

Spatial Gaussian Markov Random Fields: Modelling, Applications and Efficient Computations

Yu Ryan Yue¹ and Xiao-Feng Wang^{2*}

¹Department of Statistics and CIS, Zicklin School of Business, Baruch College, The City University of New York, New York, NY, USA

²Department of Quantitative Health Sciences / Biostatistics Section, Cleveland Clinic Lerner Research Institute, Cleveland, OH, USA

Abstract

A powerful modelling tool for spatial data is the framework of Gaussian Markov random fields (GMRFs), which are discrete domain Gaussian random fields equipped with a Markov property. GMRFs allow us to combine the analytical results for the Gaussian distribution as well as Markov properties, thus allow for the development of computationally efficient algorithms. Here we briefly review popular spatial GMRFs, show how to construct them, and outline their recent developments and possible future work.

Keywords: Gaussian Markov random fields; Markov chain Monte Carlo; spatial statistics; Cholesky factorization; Integrated nested Laplace approximation

Introduction

Gaussian Markov random fields (GMRFs) are powerful and important tools for modeling spatial data. They have been widely used in different areas of spatial statistics including disease mapping, spatial-temporal modeling and image analysis. Constructing a GMRF is straightforward: it is just a finite-dimensional random vector following a multivariate Gaussian distribution with additional conditional independence properties, hence termed as Markov. It is convenient and invaluable to combine the analytical results for the Gaussian distribution and the Markov properties, which enables us to solve a large class of statistical models. Historically, the most common method to make inference for the parameters in GMRFs has been maximum likelihood [1,2]. The behavior of maximum likelihood estimator is asymptotic in nature and its small sample behavior is often unknown. On the other hand, the Markov property has become a requirement for constructing efficient Markov chain Monte Carlo (MCMC) algorithms for GMRFs. Rue [3] showed that the Markov property makes it possible to apply numerical methods on sparse matrices. He proposed fast algorithms for sampling and evaluating the log-density of a GMRF, and conducted efficient MCMC-based inferences. Rue and Held [4] provides a comprehensive account of the main properties of GMRFs, emphasizes the strong connection between GMRFs and numerical methods for sparse matrices, and outlines various applications of GMRFs for statistical inference (e.g., spatial statistics, time-series analysis, graphical models).

More specifically, a GMRF is a Gaussian random vector $x = (x_1, \dots, x_n)'$ with Markov property: for some $i \neq j$, x_i and x_j are independent conditional on other variables x_{-ij} . It is defined over a set of discrete indexed locations connected by a graph labelled by G , which shows the conditional independence property of x . We say x is a GMRF with respect to G with mean μ and precision (inverse covariance) matrix Q if its density has the form

$$\pi(x) = (2\pi)^{-n/2} |Q|^{1/2} \exp \left\{ -\frac{1}{2} (x - \mu)' Q (x - \mu) \right\}. \quad (1)$$

The Markov property in x is conveniently encoded in matrix Q : $Q_{ij} = 0$ if and only if x_i and x_j are conditionally independent. The pattern of zero and non-zero elements in such a matrix is called its sparsity structure. The total number of non-zero elements divided by the total number of elements is called the density of the matrix. In most

cases only $O(n)$ of the n^2 entries of Q are not zeroes, so Q has a very low density and is a sparse matrix. This allows for a fast Cholesky decomposition of Q as LL' , where L is the lower triangular matrix that inherits the sparseness of Q . Therefore, only non-zero terms in L are computed and the nodes can be reordered to decrease the number of the non-zero terms. The typical cost of this decomposition depends on the dimension of the GMRF: it is $O(n)$ for one dimension, $O(n^{3/2})$ for two dimensions and $O(n^2)$ for three dimensions. Using the Cholesky triangle, it is easy to produce random samples from a GMRF, and compute log-density of (1) and marginal variances; see Rue and Held [4] for more technical details.

GMRFs have been used in a wide range of common statistical models such as generalized linear (mixed) models, generalized additive models, dynamic linear models, and spatial and spatio-temporal models, among others. Those models can be written as hierarchical models which assume a n -dimensional latent GMRF with a sparse precision matrix to be point-wise observed through n_d conditional independent data y . The estimation of the models takes advantage of modern techniques for sparse matrices. We here focus on the (two dimensional) spatial GMRFs that have been extensively used in the hierarchical analysis for spatial data [5]. We briefly review popular spatial GMRFs, show how to construct them, and outline their recent developments and possible future work.

