Part 3: Lattice processes

Last week:

- Lattice data and lattice processes
- Neighborhoods and cliques
- Conditional independence graph
- Brook's lemma
- Hammersley-Clifford's theorem
- Markov Random Fields (MRF)

Yesterday:

- Ising model (Binary MRF)
- Inference

Today

- Inference (INLA = Integrated Nested Laplace Approximations)
- Modeling
  - Germany
  - Scot's pine in Sweden (maybe)

Lattice data and lattice process

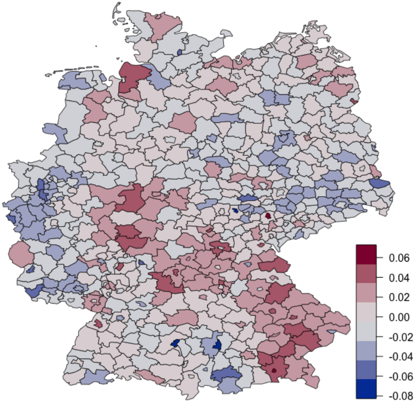
Lattice data: $Y(s_i)$: Data for discrete spatial features (districts, pixels, voxels, non-overlapping catchments)

Lattice process: $Z(s : s \in D_s)$ defined on a finite or countable subset D_s of R^d (d=2).

Want to define models 'locally' through full conditional distributions.

Germany

Sample from GMRF model with Gaussian data process

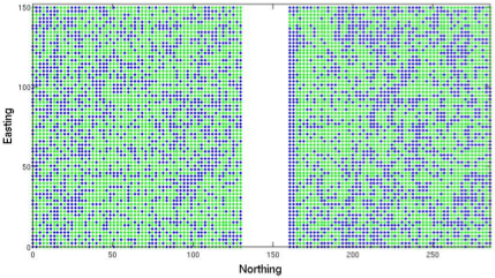


Scots Pine Data

Pedigree 56 unrelated parents, partial diallel design. Original 8160 seedlings.

Spatial location 2.2×2.2m grid, two trail sites.

Data Hight and bad(1) / good(0) branch angle of 4970 26-years-old scots pine.



INLA: Integrated Nested Laplace approximations

- A new method for doing Bayesian inference for latent Gaussian Markov Random Field models.
- Based on direct approximations (no sampling)
- R-package, `r-inla`, available at `www-r-inla.org`

Key papers

H. Rue, S. Martino & N. Chopin *Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations*, JRSS-B, 2009

Martins, T. G., Simpson, D., Lindgren, F. K. & Rue, H. *Bayesian computing with INLA : New features.* Comp. Stat. & Data Anal., 2013

Inference methods

Markov Chain Monte Carlo, MCMC

- Run a Markov chain to get samples from $\pi(x, \theta|y)$
- Can find posterior for any parameter(s) / variable(s) or function of variables
- But 1: Need many iterations / takes a long time.
- But 2: Burn-in and mixing problems, hard to detect.

Inference methods

Markov Chain Monte Carlo, MCMC

- Run a Markov chain to get samples from $\pi(x, \theta | y)$
- Can find posterior for any parameter(s) / variable(s) or function of variables
- But 1: Need many iterations / takes a long time.
- But 2: Burn-in and mixing problems, hard to detect.

Integrated Nested Laplace Approximations, INLA

- Non-sampling based numerical method.
- Fast(er) $\Rightarrow$ enables simulation studies
- But 1: For *latent Gaussian Markov Random Field (GMRF) models* with max. 15-20 non-Gaussian hyper-parameters.
- But 2: Posteriors for functions of variables and jointly more tricky.

Inference methods

Markov Chain Monte Carlo, MCMC

- Run a Markov chain to get samples from $\pi(x, \theta | y)$
- Can find posterior for any parameter(s) / variable(s) or function of variables
- But 1: Need many iterations / takes a long time.
- But 2: Burn-in and mixing problems, hard to detect.

Inference methods

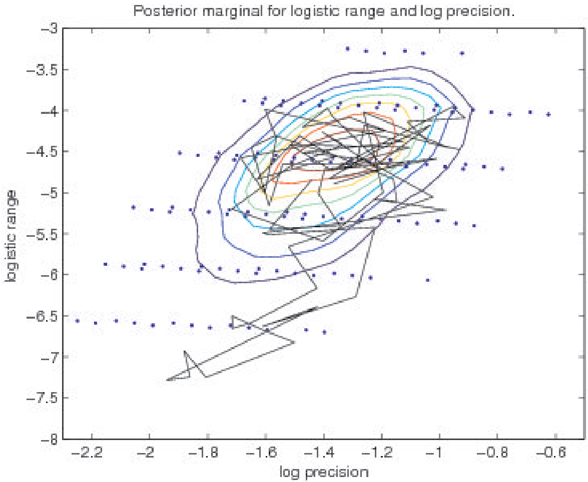
Markov Chain Monte Carlo, MCMC

- Run a Markov chain to get samples from $\pi(x, \theta|y)$
- Can find posterior for any parameter(s) / variable(s) or function of variables
- But 1: Need many iterations / takes a long time.
- But 2: Burn-in and mixing problems, hard to detect.

