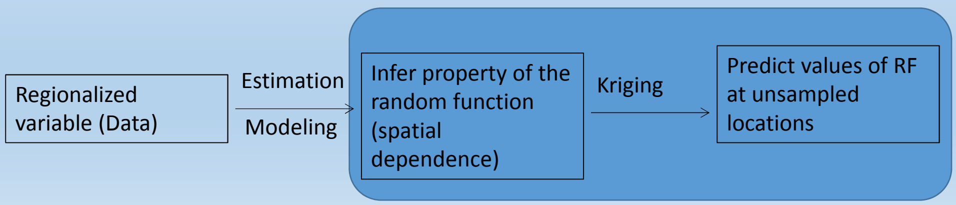
COA 616 Geostatistics in Environmental Sciences

Lecture 11 – Bayesian kriging

Wei Wu

November 22, 2016



Bayesian inference

- Bayes' theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$P(\theta|data) = \frac{P(data|\theta)P(\theta)}{P(data)} \propto P(data|\theta)P(\theta)$$

Posterior = likelihood*prior/normalizing constant

Posterior $\propto$ likelihood*prior

- Advantage of Bayesian inference

-- simplicity of interpretation of the results

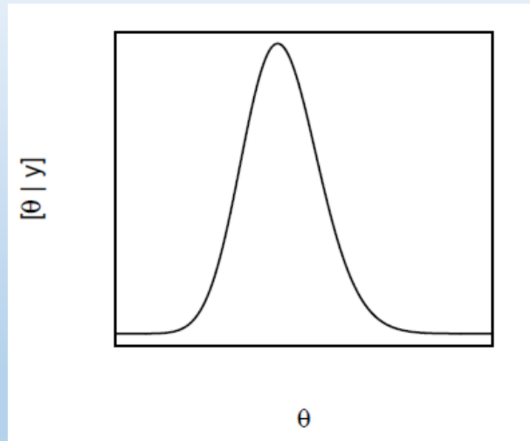
-- being consistent and ability of leaning from previous knowledge

-- estimating different sources of errors

-- data assimilation is automatic and the data need not be at the same scale and time

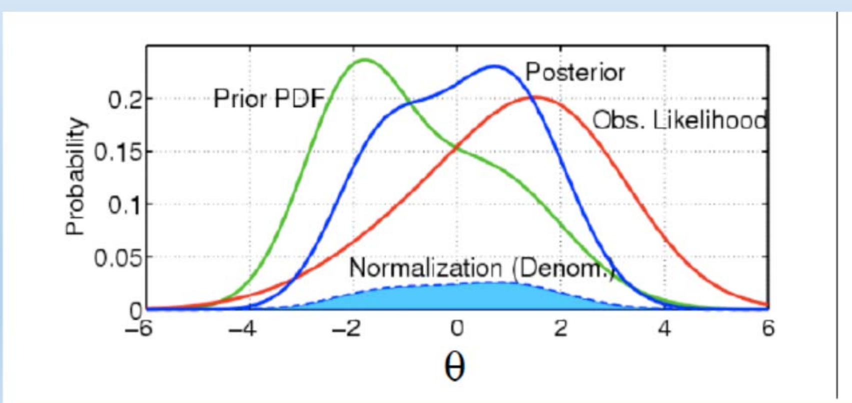
What sets Bayesian apart

- Bayesian analysis is the only branch of statistics that treats all unobserved quantities as random variables (θ).
- The data are random variables before they are observed and fixed after they have been observed. We seek to understand the probability distribution of unobserved using fixed observations, i.e. [θ|y]
- Those distribution quantify our uncertainty about θ.



Posterior

- Posterior distribution represents a balance between the information contained in the likelihood and the information contained in the prior distribution.
- An informative prior influences the posterior distribution. A vague prior exerts minimal influence on posterior distribution.