Construction of Spatial GMRFs

From definition (1), we can see that a GMRF is completely determined by its precision matrix Q : different Q 's bring different Markov properties to the field. To build matrix Q , one has to specify a particular neighbourhood structure, corresponding to the type of the spatial data. Below we show how to construct a few GMRFs that have been widely used in spatial statistics.

***Corresponding author:** Xiao-Feng Wang, Department of Quantitative Health Sciences, Cleveland Clinic Lerner Research Institute, 9500 Euclid Avenue/JJN3, Cleveland, OH 44195, USA, Tel: 216-445-7737; Fax: 216-444-8023; E-mail: wangxf6@ccf.org

Received April 22, 2014; Accepted April 24, 2014; Published April 28, 2014

Citation: Yue YR, Wang XF (2014) Spatial Gaussian Markov Random Fields: Modelling, Applications and Efficient Computations. J Biomet Biostat 5: e128. doi:10.472/2155-6180.1000e128

Copyright: © 2014 Yue YR, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

GMRFs on irregular lattices

An important type of spatial data is so called areal data, where the locations of observations are geographic regions (e.g., the states of the US) with adjacency information. For areal spatial data, the first-order GMRFs on irregular lattices are often preferred. To construct neighbourhood structure, we may define two regions are neighbors if they share a common border. Other ways to define neighbors are also possible. Between neighboring regions i and j , a Gaussian increment is defined as $x_i - x_j \sim N(0, 1/w_{ij}\tau)$, where w_{ij} are positive and symmetric weights. We can let $w_{ij} = 1$ if we believe region i equally depends on its neighbors, or let w_{ij} be, for example, the inverse Euclidean distance between region centroids if we think the neighbors somehow contribute differently. Assuming the increments are independent, the density of this GMRF is given by:

$$\pi(\mathbf{x}) \propto \tau^{(n-1)/2} \exp \left\{ -\frac{\tau}{2} \sum_{i \sim j} w_{ij} (\mathbf{x}_i - \mathbf{x}_j)^2 \right\}, \quad (2)$$

Where $i \sim j$ denotes the set of all unordered pairs of neighbors. The corresponding precision matrix \mathbf{Q} has entries

$$Q_{ij} = \tau \begin{cases} w_{i+} & \text{if } i = j, \\ -w_{ij} & \text{if } i \sim j, \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

Where $w_{i+} = \sum_{j: j \sim i} w_{ij}$, the summation over neighbors of node i on a lattice. Since the sum of each row is zero, \mathbf{Q} is singular with rank $n-1$. It is easy to see that the conditional distribution of x_i is

$$x_i | \mathbf{x}_{-i}, \tau \sim N \left(\frac{\sum_{j: j \sim i} w_{ij} x_j}{w_{i+}}, \frac{1}{\tau w_{i+}} \right), \quad (4)$$

where the conditional mean of x_i depends on its neighboring nodes x_j through weights w_{ij} and its conditional variance depends on weight sum w_{i+} .

GMRFs on regular lattices

Another important type of spatial data are point-referenced data, where the spatial locations are points with known coordinates. When the coordinates constitute a regular lattice, one can construct first and second-order GMRFs as described below.

For a regular lattice with $n=n_1 n_2$ nodes, let (i, j) denote the node in the i th row and j th column. In the interior, we can define the nearest four nodes of (i, j) as neighbors, i.e., $(i+1, j)$, $(i-1, j)$, $(i, j+1)$, $(i, j-1)$. Along the boundaries, we define the neighbors of (i, j) to be the two or three adjacent nodes. For example, the neighbors of (i, j) are $(i, j+1)$ and $(i+1, j)$ if (i, j) is the upper left corner. Without further weights, the corresponding precision matrix and the full conditionals of x_i are given in (3) and (4) with $w_{ij}=1$, respectively. By weighting the horizontal and vertical neighbors differently, this GMRF can be extended to an anisotropic model [4]. In practice the first-order GMRF models may not provide sufficient spatial smoothing as we need. To increase the smoothness of the field, we can use higher-order neighborhood structures. One way to build a second-order neighborhood structure for x_{ij} in the interior is based on the increment

