Progress this week
November 22, 2019

- Added some keywords.
- Turned introduction into complete sentences.

Previous word count: — **Current word count:**
Previous page count: — **Current page count:**

REVIEW ARTICLE

The Integrated Nested Laplace Approximation applied to Spatial Point Process Models

Kenneth Flagg and Andrew Hoegh

Montana State University, Bozeman, MT

ARTICLE HISTORY
Compiled November 22, 2019

ABSTRACT
This template is for authors who are preparing a manuscript for a Taylor & Francis journal using the LaTeX document preparation system and the `interact` class file, which is available via selected journals' home pages on the Taylor & Francis website.

KEYWORDS
INLA, spatial prediction, log-Gaussian Cox process, spatial point process

1. Introduction

Spatial prediction is a high-dimensional inference problem. When the goal of statistical modeling is to produce a graphical map of a random variable over space, the model ultimately must be able to predict that random variable at every pixel of the image. A map image will typically be at least several hundred by several hundred pixels, so in total there can easily be hundreds of thoudands of pixels requiring predictions. Thus, even when a model has only half a dozen parameters, it may include hundreds of thousands of latent variables.

Spatial point process models further complicate the situation with difficult likelihoods. *(Cite some computational papers — Baddely?)* Both maximum likelihood and Bayesian model fitting require integrating the intensity function over space, but the integral is generally not available in closed form. Many methods have been introduced including quadrature-based approximations *(cite Baddeley)*, pseudodata approaches *(cite Baddeley/Berman/Turner etc)*, and Markov chain Monte Carlo [3].

Development of the integrated nested Laplace approximation (INLA) has made accurate approximate model fitting considerably more feasible for a particular class of log-Gaussian Cox process (LGCP) models. INLA was developed to fit Bayesian hierarchical models with many latent Gaussian variables [4]. A key part of INLA's computational simplicity is that it calculates the posterior distribution of each latent Gaussian variable one at a time; that is, it provides only the posterior marginal distributions rather that the full joint distribution.

When using a LGCP for spatial mapping, two aspects make INLA a suitable approach. First, the LGCP is driven by a spatial Gaussian process, so the latent variables are Gaussian. Second, even though the latent variables are expected to exhibit spatial

CONTACT Kenneth Flagg. Email: kenneth.flagg@montana.edu

dependence, their full joint distribution is not needed. In most situations it suffices to map their predicted values, variance, and upper and lower interval bounds pointwise across space.

1.1. Log-Gaussian Cox Process

Poisson process intensity $\lambda(\mathbf{s})$ events per unit area

Model $\log \lambda(\mathbf{s}) = Z(\mathbf{s})$, $Z(\mathbf{s})$ spatial Gaussian process

Random continuous function: $Z(\mathbf{s})$ a Gaussian random variable, Mean μ Matern covariance function

Poisson process log-likelihood: $\ell(\lambda) = C - \int \lambda(\mathbf{s})d\mathbf{s} + \sum \log(\lambda(\mathbf{s_i}))$

(Need to add covariates to notation)

1.2. INLA

INLA fast approximation for marginals useful for mapping [4]

Integrated Nested Laplace Approximation

Bayesian Hierarchical models, many latent Gaussian variables, few parameters

Laplace approximation in general: $\int \exp[h(x)]\mathrm{d}x$, Taylor expansion of $h(x)$

Example from [1]

- $\mathbf{y} = (y_1, \ldots, y_n)'$ independent Gaussian observations
- $y_i \sim \mathrm{N}(\theta, \sigma^2)$
- $\theta \sim \mathrm{N}(\mu_0, \sigma_0^2)$
- $\psi = 1/\sigma^2$, $\psi \sim \mathrm{Gamma}(a, b)$
- The posterior distribution of ψ:

$$p(\psi|\mathbf{y}) \propto \frac{p(\mathbf{y}|\theta, \psi)p(\theta)p(\psi)}{p(\theta|\psi, \mathbf{y})}$$

- Laplace approximation:

$$\tilde{p}(\psi|\mathbf{y}) \propto \frac{p(\mathbf{y}|\theta, \psi)p(\theta)p(\psi)}{\tilde{p}_G(\theta^*|\psi, \mathbf{y})}$$

Repeat for θ, will depend on ψ

Provides marginal posteror for one entry at a time of a vector $\boldsymbol{\theta}$

2. Methodology

2.1. The SPDE Approach

GP has dense covariance matrix

SPDE result approximates GP with CAR and sparse covariance matrix [2]

$$(\kappa^2 - \Delta)^{\alpha/2}(\tau Z(\mathbf{s})) = W(\mathbf{s})$$

Choose nodes $\mathbf{s}_i$ to model $Z(\mathbf{s}_i)$, build a triangular mesh, autoregressive model on the nodes

Change of basis: barycentric coordinates allow sparse matrices and simple linear inerpolation

2.2. Going Off the Grid

Log-likelihood:

$$\ell(Z) = C - \int \exp\left[Z(\mathbf{s})\right] d\mathbf{s} + \sum Z(\mathbf{s_i})$$

[5]:

$$\ell(Z) \approx C - \sum_i \tilde{\alpha}_i \exp\left[\sum_j z_j \phi_j(\tilde{\mathbf{s}}_i)\right] + \sum_i \sum_j z_j \phi_j(\mathbf{s_i})$$

(Poisson distribution)

2.3. Variable Sampling Effort – Delete this subsection?

Observed a thinned process
Thinning process can be known or unknown
Scale SPDE node integration weights by thinning probabilities when known
Incorporate log-linear model for thinning probability when unknown [6]

3. Applications

3.1. Simulation Study

3.2. Data Application

Examples with data, maybe `bei` dataset or Victorville

4. Conclusion and Discussion

References

[1] M. Blangiardo and M. Cameletti, *Spatial and Spatio-temporal Bayesian Models with R-INLA*, Wiley, 2015.
[2] F. Lindgren, H. Rue, and J. 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 (2011), pp. 423–498.
[3] J. Møller and R. Waagepetersen, *Modern spatial point process modelling and inference*, Scandinavian Journal of Statistics 34 (2007), pp. 643–711.
[4] H. Rue, S. Martino, and N. Chopin, *Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations*, Journal of the royal statistical society: Series b (statistical methodology) 71 (2009), pp. 319–392.
[5] D. Simpson, J.B. Illian, F. Lindgren, S.H. Sørbye, and H. Rue, *Going off grid: Computationally efficient inference for log-gaussian cox processes*, Biometrika 103 (2016), pp. 49–70.

[6] Y. Yuan, F.E. Bachl, F. Lindgren, D.L. Borchers, J.B. Illian, S.T. Buckland, H. Rue, T. Gerrodette, *et al.*, *Point process models for spatio-temporal distance sampling data from a large-scale survey of blue whales*, The Annals of Applied Statistics 11 (2017), pp. 2270–2297.