# Gausian Markov Random Field / SPDE method

Joaquin Cavieres G.

University of Valparaíso
Ph.D (c) in Statistics

October 27, 2020

# Contents

# Introduction

# Introduction

Key concepts in Spatial statistics:

- The spatial domain is $D$, or if we can spatio-temporal observations $D \times T, T \subset \mathscr{R}$
- We call 'spatial random field' $u(\boldsymbol{s}), \boldsymbol{s} \in D$, or $(\boldsymbol{s}, t) \in D \times T$ for spatiotemporal observations.
- The observations $y_i$ can be modelling setting: $y_i = u(\boldsymbol{s}_i) + \epsilon_i$. Here $u(\cdot)$ is a structured random effect.

# Introduction

Gaussian distribution for $\boldsymbol{u} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$:

- Density function:

$$d(\boldsymbol{u}) = \frac{1}{(2\pi)^{N/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{u}-\boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{u}-\boldsymbol{\mu})\right) \tag{1}$$

- Covariance:

$$\left[\begin{array}{c} u_1 \\ u_2 \end{array}\right] \sim \mathcal{N}\left(\left[\begin{array}{c} \boldsymbol{\mu}_{u_1} \\ \boldsymbol{\mu}_{u_2} \end{array}\right], \left[\begin{array}{cc} \boldsymbol{\Sigma}_{u_1,u_1} & \boldsymbol{\Sigma}_{u_1,u_2} \\ \boldsymbol{\Sigma}_{u_2,u_1} & \boldsymbol{\Sigma}_{u_2,u_2} \end{array}\right]\right) \tag{2}$$

- Conditional distribution for each random effect $u$:

$$u_1 \mid u_2 \sim \mathcal{N}\left(\mu_{\boldsymbol{u_1}} + \boldsymbol{\Sigma}_{\boldsymbol{u_1},\boldsymbol{u_2}} \boldsymbol{\Sigma}^{-1}_{\boldsymbol{u_2},\boldsymbol{u_2}} \left(u_2 - \mu_{\boldsymbol{u_2}}\right), \boldsymbol{\Sigma}_{\boldsymbol{u_1},\boldsymbol{u_1}} - \boldsymbol{\Sigma}_{\boldsymbol{u_1},\boldsymbol{u_2}} \boldsymbol{\Sigma}^{-1}_{\boldsymbol{u_2},\boldsymbol{u_2}} \boldsymbol{\Sigma}_{\boldsymbol{u_2},\boldsymbol{u_1}}\right) \tag{3}$$

What is the problem? $\Sigma$ is a large dense matrix.

# The precision matrix

# The precision matrix

The precision matrix is the inverse of the covariance matrix: $\Sigma^{-1}$, so:

$$\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1} \tag{4}$$

- Density function:

$$d(\boldsymbol{u}) = \frac{|\boldsymbol{Q}|^{1/2}}{(2\pi)^{N/2}} \exp\left(-\frac{1}{2}(\boldsymbol{u}-\boldsymbol{\mu})^T \boldsymbol{Q}(\boldsymbol{u}-\boldsymbol{\mu})\right) \tag{5}$$

- Precision Matrix:

$$\left[\begin{array}{c} u_1 \\ u_2 \end{array}\right] \sim \mathcal{N}\left(\left[\begin{array}{c} \boldsymbol{\mu}_{u_1} \\ \boldsymbol{\mu}_{u_2} \end{array}\right], \left[\begin{array}{cc} \boldsymbol{Q}_{u_1,u_1} & \boldsymbol{Q}_{u_1,u_2} \\ \boldsymbol{Q}_{u_2,u_1} & \boldsymbol{Q}_{u_2,u_2} \end{array}\right]^{-1}\right) \tag{6}$$

- Conditional distribution for each random effect $u$

$$u_1 | u_2 \sim \mathcal{N}\left(\mu_{u_1} + \boldsymbol{Q}_{u_1,u_1}^{-1} \boldsymbol{Q}_{u_1,u_2}\left(u_2 - \mu_{u_2}\right), \boldsymbol{Q}_{u_1,u_1}^{-1}\right) \tag{7}$$

# The precision matrix

How works $\boldsymbol{Q}$?

