Spatial Predictive Process Model

John Tipton

April 8, 2014

1 Full Dimensional Model Statement

1.1 Data Model

$$\boldsymbol{y}_t = \boldsymbol{H}_t \boldsymbol{X} \boldsymbol{\beta}_t + \boldsymbol{\eta}_t + \boldsymbol{\epsilon}_t$$

1.2 Process Model

$$\boldsymbol{\beta}_t \sim N(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta) \qquad \boldsymbol{\Sigma}_\beta = \sigma_\beta^2 \boldsymbol{I}_t$$
$$\boldsymbol{\eta}_t \sim N(0, \boldsymbol{\Sigma}_\eta) \qquad \boldsymbol{\Sigma}_\eta = \sigma_\eta^2 \boldsymbol{R}(\phi) \qquad \boldsymbol{R}(\phi) = \exp\left(-\boldsymbol{D}_t/\phi\right)$$
$$\boldsymbol{\epsilon}_t \sim N(0, \boldsymbol{\Sigma}_\epsilon) \qquad \boldsymbol{\Sigma}_\epsilon = \sigma_\epsilon^2 \boldsymbol{I}_t$$

1.3 Parameter Model

$$\boldsymbol{\mu}_\beta \sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$$
$$\sigma_\beta^2 \sim IG(\alpha_\beta, \beta_\beta)$$
$$\sigma_\eta^2 \sim IG(\alpha_\eta, \beta_\eta)$$
$$\sigma_\epsilon^2 \sim IG(\alpha_\epsilon, \beta_\epsilon)$$
$$\phi \sim IG(\alpha_\phi, \beta_\phi)$$

where $\boldsymbol{I}_\beta$ is the identity matrix of size $\tau \times \tau$ where τ is the number of parameters in $\boldsymbol{\beta}_t$, $\boldsymbol{I}_t$ is the identity matrix of size $n_t \times n_t$ and n_t is the number of samples of y_t at time t and $\boldsymbol{D}_t$ is the distance matrix between locations observed at time t. Define $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_\eta + \boldsymbol{\Sigma}_\epsilon$

2 Predictive Process

For large dimensional spatial processes it can be computationally expensive to invert $\boldsymbol{\Sigma}$. This motivates the use of a predictive process $\tilde{\boldsymbol{\eta}}$ to approximate $\boldsymbol{\eta}$ over a set of knots $\tilde{S}$ where $\tilde{\boldsymbol{\eta}} = \boldsymbol{r}(\boldsymbol{s}, \boldsymbol{s}^*|\sigma_\eta^2, \phi)^T \boldsymbol{R}^{*-1}(\boldsymbol{s}^*, \boldsymbol{s}^*|\sigma_\eta^2, \phi)\boldsymbol{\eta}*$. The covariance between the desired locations $\boldsymbol{s} \in S$ and the set of knots $\boldsymbol{s}^* \in \tilde{S}$ is $\boldsymbol{r}(\boldsymbol{s}, \boldsymbol{s}^*|\sigma_\eta^2, \phi)$. The covariance matrix over the set of knots is $\boldsymbol{R}^{*-1}(\boldsymbol{s}^*, \boldsymbol{s}^*|\sigma_\eta^2, \phi)$. The lower dimensional $\boldsymbol{\eta}^* \sim \text{MVN}(\boldsymbol{0}, \boldsymbol{R}^{*-1}(\boldsymbol{s}^*, \boldsymbol{s}^*|\sigma_\eta^2, \phi))$.

3 Posterior

$$\prod_{t=1}^{T} [\boldsymbol{\beta}_t, \boldsymbol{\mu}_\beta, \sigma_\beta^2, \sigma_\eta^2, \sigma_\epsilon^2, \phi|\boldsymbol{y}_t] \propto \prod_{t=1}^{T} [\boldsymbol{y}_t|\boldsymbol{\beta}_t, \sigma_\eta^2, \phi, \sigma_\epsilon^2][\boldsymbol{\beta}_t|\boldsymbol{\mu}_\beta, \sigma_\beta^2][\boldsymbol{\mu}_\beta][\sigma_\beta^2][\sigma_\eta^2][\sigma_\epsilon^2][\phi]$$

