INLA for GMRFs (e.g. GLMMs, Spatial Models) with An Example of Leukemia Cases

Frances Lin

June 2022

Background and Introduction

The steps involving the Bayesian inference may appear easy and straightforward:

- updating prior beliefs about the unknown parameters and
- obtaining the posterior distribution for the parameters.

However, this is much harder to do in practice since solutions in closed-form may not always be determined.

- MCMC (Markov chain Monte Carlo) was introduced and represented a breakthrough in Bayesian inference in the early 1990s.
- Tools such as `WinBugs` (Spiegelhalter et al., 1995), `JAGS` (Plummer, 2016), and `stan` (Stan Development Team, 2015) have also been developed, and
- Bayesian statistics has gained popularity in many fields.

Background and Introduction

However, MCMC methods

- not only can be computationally demanding (i.e. requires a large amount of CPU),
- but also present convergence issues.

INLA (integrated nested Laplace approximation) is a fast alternative to MCMC for Bayesian inference. INLA

- can be applied to a very flexible class of models named LGMs (latent Gaussian models), which ranges from GLMMs (generalized linear mixed models) to time-series, spatial and spatio-temporal models.
- allows for faster and more accurate inference without trading speed for accuracy, and
- is accessible through the **R** package `R-INLA` (Citation).

Applications

INLA have found spatial or spatio-temporal applications in a wide variety of fields such as environment, ecology, disease mapping, public health, cancer research, energy, economics, risk analysis, etc.

Selected examples include:

- environmental risk factors to liver fluke in cattle (Innocent et al., 2017);
- polio-virus eradication in Pakistan (Mercer et al., 2017);
- socio-demographic and geographic impact of HPV vaccination (Rutten et al., 2017);
- topsoil metals and cancer mortality (Lopez-Abente et al., 2017);
- probabilistic prediction of wind power (Lenzi et al., 2017);
- applications in spatial econometrics (Bivand et al., 2014; Gomez-Rubio et al., 2015; Gomez-Rubio et al., 2014);
- predicting extreme rainfall events in space and time (Opitz et al., 2018), etc.

Outline

- Key Components
  0. Bayesian Inference
  1. Latent Gaussian Models
  2. Additive Models
  3. Gaussian Markov Random Fields
  4. Additive Models and Gaussian Markov Random Fields
  5. Laplace Approximations
- INLA
- INLA-SPDE (Stochastic Partial Differential Equations) Approach
- Discussion
- A Spatial Example of Leukemia Cases using the package `R-INLA`

0. Bayesian Inference

The posterior distribution is proportional to the likelihood function multiples by the prior distribution

$$f(\theta|y) = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)d\theta} \propto p(y|\theta)p(\theta),$$

where $p(y|\theta)$ is the likelihood function, $p(\theta)$ is the prior, and $\int p(y|\theta)p(\theta)d\theta$ is the normalizing constant.

- Based on the posterior distribution, relevant statistics for the parameters of interest (e.g. marginal distribution, means, variances, and credibility intervals) can be obtained.
- However, the integral is generally intractable in closed-form, thus requiring the use of numerical methods such as MCMC.

1. Latent Gaussian Models

The latent Gaussian models (LGMs) is a class of three-stage Bayesian hierarchical models. It involves the following stages:

1. Observations y is assumed to be conditionally independent, given a latent Gaussian random field x and hyperparameter θ_1

$$y|x, \theta_1 \sim \prod_{i \in I} p(y_i|x_i, \theta_1). \qquad \textit{likelihood}$$

2. The latent field $x|\theta_2$ is assumed to be a GMRF (Gaussian Markov random field) with a sparse precision matrix Q

$$x|\theta_2 \sim p(x|\theta_2) = N(\mu(\theta_2), Q^{-1}(\theta_2)), \qquad \textit{latent field}$$

   where $Q = \Sigma^{-1}$ is the precision matrix and θ_2 is a hyperparameter.

3. The hyperparameters of the latent field that are not necessarily Gaussian are assumed to follow a prior distribution

$$\theta = (\theta_1, \theta_2) \sim p(\theta), \qquad \textit{hyperpriors}$$

   where $p(\cdot)$ is a known distribution.

1. Latent Gaussian Models - Cont.

Then, the posterior distribution, structured in a hierarchical way, becomes

$$\begin{aligned} p(x, \theta | y) &\propto p(y | x, \theta) p(x, \theta) \\ &\propto \prod_{i \in I} p(y_i | x_i, \theta) p(x | \theta) p(\theta). \end{aligned}$$

For computational reasons and to ensure accurate approximations, the following assumptions hold:

1. Each observation y_i depends only on one component of the latent field x_i, and most components of x will not be observed.
2. The distribution of the latent field x is Gaussian and is close to a Gaussian Markov random field (GMRF) when the *dim* of n is high (10^3 to 10^5).
3. The number of hyperparameters θ is small (~ 2 to 5 but < 20).

2. Additive Models