- Let $W$ be a matrix with elements $W_{ij} = -Q_{ij}/Q_{ii}$ for $j \neq i$ and $W_{ii} = 0 \ \forall \ i$
- The conditional distribution for $u_i|\boldsymbol{u}$ is:

$$(u_i|u_j, j \neq i) \sim \mathcal{N}\left(\mu_i + \sum_{j \neq i} W_{ij}\left(u_j - \mu_j\right), 1/Q_{i,i}\right) \tag{8}$$

- $i$( row number ) of $W$ are the weights that influencing $u$ in the conditional distribution for $u_i$ and the conditional variance is $1/Q_{i,i}$

# Gaussian Markov Random Field `GMRF`

# Gaussian Markov Random Field GMRF

A gaussian random field (GRF) is a gaussian markov random field (GMRF) if:

- $Q_{ij} = 0$
- If $Q_{ij} \neq 0$ and $i \neq j$, $u_j$ is called neighbour of $u_i$
- The set of neighbours to a $u_i$ is denoted by:

$$\mathbb{N}_i = j : Q_{ij} \neq 0, j \neq i$$

- The relation in the neighbourhood is symemetric: $j \in \mathbb{N}_i \Leftrightarrow i \in \mathbb{N}_j$ since $\boldsymbol{Q}$ is symmetric too.
- The full conditional distribution for $u_i$ depend only on the rest of random effect in the neighbourhood $\mathbb{N}_i$, $u_j : j \in \mathbb{N}_i$. This is the Markov property for a spatial random field

# Gaussian Markov Random Field GMRF

Generally in geostatistical situations, the prior and the posterior field have sparse precisions. What means that?

- If $\boldsymbol{u}$ is a Markov field the presicion matrix $\boldsymbol{Q}$ is sparse
- If we applying a Cholesky factorization $\boldsymbol{Q} = \boldsymbol{R}^T\boldsymbol{R}$ it will be sparse if $\boldsymbol{Q}$ is sparse.

A $\mathrm{GMRF}$ it would be computationally feasible if we replace operations what use $\Sigma$ with operations usgin $\boldsymbol{Q}$ or $\boldsymbol{R}$

# Gaussian Markov Random Field GMRF

The idea is use the GRF instead of full covarian useful, sparse and positive definite $\boldsymbol{Q}$ -matrices

Exist general methods for estimate presicion values for a given neighbourhood structures, but are inefficient for large neighbourhood. Another method is the proposal by Rue and Tjelmeland (2002) but also is computationally demanding numerical optimisation.

# Stochastic partial differential equations (SPDE)

# Matérn family (covariance)

We comeback to deffinition of gaussian random field but now formally:

Let $s$ a location in a particular area $D$ and $u(s)$ is the random effect (spatial) at that location. $\boldsymbol{u}(s)$ is a stochastic process with $\boldsymbol{s} \in \boldsymbol{D}$ and $\boldsymbol{D} \subset \mathbb{R}^d$ is the spatial domain where are measured the observations. $u(s_i)$ is a realization de $u(s)$ where $i = 1, \ldots, n$ locations. We asumme that $u(s)$ has a multivariate Gaussian distribution (GRF), continuous over the space indexed by $\boldsymbol{s}$ and defined by the mean and the covariance ([Cressie, 1993])

# Matérn family (covariance)

## Continuous domain spatial

- Matérn covariance familiy on $\mathbf{s} \in \mathbb{R}^d$ :

$$
\begin{aligned}
r_M(\boldsymbol{s}_1, \boldsymbol{s}_2) &= C\left(u\left(\boldsymbol{s}_1, u\left(\boldsymbol{s}_2\right)\right)\right) \\
&= \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\kappa \left\|\boldsymbol{s}_2 - \boldsymbol{s}_1\right\|\right)^{\nu} K_{\nu}\left(\kappa \left\|\boldsymbol{s}_2 - \boldsymbol{s}_1\right\|\right)
\end{aligned}
$$

with scale $\kappa > 0$, shape/smoothness $\nu > 0$ and $K_\nu$ a modified Bessel function.

Fields with Matérn covariances are solutions of the SPDE method (Whittle, 1954,1963) base on the Laplacian $\triangle = \nabla^T \nabla$ :

$$
\left(\kappa^2 - \Delta\right)^{\alpha/2} u(\boldsymbol{s}) = \mathscr{W}(\boldsymbol{s})
$$

where $\mathscr{W}(\boldsymbol{s})$ is the spatial white noise, $\alpha = \nu + d/2$ and

$$
\sigma^2 = \frac{\Gamma(\nu)}{\Gamma(\alpha)(4\pi)^{d/2}\kappa^{2\nu}\tau^2}
$$

# SPDE

## Hilbert space approximation

A finite Hilbert space uses a set of $N$ basis functions $\{\psi_k\}$ and weights $\{w_k\}$ for that:

$$\boldsymbol{u}(\boldsymbol{s}) = \sum_{k=1}^{n} \psi_k(\boldsymbol{s}) w_k$$

where $\psi(\cdot)$ are deterministic basis functions and $\{u_1, \ldots, u_n\}$ is a vector of weights that is chosen so that the distribution of the functions $\boldsymbol{u}(\boldsymbol{s})$ approximates the distribution of solutions to the SPDE on the domain

# SPDE

## Construction of the $Q = \Sigma^{-1}$

To obtain a Markov structure we using piecewise polynomial basis functions with compact support (essentially it's a Finite Element method)

