

$$\{y(\mathbf{s}_i); \mathbf{s}_i \in \mathcal{D} \subseteq \mathbb{R}^d\} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

♦ **Gaussian processes (GP)** and **Gaussian random fields (GRF)** are used interchangeably in two-dimensional space $d = 2$. The mean vector is $\boldsymbol{\mu} := (\mu(\mathbf{s}_i))$, and the covariance matrix is $\boldsymbol{\Sigma} := (C(\mathbf{s}_i, \mathbf{s}_j)) > 0$.

Weak Stationarity

- The mean function $\mu(\cdot)$ is constant.
- The covariance function $C(\cdot, \cdot)$ depends only on the relative position of two locations, not on their absolute positions in the input space. The process is **isotropic** if $C(\cdot, \cdot)$ depends only on the magnitude of distance, regardless of direction.

SpatialDE fits response **GP** models to the gene expression levels of each gene across a finite set of N spatial locations. The spatial component of $\boldsymbol{\Sigma} = \sigma_s^2 \cdot \mathbf{K} + \sigma_e^2 \cdot \mathbf{I}$ employs the squared exponential kernel function, which is also the **Matérn kernel** with $\nu \rightarrow +\infty$. However, it is challenged by computational complexity $\mathcal{O}(N^3)$.

$$K_{ij} = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu} \delta_{ij}}{l} \right)^\nu \cdot K_\nu \left(\frac{\sqrt{2\nu} \delta_{ij}}{l} \right) \text{ and denote } \kappa = \frac{\sqrt{2\nu}}{l}, \nu, l > 0,$$

where $K_\nu(\cdot)$ is the modified Bessel function of the second kind and $\delta_{ij} = \|\mathbf{s}_i - \mathbf{s}_j\|$.

nnSVG proposes four modifications:

1. It uses the exponential kernel function, which is also the Matérn kernel with $\nu = 1/2$;
2. It uses a gene-specific length scale l in the kernel;
3. It uses a scalable approximation of $\boldsymbol{\Omega}$ based on **nearest-neighbor Gaussian processes (NNGP)** in $\mathcal{O}(Nm^3)$, where m is the number of NNs, resulting in a sparse precision matrix $\tilde{\boldsymbol{\Omega}} = \tilde{\boldsymbol{\Sigma}}^{-1}$;
4. The degrees of freedom of the asymptotic chi-squared distribution are 2 instead of 1 for significance testing.

♦ **Markov random fields (MRF)** exhibit the Markov property and can implement the MCMC sampling or fast direct numerical methods, typically in $\mathcal{O}(N^{3/2})$.

$$f(y(\mathbf{s}_i) = y_i \mid y(\mathbf{s}_j) = y_j, j \neq i) = f(y(\mathbf{s}_i) = y_i \mid y(\mathbf{s}_j) = y_j, j \in \partial i) \text{ for } i = 1, \dots, N,$$

where ∂i are the indicators of a set of neighbors to location \mathbf{s}_i . Thus, $\Omega_{ij} = 0$ if $j \notin \{\partial i \cup i\}$.

♦ The **integrated nested Laplace approximation (INLA)** method has gained popularity in spatial data analysis primarily due to its innovative approach to approximating the posterior distributions of latent Gaussian models:

1. The **GMRF** provides a computationally efficient and flexible framework for modeling spatial dependencies;
2. The **stochastic partial differential equation (SPDE)** approach extends this flexibility to continuous spatial processes and enables INLA to handle non-stationary processes.

◆ Hierarchical GP Regression Model

$$\mathbf{y}(\mathbf{s}) := (y(\mathbf{s}_1), \dots, y(\mathbf{s}_N))^\top = \boldsymbol{\mu}(\mathbf{s}) + \mathbf{b}(\mathbf{s}) + \boldsymbol{\epsilon}(\mathbf{s}) \quad \text{for each gene}$$

$$\mathbf{b}(\mathbf{s}) \sim \mathcal{N}_N(\mathbf{0}, \sigma_s^2 \cdot \mathbf{K}) \quad \perp \quad \boldsymbol{\epsilon}(\mathbf{s}) \sim \mathcal{N}_N(\mathbf{0}, \sigma_\epsilon^2 \cdot \mathbf{I})$$