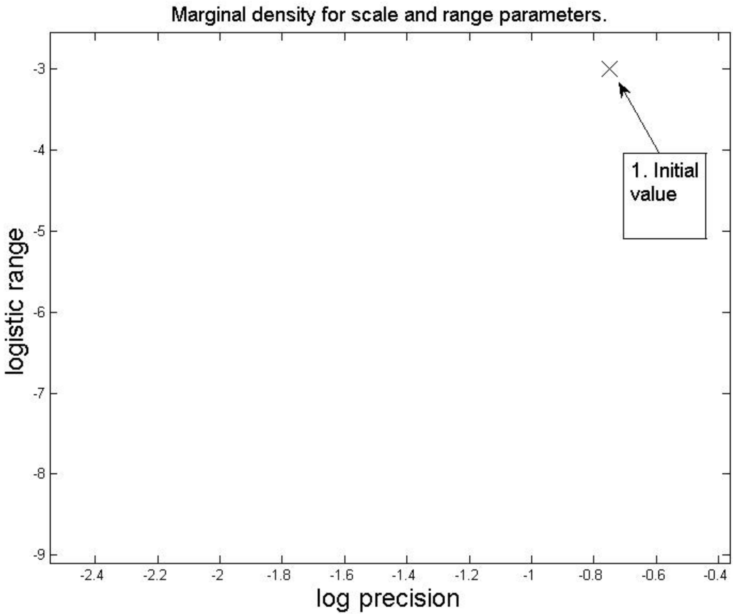
Integrated Nested Laplace Approximations, INLA

- Non-sampling based numerical method.
- Fast(er) $\Rightarrow$ enables simulation studies
- But 1: For *latent Gaussian Markov Random Field (GMRF) models* with max. 15-20 non-Gaussian hyper-parameters.
- But 2: Posteriors for functions of variables and jointly more tricky.

