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 believes about the unknown parameters with observed data and obtaining the posterior distribution for the parameters. However, this is much harder to do in practice since solutions in closed-form may not be determined.

The simulation-based inference through the idea of MCMC (Markov chain Monte Carlo) was then introduced and represented a breakthrough in Bayesian inference (Robert & Casella, 1999) in the early 1990s. MCMC tools such as `WinBugs` (Spiegelhalter et al., 1995), `JAGS` (Plummer, 2016), and `stan` (Stan Development Team, 2015) have also been developed. Bayesian statistics has gained popularity in many fields. However, these MCMC methods, based on sampling, not only are 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 a class of LGMs (latent Gaussian models) that does not require sampling. I.e. INLA can be applied to a very wide and flexible class of models ranging from GLMM to spatial and spatio-temporal models, named LGMs. INLA allows for faster and more accurate inference without trading speed for accuracy, and it can be accessed through the **R** package `R-INLA`.

Notation

0. Bayesian Inference

The posterior distribution is proportional to the likelihood function multiples by the prior distribution **(Need citation)**

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

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 such as marginal distribution, means, variances, quantiles, credibility intervals, etc. can be obtained (Rue et al., 2017).

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-state Bayesian hierarchical models. It involves the following stages:

Observations y is assumed to be conditionally independent, given a latent Gaussian random field x (joint distribution of all parameters in the linear predictor) and hyperparameters θ

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

The versatility of the model class can be specified through the unobserved latent Gaussian field

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

where $Q = \Sigma^{-1}$ is the precision matrix, which captures the underlying dependence structure of the data. Latent field can be, for example, covariates, unstructured random effects (e.g. white noise), structure random effects (e.g. temporal dependency, spatial dependency, smoothness terms).

The hyperparameters of the latent field that is not necessarily Gaussian

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

contains the prior knowledge of the parameters and controls the likelihood of the data and/or the latent Gaussian field. Hyperparameters can be, for example, variance of observation noise, variance of the unstructured random field, range of a structured random effect.

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

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

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

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 n is high (10^3 to 10^5).

3. The number of hyperparameters θ is small (~ 2 to 5 but < 20).

2. Additive Models

LGMs are an umbrella class that generalizes the large number of related additive and/or generalized linear models. Consider the generalized linear model setup, for example,

$$y \sim \prod_i^N p(y_i|x_i, \theta)$$

$$g(\mu_i) = \eta_i = \alpha + \sum_j \beta_j z_{ji} + \sum_k f_k(w_{ki}) + \varepsilon_i,$$

where α is the overall intercept, $\{\beta_j\}$ are linear effects of fixed covariates z, and $\{f_k\}$, which are used to represent specific Gaussian processes, are nonlinear/smooth effects of some covariates w. Examples of model components f_k include random effects with different types of correlations, spatially or temporally correlated effects, smoothing and stochastic spline, and measure errors.

For GLMs (generalized linear models), the expression becomes $\alpha + \sum_j \beta_j z_j$ (i.e. $f(\cdot) = 0$). For GAMs (generalized additive models), the expression becomes $\alpha + \sum_k f_k(w_k)$.

The model is a LGM iff the joint distribution of

$$x = (\eta, \alpha, \beta, f(\cdot))$$

is Gaussian. I.e.

$$x = (\eta, \alpha, \beta, f(\cdot))|\theta \sim N(\mu(\theta), Q^{-1}(\theta)).$$

This can be achieved by putting Gaussian priors on the intercept and the parameter of the fixed effects.

If we further assume conditional independence of x, then this latent field x is a Gaussian Markov random field.

3. Gaussian Markov Random Fields

GMRFs

The Markov assumption in the GMRFs results in a sparse precision matrix.

4. Additive Models and Gaussian Markov Random Fields

5. Laplace Approximations

INLA

Applications

The `R` package `R-INLA` have found spatial applications in a wide variety of fields such as environment, ecology, disease mapping, medical imaging, public health, cancer research, energy, economics, risk analysis, etc.

Some selected examples include: environmental risk factors to liver Fluke in cattle (Innocent et al., 2017); modelling recovering fish populations (Boudreau et al., 2017); polio-virus eradication in Pakistan (Mercer et al., 2017); cortical surface fMRI data (Mejia 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) with spatially misaligned data; ethanol and gasoline pricing (Laurini, 2017); applications in spatial econometrics (Bivand et al., 2014; Gomez-Rubio et al., 2015; Gomez-Rubio et al., 2014); probabilistic prediction of wind power (Lenzi et al., 2017); modeling landslides as point processes (Lombardo et al., 2018); predicting extreme rainfall events in space and time (Opitz et al., 2018), etc.

Reference

Bakka, H., Rue, H., Fuglstad, G. A., Riebler, A., Bolin, D., Illian, J., Krainski, E., & Lindgren, F. (2018). Spatial modeling with R-INLA: A review. Wiley Interdisciplinary Reviews: Computational Statistics, 10(6), e1443.

Bolin, D. (2015). *Lecture 1: Introduction Gaussian Markov random fields.* Personal Collection of D. Bolin, Chalmers University of Technology, Gothenburg Sweden.

Morrison, K. (2017). A gentle INLA tutorial. Precision Analytics. https://www.precision-analytics.ca/articles/a-gentle-inla-tutorial/.

Rue, H., Riebler, A., Sørbye, S. H., Illian, J. B., Simpson, D. P., & Lindgren, F. K. (2017). Bayesian computing with INLA: a review. Annual Review of Statistics and Its Application, 4, 395-421.