Conjugate Nearest Neighbor Gaussian Process (NNGP) models

Gaussian Process Model

$$y(\mathbf{s}_i)=\mathbf{x}(\mathbf{s}_i)^{\top}\boldsymbol{\beta}+w(\mathbf{s}_i)+\epsilon(\mathbf{s}_i)$$

Spatial linear mixed effects model, where $\epsilon(\mathbf{s}_i)\sim N\left(0,\tau^2\right)$, $w(\mathbf{s}_i)$ is the location specific random effects.

Gaussian Process models provides that $w(\mathbf{s})\sim GP(0,C(\cdot,\cdot\mid\boldsymbol{\theta}))$ where $C(\cdot,\cdot\mid\boldsymbol{\theta})$ is a covariance function. Therefore, $\mathbf{w}=(w(\mathbf{s}_1),\ldots,w(\mathbf{s}_n))^{\top}\sim N(0,\mathbf{C}(\boldsymbol{\theta}))$.

Finally, we can wirte a Hierarchical model:

$$\mathbf{y}\sim N\left(\mathbf{X}\beta+\mathbf{w},\tau^2\mathbf{I}\right),\mathbf{w}\sim N(\mathbf{0},\mathbf{C}(\boldsymbol{\theta}))$$

We can also write it as marginal model:

$$\mathbf{y}\sim N\left(\mathbf{X}\boldsymbol{\beta},\mathbf{C}(\boldsymbol{\theta})+\tau^2\mathbf{I}\right)$$

Conjugate NNGP model

The conjugate NNGP algorithm is an implementation of the response NNGP model which fixes certain spatial covariance parameters leading to exact (MCMC-free) posterior Bayesian inference. Under the full-GP specification, the covariance function for $y(s)$ is specified as

$$\Sigma(\mathbf{s}_i,\mathbf{s}_j)=C(\mathbf{s}_i,\mathbf{s}_j\mid\boldsymbol{\theta})+\tau^2\delta(\mathbf{s}_i,\mathbf{s}_j)$$

We can write $C(\mathbf{s}_i,\mathbf{s}_j\mid\boldsymbol{\theta})=\sigma^2(R(\mathbf{s}_i,\mathbf{s}_j\mid\boldsymbol{\phi}))$ where σ^2 is the marginal variance, and R is the correlation function parameterized by $\boldsymbol{\phi}$, i.e., $\boldsymbol{\theta}=\left(\sigma^2,\boldsymbol{\phi}\right)$. We can also write $\tau^2=\alpha\sigma^2$.

Finally, we rewrite covariance function as:

$$\Sigma(\mathbf{s}_i,\mathbf{s}_j)=\sigma^2(R(\mathbf{s}_i,\mathbf{s}_j\mid\boldsymbol{\phi})+\alpha\delta(\mathbf{s}_i,\mathbf{s}_j))$$

To solve computational cost problems when data is observed at a large number of locations, the conjugate NNGP model fixes $\boldsymbol{\phi}$ and α (selected by grid search and k-fold corss valdation), and generates the NNGP covariance function $\widetilde{M}(\cdot,\cdot\mid\alpha,\phi)$ approximating $R(\cdot,\cdot\mid\phi)+\alpha\delta(\cdot,\cdot)$.

Finally, we can write a marginal model:

$$\mathbf{Y}\sim N\left(\mathbf{X}\boldsymbol{\beta},\sigma^2\widetilde{\mathbf{M}}\right)$$

where $\widetilde{\mathbf{M}}=\widetilde{\mathbf{M}}(\phi,\alpha)$ is a fixed covariance matrix, which is sparse and ensures all the posterior distributions and moments can be evaluated efficiently.

Therefore, we obtain a standard Bayesian linear model with only unknowns $\boldsymbol{\beta},\sigma^2$. Using a Normal-Inverse-Gamma prior for $\boldsymbol{\beta},\sigma^2$ leads to conjugate Normal-Inverse-Gamma posterior distributions.

covariance function $\sigma^2\times\exp(-h/\phi)$