Approximate Inference for Hierarchical Gaussian Markov Random Field Models

Sara Martino and Håvard Rue
Norwegian University of Science and Technology, Trondheim, Norway

Abstract: Approximate deterministic inference is produced for posterior marginals for hierarchical GMRF models without doing any simulation at all.

Theory

Hierarchical GMRF

A hierarchical GMRF is a model characterised by several stages of observables and parameters, typically:

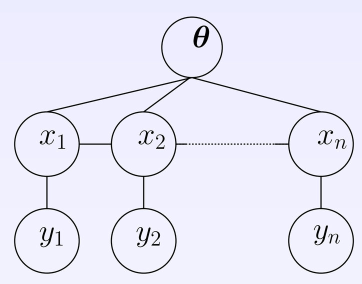


Fig 1: graph

- **Distribution for the data**

$$\pi(\boldsymbol{y} \mid x) = \prod \pi(y_i \mid x_i)$$

- **Distribution for latent variables (GMRF)**

$$\pi(\boldsymbol{x} \mid \boldsymbol{\theta}) \propto \left(-\frac{1}{2}\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x}\right)$$

where $\boldsymbol{Q}$ is typically a sparse matrix

- **Distribution for hyperparameters** $\pi(\boldsymbol{\theta})$

The posterior is then

$$\pi(\boldsymbol{x}, \boldsymbol{\theta} \mid \boldsymbol{y}) \propto \exp\left(-\frac{1}{2}\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} + \sum \log \pi(y_i|x_i)\right) \pi(\boldsymbol{\theta}) \qquad (1)$$

Goal of the analysis: posteriors for the latent field

$$\pi(x_i \mid \boldsymbol{y}) = \int_{\boldsymbol{\theta}} \int_{\boldsymbol{x}_{-i}} \pi(\boldsymbol{x}, \boldsymbol{\theta} \mid \boldsymbol{y})\, d\boldsymbol{x}_{-i}\, d\boldsymbol{\theta}$$

and for hyperparameters $\pi(\boldsymbol{\theta} \mid \boldsymbol{y})$.

Analytical integration $\pi(\boldsymbol{x}, \boldsymbol{\theta} \mid \boldsymbol{y})$ is usually not possible, MCMC it the common answer to inferential problems, **BUT** MCMC can take unreasonable long time to obtain precise estimate.

We propose a **deterministic approximate** alternative, computed almost instant compared to the time any MCMC needs to reach a decent accuracy [1].

GMRF-approximation

Build a Gaussian approximation to the full conditional for $\boldsymbol{x}$

$$\tilde{\pi}_G(\boldsymbol{x} \mid \boldsymbol{\theta}, \boldsymbol{y}) \propto \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T (\boldsymbol{Q} + \mathrm{diag}(\boldsymbol{c}))(\boldsymbol{x} - \boldsymbol{\mu})\right) \qquad (2)$$

by matching the mode and curvature at the mode.

- **Approximate posterior for $\boldsymbol{\theta}$**

$$\pi(\boldsymbol{\theta} \mid \boldsymbol{y}) = \frac{\pi(\boldsymbol{x}, \boldsymbol{\theta} \mid \boldsymbol{y})}{\tilde{\pi}_G(\boldsymbol{x} \mid \boldsymbol{\theta}, \boldsymbol{y})}.$$

rhs evaluated at the modal value of $\boldsymbol{x}$ for each $\boldsymbol{\theta}$. This is the Laplace approximation.

- **Approximate posterior for x_i**

$$\hat{\tilde{\pi}}(x_i \mid \boldsymbol{y}) = \sum_j \tilde{\pi}_G(x_i \mid \boldsymbol{\theta}_j, \boldsymbol{y})\, \tilde{\pi}(\boldsymbol{\theta}_j \mid \boldsymbol{y})\, \Delta_j$$

All computations are feasible and fast due to the sparseness of the precision matrix $\boldsymbol{Q}$

Examples

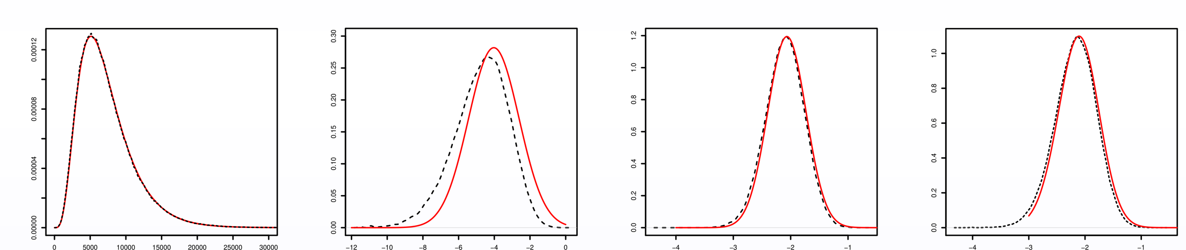
The plots show the comparison between the GMRF approximation (**red**) and the density estimation obtained from a very long MCMC run (**black**)

Time series

Number of occurrences of rainfall over 1mm in the Tokyo area for each day during 1983-84

