-parametric family of distributions, developed from Dirichlet processes.

AMS 2000 subject classifications: Primary 36E05; secondary 62G99.

Keywords: Hyper Markov law, stick-breakingmeasure, non-parametric prior, decomposable graphical distribution, covariance selection.

Full Text: PDF

Heinz, Daniel S., Building hyper Dirichlet processes for graphical models, Electronic Journal of Statistics, 3, (2009), 290-315 (electronic). DOI: 10.1214/08-EJS269.

References

Bush, C.A. and MacEachern, S.N. (1996). A semiparametric Bayesian model for randomized block designs. Biometrika, 83 275–285.

Carvalho, C., Massam, H. and West, M. (2007). Simulation of hyper-inverse Wishart distributions in graphical models. Biometrika, 94 647–659. http://ftp.stat.duke.edu/WorkingPapers/05-03.html. MR2410014

Dawid, A.P. and Lauritzen, S.L. (1993). Hyper Markov laws in the statistical analysis of decomposable graphical models. The Annals of Statistics, 21 1272–1317. MR1241267

Dempster, A. (1972). Covariance selection. Biometrics, 28 157–175.

Dobra, A., Hans, C., Jones, B., Nevins, J., Yao, G. and West, M. (2004). Sparse graphical models for exploring gene expression data. Journal of Multivariate Analysis, 90 196–212. MR2064941

Escobar, M.D. and West, M. (1995). Bayesian density estimation and inference using mixtures. Journal of the American Statistical Association, 90 577–588. MR1340510

Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. The Annals of Statistics, 1 209–230. MR0350949

Ghosh, J.K. and Ramamoorthi, R. (1995). Consistency of Bayesian inference for survival analysis with or without censoring. In Analysis of Censored Data (H. Koul and J. Deshpande, eds.). 95–104. MR1483342

Giudici, P. and Green, P. (1999). Decomposable graphical gaussian model determination. Biometrika, 86 785–801. MR1741977

Griffin, J.E. and Steel, M.F.J. (2006). Order-based dependent Dirichlet processes. Journal of the American Statistical Association, 101. MR2268037

Holland, P.W. and Leinhardt, S. (1981). An exponential family of probability distributions. Journal of the American Statistical Association, 76 33–50. MR0608176

Kim, Y. and Lee, J. (2001). On posterior consistency of survival models. The Annals of Statistics, 29 668–686. MR1865336

Letac, G. and Massam, H. (2007). Wishart distributions for decomposable graphs. The Annals of Statistics, 35 1278–1323. MR2341706

Liu, J. and Massam, H. (2006). The conjugate prior for discrete hierarchical log-linear models. http://www.citebase.org/abstract?id=oai:arXiv. org:math/0609100.

Marrelec, G. and Benali, H. (2006). Asymptotic Bayesian structure learning using graph supports for Gaussian graphical models. Journal of Multivariate Analysis, 97 1451–1466. MR2256161

Pattison, P. and Wasserman, S. (1999). Logit models and logistic regression for social networks: II. multivariate relations. British Journal of Mathematical and Statistical Psychology, 52 169–193.

Pievatolo, A. and Rotondi, R. (2000). Analysing the interevent time distribution to identify seismicity phases: A Bayesian nonparametric approach to the multiple changepoint problem. Applied Statistics, 49 543–562. MR1824558

Robins, G., Pattison, P. and Wasserman, S. (1999). Logit models and logistic regressions for social networks: III. valued relations. Psychometrika, 64 371–394. MR1720089

Roverato, A. and Whittaker, J. (1998). The Isserlis matrix and its application to non-decomposable graphical Gaussian models. Biometrika, 85 711–725. MR1665842

Schervish, M.J. (1995). Theory of Statistics. Springer-Verlag, New York. MR1354146

Sethuraman, J. (1994). A constructive definition of Dirichlet measures. Statistica Sinica.

Speed, T. and Kiiveri, H. (1986). Gaussian Markov distributions over finite graphs. The Annals of Statistics, 14 138–150. MR0829559

Strauss, D. and Ikeda, M. (1990). Pseudolikelihood estimation for social networks. Journal of the American Statistical Association, 85 204–212. MR1137368

Susarla, V. and Ryzin, J.V. (1976). Nonparametric Bayesian estimation of survival curves from incomplete observations. Journal of the American Statistical Association, 71 897–902. MR0436445

Wasserman, S. and Pattison, P. (1996). Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and p*. Psychometrika, 61 401–425. MR1424909