Common applications of Bayesian hierarchical models

The usefulness of hierarchical modeling goes beyond the obvious application to hierarchically grouped data. In fact, with exception of the basic statistical models, most Bayesian models are hierarchical. Many complex systems and processes can be approximated using multiple levels (or layers) of simpler models. The common assumption that parameters are conditionally independent of other parameters or data more than one level of hierarchy away greatly simplifies computation, making hierarchical models accessible. Below, we briefly discuss some common applications of hierarchical models and provide references for further investigation.

Mixed effects models

Mixed effects modeling is the frequentist approach for grouped, correlated data we studied in Lesson 11. This class of models is very popular in biology and social science applications.

See Gelman and Hill (2006); Gelman et al. (2014).

Mixture models

Mixture models provide a nice way to build nonstandard probability distributions from simper distributions, as well as to identify unlabeled clusters/populations in the data. Mixture models can be formulated hierarchically, allowing us to estimate unobserved (latent) variables in a technique called data augmentation. We briefly explore mixture models in the honors section of Lesson 11.

See Frühwirth-Schnatter (2006); Gelman et al. (2014).

Generalized linear models (GLMs)

These models generalize normal linear regression models in the sense that the likelihood belongs to a more general class of distributions. The binomial and Poisson regressions encountered in this course are examples of GLMs. Data augmentation techniques similar to those used for mixture models make GLMs amenable to hierarchical modeling.

See Agresti (2013); Gelman et al. (2014).

Time series data

The models used in this course are often inappropriate for time series data (observations collected over time) due to the correlation between them. Such autocorrelation can be accounted for by introducing a hierarchical structure with parameters that evolve in time. Examples include the popular state space and hidden Markov models.

See Prado and West (2010); Harrison and West (1999).

Spatial data

Just as observations collected across time are often correlated, observations from distinct spatial locations can exhibit dependence. For example, we might expect a measurement at location x to be more similar to measurement y five meters away than to measurement z 100 meters away. State space models and nonparametric models for response surfaces are common for spatial data.

See Banerjee et al. (2014).

Neural networks

Neural networks and deep learning have become a primary tool in machine learning. They involve layers of "neurons" that separate inputs from outputs, allowing nonlinear relationships. These intermediate nodes can be thought of as latent variables in a hierarchical probabilistic model, although Bayesian inference of neural networks is uncommon.

See Hastie et al. (2009) and this link to an online book introducing neural networks and deep learning.

Nonparametric methods

Bayesian nonparametric models move beyond inference for parameters to inference for functions and distributions. Finite-dimensional representations of the necessary priors often appear as hierarchical models. Two of the most popular nonparametric priors are the Gaussian process prior (typically used as a prior on continuous functions), and Dirichlet process prior (as a prior on probability distributions).

See Gelman et al. (2014); Hjort et al. (2010).

References

A. Agresti. *Categorical Data Analysis*. Wiley Series in Probability and Statistics. Wiley, 2013. ISBN 9780470463635. URL https://books.google.com/books?id=UOrr47-2oisC.

Sudipto Banerjee, Bradley P. Carlin, and Alan E. Gelfand. *Hierarchical modeling and analysis for spatial data*. Crc Press, 2014.

Sylvia Frühwirth-Schnatter. *Finite mixture and Markov switching models*. Springer Science & Business Media, 2006.

Andrew Gelman and Jennifer Hill. *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press, 2006.

Andrew Gelman, John B. Carlin, Hal S. Stern, and Donald B. Rubin. *Bayesian data analysis*, volume 2. Chapman & Hall/CRC Boca Raton, FL, USA, 2014.

Jeff Harrison and Mike West. *Bayesian forecasting & dynamic models*. Springer New York, 1999.

T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*. Springer Series in Statistics. Springer New York, 2009. ISBN 9780387848587. URL https://books.google.com/books?id=tVIjmNS3Ob8C.

Nils Lid Hjort, Chris Holmes, Peter Müller, and Stephen G. Walker. *Bayesian nonparametrics*, volume 28. Cambridge University Press, 2010.

Raquel Prado and Mike West. *Time series: modeling, computation, and inference*. CRC Press, 2010.