Bayesian Model Selection

An Alternative to AIC for Ecological Inference

Devin S. Johnson, Ph.D.
National Marine Mammal Laboratory, NOAA
Seattle, WA

Email: Devin.Johnson@NOAA.gov

Alaska Fisheries Science Center
NATIONAL MARINE FISHERIES SERVICE – NOAA FISHERIES

Benefits of a Bayesian MCMC approach

- Allows inclusion of expert knowledge in selection of regression covariates
  - AIC / BIC methods treat all covariates equally
- Selection over large model spaces
  - Parameters in each model must be separately estimated for AIC methods
- A sample from the joint model and parameter posterior is obtained
- Straightforward extension to generalized linear mixed models

Generalized Linear Mixed Models

Data Model

$$Y(\mathbf{s})|Z(\mathbf{s}) \sim \text{i.i.d. } P(\ell^{-1}\{Z(\mathbf{s})\}),$$

where

$$E[Y(\mathbf{s})|Z(\mathbf{s})] = \ell^{-1}\{Z(\mathbf{s})\}$$

Parameter model

$$\mathbf{Z} = (Z(\mathbf{s}_1), \ldots, Z(\mathbf{s}_n))' \sim N_n(\mathbf{X}\boldsymbol{\beta}; \boldsymbol{\Sigma})$$

where $\boldsymbol{\Sigma}$ is defined by a covariance model (geostatistical, CAR, AR,...)

Covariance function

$$\text{Cov}\{Z(\mathbf{s}_i),\ Z(\mathbf{s}_j)\} = \sigma^2 \rho(\mathbf{h}; \boldsymbol{\phi})$$

$$\text{Var}\{Z(\mathbf{s})\} = \sigma^2$$

where,

- $\mathbf{h} = \mathbf{s}_i - \mathbf{s}_j$
- σ^2 is the sill ($0 < \sigma^2 < \infty$)
- $\boldsymbol{\phi}$ are the spatial correlation parameters
- $\rho(\mathbf{h}; \boldsymbol{\phi})$ is a nonnegative correlation function
  e.g. $\rho(\mathbf{h}; \boldsymbol{\phi}) = \exp\left\{-(\mathbf{h}'\boldsymbol{\phi}\mathbf{h})^{1/2}\right\}$

Bayesian model selection

- Model incorporated as another parameter, M with sample space $\mathcal{M} = \{m_0, \ldots, m_K\}$
- For each m_k we have $\boldsymbol{\vartheta}_k = (\boldsymbol{\beta}_k, \sigma^2, \boldsymbol{\phi}, \mathbf{Z})$
- Inference for the model can be made through the posterior model probability (PMP)

$$P(m_k|\mathbf{Y}) \propto \int P(\mathbf{Y}|\boldsymbol{\vartheta}_k, m_k) P(\boldsymbol{\vartheta}_k|m_k) P(m_k) d\boldsymbol{\vartheta}_k$$
$$= P(\mathbf{Y}|m_k) P(m_k)$$

Model prior distribution

A classic model prior is derived by treating inclusion of the p coefficients as a series of independent Bernoulli trials with probability π_j. The result is the following prior

$$P(m_k) = \prod_{j=1}^{p} \pi_j^{I_{k_j}} (1 - \pi_j)^{1 - I_{k_j}},$$

where I_{k_j} is the indicator that $\beta_j \neq 0$

Fish abundance in the Appalachian region

- In 1994 – 1995, $n = 119$ stream sites were sampled by EPA in Mid-Atlantic highlands region of U.S. (MAHA)
- $Z(\mathbf{s})$ = Abundance of pollution intolerant fish – important indicators of stream health
- Environmental Covariates (π_j):

Strahler order (0.75)	Road density (0.5)
Elevation (0.75)	% watershed disturbed (0.5)
Watershed area (0.75)	Habitat quality index (0.5)
	% fish cover (0.5)
	Dissolved O_2 conc. (0.5)
	% fine sediments (0.5)

Model chain summary

		PMP				
Covariate	PIP	0.12	0.06	0.05	0.05	0.04
Strahler order	0.88	●	●	●	●	●
Elevation	0.29				●	
Area	0.43		●			
Road density	0.38			●		●
% Disturbance	0.79	●	●		●	●
Habitat quality	0.74	●	●	●	●	●
Dissolved O_2	0.15					
% Fish cover	0.10					
% Fine sed.	0.13					

Marginal coefficient posterior distributions