Spatial Reading Group

Optional Subtitle

February 13, 2017

Outline

Outline

First Slide Title

Optional Subtitle

- My first point.
- My second point.

Outline

Second Slide Title

- First item.

Second Slide Title

- First item.
- Second item.

Second Slide Title

- First item.
- Second item.
- Third item.

Second Slide Title

- First item.
- Second item.
- Third item.
- Fourth item.

Second Slide Title

- First item.
- Second item.
- Third item.
- Fourth item.
- Fifth item.

Second Slide Title

- First item.
- Second item.
- Third item.
- Fourth item.
- Fifth item. Extra text in the fifth item.

Kriging

- Suppose our objective is to predict the value of the signal at an arbitrary location $S(x)$.
- Note that $(S(x), Y)$ is multivariate Gaussian with mean vector $\mu\mathbf{1}$ and covariance matrix

$$\begin{pmatrix} \sigma^2 & \sigma^2 r^T \\ \sigma^2 r & \sigma^2 V \end{pmatrix},$$

where r is a vector with elements $r_i = \rho(\|x - x_i\|)$ and $V = \sigma^2 R + \tau^2 I$.

Kriging

- Conditional mean and variance:

$$E(S(x)|Y) = \mu + r^T V^{-1}(Y - \mu\mathbf{1}),$$

$$Var(S(x)|Y) = \sigma^2(1 - r^T V^{-1} r).$$

- Two types:
  - Ordinary kriging: replace μ by its weighted least squares estimator

    $$\hat{\mu} = (\mathbf{1}^T V^{-1}\mathbf{1})^{-1}\mathbf{1}^T V^{-1} Y.$$

  - Simple kriging: replace μ by $\hat{\mu} = \bar{y}$.
- Both kriging predictors can be expressed as a linear combination: $\hat{S}(x) = \sum_{i=1}^n a_i(x) Y_i$, but $\sum_{i=1}^n a_i(x) = 1$ only for ordinary kriging.

Maximum Likelihood Estimation

- Gaussian model

$$Y \sim N(D\beta, \sigma^2 R(\phi) + \tau^2 I)$$

  with covariates matrix $D_{n\times p}$, regression coefficients β , covariance of a parametric model for $S(x)$, and nugget variance τ^2.

- The log-likelihood function is

$$\begin{aligned} L(\beta, \tau^2, \sigma^2, \phi) = & -0.5\{n\log(2\pi) + log\{|(\sigma^2 R(\phi) + \tau^2 I)|\} \\ & + (y - D\beta)^T(\sigma^2 R(\phi) + \tau^2 I)^{-1}(y - D\beta)\} \end{aligned}$$

- Let $\nu^2 = (\tau^2/\sigma^2)$, $V = R(\phi) + \nu I$, then $L(\beta, \tau^2, \sigma^2, \phi)$ is maximized at

$$\hat{\beta}(V) = (D^T V^{-1} D)^{-1} D^T V^{-1} y \qquad (1)$$

$$\hat{\sigma}^2(V) = n^{-1}\{y - D\hat{\beta}(V)\}^T V^{-1}\{y - D\hat{\beta}(V)\} \qquad (2)$$

Maximum Likelihood Estimation

- Plug (1) and (2) into $L(\beta, \tau^2, \sigma^2, \phi)$ and obtain the concentrated log-likelihood:
  $L_0(\nu^2, \phi) = -0.5\{n \log(2\pi) + n \log \hat{\sigma}^2(V) + log|V| + n\}$
- Optimize $L_0(\nu^2, \phi)$ numerically with respect to ν and ϕ; back substitution to obtain $\hat{\sigma}^2$ and $\hat{\beta}$
- Re-parameterisation of V can be used to obtain more stable estimation, e.g the ratio σ^2/ϕ is more stable than σ^2 and ϕ
- Computational tool: profile log-likelihood:
  Assume a model with parameters (α, ψ),

$$L_p(\alpha) = L(\alpha, \hat{\psi}(\alpha)) = \max_{\psi}(L(\alpha, \psi))$$

Maximum Likelihood Estimation

- Non-Gaussian data:
  (1): transformation to Gaussian (2) generalized linear model
- (1) E.g. Box-Cox transformation; denote the transformed responses $Y^* = (Y_1^*, ..., Y_n^*)$, and fit a Gaussian model

$$Y^* \sim N(D\beta, \sigma^2\{R(\phi) + \tau^2 I\})$$

  Computationally demanding, transformation may impede scientific interpretation
- (2) Generalized linear model