For a domain two dimensional we use picewise linear basis functions defined by a triangulation of the domain of interest

$$\boldsymbol{Q} = \tau^2 \left(\kappa^4 \boldsymbol{C} + 2\kappa^2 \boldsymbol{G}_1 + \boldsymbol{G}_2\right)$$

where by default $\alpha = 2$ so that the elements of $\boldsymbol{Q}$ have explicit expressions as functions of $\kappa$ and $\tau$. Assigning the Gaussian distribution $\boldsymbol{u} \sim \mathscr{N}(\mathbf{0}, \boldsymbol{Q})$ now this is approximative solutions to the SPDE (in a stochastically weak sense) [Lindgren et al., 2015]

More references: [Rue and Held, 2005], [Lindgren et al., 2011], [Blangiardo and Cameletti, 2015],

# SPDE

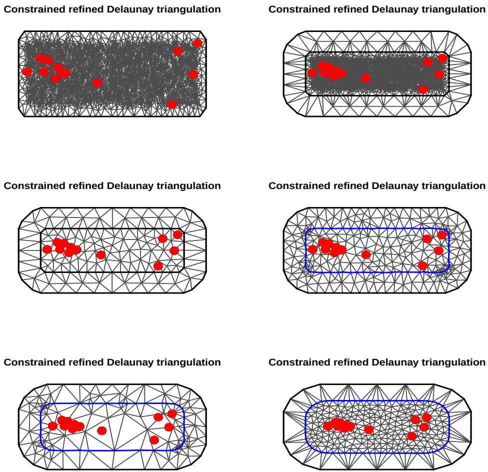


Figure: Differents SPDE/GMRF triangulations

# Hierarchical Models

Hierarchical Bayesian Model in INLA ([Rue et al., 2009], [Bakka et al., 2018])

- We need three principal components:
  - Parameters $\Rightarrow \boldsymbol{\theta}$
  - Spatial field ("latent field" or "gaussian field" or "gaussian random field") $\Rightarrow \boldsymbol{u}$
  - Data $\Rightarrow \boldsymbol{y}$
- Priors for the parameters: $\pi(\boldsymbol{\theta})$
- The spatial field is generally a Gaussian process (GP) or Gaussian Markov Random Field (GMRF) and the density function is $\pi(\boldsymbol{u} \mid \boldsymbol{\theta})$ (conditionally on the parameters $\boldsymbol{\theta}$ )
- The data are independent for a set of locations with density function: $\pi(\boldsymbol{y} \mid \boldsymbol{u}, \boldsymbol{\theta})$
- We want estimate $\pi(\boldsymbol{\theta} \mid \boldsymbol{y})$ and $\pi(\boldsymbol{u} \mid \boldsymbol{y})$ (without MCMC method)

Hierarchical model (Bayesian)

$$
\begin{aligned}
\boldsymbol{\theta} &\sim \boldsymbol{\theta} && \text{Hyperparameters} && (9)\\
\boldsymbol{u} \mid \boldsymbol{\theta} &\sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q}(\boldsymbol{\theta})^{-1}) && \text{Latent gaussian field} && (10)\\
\eta_i &= \sum_j c_{ij} x_j && && \\
y_i \mid \boldsymbol{u}, \boldsymbol{\theta} &\sim \prod_i \pi(y_i \mid \eta_i, \boldsymbol{\theta}) && \text{Observations} && (11)
\end{aligned}
$$

where $\boldsymbol{Q}(\theta)$ is the presicion matrix, $\boldsymbol{u}$ is the latent gaussian field and $\eta(\boldsymbol{u}) =$ A $\boldsymbol{u}$, where the matrix $\boldsymbol{A}$ maps the latent variable vector $\boldsymbol{u}$ to the predictors $\eta_i = \eta_i(\boldsymbol{u})$ associated to the observations $y_i$

# References I


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

Blangiardo, M. and Cameletti, M. (2015).
*Spatial and spatio-temporal Bayesian models with R-INLA*.
John Wiley & Sons.

Cressie, N. A. (1993).
Spatial prediction and kriging.
*Statistics for Spatial Data (Cressie NAC, ed). New York: John Wiley & Sons*, pages 105–209.

Lindgren, F., Rue, H., et al. (2015).
Bayesian spatial modelling with r-inla.
*Journal of Statistical Software*, 63(19):1–25.

# References II


Lindgren, F., Rue, H., and Lindström, J. (2011).
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(4):423–498.

Rue, H. and Held, L. (2005).
*Gaussian Markov random fields: theory and applications*.
CRC press.

Rue, H., Martino, S., and Chopin, N. (2009).
Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations.
*Journal of the royal statistical society: Series b (statistical methodology)*, 71(2):319–392.