- Latent GP: $\mathcal{N}_N(\mathbf{y}(\mathbf{s}) \mid \boldsymbol{\mu}(\mathbf{s}) + \mathbf{b}(\mathbf{s}), \sigma_\epsilon^2 \cdot \mathbf{I}) \times \mathcal{N}_N(\mathbf{b}(\mathbf{s}) \mid \mathbf{0}, \sigma_s^2 \cdot \mathbf{K}) \times \pi(\sigma_s^2, l, \sigma_\epsilon^2)$
- Response GP: $\mathcal{N}_N(\mathbf{y}(\mathbf{s}) \mid \boldsymbol{\mu}(\mathbf{s}), \sigma_s^2 \cdot \mathbf{K} + \sigma_\epsilon^2 \cdot \mathbf{I}) \times \pi(\sigma_s^2, l, \sigma_\epsilon^2)$

← A latent spatial process capturing spatial dependence

◆ SPDE Representation of a Continuous GRF (consider $b(\mathbf{s})$ above)

A GRF $y(\mathbf{s})$ with the Matérn covariance $\boldsymbol{\Sigma} = \sigma_s^2(\nu, \kappa, d) \cdot \mathbf{K}$ is a solution to the SPDE:

$$(\kappa^2(\mathbf{s}) - \Delta)^{\frac{\alpha}{2}} (\gamma(\mathbf{s}) \cdot y(\mathbf{s})) = \mathcal{W}(\mathbf{s}), \quad \mathcal{W}(\mathbf{s}) \stackrel{\text{innovation}}{\text{i.i.d.}} \mathcal{N}(0, 1), \quad \alpha = \nu + \frac{d}{2} \in \mathbb{N}, \quad \Delta f(\mathbf{s}) := \nabla^2 f(\mathbf{s}) := \sum_{i=1}^d \frac{\partial^2}{\partial s_i^2} f(\mathbf{s})$$

[ν] - 1 times
mean square differentiable

κ , α , and γ determine the correlation range, smoothness, and marginal variance of $y(\mathbf{s})$.

$$0 < \alpha \leq 2, \quad \gamma \propto 1 / \sigma_s$$

The "R-INLA" package constructs a Matérn SPDE model with the spatial scale $\kappa = \sqrt{8\nu} / \rho$, where the spatial range is $\rho = 2l$.

The penalized complexity (PC) prior, $\pi(\sigma_s, \rho) = \frac{1}{2} d \lambda_\rho \rho^{-1-d/2} \exp \left\{ -\lambda_\rho \rho^{-d/2} \right\} \cdot \lambda_{\sigma_s} \exp \left\{ -\lambda_{\sigma_s} \sigma_s \right\}$

$$\mathbb{P}(\sigma_s > \sigma_0) = p_{\sigma_s} \quad \mathbb{P}(\rho < \rho_0) = p_\rho$$

◇ Non-stationarity is achieved when $\kappa^2(\mathbf{s})$ and/or $\gamma(\mathbf{s})$ are non-constant for $\mathbf{s} \in \mathcal{D} \subsetneq \mathbb{R}^d$.

$$\ln \kappa^2(\mathbf{s}) = \sum_i \beta_i^{(\kappa^2)} \cdot B_i^{(\kappa^2)}(\mathbf{s}) \quad \text{and} \quad \ln \gamma(\mathbf{s}) = \sum_i \beta_i^{(\gamma)} \cdot B_i^{(\gamma)}(\mathbf{s})$$

◆ As for a GMRF, the result $y(\mathbf{s}) = \sum_{q=1}^{n \approx N} w_q \cdot \psi_q(\mathbf{s})$, given $(w_1, \dots, w_n)^\top \sim \mathcal{N}_n(\mathbf{0}, \tilde{\boldsymbol{\Omega}}_w^{-1})$,
sparse

is a basis function representation

with piecewise linear basis functions $\psi_q(\mathbf{s})$, and Gaussian weights w_q with Markov dependences determined by a general triangulation of the bounded domain \mathcal{D} (finite element methods).