The model:

$$\begin{aligned} y_i \mid x_i &\sim \mathrm{Bin}\{n_i, p_i\};\ i = 1, \ldots, 366 \\ p_i &= \exp(x_i)/(1 + \exp(x_i)); \\ \boldsymbol{x} \mid \kappa &\quad \text{circular RW2 with precision } \kappa \boldsymbol{Q} \\ \kappa &\sim \mathrm{Gamma}(a, b) \end{aligned}$$



(a) $\pi(\kappa \mid \boldsymbol{y})$ (b) $\pi(x_i \mid \boldsymbol{y}, \kappa = 10)$ (c) $\pi(x_i \mid \boldsymbol{y}, \kappa = 1000)$ (d) $\pi(x_i \mid \boldsymbol{y})$

Fig 2: posteriors for the Tokyo example

Model for disease mapping [2]

The model:

$$y_i \mid \eta_i \sim \mathrm{Pois}(E_i \exp(\eta_i)), : i = 1, \ldots, n$$

where E_i is the expected number of cases in region i, and $\eta_i = u_i + v_i$ is the relative risk.
Spatial component

$$\pi(\boldsymbol{u} \mid \kappa_u) \propto \kappa_u^{(n-1)/2} \exp\left(-\frac{\kappa_u}{2} \sum_{i \sim j} (u_i - u_j)^2\right)$$

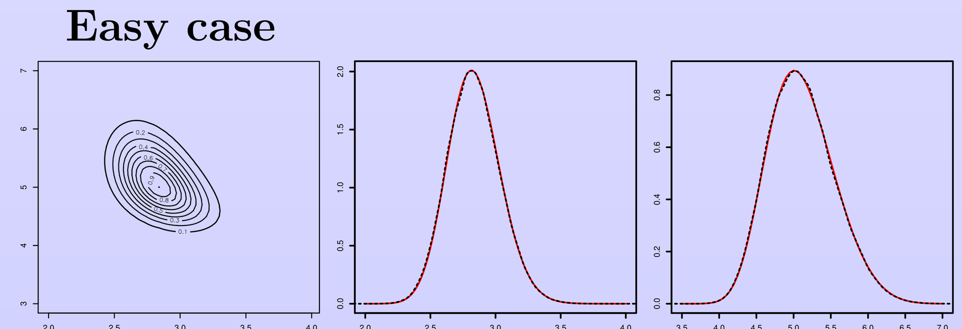
Random component $\boldsymbol{v} \sim N(0, \kappa_v^{-1} \boldsymbol{I})$
Gamma priors for the hyperparameters (κ_u, κ_v)

The results are presented for two data sets

- **Easy case** Oral cavity cancer mortality for males in Germany (1986-1990). It's considered "easy" since the observed counts are high
- **Hard case** Insulin dependent Diabetes Mellitus in Sardinia. This is a sparse data set, which makes the Poisson distribution different from a Gaussian.

Approximating $\pi(\theta \mid \boldsymbol{y})$

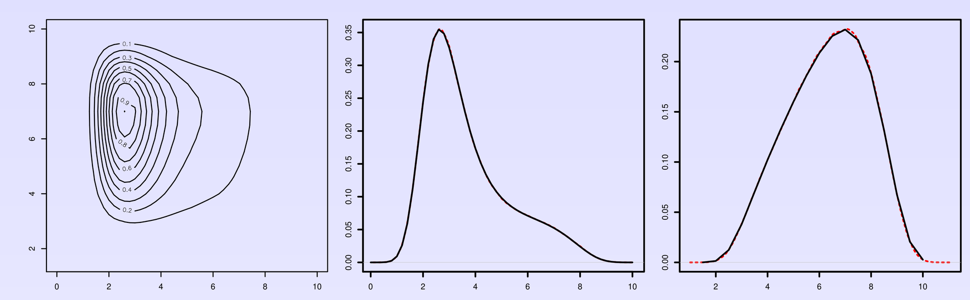
Easy case



a) $\pi(\kappa_v, \kappa_u \mid \boldsymbol{y})$ b) $\pi(\kappa_u \mid \boldsymbol{y})$ c) $\pi(\kappa_u \mid \boldsymbol{y})$

Fig 3: posteriors for the precision parameters

Hard case

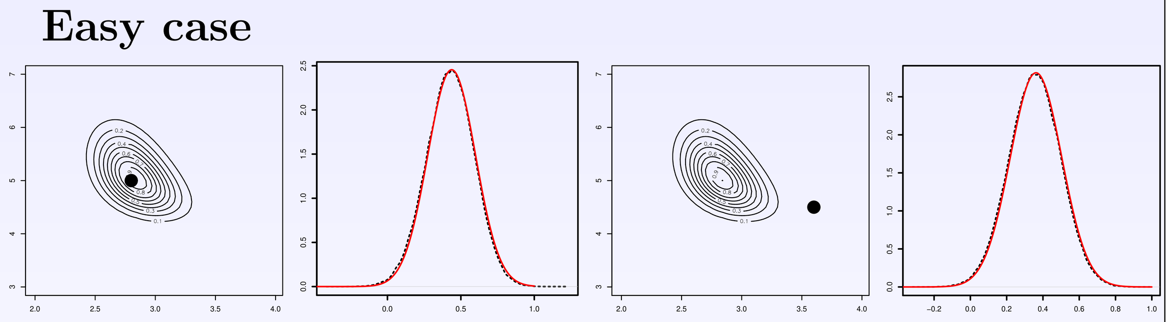


(a) $\pi(\kappa_v, \kappa_u \mid \boldsymbol{y})$ (b) $\pi(\kappa_u \mid \boldsymbol{y})$ (c) $\pi(\kappa_u \mid \boldsymbol{y})$

