An introduction to SPDE

Yawen Guan

Department of Statistics
Colorado State University

August 31, 2023

Overview

- GP models the covariance matrix. A dense large matrix hinders computation.
- GMRF models the precision matrix. It has sparse matrix allowing faster computation.
- GP is flexible. It can be used to model geostatistical or areal data, while GMRF only works for areal data (lattice or county).
- SPDE approximation creates a nice bridge between GP and GMRF.

What is SPDE

SPDE stands for stochastic partial differential equation, that is

$$\text{PDE} + \text{random noise}$$

- PDEs are deterministic functions. They are often used by mathematicians and engineers to describe natural phenomena.
- It is difficult to use PDEs to model every aspect of a phenomenon. Therefore, it is useful to add random noise to PDEs.

Link between SPDE and GP (Lindgren et al. 2011)

- The solution of a particular SPDE follows a GP with a specific covariance function

$$\mathcal{L}w(\mathbf{s}) = \mathcal{W}(\mathbf{s})$$

  - $w(\mathbf{s})$ represents a spatial process in a continuous spatial domain, and is called the solution of the SPDE.
  - $\mathcal{W}(\mathbf{s})$ is a Gaussian white noise process.
  - $\mathcal{L}$ is a differential operator.
  - An example of $\mathcal{L}$ is the 1st derivatives of $w(\mathbf{s})$ with respect to space. For $\mathcal{R}^2$, $\mathbf{s} = (s_x, s_y)'$, 1st derivatives is $\frac{\partial w(\mathbf{S})}{\partial s_x}$ and $\frac{\partial w(\mathbf{S})}{\partial s_y}$.
  - Another example of $\mathcal{L}$ is the Laplace operator

$$\triangle = \frac{\partial^2 w(\mathbf{s})}{\partial x^2} + \frac{\partial^2 w(\mathbf{s})}{\partial y^2}$$

  - The choice of $\mathcal{L}$ determines the covariance structure of the solution $w(\mathbf{s})$.

Matérn covariance family

A particular SPDE $<=>$ GP with a specific covariance function

- Only a few correspondences between the form of SPDE and GP with a given covariance function are known.
- Remember the Matérn covariance function, which has the form

$$C(\mathbf{s}_i, \mathbf{s}_j; \boldsymbol{\theta}) = \frac{\sigma^2}{\Gamma(\nu) 2^{(\nu-1)}} \left(\frac{||\mathbf{s}_i - \mathbf{s}_j||}{\phi} \right)^{\nu} K_{\nu} \left(\frac{||\mathbf{s}_i - \mathbf{s}_j||}{\phi} \right)$$

  - K_ν is the modified Bessel function of the second kind,
  - $\nu > 0$ is the smoothness parameter,
  - ϕ is the range parameter,
  - σ^2 is the marginal variance.

- We have seen two special cases of Matérn covariance functions:
  - Exponential covariance function is Matérn with $\nu = 0.5$,
  - Gaussian covariance function is Matérn with $\nu = \infty$,

One configuration

- We can overlay a grid in the continuous space. Let the vector $\mathbf{w}$ be a GMRF defined on a 2-dimensional grid with Gaussian full conditionals, with:

$$\begin{aligned} E(w_{ij}|\mathbf{w}_{-ij}) &= \frac{1}{a}(w_{i-1,j} + w_{i+1,j} + w_{i,j-1} + w_{i,j+1}) \\ \mathrm{Var}(w_{ij}|\mathbf{w}_{-ij}) &= 1/a \end{aligned}$$

  - ij is the 2-dimensional lattice index,
  - $\mathbf{w}_{-ij}$ contains the closest 4 vertices of (i,j).
  - $\mathbf{w}$ is a discrete version of the solution to the SPDE $(\kappa^2 - \triangle)w(\mathbf{s}) = \mathcal{W}(\mathbf{s})$, which corresponds to a GP with Matérn covariance function with $\nu = 1$

Advantages:

- $\mathbf{w}$ is a GMRF defined locally and has a sparse precision matrix.
- $\mathbf{w}$ is a discrete version of $w(\mathbf{s})$ so Q^{-1} is close to Σ.
- By selecting the form of SPDE, we are explicitly modeling a GP.

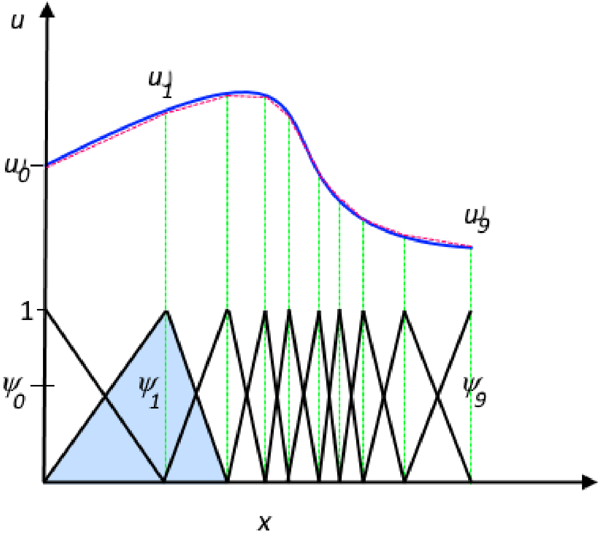
What if data is not on a grid?

