Bayesian Kriging - Part II

Applied Spatial Statistics

Bayesian Kriging

- In the previous lecture we introduced the basic idea of Bayesian statistics
- We considered a few very simple cases
- In this lecture we discuss the computational methods needed to apply Bayesian methods to harder problems
- We then conduct a Bayesian analysis of a geostatistical model

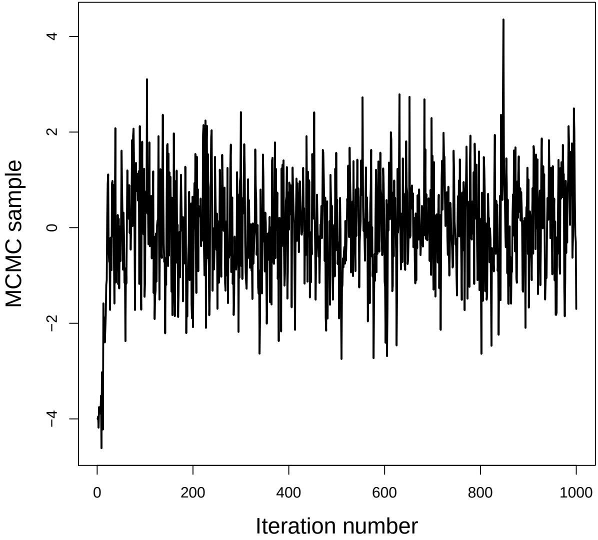
Bayesian computing

- When the model has many parameters it is difficult to summarize the posterior
- For example, our spatial model had $p + 1$ regression coefficients $\boldsymbol{\beta}$ and several covariance parameters $\boldsymbol{\theta}$
- Markov Chain Monte Carlo (MCMC) methods draw samples from the posterior to **approximate** the posterior
- Say $\beta_1^{(1)}, ..., \beta_1^{(S)}$ are S samples from the posterior of β_1
- Then a histogram of the S samples, approximates the posterior distribution
- The sample mean approximates the posterior mean, etc.

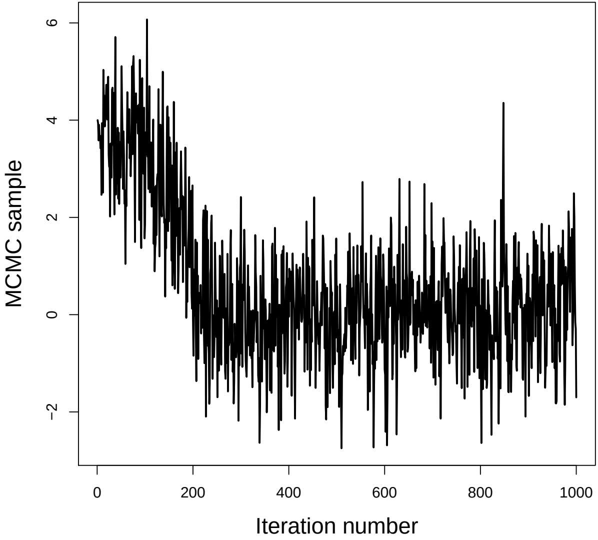
Bayesian computing

- As with optimization, MCMC begins with an initial value for all parameters
- It then makes draws from the posterior using Gibbs or Metropolis-Hastings algorithms
- These samples are correlated from one sample to the next
- Ideally this chain converges to the posterior distribution
- It converges to a distribution, not a point, so the trace plot should resemble a bar code/caterpillar

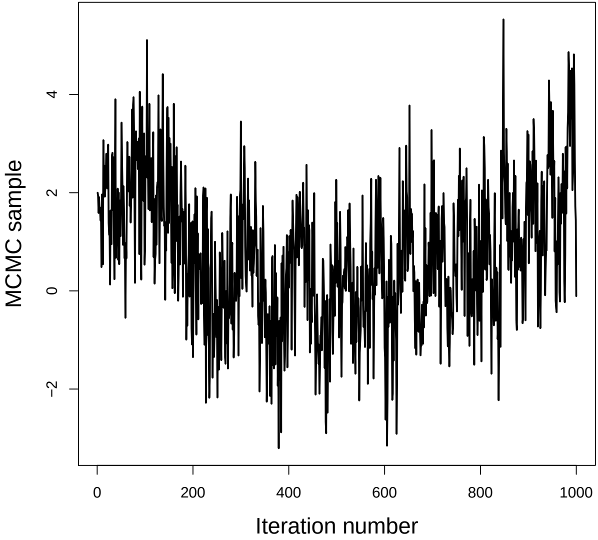
Great convergence



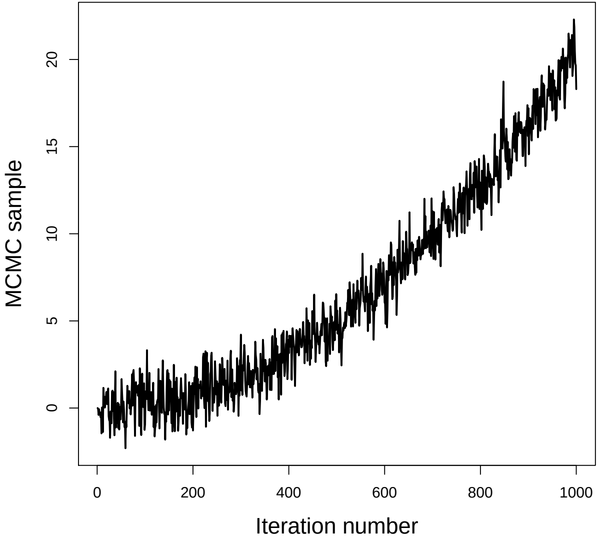
Good convergence



Questionable convergence



Poor convergence



Bayesian computing

- You discard the first T samples (say T=5k) as burn-in
- The remaining $S - T$ samples (say S=25k) are used to approximate the posterior distribution
- There are many general packages: `OpenBUGS`, `JAGS`, `NIMBLE`, `STAN`, `INLA`,...
- We will use the `spBayes` package in `R` because it is tailored to spatial models

Bayesian Kriging model

The exponential covariance model is the same as before,

$$\mathbf{Y} \sim \text{Normal}\{\mathbf{X}\beta, \Sigma(\boldsymbol{\theta})\}$$

where $\boldsymbol{\beta} = (\beta_0, ..., \beta_p)^T$ and $\boldsymbol{\theta} = (\sigma^2, \tau^2, \phi)$

- A Bayesian analysis requires priors
- Typically, $\boldsymbol{\beta} \sim \text{Normal}(0, c^2 I_{p+1})$ for large c
- A common prior for the range is $\phi \sim \text{Uniform}(0, d)$ where d is the extent of the spatial domain
- For technical reasons, variances usually have inverse gamma priors $\sigma^2, \tau^2 \sim \text{InvGamma}(a, b)$, for small a and b
- These priors are uninformative, but if you have solid prior information you should use it!

Bayesian Kriging

- MCMC produces S posterior samples of the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$
- In Kriging we fixed these parameters at an estimated value and made predictions
- In the Bayesian setting, we can make a prediction for each sample, and thus account for this uncertainty
- The draws $Y_0^{(1)}, ..., Y_0^{(S)}$ are from the posterior predictive distribution
- The mean of these S predictions can be used as the estimate, and the standard deviation quantifies uncertainty