4 Full Conditionals

4.1 Full Conditional for $\boldsymbol{\beta}_t$

For $t = 1, \ldots, T$,

$$[\boldsymbol{\beta}_t | \cdot] \propto [\boldsymbol{y}_t | \boldsymbol{\beta}_t, \sigma_\eta^2, \sigma_\epsilon^2, \phi][\boldsymbol{\beta}_t | \boldsymbol{\mu}_\beta, \sigma_\beta^2]$$

$$\propto e^{-\frac{1}{2}(\boldsymbol{y}_t - \boldsymbol{H}_t \boldsymbol{X} \boldsymbol{\beta}_t)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t - \boldsymbol{H}_t \boldsymbol{X} \boldsymbol{\beta}_t)} e^{-\frac{1}{2}(\boldsymbol{\beta}_t - \boldsymbol{\mu}_\beta)^T \boldsymbol{\Sigma}_\beta^{-1}(\boldsymbol{\beta}_t - \boldsymbol{\mu}_\beta)}$$

$$\propto e^{-\frac{1}{2}\{\boldsymbol{\beta}_t^T(\boldsymbol{X}^T \boldsymbol{H}_t^T \boldsymbol{\Sigma}^{-1} \boldsymbol{H}_t \boldsymbol{X} + \boldsymbol{\Sigma}_\beta^{-1})\boldsymbol{\beta}_t - 2\boldsymbol{\beta}_t^T(\boldsymbol{X}^T \boldsymbol{H}_t^T \boldsymbol{\Sigma}^{-1} \boldsymbol{y}_t + \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\mu}_\beta)\}}$$

which is Normal with mean $\boldsymbol{A}^{-1}\boldsymbol{b}$ and variance $\boldsymbol{A}^{-1}$ where

$$\boldsymbol{A}^{-1} = (\boldsymbol{X}^T \boldsymbol{H}_t^T \boldsymbol{\Sigma}^{-1} \boldsymbol{H}_t \boldsymbol{X} + \boldsymbol{\Sigma}_\beta^{-1})^{-1}$$

$$\boldsymbol{b} = (\boldsymbol{X}^T \boldsymbol{H}_t^T \boldsymbol{\Sigma}^{-1} \boldsymbol{y}_t + \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\mu}_\beta)$$

4.2 Full Conditional for $\boldsymbol{\mu}_\beta$

$$[\boldsymbol{\mu}_\beta | \cdot] \propto \prod_{t=1}^{T} [\boldsymbol{\beta}_t | \boldsymbol{\mu}_\beta, \sigma_\beta^2][\boldsymbol{\mu}_\beta]$$

$$\propto e^{-\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{\beta}_t - \boldsymbol{\mu}_\beta)^T \boldsymbol{\Sigma}_\beta^{-1}(\boldsymbol{\beta}_t - \boldsymbol{\mu}_\beta)} e^{-\frac{1}{2}(\boldsymbol{\mu}_\beta - \boldsymbol{\mu}_0)^T \Sigma_0^{-1}(\boldsymbol{\mu}_\beta - \boldsymbol{\mu}_0)}$$

$$\propto e^{-\frac{1}{2}(\boldsymbol{\mu}_\beta^T(T\boldsymbol{\Sigma}_\beta^{-1} + \boldsymbol{\Sigma}_0^{-1})\boldsymbol{\mu}_\beta - 2\boldsymbol{\mu}_\beta^T(\sum_{t=1}^{T} \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\beta}_t + \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\mu}_0))}$$

which is multivariate normal with mean $(T\boldsymbol{\Sigma}_\beta^{-1} + \boldsymbol{\Sigma}_0^{-1})^{-1}(\sum_{t=1}^{T} \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\beta}_t + \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\mu}_0)$ and variance $(T\boldsymbol{\Sigma}_\beta^{-1} + \boldsymbol{\Sigma}_0^{-1})^{-1}$

4.3 Full Conditional for σ_β^2

$$[\sigma_\beta^2 | \cdot] \propto \prod_{t=1}^{T} [\boldsymbol{\beta}_t | \boldsymbol{\mu}_\beta, \sigma_\beta^2][\sigma_\beta^2]$$

$$\propto (\prod_{t=1}^{T} |\boldsymbol{\Sigma}_\beta|^{-\frac{1}{2}}) e^{-\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{\beta}_t - \boldsymbol{\mu}_\beta)^T \boldsymbol{\Sigma}_\beta^{-1}(\boldsymbol{\beta}_t - \boldsymbol{\mu}_\beta)} (\sigma_\beta^2)^{-(\alpha_\beta + 1)} e^{-\frac{\beta_\beta}{\sigma_\beta^2}}$$

$$\propto (\sigma_\beta^2)^{-(\alpha_\beta + \frac{\sum_{t=1}^{T} n_t}{2} + 1)} e^{-\frac{1}{\sigma_\beta^2}(\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{\beta}_t - \boldsymbol{\mu}_\beta)^T(\boldsymbol{\beta}_t - \boldsymbol{\mu}_\beta) + \beta_\beta)}$$

which is $\text{IG}(\alpha_\beta + \frac{\sum_{t=1}^{T} n_t}{2}, \frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{\beta}_t - \boldsymbol{\mu}_\beta)^T(\boldsymbol{\beta}_t - \boldsymbol{\mu}_\beta) + \beta_\beta)$ since the determinant $|\boldsymbol{\Sigma}_\beta| = (\sigma_\beta^2)^{n_t}$ and $\boldsymbol{\Sigma}_\beta^{-1} = \frac{1}{\sigma_\beta^2}\boldsymbol{I}_t$