$$(x_{i+1,j} + x_{i-1,j} + x_{i,j+1} + x_{i,j-1}) - 4x_{ij}$$

which is the sum of second-order differences in vertical and horizontal

directions. For the nodes on or near the boundaries we need different increments; see Yue and Speckman [6] for details. By assuming all the increments follow independent Gaussian distributions, we build a second-order GMRF on a regular lattice. The coefficients of the corresponding \mathbf{Q} can be found by expanding the quadratic terms of those increments. Using graphical notation, the conditional distribution of x_{ij} in the interior has mean and precision

$$E(x_{ij} | \mathbf{x}_{-ij}) = \frac{1}{20} \begin{pmatrix} \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ \\ \circ & \circ & \bullet & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \circ & \circ \\ 8 \circ & \bullet & \circ & \bullet & \circ & -2 \circ & \circ & \circ & \circ & -1 \bullet & \circ & \circ & \circ & \bullet \\ \circ & \circ & \bullet & \circ & \circ & \circ & \bullet & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ \end{pmatrix},$$

$$E(x_{ij} | \mathbf{x}_{-ij}) = \frac{1}{20\tau},$$

where the locations denoted by a ' \bullet ' represent the neighbors that x_{ij} depends on and the number in front of each grid denotes the weight given to the corresponding ' \bullet ' locations. This second-order GMRF is closely related to thin-plate spline [7], which has been widely used as a spatial smoother [1,8]. Yue and Speckman [6] showed that this GMRF can also be derived by discretizing the penalty function of the thin-plate spline.

Continuous indexed GMRFs

The GMRFs mentioned above are all discrete indexed in nature and thus only work appropriately for the spatial data on lattices. If the observations are measured at irregularly-spaced locations, one often bins the locations to a regular lattice first and then apply the GMRF to the summary statistics calculated for the grids. As a result, we lose information from binning process and the the spatial resolution of the lattice significantly affects the inference. Another problem is that a regular GMRF can only capture the smoothness of a spatial field, but not its structure of spatial correlations, because it defines the precision matrix (not covariance matrix). Thus, we cannot make inference regarding the correlation between two locations.

To address the issues mentioned above, Lindgren et al. [9] derived a new class of continuous indexed GMRFs, which are explicit mappings of Matérn Gaussian fields that have been extensively used in statistical modeling of spatial data. Letting $x(\mathbf{u})$ be a realization of spatial field x at location $\mathbf{u} \in \mathbb{R}^2$, such GMRFs are derived by solving a stochastic partial differential equation (SPDE)

$$(\kappa^2 - \Delta)^{\alpha/2} [\tau x(\mathbf{u})] = W(\mathbf{u}), \quad (6)$$

where $\Delta = \partial^2 / \partial u_1^2 + \partial^2 / \partial u_2^2$ is the two-dimensional Laplacian operator, $\kappa > 0$ is the spatial scale, $\alpha > 0$ controls the smoothness of the realizations, $\tau > 0$ controls the variance, and w is the spatial Gaussian white noise. The link to the Matérn smoothness ν and variance σ^2 is $\nu = \alpha - d/2$ and $\sigma^2 = \Gamma(\nu) \Gamma(\alpha) (4\pi)^{d/2} \kappa^{2\nu} \tau^2$, where d is the spatial dimension. A measure of the spatial range can be empirically derived as $\sqrt{8\nu / \kappa}$. Lindgren et al. [9] solve the SPDE (6) using finite element method. They first approximate $x(\mathbf{u})$ with piecewise linear basis functions defined on a triangular domain, and then turn (6) into a system of linear equations to derive the GMRF solution.

As shown in Lindgren et al. [9], the derived GMRF is the best piecewise linear approximation to the continuous solution to the SPDE given a triangulation. Since it is a GMRF representation of Matérn fields, it allows us to capture both spatial correlation and spatial

smoothness in a spatial process. Another exciting aspect of such models is their flexibility. There is no conceptual or computational barrier to extending them to being, for example, non-stationary, multivariate and spatial-temporal GMRFs. It is even possible to construct them on the sphere and other manifolds.

Extensions on GMRFs

Adaptive GMRFs

The smoothness of a GMRF is determined by the scale of increments, which is often invariant across the space. It means that the GMRF provides the same amount of smoothing at every location. This, however, limits the GMRF to stationary spatial processes only. We here present a few adaptive GMRFs that are recently developed to deal with non-stationary spatial processes.

Brezger et al. [10] extended the first-order GMRF in (2) to being adaptive by letting weights w_{ij} vary with locations. More specifically, they take independent gamma priors on each w_{ij} , that is $w_{ij} \sim \text{Gamma}(1/2, 1/2)$, making the marginal distribution of difference $(x_i - x_j)$ a Cauchy distribution. This approach allows varying strength of interactions between neighboring sites $i \sim j$, but without spatial prior structure.

