Gaussian Markov Random Fields: Theory and Applications, by Håvard RUE and Leonhard HELD, Boca Raton, FL: Chapman & Hall/CRC, 2005, ISBN 1-58488-432-0, xvi + 263 pp., $79.95.

When observations are clearly not dependent, it is necessary to use statistical models that capture the dependence. If the data are observed over time, then one of the simplest models for dependent observations is the autoregressive process

$$x_t = \phi x_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathrm{N}(0, \sigma^2), |\phi| \le 1.$$

This is a Markov process, because the dependence is modeled on the basis of relationships between neighboring observations in time. The assumption that the variation between neighboring observations is Normally distributed allows the theory of the multivariate Normal distribution to be used in inference for ϕ and in the development of diagnostics for these models.

If the observations are dependent over space, the benefits from having them ordered along a line are lost, and the assumption of Normality becomes central to any analysis. For example, observations of pollution in one county may be dependent on observations in neighboring counties, but conditionally independent of observations in counties beyond those. Each county can be represented by a vertex in a graph and edges correspond to counties that are neighbors. Given such a graph, if observations at any two vertices *not* connected by a single edge are independent conditional on all other observations and all of the observations are Normal, then the distribution of the observations is a Gaussian Markov random field (GMRF).

A GMRF is specified by its conditional means and its precision matrix $\mathbf{Q}$, which is the inverse of the covariance matrix $\mathbf{\Sigma}$. The Markov assumption ensures that all elements of $\mathbf{Q}$ that correspond to nonneighboring points are 0.

Furthermore, if i and j are neighboring locations, then the correlation of observations at those points conditional on all other observations is $Q_{ij}/\sqrt{Q_{ii}Q_{jj}}$, and the conditional mean and precision of any single observation can also be expressed as simple functions of nonzero elements of $\mathbf{Q}$. Since GMRF's can be specified through Normal conditionals defined by sparse matrices, this allows certain computations to run much more quickly than they would in the case of methods based on the dense matrix $\boldsymbol{\Sigma}$. Throughout the book, Normal distributions are specified by precisions, rather than by variances.

The authors assume that the reader has a good working knowledge of conditional probability and of the multivariate Normal distribution at the level of Hogg, McKean, and Craig (2005). Beyond this, they outline or discuss in detail the technical aspects of fitting GMRF's, with a strong emphasis on only those details that are needed to make practical use of the methods. More formal and mathematical results are referenced in the endnotes. Their style is terse and if some detail seems to be missing, it can often be found upon rereading. If not, then the references and endnotes will point to where those details can be found. Theory is usually presented first, followed by a detailed investigation of one or more complex examples.

After a brief and very useful introduction, Chapter 2 contains most of the general results about GMRF's. It begins with a presentation of notation necessary to describe conditional distributions relative to graphical dependence networks. The use of the precision matrix in model specification, inference, and simulation is then discussed, and a section is dedicated to a general review of solvers for sparse systems of linear equations. Since the speed of precision-matrix-based methods depends on the speed of these solvers, this discussion is central to the practical use of these methods, and any text on inference that requires these methods should include a section of this kind. The use of toroidal boundary conditions on the data is discussed, showing how they can lead to further speed increases through use of cyclic precision matrices. Finally, several methods for ensuring that $\mathbf{Q}$ is positive definite are compared.

The three remaining chapters apply the basic techniques to increasingly complex models. Whereas the third chapter is accessible and self-contained, the authors admit that the last two chapters cover areas where research is ongoing and suggest that they are best used by those with prior knowledge of those areas.

Chapter 3 presents GMRF's in which $\mathbf{Q}$ is not of full rank. This can result from a linear constraint on conditional means, but also arises in cases where the data are first or second differences between observations at neighboring sites. For example, GMRF methods can be used to estimate parameters for random walks, based upon their paths. After fitting a random walk on the line, the authors go on to random walks on lattices, to random walks in continuous time observed at discrete random times, and to second order random walks. Very nice graphics are used to describe sums with a simple geometrical structure that would be hard to express with formal notation.

The fourth chapter discusses hierarchical models, in which GMRF's are used to model dependencies between parameters. As an example, one could observe a sequence Y_i of Bernoulli(p_i) random variables that are not independent. To model the dependence, one could assume that $X_i = \text{logit}(Y_i)$ arises from an autoregressive model and further assume that the variance of the random terms in that autoregressive process is random with some suitable distribution. Fitting a model of this kind requires use of Markov chain Monte Carlo (MCMC) techniques, which are briefly discussed. The focus of these discussions is not on convergence (which is referenced elsewhere), but to the blocking of parameters to increase the efficiency of the MCMC algorithms. Other examples include the fitting of a model in which monthly numbers of driver deaths on British roads are modeled as

$$Y_i \sim \text{N}(s_i + t_i, 1/\kappa_y),$$

where s and t are models for seasonal variation based on random walks and κ is the precision.

After considering Normal responses, various extensions to non-Normal responses are given. Rents in Munich are fitted to a mixture Normal distribution, in which the precision parameter is Gamma-distributed so as to produce a response with a t-distribution. A logistic model is fitted to rainfall data via a model that samples from a Normal and a Kolmogorov–Smirnov distribution. A model with Poisson response is fitted by replacing a complicated Poisson-based conditional distribution with a Normal distribution via a Taylor series approximation. In the second half of the last chapter, new methods are presented for improving this approximation, based on incorporating aspects of the Taylor series remainder. All of these examples, although mathematically complex, are presented with an emphasis on how certain fitting strategies can lead to increases in computing speed.

The final new topic is the use of GMRF's to approximate the Gaussian random fields used in geostatistics. Unlike GMRF's, these models are continuous and have covariance functions defined on the entire real line. Since these covariance functions concentrate most of the dependence within a finite region, they can be approximated by a random process when the extremely weak long-distance dependence is ignored. Great savings in speed are obtained by using a toroidal GMRF, in which the underlying graph has a toroidal structure. Since this implies that data in a square window tiles the plane with replicates of itself, this model should only be used with great caution.

Overall, the book is an excellent, clearly written introduction to a class of advanced models for use by those with a solid background in probability and statistics. The book has one major flaw, which is related to its production. Although the page layout, in general, is nicely done, major mistakes were made in reproducing some figures. For example, there are many maps of Germany with multiple regions in which each region has a gray scale value related to a response. The printing technology used for this book so muddies these grays together that it is almost impossible to compare the observed responses to fitted ones, except at the most extreme divergences. This could have been improved by using different printing technologies or by printing the maps in color.

Jeffrey D. PICKA
University of New Brunswick

REFERENCE

Hogg, R. V., McKean, J., and Craig, A. T. (2005), *Introduction to Mathematical Statistics* (6th ed.), Upper Saddle River, NJ: Prentice–Hall.