- Lindgren et al. (2011) deploys a triangular mesh over the data location and defines basis functions over the mesh.
- This allows an approximation to the surface of $w(\mathbf{s})$ using a basis representation,

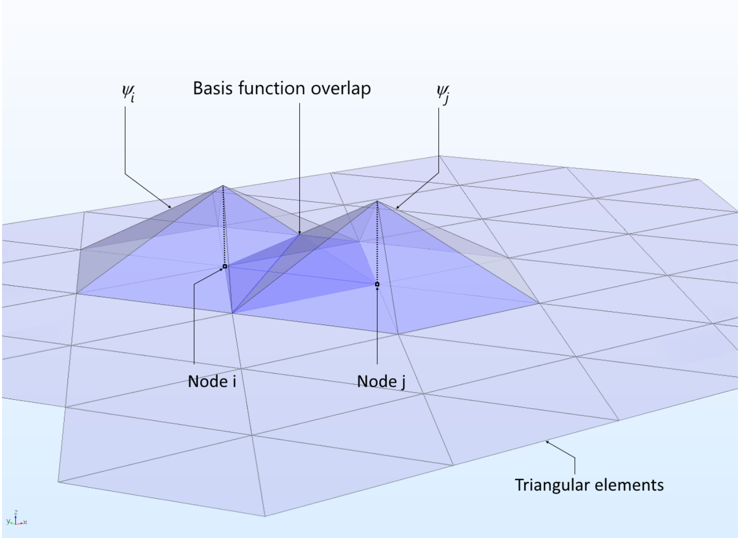
$$w(\mathbf{s}) = \sum_{j=1}^{m} \psi_j(\mathbf{s}) w_j$$

  - j = 1,. . . ,m are the vertices in the mesh.
  - ψ_j are piecewise linear basis functions with each ψ_j equals to 1 at a vertex j and 0 at all other vertices.
  - w_j is the j^{th} element of $\mathbf{w}$ when $\mathbf{w}$ is represented as a long vector.
  - Basis functions are evaluated at all data locations. The values are stored in a matrix A, with $A_{i,j} = \psi_j(\mathbf{s}_i)$.
  - A is the projection matrix that maps vertices to data locations.

Basis functions in 1D



Basis functions in 2D



Summarizing SPDE approach

The approach to model spatial data with SPDE is similar to the modeling done previously.

For example, we can use a SLM to model a Gaussian response $Y(\mathbf{s})$,

$$Y(\mathbf{s}) = X(\mathbf{s})^T \boldsymbol{\beta} + w(\mathbf{s}) + \epsilon(\mathbf{s}), \epsilon(\mathbf{s}) \overset{iid}{\sim} N(0, \tau^2), \tag{1}$$

$$w(\mathbf{s})|\boldsymbol{\theta} \sim GP(0, C(\mathbf{s}_i, \mathbf{s}_j; \boldsymbol{\theta})).$$

Instead of calculating the covariance matrix, we replace the Gaussian process with its GMRF version $\mathbf{w}$ on a mesh,

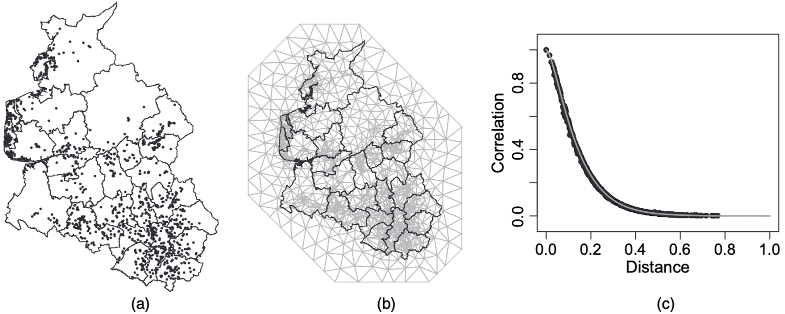
$$\mathbf{w} \sim MVN(\boldsymbol{\mu}_w, Q_w^{-1}).$$

$\mathbf{w}$ is defined on the vertices of a mesh which may not coincide with data locations. The basis function evaluation matrix A maps the vertices to the data locations. Hence (1) is effectively replaced by

$$\mathbf{y} = X\boldsymbol{\beta} + A\mathbf{w} + \boldsymbol{\epsilon}$$

Example from Lindgren et al. (2011)

Given the data locations (fig a), a triangular mesh is defined for the area of interest (fig b). Typically, the mesh has a buffer zone outside of data locations to reduce boundary effects. For this example, the empirical correlation function calculated from the GMRF (dots) matches the theoretical correlation function (line).



In summary

- To use the SPDE approach, our hierarchical model for Gaussian or non-Gaussian response stays the same. This is nice since we can keep the same models.
- In SPDE approach, the multivariate normal density function for the spatial random effects is evaluated differently. Instead of using the covariance matrix, it uses the precision matrix Q^{-1}. Computation utilizing the GMRF is all taken care of in the R-INLA packages.
- Cons: only a few correspondence between the form of SPDE and GP with a given covariance function are known. For example, we can find the SPDE corresponding to Matérn GP with smoothness equal to an integer value. But, for a more general covariance function, the SPDE approach may not be applicable as the corresponding SPDE is not known.