BYM2, Blending Error, and Covariate Bias

Background

Spatial Models in Epidemiology

The use of spatial models in epidemiology often tries to link observed spatial covariates with a health outcomes while adjusting for unobserved covariates that may or may not be spatially correlated. By adjusting for this unobserved correlation, the association between observed covariates and the outcome of interest may be better understood. Not accounting for this spatial correlation could lead to biased estimates of the parameter of interest or underestimation of the standard errors/posterior standard deviations(cite).

On the other hand adding a spatially correlated random effect can also have negative consequences when the underlying process does not have spatially correlated error terms. The spatially correlated error term can account for much of the variation that the covariates would normally and there for bias results for those parameters towards the null. Justification for the use (or non use) of spatial terms can often be difficult to argue and their effects often have a non trivial effect on the outcome of the model at hand.

BYM2 and the PC prior

In an analysis done by Riebler et. al.(cite) the authors propose a new model that is a modification of the BYM model that takes into account scaling so that both overdispersion and spatial random effects may be included and that they are on a similar scale. To illustrate this let us assume that

$$y_i|\theta_i \sim Poisson(E_i\theta_i)$$

Where y_i is the observed count of some noncontagious disease for an area, θ_i is the relative risk for the area and E_i is the expected value for a particular area. $log(\theta_i)$ may the be modeled using the form

$$log(\theta_i) = \beta \cdot x_i^T + b_i$$

Where x_i is a vector of explanatory covariates adn β is an array of parameters that account for the relationships between the covariates and the relative risk. b_i is the random effect term that accounts for the error in the linear relationship between x_i and θ_i. It is in this term that we may also account for the spatial relationships between data points. Dean et al(cite) proposes that the functional form of the term b may be written as

$$b = \frac{1}{\sqrt{\tau}}\left(\sqrt{1-\phi}v + \sqrt{\phi}u\right)$$

Where τ is the precision of the random effects, v is the error that occurs from simple overdispersion, u is the error that may be attributed to unobserved spatial correlation and ϕ determines how much each type of error contributes to the final estimate. One difficulty with this model is that it may falsely attribute variation of the errors to the spatial component u when in actuality it is not part of the underlying process. This over-complicates the model and could lead to biased parameter estimates. In addition setting hyperpriors on ϕ and τ becomes difficult across different scenarios as it is dependent on the structure of the spatial graph at hand. A way to adjust for this is proposed by Riebler et. al.(cite) in which b is modeled using a modified $u_\star$ in which

$$u_\star \sim \mathcal{N}(0, Q_\star^-)$$

$Q_\star^-$ calculation is shown in detail in Riebler et. al.(cite). The benefit of this model is that it allows generalized specifications of hyperpriors such that they may be applied to many models despite their geographic structure. In turn, this leads to the application of the penalized complexity (PC) prior. The PC prior allows for a model to include multiple terms however penalizes the model for its inclusion when it is unwarranted. In our case this permits the inclusion of a spatial term that will be essentially removed (i.e. $\phi = 0$) if their is not enough evidence for its inclusion in the data. From here on we will refer to the Dean et al. model with scaling and the PC prior as the BYM2 model. A more detailed description of the PC prior and its applications may be found in Simpson et al.(cite).

Testing the BYM2 model

Past Research

In the proposal of the BYM2 Riebler et. al.(cite) test how accurately the model is able to decompose to a simpler model (where either $\phi = 0$ or $\phi = 1$) when the data was simulated using such a specification. The BYM2 model was shown to perform well under both situations and perform as well or better than other spatial models at calculating τ. While this testing shows that the model is able to decompose well when their are no true or only spatial effects we would expect most model to have a blending of the two.

In order to test the ability of the BYM2 model to accurately maintain mixture of both spatial and non spatial components testing an array of values of ϕ where $\phi \in [0, 1]$. In this way we may test the ability of the model to maintain the proportionality of the two random effects u and v with out unnecessarily

decomposing them. This is especially important for cases where $\phi \approx .5$ and there is equal division of contribution to relative risk from both u and v.

In addition testing the BYM2 models ability to correctly asses covariate effects in the presence of different values of ϕ must also be assessed. As most epidemiologists are concerned with the correlation between the observed variables and the outcome, it is essential that the BYM2 model is able to accurately recover the true β parameters for a given covariate when the correct model form is used to estimate the parameters.

Methods

In order to test the BYM2 model a set of simulated data was created using the following form

$$log(\theta) = \beta \cdot x^T + (\frac{1}{\sqrt{\tau}}(\sqrt{1-\phi}\ v + \sqrt{\phi}\ u_\star))$$

For testing how well the BYM2 model is able to accurately reproduce ϕ we specify the variables such that β is a vector with one element 0, x is a matrix with a single column which is $N_{iid}(0,1)$, $\frac{1}{\sqrt{\tau}} = .5$, $v \sim N(0,1)$, $u_\star \sim N(0, Q_\star^-)$. We vary values of ϕ by increments of .001 from .001 to 1. In this way we can look at how the estimated values of ϕ change as we approach critical points of 0 (no spatial errors), 1 (only spatial errors), and .5 (errors that have both overdispersion and spatial components).

To test the effectiveness of the BYM2 model to return accurate estimates of covariates a similar form is used to generate data with the following adjustments. β is a vector of parameters with values $[0, 1, -1]$ and x is a matrix of values with a row for every observation and 3 columns for each beta. The elements within the matrix are created such that $x \sim N(0,1)$.