Fitting Gaussian Markov random fields to Gaussian fields

Håkon Tjelmeland,
Department of Mathematical Sciences,
Norwegian University of Science and Technology,
Trondheim, Norway.
E-mail: haakont@stat.ntnu.no

1 Introduction

In Bayesian spatial and spatio-temporal models, one often needs to represent a spatial variable, x say, by a stationary Gaussian field. We assume the Gaussian field to be represented on a finite lattice, $\Lambda = \{(i,j)|i = 0,\ldots,n_r-1, j = 0,\ldots,n_c-1\}$. If one or several realisations of x are available, from which the mean and covariance functions can be estimated, it is common to reinsert the estimated functions in the Bayesian model. If no realisations of x are available, one in stead typically uses prior information to specify parametric forms for the mean and covariance functions. Bayesian spatial models often have to be analysed via Markov chain Monte Carlo (MCMC) techniques, see, for example, Gilks et al. (1996). To be able to specify efficient MCMC procedures, it is often preferable to let the posterior distribution have a Markov property. It is common to ensure this by choosing a Markov prior field for x. The observation likelihood usually does not destroy this Markov property.

The above scenario requires that we specify x as a Gaussian field with a Markov property. This subclass of Gaussian fields is called Gaussian Markov random fields (GMRF) or conditional autoregressive models (Cressie, 1993; Besag and Kooperberg, 1995). Thus, one must be able to fit a GMRF to a given correlation function. This is the problem discussed in this paper, see Rue and Tjelmeland (1999) for a more thorough treatment.

2 Gaussian random fields and Gaussian Markov random fields

A general Gaussian field on the lattice Λ is a multivariate Gaussian variable with a given mean μ and covariance matrix Σ. The Gaussian field is a GMRF if it has the Markov property

$$\pi(x_{ij}|x_{\Lambda\setminus\{(i,j)\}}) = \pi(x_{ij}|x_{\partial(i,j)}), \tag{1}$$

where $\partial(i,j) \subseteq \Lambda \setminus \{(i,j)\}$ denotes the set of neighbours of (i,j). Typically, $\partial(i,j)$ is a set of the sites closest to (i,j). The Markov property defined by the neighbourhood system $\{\partial(i,j);(i,j) \in \Lambda\}$ induces a zero-pattern structure in the precision matrix $Q = \Sigma^{-1}$ by

$$(k,l) \notin \partial(i,j) \quad \Rightarrow \quad Q_{(i,j),(k,l)} = 0. \tag{2}$$

GMRFs have several computational advantages compared to general Gaussian fields (Rue, 2001). The parameters of GMRFs have interpretations via conditional mean and variances. However, unlike the mean and correlation functions of Gaussian fields, the GMRF parameters have no clear unconditional interpretations.

Let $\gamma(k,l)$ denote the covariance function for a homogeneous Gaussian field, defined as the covariance between $x_{0,0}$ and $x_{k,l}$, and let $\rho(k,l) = \gamma(k,l)/\gamma(0,0)$ and $f(x)$ be the corresponding correlation and density functions, respectively. Let $\widetilde{\gamma}$, $\widetilde{\rho}$ and $\widetilde{f}$ denote corresponding quantities for a GMRF. The number of parameters in a GMRF with a $(2m+1)\times(2m+1)$ neighbourhood, when imposing rotation and reflection invariance, is $m(m+1)/2+m+1$, see Figure 1 for $m=2$. We set $\theta_{kl} = Q_{(0,0),(k,l)}$ and let θ be a vector of all parameters of the GMRF, so $\widetilde{\gamma}$, $\widetilde{\rho}$ and $\widetilde{f}$ become functions of θ. As it is straightforward to fit exactly the mean and variance functions, we can limit the attention to Gaussian fields and GMRFs with zero mean and unit variance. Thus, our aim defined above can be rephrased to the following. For a given correlation function $\rho(k,l)$, and thereby a Gaussian density f, and a specified neighbourhood size m, identify the parameter vector, θ, that makes $\widetilde{f}$ and f as similar as possible.