Fig 4: posteriors for the precision parameters

Approximating $\pi(x_i \mid \theta, \boldsymbol{y})$

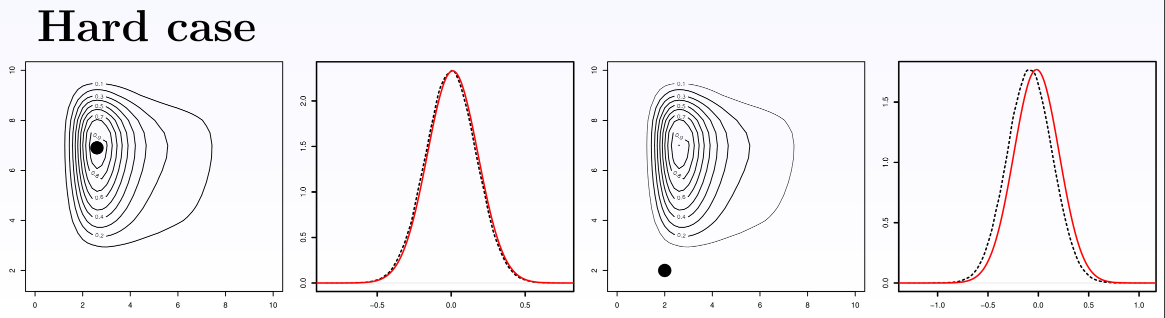
Easy case



(a) $\pi(u_i|\boldsymbol{\theta}, \boldsymbol{y})$ (b) $\pi(u_i|\boldsymbol{\theta}, \boldsymbol{y})$

Fig 5: posteriors spatial component for fixed hyperparameter

Hard case



(a) $\pi(u_i|\boldsymbol{\theta}, \boldsymbol{y})$ (b) $\pi(u_i|\boldsymbol{\theta}, \boldsymbol{y})$

Fig 6: posteriors spatial component for fixed hyperparameter

Approximating $\pi(x_i \mid \boldsymbol{y})$

Easy case **Hard case**

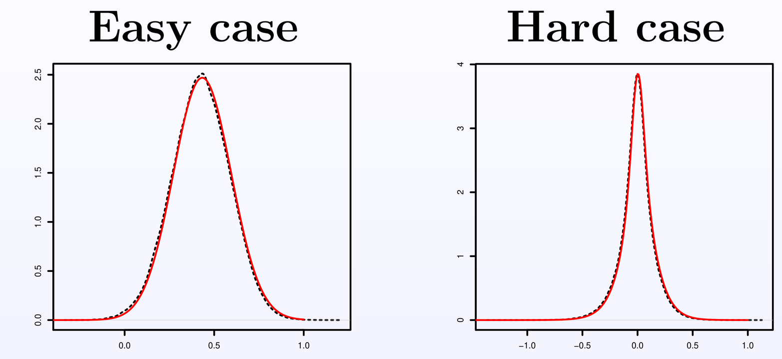


Fig 7: posteriors for one node in the spatial component

New improved approximation

At the cost of some more computations it is possible to go beyond the Gaussian approximation, thus getting more precise estimates in those cases where the Gaussian approximation is not enough, for example those in Figure 2(b) and Figure 6(b). The figure below shows the improved approximation (**green**) compared to the Gaussian approximation (**red**) and the MCMC based estimate (**black**).

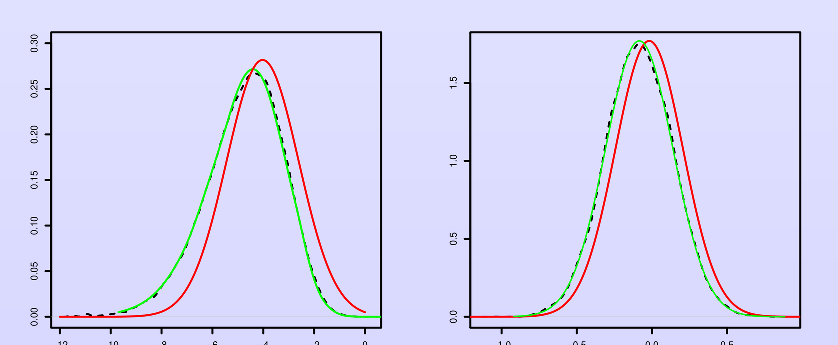


Fig 8: improved approximation