Probability [y|θ]

Θ is known as to be ½. Probability of number of whites conditional on three draws

$$p(y \mid n, p) = \binom{n}{y} p^{y} (1-p)^{n-y}$$

$y =$ Number of whites	$[y\|\boldsymbol{\theta}]$
0	.125
1	.375
2	.375
3	.125
$\sum_{i=1}^{4} [y\|\boldsymbol{\theta}_i] =$	1

Likelihood [y|θ]

Probability of two whites on three draws conditional on θ_i

Parameter	Likelihood $[y\|\boldsymbol{\theta}_i]$
$\theta_1 = 5/6$	.347
$\theta_2 = 1/2$	.375
$\theta_3 = 1/6$	.069
$\sum_{i=1}^{3} [y\|\boldsymbol{\theta}_i] =$	.791

Posterior [θ|y]

Probability of two whites on three draws conditional on θ_i

Probability of θ_i conditional on two whites on three draws

Parameter	Prior $[\theta_i]$	Likelihood $[y\|\theta_i]$	Joint $[y\|\theta_i]\,[\theta_i]$	Posterior $\frac{[y\|\theta_i][\theta_i]}{[y]} = [\theta_i\|y]$
θ_1	0.333	0.347	0.115	0.439
θ_2	0.333	0.375	0.125	0.474
θ_3	0.333	0.069	0.023	0.087
		$[y] = \sum_{i=1}^{3} [y\|\theta_i]\,[\theta_i] =$	0.261	$\sum_{i=1}^{3} [\theta_i\|y] = 1$

Bayesian spatial prediction

In the conventional geostatistical approaches for interpolation (kriging), the covariance (semivariance) structure is estimated first and the parameters for the structured are assumed to be fixed, and then the estimated covariance is used for interpolation. Assuming fixed parameters ignore the effect of the uncertainty in the convariance structure on subsequent predictions.

A Bayesian approach to interpolation of spatial processes will provide a general methodology for taking into account the uncertainty about parameters on subsequent predictions.

Bayesian Kriging

Instead of estimating the parameters for semivariances, we put a prior distribution on them, and update the distribution using the data.

Model: $(Z \mid \theta) \sim N(\beta, \sigma^2 C(\phi) + \tau^2 I)$

Prior: $f(\theta) = f(\beta) f(\sigma^2) f(\phi) f(\tau^2)$

Posterior: $f(\beta \mid Z = z) \propto f(\beta) \iiint f(z \mid \theta) f(\sigma^2) f(\phi) f(\tau^2) d\sigma^2 d\phi d\tau^2$

The Bayesian approach generalizes automatically to the case in which the variogram parameters are unknown, whereas the classical approach essentially makes the assumption that these are known.

Bayesian modeling of stationary processes

The basic model:

$$Y(s) = \mu(s) + \omega(s) + \varepsilon(s)$$

where the mean $\mu(\mathrm{s}) = \mathrm{X}(\mathrm{s})\beta$. The residuals has two components:

$\omega(\mathrm{s})$ a correlated error term, and $\varepsilon(\mathrm{s})$ an uncorrelated error term.

$\omega(\mathrm{s})$ introduces the partial sill σ^2 and range ϕ.

$\varepsilon(\mathrm{s})$ introduces the nugget effect τ^2. The nugget effect represents measurement error and/or microscale variability.

Priors

Independent priors are usually chosen:

$$p(\theta) = p(\beta)\, p(\sigma^2)\, p(\tau^2)\, p(\phi)$$

Good candidates are multivariate normal for β, and inverse gamma for σ^2 and τ^2. The prior ϕ depends on what ϕ represents. If ϕ is the inverse of the range, a gamma prior is commonly used.

Generally we prefer relatively noninformative priors. For β we use a flat that will be a proper posterior.

However, for the covariance parameters improper priors lead to improper posteriors.

geoR

Krige.bayes is a generic function for Bayesian geostatistical analysis of (transformed) Gaussian where predictions take into account the parameter uncertainty.

It can be set to run conventional kriging methods which use known parameters or plug-in estimates. However, the functions krige.conv and ksline are preferable for prediction with fixed parameters.