3 Fitting a GMRF by matching the correlation function

To fit $\widetilde{f}$ to a given f, the perhaps most natural choice is to minimise, with respect to θ, the Kullback-Leibler discrepancy between them. If we let the densities f and $\widetilde{f}$ be defined on a torus, the computation of the Kullback-Leibler discrepancy is especially efficient and numerical minimisation becomes feasible, see Rue and Tjelmeland (1999) for details. However, it can be shown that Kullback-Leibler matches exactly all correlations inside $\partial(0,0) \cup \{(0,0)\}$ and lets the remaining ones be determined by inversion of the precision matrix. As a result, $\widetilde{\rho}$ is often very different from ρ for larger lags.

The Kullback-Leibler fits the correlation function $\widetilde{\rho}$ to ρ using only lags in $\partial(0,0)$. We propose to extend this idea by including also lags outside $\partial(0,0)$ by minimising

$$\mathcal{D}(f,\widetilde{f}) = \sum_{(k,l)\in\Lambda\setminus\{00\}}\sum (\rho(k,l) - \widetilde{\rho}(k,l))^2 W(k,l,\rho(k,l)), \tag{3}$$

where $W(k,l,\rho(k,l)) \geq 0$ is a weight function. The Kullback-Leibler discrepancy corresponds to setting $W(k,l,\rho(k,l))$ to some positive figure whenever $(k,l) \in \partial(0,0) \cup \{(0,0)\}$ and zero otherwise. We propose to use

$$W(k,l,\rho(k,l)) = \frac{1}{2\pi d(k,l)}, \tag{4}$$

where $d(k,l)$ is the Euclidean distance between $(0,0)$ and (k,l). With this choice the contribution from each radii has roughly the same weight and the fitted GMRF should have good average properties. By defining f and $\widetilde{f}$ on a torus, efficient computation of $\mathcal{D}(f,\widetilde{f})$ is possible, again see Rue and Tjelmeland (1999) for details. We have fitted the exponential, Gaussian, spherical and Matern correlation functions for ranges up to 50 lags and for neighbourhoods from 3×3 to 9×9. Some results are given in Figure 2 and more can be found in Rue and Tjelmeland (1999).

Similar to the GMRF on the line, the correlation function of a GMRF on a torus consists of exponential and exponential decaying sin's, see for example Besag and Kooperberg (1995). The exponential correlation function is therefore the easiest case, giving an almost perfect match for neighbourhood sizes larger than 3×3. The Gaussian and spherical correlation functions represent more difficult cases, as they consist of exponentially decaying sin's of all orders. Despite this fact, the fit is surprisingly good, even with 5×5 neighbourhoods.

Table 1 gives the upper right part (compare with Figure 1) of the coefficients in the fitted GMRF, with a 5×5 neighbourhood, for the three correlation functions in Figure 2. We find the parameters counter intuitive, when we take into account that the correlation functions we are matching are all positive. Even when using interpretations of the parameters as conditional mean and variances, the fitted parameter values are very hard to understand intuitively. Thus, we suggest that modelling of Gaussian spatially varying variables with a Markov property, should be done through a correlation function and then fit a GMRF with a prescribed neighbourhood using $\mathcal{D}(f,\widetilde{f})$. For the fitting process to be computationally feasible (for a large grid) it must be done for fields defined on a torus. However, when the fitted GMRF parameters have been obtained, they may be used also for GMRFs with other boundary conditions.

Table 1: Coefficients in the fitted GMRF using the $\mathcal{D}(f, \widetilde{f})$ criterion and a 5×5 neighbourhood for exponential, Gaussian and spherical correlation functions with range $r = 30$. The tables show the upper right part of Figure 1.

Exponential

3.1918	−1.8615	1.2606
−4.1667	−0.1940	−1.8615
14.526	−4.1667	3.1918

Gaussian

34.556	−4.9243	−1.9642
−290.25	108.55	−4.9243
635.79	−290.25	34.556

Spherical

24.901	−15.233	9.2614
−20.662	−0.3277	−15.233
69.179	−20.662	24.901

References

Besag, J. and Kooperberg, C. (1995). On conditional and intrinsic autoregression, *Biometrika* **82**(4): 733–746.

Cressie, N. A. C. (1993). *Statistics for spatial data*, 2 edn, John Wiley, New York.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in practice*, London: Chapman & Hall.

Rue, H. (2001). Fast sampling of Gaussian Markov random fields, *Journal of the Royal Statistical Society, Series B* **63**(2).

Rue, H. and Tjelmeland, H. (1999). Fitting Gaussian Markov random fields to Gaussian fields, *Statistics no. 16*, Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway. Available from http://www.math.ntnu.no/preprint/.

RESUME

Cet article considère la tâche suivante, souvent produite dans les modèles bayésiens spatiaux: construire un champ gaussien markovien homogène (GMRF) sur un treillis avec une fonction de corrélation donnée. La distance de Kullback-Leibler manque souvent dans cette tâche, ce qui donne un grave comportement indésirable de la fonction de corrélation pour des distances en dehors du voisinage. Nous proposons un nouveau critère qui résout cette difficulté, et nous démontrons que des GRMFs avec de petits voisinages peuvent très bien approximer des champs gaussiens, même avec de longues corrélations.

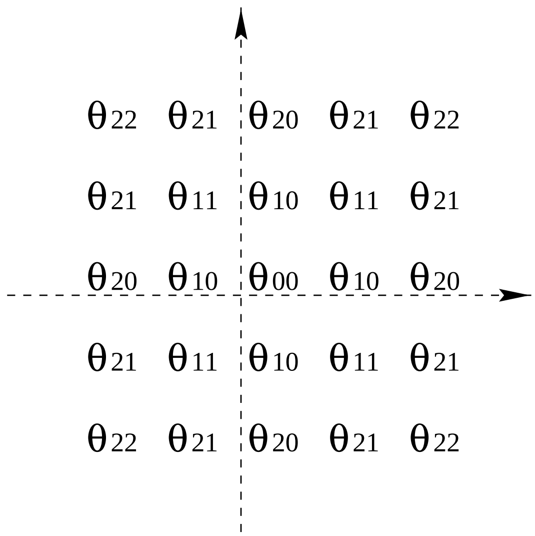

Figure 1: The 6 parameters for a GMRF with a 5×5 neighbourhood.

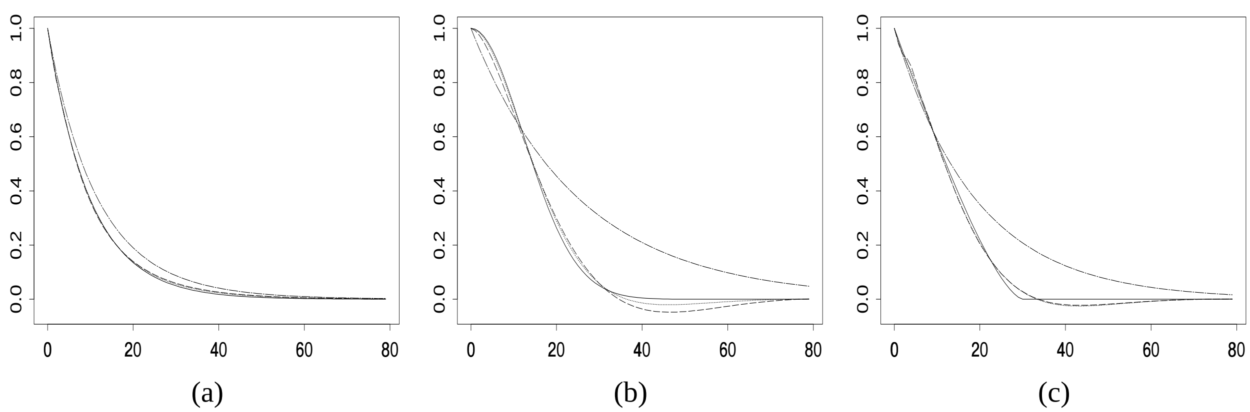

Figure 2: Target correlation function (solid) and correlation functions for fitted GMRFs with neighbourhood sizes 3×3 (dash-dotted), 5×5 (dashed) and 7×7 (dotted). Target correlation functions are (a) exponential, (b) Gaussian and (c) spherical, all with correlation range $r = 30$.