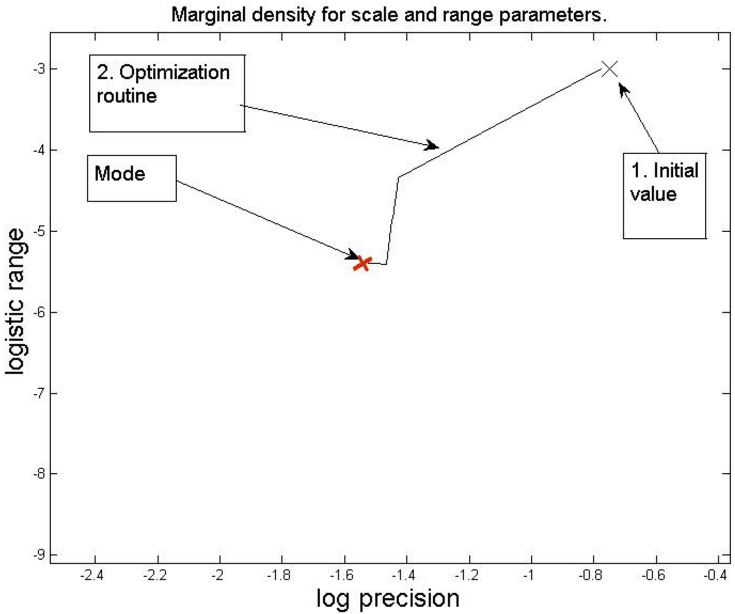
MCMC vs INLA, 2D example



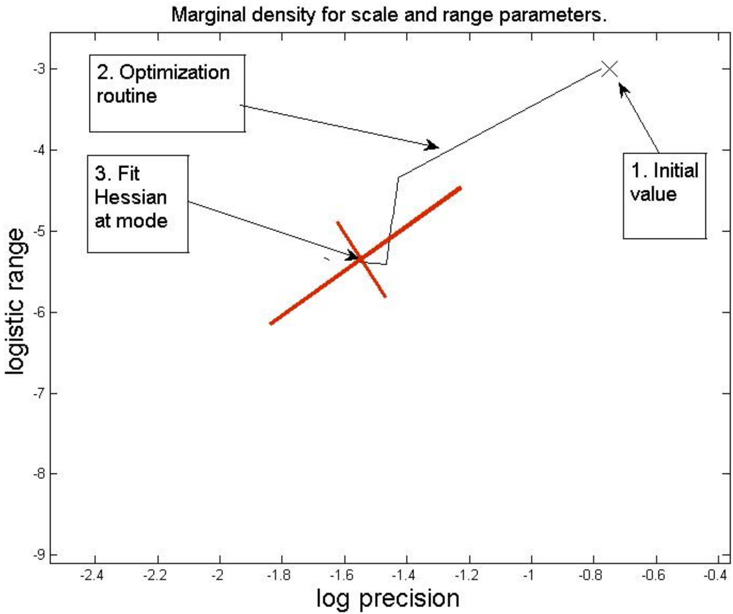
MCMC vs INLA, 2D example



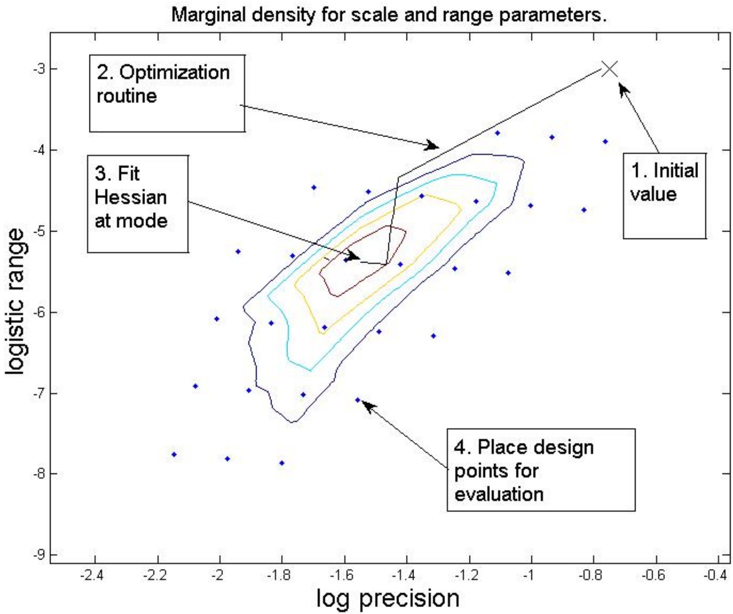
MCMC vs INLA, 2D example



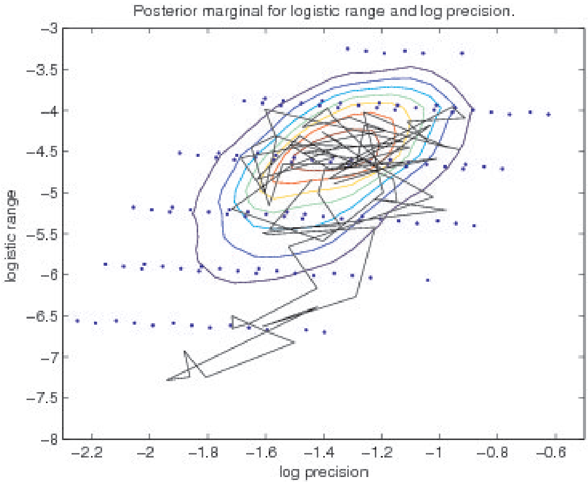
MCMC vs INLA, 2D example



MCMC vs INLA, 2D example



MCMC vs INLA, 2D example



Integrated nested Laplace approximation

Hierarchical model:

1. Data: $y \sim \pi(z|y,\theta) = \prod_{i=1}^{n} \pi(z_i|y_i)$
2. Latent field: $y \sim \pi(x|\theta)$, $y \sim N(\mu, Q^{-1})$, GMRF.
3. Hyper-parameters: $\theta \sim \pi(\theta)$, only few non-Gaussian.

Want to find $\pi(\theta|z)$, $\pi(y_i|z)$,....

Integrated nested Laplace approximation

Hierarchical model:

1. Data: $y \sim \pi(z|y,\theta) = \prod_{i=1}^{n} \pi(z_i|y_i)$
2. Latent field: $y \sim \pi(x|\theta)$, $y \sim N(\mu, Q^{-1})$, GMRF.
3. Hyper-parameters: $\theta \sim \pi(\theta)$, only few non-Gaussian.

Want to find $\pi(\theta|z)$, $\pi(y_i|z)$,....

Ideas

1. Laplace approximation
2. Numerical integration

Integrated nested Laplace approximation

Hierarchical model:

1. Data: $y \sim \pi(z|y,\theta) = \prod_{i=1}^{n} \pi(z_i|y_i)$
2. Latent field: $y \sim \pi(x|\theta)$, $y \sim N(\mu, Q^{-1})$, GMRF.
3. Hyper-parameters: $\theta \sim \pi(\theta)$, only few non-Gaussian.

Want to find $\pi(\theta|z)$, $\pi(y_i|z)$,....