4.4 Full Conditional for σ_η^2

$$[\sigma_\eta^2 | \cdot] \propto \prod_{t=1}^{T} [\boldsymbol{y}_t | \boldsymbol{\beta}_t, \sigma_\eta^2, \phi, \sigma_\epsilon^2][\sigma_\eta^2]$$

$$\propto (\prod_{t=1}^{T} |\boldsymbol{\Sigma}|^{-\frac{1}{2}}) e^{-\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{y}_t - \boldsymbol{H}_t \boldsymbol{X} \boldsymbol{\beta}_t)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t - \boldsymbol{H}_t \boldsymbol{X} \boldsymbol{\beta}_t)} (\sigma_\eta^2)^{-\alpha_\eta + 1} e^{-\frac{\beta_\eta}{\sigma_\eta^2}}$$

$$\propto |\boldsymbol{\Sigma}|^{-\frac{T}{2}} e^{-\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{y}_t - \boldsymbol{H}_t \boldsymbol{X} \boldsymbol{\beta}_t)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t - \boldsymbol{H}_t \boldsymbol{X} \boldsymbol{\beta}_t)} (\sigma_\eta^2)^{-\alpha_\eta + 1} e^{-\frac{\beta_\eta}{\sigma_\eta^2}}$$

which can be sampled using a Metropolis-Hastings step

4.5   Full Conditional for σ_ϵ^2

$$
\begin{aligned}
[\sigma_\epsilon^2|\cdot] &\propto \prod_{t=1}^{T}[\boldsymbol{y}_t|\boldsymbol{\beta}_t,\sigma_\eta^2,\phi,\sigma_\epsilon^2][\sigma_\epsilon^2] \\
&\propto (\prod_{t=1}^{T}|\boldsymbol{\Sigma}|^{-\frac{1}{2}})e^{-\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{y}_t-\boldsymbol{H}_t\boldsymbol{X}\boldsymbol{\beta}_t)^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t-\boldsymbol{H}_t\boldsymbol{X}\boldsymbol{\beta}_t)}(\sigma_\epsilon^2)^{-\alpha_\epsilon+1}e^{-\frac{\beta_\epsilon}{\sigma_\epsilon^2}} \\
&\propto |\boldsymbol{\Sigma}|^{-\frac{T}{2}}e^{-\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{y}_t-\boldsymbol{H}_t\boldsymbol{X}\boldsymbol{\beta}_t)^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t-\boldsymbol{H}_t\boldsymbol{X}\boldsymbol{\beta}_t)}(\sigma_\epsilon^2)^{-\alpha_\epsilon+1}e^{-\frac{\beta_\epsilon}{\sigma_\epsilon^2}}
\end{aligned}
$$

which can be sampled using a Metropolis-Hastings step

4.6   Full Conditional for ϕ

$$
\begin{aligned}
[\phi|\cdot] &\propto \prod_{t=1}^{T}[y_t|\beta_t,\sigma_\eta^2,\phi,\sigma_\epsilon^2][\phi] \\
&\propto \prod_{t=1}^{T}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}e^{-\frac{1}{2}(\boldsymbol{y}_t-\boldsymbol{H}_t\boldsymbol{X}\boldsymbol{\beta}_t)^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t-\boldsymbol{H}_t\boldsymbol{X}\boldsymbol{\beta}_t)}\phi^{-\alpha_\phi+1}e^{-\frac{\beta_\phi}{\phi}} \\
&\propto |\boldsymbol{\Sigma}|^{-\frac{T}{2}}e^{-\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{y}_t-\boldsymbol{H}_t\boldsymbol{X}\boldsymbol{\beta}_t)^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_t-\boldsymbol{H}_t\boldsymbol{X}\boldsymbol{\beta}_t)}\phi^{-\alpha_\phi+1}e^{-\frac{\beta_\phi}{\phi}}
\end{aligned}
$$

which can be sampled using a Metropolis-Hastings step

5   Posterior Predictive Distribution

The posterior predictive distribution for $\boldsymbol{y}_t$ is sampled a each MCMC iteration k by

$$
\boldsymbol{y}_t^{(k)} \sim N(\boldsymbol{H_t X \beta}_t^{(k)}, \boldsymbol{\Sigma}^{(k)})
$$