Similarly, Yue and Speckman [6] extended the second-order GMRF in (5) to being non-stationary by spatially adaptive modeling its conditional variance, i.e., $\text{Var}(x_{ij} | \mathbf{x}_{-ij}) = (20\tau_{ij})^{-1}$. Thus, a small value of τ_{ij} (large variance) leads to less smoothing, appropriate when x_{ij} shows increased local variation. Then, let $\tau_{ij} = \tau\gamma_{ij}$, so that τ is the global scale parameter and γ_{ij} controls the local smoothing of the field. Finally, a first-order spatial GMRF is taken as a prior on $\log(\gamma_{ij})$ to make the smoothing vary over the space. The resulting GMRF is able to capture both local and global features of a spatial process while retains the nice Markov properties for computation. Yue et al. [11] proposed a similar adaptive GMRF but with independent gamma priors on γ_{ij} .

The SPDE models can be generalized to be non-stationary in several ways. Lindgren et al. [9] let spatial scale κ and precision τ in (6) to depend on the coordinate \mathbf{u} , and allow them vary slowly through log linear models. Therefore, one can analyze how the spatial non-stationarity depends on certain covariates. Bolin and Lindgren [9] used a nested SPDE approach to develop a large class of non-stationary covariance functions, such as oscillating covariance functions. Fuglstad et al. [12] introduced a new class of non-stationary spatial GMRFs with varying local anisotropy by adding a diffusion matrix that varies with position to the Δ operator in (6).

The adaptive GMRF models mentioned above have been successfully applied to the function magnetic resonance imaging (fMRI) data [10,13], the precipitation data [6,14], the global ozone data [15] and the neuroimaging meta-analysis [11].

Other extensions

Spatial-temporal GMRF models are extensions of spatial GMRF models to account for additional temporal variation. Think about a sequence of T graphs in time and let x_{it} denote the i th node in t th graph. A common extension to a spatial-temporal GMRF is to take into account temporal neighbors in addition to spatial neighbors. The temporal neighbors can be the same nodes in the previous and next graphs, that is, x_{it-1} and x_{it+1} . The precision matrix of the corresponding GMRF can be written as $\mathbf{Q} = \mathbf{Q}_T \otimes \mathbf{Q}_s$ (or $\mathbf{Q}_s \otimes \mathbf{Q}_T$), where \mathbf{Q}_T and \mathbf{Q}_s are precision matrices in time and space, respectively. We can use conditional autoregressive process (Besag, 1974) [16] to model

temporal correlations and the spatial GMRFs presented above to model spatial correlations. Then, both \mathbf{Q}_T and \mathbf{Q}_s are sparse matrices, making \mathbf{Q} sparse as well. Cameletti et al. [17] employed a model of this type to perform spatial-temporal analysis on particulate matter concentration data.

It is possible to build a multivariate GMRF by extending the SPDE framework considered above [18]. The idea is to replace a single SPDE with a system of SPDEs. The co-variances matrices constructed with this approach are automatically symmetric positive definite. The sparse precision matrix of this multivariate GMRF facilitates its implementation to large data sets. Hu et al. [19] have shown that these models are related (but not equivalent) to the multivariate Matérn fields constructed by.

Implementation of GMRFs

The GMRFs are often implemented in a class of structured additive regression models, which are quite flexible and have been extensively used [20]. In these models, the response variable y_i is assumed to belong to an exponential family, where the mean μ_i is linked to a structured additive predictor η_i through a link function $g(\cdot)$, so that $g(\mu_i) = \eta_i$. The predictor η_i accounts for various covariate effects in an additive way:

$$\eta_i = \alpha + \sum_{j=1}^{n_f} f_j(u_{ji}) + \sum_{k=1}^{n_\beta} z_{ki} \beta_k + \varepsilon_i, \quad (7)$$

Where f_j 's are unknown functions of the covariates u_{ji} , β_k 's represent the linear effect of z_k , and the ε_i 's are error terms. The structured additive regression models cover a large class of spatial and spatiotemporal models, geostatistical and geosadditive models. The spatial GMRFs can be used to model f_j when the covariate u_{ji} are spatial locations. The model (7) can handle various types of responses (e.g., continuous, binary and count data) and very different forms that the unknown f_j can take (e.g., random, nonlinear and spatial effects). As a matter of fact, most Bayesian models used in spatial statistics are of this form [1].