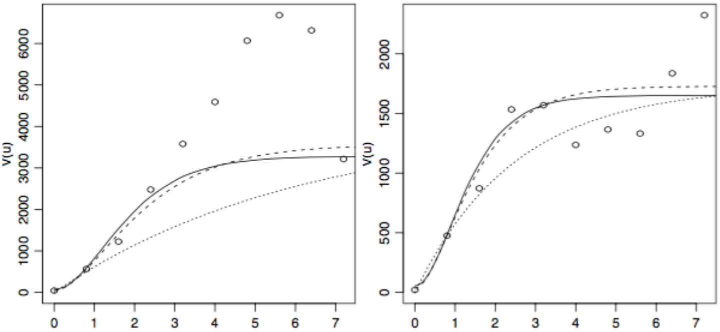
$$L(\theta|S) = \prod_{i=1}^{n} f_i(y_i|S, \theta)$$

$$L(\theta, \phi) = \int_S \prod_{i=1}^{n} f_i(y_i|s, \theta) g(s|\phi) ds$$

  Involve high dimensional integration;need MCMC/Hierarchical likelihood/Generalized estimating equations

Maximum Likelihood Estimation (An Example)

Model with constant mean							
Model	$\hat{\mu}$			$\hat{\sigma}^2$	$\hat{\phi}$	$\hat{\tau}^2$	logL
$\kappa = 0.5$	863.71			4087.6	6.12	0	−244.6
$\kappa = 1.5$	848.32			3510.1	1.2	48.16	−242.1
$\kappa = 2.5$	844.63			3206.9	0.74	70.82	−242.33
Model with linear trend							
Model	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\sigma}^2$	$\hat{\phi}$	$\hat{\tau}^2$	logL
$\kappa = 0.5$	919.1	−5.58	−15.52	1731.8	2.49	0	−242.71
$\kappa = 1.5$	912.49	−4.99	−16.46	1693.1	0.81	34.9	−240.08
$\kappa = 2.5$	912.14	−4.81	−17.11	1595.1	0.54	54.72	−239.75



Preferential Sampling

The Problem

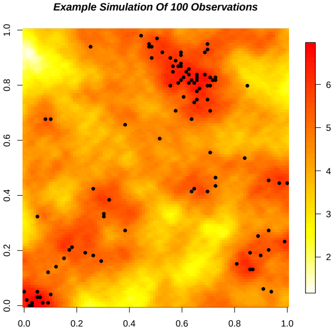
- So far we have assumed the sampling locations X are fixed, or assumed known.
- What if the sampling locations depend on the underlying field S?

Example

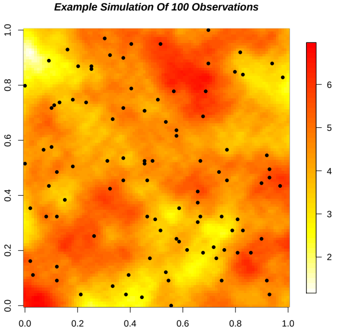
- Pollution data from measuring stations
- Ocean temperature data from marine mammals
- Lead concentration in Galicia (to be shown)

Preferential Sampling

Figure: Example of a single realisation of S and corresponding 100 sampling locations selected using a spatial Poisson Process with intensity $\lambda(x) = \exp(\beta S(x))$.



(a) Example of 100 preferentially sampled locations ($\beta = 2$)



(b) Example of 100 non-preferentially sampled locations ($\beta = 0$)

Preferential Sampling

Solution

- We must account for the dependence between X and S.

$$L(\boldsymbol{\theta}) = \int [X, Y, S] \, \mathrm{d}S. \tag{3}$$

- Diggle et al. 2010 - Monte Carlo
- Integrated Nested Laplace Approximation (INLA) - Joe
- Template Model Builder - Danny

Preferential Sampling

Results

Model	Parameter	Standard MLE	*TMB*
Preferential	Bias	(0.77, 1.36)	(0.41, 0.94)
Preferential	Root-mean-square error	(0.86, 1.40)	(0.60, 1.05)

Table: Comparison of approximate 95% confidence intervals for the root-mean-square errors and bias between standard MLE and *TMB* over 50 independent simulations for preferential ($\beta = 2$) at location $x_0 = (0.49, 0.49)$.

Summary

- The first main message of your talk in one or two lines.
- The second main message of your talk in one or two lines.
- Perhaps a third message, but not more than that.

- Outlook
  - Something you haven't solved.
  - Something else you haven't solved.

For Further Reading I

A. Author.
Handbook of Everything.
Some Press, 1990.

S. Someone.
On this and that.
Journal of This and That, 2(1):50–100, 2000.