Ideas

1. Laplace approximation

$$\pi(\theta|z) = \frac{\pi(\theta, y|z)}{\pi(y|z,\theta)} \approx \frac{\pi(\theta, y|z)}{\hat{\pi}(y|z,\theta)}$$

   where $\hat{\pi}(y|z,\theta)$ Gaussian approximation to $\pi(y|z,\theta)$

2. Numerical integration

Integrated nested Laplace approximation

Hierarchical model:

1. Data: $y \sim \pi(z|y,\theta) = \prod_{i=1}^{n} \pi(z_i|y_i)$
2. Latent field: $y \sim \pi(x|\theta)$, $y \sim N(\mu, Q^{-1})$, GMRF.
3. Hyper-parameters: $\theta \sim \pi(\theta)$, only few non-Gaussian.

Want to find $\pi(\theta|z)$, $\pi(y_i|z)$,....

Ideas

1. Laplace approximation

$$\pi(\theta|z) = \frac{\pi(\theta, y|z)}{\pi(y|z,\theta)} \approx \frac{\pi(\theta, y|z)}{\hat{\pi}(y|z,\theta)}$$

   where $\hat{\pi}(y|z,\theta)$ Gaussian approximation to $\pi(y|z,\theta)$
   $\pi(z_i|y_i,\theta)$ Gaussian $\Rightarrow$ $\pi(t|z,\theta)$ Gaussian

2. Numerical integration

Integrated nested Laplace approximation

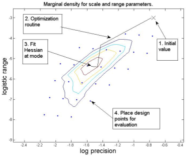
Hierarchical model:

1. Data: $y \sim \pi(z|y,\theta) = \prod_{i=1}^{n} \pi(z_i|y_i)$
2. Latent field: $y \sim \pi(x|\theta)$, $y \sim N(\mu, Q^{-1})$, GMRF.
3. Hyper-parameters: $\theta \sim \pi(\theta)$, only few non-Gaussian.

Want to find $\pi(\theta|z)$, $\pi(y_i|z)$,....

Ideas

1. Laplace approximation



2. Numerical integration

INLA-Integrated Laplace

Want to find $\pi(\theta|z)$, $\pi(y_i|z)$,....

Want to find $\pi(\theta|z)$, $\pi(y_i|z)$,....

$$\pi(y_i|z) = \int \pi(y_i|\theta, z)\ \pi(\theta|z)\ d\theta$$

INLA-Integrated Laplace

Want to find $\pi(\theta|z)$, $\pi(y_i|z)$,....

$$\pi(y_i|z) = \int \pi(y_i|\theta, z)\ \pi(\theta|z)\ d\theta$$

- $\pi(\theta|z) \Rightarrow$ Laplace approximations

INLA-Integrated Laplace

Want to find $\pi(\theta|z)$, $\pi(y_i|z)$,....

$$\pi(y_i|z) = \int \pi(y_i|\theta, z)\ \pi(\theta|z)\ d\theta$$

- $\pi(\theta|z) \Rightarrow$ Laplace approximations
- $\pi(y_i|\theta, z) \Rightarrow$ Laplace approximations
  $\pi(y_i|\theta, z) \approx \frac{\pi(y_i, y_{-i}|\theta, y)}{\hat{\pi}(y_{-i}|\theta, y)}$

INLA-Integrated Laplace

Want to find $\pi(\theta|z)$, $\pi(y_i|z)$,....

$$\pi(y_i|z) = \int \pi(y_i|\theta, z)\ \pi(\theta|z)\ d\theta$$

- $\pi(\theta|z) \Rightarrow$ Laplace approximations
- $\pi(y_i|\theta, z) \Rightarrow$ Laplace approximations
  $\pi(y_i|\theta, z) \approx \frac{\pi(y_i, y_{-i}|\theta, y)}{\hat{\pi}(y_{-i}|\theta, y)}$
- $\int\ \mathrm{d}\theta \Rightarrow$ Numerical integration

INLA or MCMC?

- Several papers that compare MCMC and INLA (Holand et al 2013 for animal models). Almost without exceptions, for a given computation time INLA gives most accurate estimates.
- NOTE: INLA sometimes works poorly for binary data. (Gaussian in mode is a poor approximation to $\pi(x|\theta, y)$).

Deviance information criterion (DIC)

- Measure for model fit, penalized for number of parameters.
- Introduced by Spiegelhalter et al. (2002)
- A generalization of the AIC (Akaike information criterion) and BIC (Bayesian information criterion) for hierarchical models.
- Computable in MCMC and INLA

DIC

Let z be data, and ϕ parameters ($\phi = (a, \beta, \sigma_\epsilon^2, \sigma_a^2)$. Deviance: $D = log(\pi(z|\phi))$

$$DIC = p_D + \bar{D}$$

- Mean deviance: Expected deviance: $\bar{D} = E_{\phi|z}(D(z|\phi))$
- Degrees of freedom: original $p_D = \bar{D} - D(\phi_{mean})$. In INLA $p_D = \bar{D} - D(\phi_{median})$.

DIC: The smaller the better model.