To fit the spatial models of form (7), one can perform standard MCMC methods to simulate the posterior distributions. Using GMRFs as priors in the Bayesian hierarchical models, it is feasible to construct accurate GMRF approximations to the full conditionals (if they are not Gaussian), based on which efficient blockwise Metropolis-Hastings algorithms can be used to explore the posteriors [4]. One is able to take advantage of the sparseness of GMRFs and implement the fast Cholesky factorization algorithms to speed up the MCMC algorithms. This can now be easily done using an R package named spam developed by Furrer and Sain [21]. Recently, Rue et al. [22] has introduced a novel Bayesian inference tool based on integrated nested Laplace approximations (INLA). The INLA method can directly compute very accurate approximations to the posterior marginal distributions for latent Gaussian models. It is much faster than MCMC and can be easily implemented using R-INLA package (<http://www.r-inla.org>). The package contains a variety of popular GMRFs that are readily to be used.

References

1. Cressie N (1993) Statistics for Spatial Data. New York: Wiley-Interscience.
2. Richardson S, Guillemeau C, Lasserre V (1992) Spatial linear models with autocorrelated error structure. The Statistician 41: 539-557.
3. Rue H (2001) Fast sampling of Gaussian Markov random fields. Journal of the Royal Statistical Society Series B 65: 325-338.

4. Rue H, Held L (2005) Gaussian Markov Random Fields: Theory and Applications, vol. 104 of Monographs on Statistics and Applied Probability. London: Chapman & Hall.
5. Banerjee S, Gelfand AE, Carlin BP (2004) Hierarchical modeling and analysis for spatial data. CRC Press.
6. Yue Y, Speckman PL (2010). Nonstationary spatial Gaussian Markov random fields. *Journal of Computational and Graphical Statistics* 19: 96–116.
7. Wahba G (1990) Spline Models for Observational Data. Philadelphia: SIAM [Society for Industrial and Applied Mathematics].
8. Nychka DW (2000) Spatial-process estimates as smoothers. In *Smoothing and regression: approaches, computation, and application*. MGA Schimek, ed. John Wiley & Sons.
9. Lindgren F, Rue H, Lindström J (2011) An explicit link between gaussian fields and Gaussian markov random fields: the stochastic partial differential equation approach (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73: 423-498.
10. Brezger A, Fahrmeir L, Hennerfeind A (2007) Adaptive Gaussian Markov random fields with applications in human brain mapping. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 56: 327-345.
11. Yue YR, Lindquist MA, Loh JM (2012) Meta-analysis of functional neuroimaging data using bayesian nonparametric binary regression. *Annals of Applied Statistics* 6: 697-718.
12. Fuglstad GA, Lindgren F, Simpson D, Rue H (2014). Exploring a new class of nonstationary spatial Gaussian random fields with varying local anisotropy. *Statistica Sinica*.
13. Yue Y, Loh JM, Lindquist MA (2010). Adaptive spatial smoothing of fMRI images. *Statistics and Its Interface* 3: 3-13.
14. Fuglstad GA, Lindgren F, Simpson D, Rue H (2013) Non-stationary spatial modelling with applications to spatial prediction of precipitation. Tech. rep., Norwegian University of Science and Technology.
15. Bolin D, Lindgren F (2011). Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. *The Annals of Applied Statistics* 5: 523-550.
16. Besag J (1974). Spatial interaction and the statistical analysis of lattice systems (with discussion). *Journal of the Royal Statistical Society, Series B: Methodological* 36: 192-236.
17. Cameletti M, Lindgren F, Simpson D, Rue H (2013) Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *ASTA Advances in Statistical Analysis* 97: 109-131.
18. Hu X, Simpson D, Lindgren F, Rue H (2013) Multivariate Gaussian Random Fields Using Systems of Stochastic Partial Differential Equations. Tech. rep., Norwegian University of Science and Technology.
19. Gneiting T, Kleiber W, Schlather M (2010) Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association* 105: 1167-1177.
20. Fahrmeir L, Tutz G (2001) Multivariate Statistical Modeling based on Generalized Linear Models. Berlin: Springer.
21. Furrer R, Sain SR (2010) spam: A sparse matrix r package with emphasis on mcmc methods for gaussian markov random fields. *Journal of Statistical Software* 36: 1-25.
22. Rue H, Martino S, Chopin N (2009) Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations