

Take Me Out to (Analyze) the Ballgame

Chris Comiskey

March 27, 2017

Contents

1	Method 3 - INLA and SPDEs	1
1.1	Matern Covariance	1
1.2	Integrated Nested Laplace Approximations (INLA)	1
1.2.1	Gaussian Markov Random Fields	1
1.2.2	INLA	3
1.3	Stochastic Partial Differential Equations (SPDE)	3
1.3.1	Stochastic Calculus	3
1.3.2	Linking INLA and SPDEs	3
1.4	Bayesian Inference with INLA and SPDEs: R-INLA	4
A	Blah Blah Blah	6

1 Method 3 - INLA and SPDEs

1.1 Matern Covariance

1.2 Integrated Nested Laplace Approximations (INLA)

1.2.1 Gaussian Markov Random Fields

Approximate Bayesian Inference for Hierarchical Gaussian Markov Random Field Models [Rue and Martino, 2007]

- \mathcal{O} definition, Wikipedia

$$f(x) = \mathcal{O}(g(x)) \text{ as } x \rightarrow \infty \text{ iff } \exists M, x_0 \text{ such that } |f(x)| \leq M|g(x)| \text{ for all } x \geq x_0$$

- Gaussian Markov Random Field

- $\mathbf{x} = \{x_i : i \in \mathcal{V}\}$
- “ \mathbf{x} is a $n = |\mathcal{V}|$ -dimensional Gaussian random vector with additional conditional independence/-Markov properties”
- $\mathcal{V} = \{1, \dots, n\}$
- Graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, with vertices, edges.
- “Two nodes, x_i and x_j are conditionally independent given the remaining elements of \mathbf{x} , if and only if $\{i, j\} \notin \mathcal{E}$.”
- “Then \mathbf{x} is a GMRF with respect to \mathcal{G} .”
- “The edges in \mathcal{E} are in one-to-one correspondence with the non-zero elements of the precision matrix of \mathbf{x} , \mathbf{Q} , in the sense that $\{i, j\} \in \mathcal{E}$ if and only if $Q_{ij} \neq 0$ for $i \neq j$.”
- “When $\{i, j\} \in \mathcal{E}$ we say that i and j are neighbours, which we denote $i \sim j$.”

- Nice description of hierarchical model architecture... “One of the main areas of application for GMRFs is that of (Bayesian) hierarchical models. A hierarchical model is characterised by **several stages** of observables and parameters. The **first stage**, typically, consists of distributional assumptions for the observables conditionally on latent parameters. For example if we observe a time series of counts y , we may assume, for $y_i, i \in D \subset \mathbb{R}^?$ a Poisson distribution with unknown mean λ_i . Given the parameters of the observation model, we often assume the observations to be conditionally independent. The **second stage** consists of a prior model for the latent parameters λ_i or, more often, for a particular function of them. For example, in the Poisson case we can choose an exponential link and model the random variables $x_i = \log(\lambda_i)$. At this stage **GMRFs provide a flexible tool to model the dependence** between the latent parameters and thus, implicitly, the dependence between the observed data. This dependence can be of various kind, such as temporal, spatial, or even spatiotemporal. The **third stage** consists of prior distributions for the unknown hyperparameters. These are typically precision parameters in the GMRF.”

- “We propose a deterministic alternative to MCMC based inference... computed almost instant[ly]... proves to be quite accurate.”
- Oops. **Gaussian Markov Random Field \neq Gaussian Random Field**;
GMRF has different properties; think graph theory edges, and non-edge-connected nodes have zero correlation, hence sparse matrices for INLA

1.2.2 INLA

Approximate Bayesian Inference for Hierarchical Gaussian Markov Random Field Models [Rue and Martino, 2007] ...More to this paper, on INLA

1.3 Stochastic Partial Differential Equations (SPDE)

1.3.1 Stochastic Calculus

Stochastic Differential Equation (wikipedia)

- “A heuristic but helpful interpretation of the stochastic differential equation (of continuous time stochastic process X_t) is that in a small time interval of length δ the stochastic process X_t changes its value by an amount that is normally distributed (for example) with expectation $\mu(X_t, t)\delta$ and variance $\sigma(X_t, t)^2\delta$ and is independent of the past behavior of the process.”

Stochastic Calculus [Mao, 2007]

- Integral of random process is another random process. Random process not integrable in traditional sense, so stochastic calculus created, by Ito, Ito Calculus. Use Brownian motion as some sort of reference point.

$$Y_t = \int_0^t H_s dX_s$$

Integrand and integrator are stochastic processes.

1.3.2 Linking INLA and SPDEs

An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach [Lindgren et al., 2011]

- “A GMRF is a discretely indexed Gaussian field \mathbf{x} , where the full conditionals” depend only on a small set of symmetric neighbor relationships, which yields sparse matrices that lend themselves to approximations—INLA
- The GMRF “computational gain comes from the fact that the zero pattern of the precision matrix \mathbf{Q} (the inverse covariance matrix) relates directly to the notion of neighbours...”
- Matrix **Q: the inverse covariance matrix**
- “The result is a basis function representation with piecewise linear basis functions, and Gaussian weights with Markov dependencies determined by a general triangulation of the domain.” [Lindgren et al., 2011]
 - Recall “basis” from linear algebra, where a set of linearly independent vectors span a space
 - function space - space of *functions*
 - basis functions - a set of functions from which can build any function in the function space

- Matern (2):

$$r(\mathbf{u}, \mathbf{v}) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} (\kappa \|\mathbf{u} - \mathbf{v}\|)^\nu K_\nu(\kappa \|\mathbf{u} - \mathbf{v}\|)$$

- scaling parameter $\kappa = \rho$ range parameter
- Empirically derived: $\rho = \sqrt{8\nu}/\kappa$

- Linear fractional SPDE:

$$(\kappa^2 - \Delta)^{\alpha/2} x(\mathbf{u}) = \mathcal{W}(\mathbf{u})$$

- $\alpha = \nu + d/2$

- Δ is Laplacian:

$$\Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$$

- Marginal Variance:

$$\sigma^2 = \frac{\Gamma(\nu)}{\Gamma(\nu + d/2)(4\pi)^{d/2}\kappa^{2\nu}}$$

1.4 Bayesian Inference with INLA and SPDEs: R-INLA

Bayesian Spatial Modelling with R-INLA [Lindgren and Rue, 2015]

- “...as discussed in Lindgren et al. (2011), one can express a large class of random field models as solutions to continuous domain stochastic partial differential equations (SPDEs), and write down explicit links between the parameters of each SPDE and the elements of precision matrices for weights in a discrete basis function representation.”
- “As discussed in the introduction, **an alternative to traditional covariance based modelling is to use SPDEs**, but carry out the practical computations using Gaussian Markov random field (GMRF) representations. This is done by **approximating the full set of spatial random functions with weighted sums of simple basis functions**, which allows us to hold on to the continuous interpretation of space, while the **computational algorithms** only see discrete structures with Markov properties. Beyond the main paper Lindgren et al. (2011), this is further discussed by Simpson, Lindgren, and Rue (2012a,b).”
- “The simplest model for (spatial field) $x(\mathbf{s})$ currently implemented in R-INLA is the SPDE/GMRF version of the stationary Matern family, obtained as the stationary solutions to

$$(\kappa^2 - \Delta)^{\alpha/2}(\tau x(\mathbf{s})) = \mathcal{W}(\mathbf{s}), \mathbf{s} \in \Omega$$

where

- Δ is the Laplacian
- κ is the spatial scale parameter
- α controls the smoothness of the realisations
- τ controls the variance
- Ω is the spatial domain
- $\mathcal{W}(\mathbf{s})$ is a Gaussian spatial white noise process

Whittle (1954, 1963) shows stationary solutions on \mathbb{R}^d have Matern covariances,

$$\text{COV}(x(\mathbf{0}), x(\mathbf{s})) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} (\kappa\|\mathbf{s}\|)^{\nu} K_{\nu}(\kappa\|\mathbf{s}\|)$$

The parameters in the two formulations are coupled so that the Matern smoothness is $\nu = \alpha - d/2$ and marginal variance is

$$\sigma^2 = \frac{\Gamma(\nu)}{\Gamma(\alpha)(4\pi)^{d/2}\kappa^{2\nu}\tau^2}$$

Exponential covariance: $\nu = 1/2$; (i) for $d = 1 \rightarrow \alpha = 1$, (ii) for $d = 2 \rightarrow \alpha = 3/2$ ”

- “The models discussed in Lindgren et al. [2011] and implemented in R-INLA are built on a basis representation

$$x(\mathbf{s}) = \sum_{k=1}^n \psi_k(\mathbf{s}) x_k$$

where

- $\psi_k(\cdot)$ are deterministic basis functions, and
- the joint distribution of the weight vector $\mathbf{x} = \{x_1, \dots, x_n\}$ is chosen so that the distribution of the functions $x(\mathbf{s})$ approximates the distribution of solutions to the SPDE on the domain.”
- piecewise polynomial basis functions
- use Finite Element Method - project the SPDE onto the basis representation
-

A Blah Blah Blah

References

- Håvard Rue and Sara Martino. Approximate bayesian inference for hierarchical gaussian markov random field models. Journal of statistical planning and inference, 137(10):3177–3192, 2007.
- Xuerong Mao. Stochastic differential equations and applications. Elsevier, 2007.
- Finn Lindgren, Håvard Rue, and Johan Lindström. An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73(4):423–498, 2011.
- Finn Lindgren and Håvard Rue. Bayesian spatial modelling with r-inla. Journal of Statistical Software, 63(19), 2015.