The Bayesian Hierarchical model and glossary

Zhe Sha

July 13, 2017

1 The Bayesian hierarchical model

In this section, we introduce the Bayesian hierarchical model for predicting the global GIA process.

We assume that the true GIA process is a real-valued spatial process continues on the sphere and denote it by $\boldsymbol{Y} : \mathbb{S}^2 \mapsto \mathbb{R}$. We use one of the GIA solution, say from one of the *ice6g* models, as the prior mean of the true process and denote it by $\boldsymbol{\mu} : \mathbb{S}^2 \mapsto \mathbb{R}$. Then the residuals between the true process and forward model solutions can be modelled as a stationary Gaussian process on the sphere

$$\boldsymbol{X} := \boldsymbol{Y} - \boldsymbol{\mu} \sim \mathcal{GP}(\mathbf{0}, \kappa(\boldsymbol{\theta})) \tag{1.1}$$

where $\kappa(\boldsymbol{\theta})$ defines the covariance function with hyper parameters $\boldsymbol{\theta}$.

In order to assess the bias and uncertainties in the *ice6g* solutions, we use the GPS observations to update the GIA process. The GPS data are the yearly trends of vertical movements in millimetre at the observed locations. These observations can be regarded as a linear map of the GIA process with measurement errors

$$Z_i = \boldsymbol{\mathcal{A}}_i \boldsymbol{Y} + \varepsilon_i, \ i = 1, \dots N. \tag{1.2}$$

where $\boldsymbol{\mathcal{A}}_i$ is the linear operator that maps the GIA process to the i^{th} GPS observation and ε_i are assumed to be independent Gaussian errors $\mathcal{N}(0, e_i^2)$. In practice, e_i^2 can usually be estimated from raw GPS data and therefore we fix them at given values in the model.

Denote the linear operator for the GPS observation vector $\boldsymbol{Z}$ by

$$\boldsymbol{\mathcal{A}} = \begin{bmatrix} \boldsymbol{\mathcal{A}}_1 \\ \vdots \\ \boldsymbol{\mathcal{A}}_N \end{bmatrix}$$

Then we can write equation 1.2 into the vector form

$$\boldsymbol{Z} = \boldsymbol{\mathcal{A}}\boldsymbol{Y} + \boldsymbol{\varepsilon} \tag{1.3}$$

Now combining equations 1.1 and 1.3, the system simplifies to

$$\tilde{\boldsymbol{Z}} = \boldsymbol{Z} - \boldsymbol{\mathcal{A}}\boldsymbol{\mu} = \boldsymbol{\mathcal{A}}\boldsymbol{X} + \boldsymbol{\varepsilon} \tag{1.4}$$

Hence, our final model is

$$\begin{cases} \tilde{\boldsymbol{Z}} = \boldsymbol{\mathcal{A}}\boldsymbol{X} + \boldsymbol{\varepsilon},\ \boldsymbol{\varepsilon} \sim \mathcal{N}(\boldsymbol{0}, \mathrm{diag}(e_1^2, e_2^2, \ldots, e_N^2)) \\ \boldsymbol{X} \sim \mathcal{GP}(\boldsymbol{0}, \kappa(\boldsymbol{\theta})) \\ \boldsymbol{\theta} \sim \boldsymbol{\pi}(\boldsymbol{\theta}) \end{cases} \tag{1.5}$$

where $\boldsymbol{\pi}(\boldsymbol{\theta})$ is the prior distribution for the hyper parameters.

2 GMRF Approximation

Suppose we would like to predict the GIA process on a set of grid points $\boldsymbol{S} = \{s_i : i = 1, \ldots, m\}$ with a given resolution. The Gaussian process model can be computationally expensive for large scale inference since the Bayesian update scales as $\mathcal{O}(m^3)$ mainly due to the inverse of a dense covariance matrix.

At the same time, Gaussian Markov random field is often used for modelling discrete spatial unit. The covariance structure is defined through its inverse, the precision matrix, which is usually sparse and thus has nice computational properties.

The Gaussian process with Matérn covariance function can be treated as solutions to a class of stochastic partial differential equations (**?**) which can then be approximated by GMRF using finite element methods.

Denote by $\tilde{\boldsymbol{X}}$ the GMRF approximation of $\boldsymbol{X}$ on a given triangulation of the sphere with piecewise linear basis functions $\{\boldsymbol{\phi}_i\}_{i \in \mathbb{N}}$, then given any location $s \in \mathbb{S}^2$

$$\boldsymbol{X}(s) \approx \boldsymbol{\phi}_i(s)^T \tilde{\boldsymbol{X}} \tag{2.1}$$

and for a given set $\boldsymbol{S}$ of locations, we have

$$\boldsymbol{X}(\boldsymbol{S}) \approx \boldsymbol{C}(\boldsymbol{S})\tilde{\boldsymbol{X}} \tag{2.2}$$

where the matrix $\boldsymbol{C}$ contains basis functions for all locations.

Now with the GMRF approximation, our model becomes

$$\begin{cases} \tilde{\boldsymbol{Z}} = \boldsymbol{\mathcal{A}}\boldsymbol{C}\tilde{\boldsymbol{X}} + \boldsymbol{\varepsilon},\ \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \mathrm{diag}(e_1^2, e_2^2, \ldots, e_N^2)) \\ \tilde{\boldsymbol{X}} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q}^{-1}(\boldsymbol{\theta})) \\ \boldsymbol{\theta} \sim \boldsymbol{\pi}(\boldsymbol{\theta}) \end{cases} \tag{2.3}$$

where $\boldsymbol{Q}$ is the precision matrix of the GMRF approximation.