Modeling large spatial datasets

Alan Gelfand\(^1\) and Andrew O. Finley\(^2\)

\(^1\) Department of Statistical Science, Duke University, Durham, North Carolina.
\(^2\) Department of Forestry, Michigan State University, Lansing, Michigan.

September 9, 2014

Computational issues

Computational issues.

- We need to evaluate
  \[-\frac{1}{2} \log \det(C_w(\theta)) - \frac{1}{2} w^\top C_w(\theta)^{-1} w\]

- What if \(n\) is LARGE? How do we tackle \(C_w(\theta)^{-1}\) and \(\det(C(\theta))\)?

Some approaches

  \(S^* = \{s^*_1, s^*_2, \ldots, s^*_n\}\): a set of “knots”.

- Smoothing causes loss in variability:
  \[w(s) \approx w_{\text{Krig}}(s) = \sum_{j=1}^{n^*} k(s - s^*_j, \theta_1) u_j, \quad u_j \overset{iid}{\sim} N(0, 1).
  \]

- No easy way to quantify this difference with kernel convolutions.

Hierarchical spatial model

\[p(\theta) \times p(\Psi) \times N(\beta | \mu_\beta, \Sigma_\beta) \times N(w | 0, C_w(\theta)) \times \prod_{i=1}^n N(y_i(s_i) | X(s_i)^\top \beta + w(s_i), D(\Psi))\]

- regression slopes
- spatial random effects from Gaussian process
- nonspatial variability (nugget)
- spatial process parameters (spatial variance, range, smoothness) and .

Low rank Gaussian process

- Call \(w(s) \sim GP_{m}(0, C_w(\theta))\) the parent process
- For \(S^* = \{s^*_1, s^*_2, \ldots, s^*_n\}\), let \(C_w(\theta) = \left\{C_w(s^*_j, s^*_j)\right\}\) :
  \[w^* = (w(s^*_1)^\top, w(s^*_2)^\top, \ldots, w(s^*_n)^\top)\] \(\sim N(0, C_w(\theta))\)

- The predictive process derived from \(w(s)\) is:
  \[w(s) = E[w(s) | w^*] = \text{cov}(w(s), w^*)^{-1} \text{var}(w^*)^{-1} w^*.
  \]

- \(w(s)\) is a degenerate Gaussian process delivering dimension-reduction.
Hierarchical predictive process models

\[ w(s) \sim w^*(s) = \begin{pmatrix} w(s_1) & \cdots & w(s_n) \end{pmatrix}^\top \]

Low rank interpolation

\[ w^* = \begin{pmatrix} w(s_1) \cdots w(s_n) \end{pmatrix}^\top \]

Hierarchical predictive process models

\[ p(\theta) \times p(\Psi) \times N(\beta \mid \mu_\beta, \Sigma_\beta) \times N(w^* \mid \mathbf{0}, C^*_w(\theta)) \]
\[ \times \prod_{i=1}^n N_m(y(s_i) \mid X(s_i)^\top \beta + \tilde{w}(s_i), D(\Psi)) \].

Systemic under-estimation:

Systematic under-estimation

\[ \text{var}\{w(s)\} = \text{var}\{E[w(s) \mid w^*]\} + \mathbb{E}\{\text{var}[w(s) \mid w^*]\} \]
\[ \geq \text{var}\{E[w(s) \mid w^*]\} = \text{var}\{\tilde{w}(s)\}. \]

Orthogonal decomposition:

\[ \text{var}\{w(s)\} = \text{var}\{\tilde{w}(s)\} + \text{var}\{w(s) - \tilde{w}(s)\} \]

\[ \tilde{z}(s) = w(s) - \tilde{w}(s) \sim GP(0, C_\tilde{z}(s_1, s_2; \theta_1)) \; ; \]
\[ C_\tilde{z}(s_1, s_2; \theta_1) = C(s_1, s_2; \theta_1) - c(s_1; \theta_1)'C'(\theta_1)^{-1}c(s_2; \theta_2). \]