Spatial Modeling for Areal Data: Spatial Autoregression

Sudipto Banerjee

University of California, Los Angeles, USA

Hierarchical GLM for Spatial disease mapping

- At unit (region) i, we observe response y_i and covariate x_i
- $g(E(y_i)) = x_i^\top \beta + w_i$ where $g(\cdot)$ denotes a suitable link function

$$p_2(\beta, \tau_w, \rho) \times N(w \,|\, 0, \tau_w(D - \rho A)) \times \prod_{i=1}^{n} p_1(y_i \,|\, x_i^\top \beta + w_i)$$

- p_1 denotes the density corresponding to the link $g(\cdot)$

- CAR model:

$$w_i \,|\, w_{-i} \sim N\left(\frac{\rho}{n_i}\sum_{j\,|\,i\sim j} w_j, \tau_w n_i\right)$$

- Apply Brook's Lemma to obtain joint density for w
- $w = (w_1, w_2, \ldots, w_k)^\top \sim N(0, \tau_w(D - \rho A))$ where $D = \text{diag}(n_1, n_2, \ldots, n_k)$
- $\rho = 1 \Rightarrow$ Improper distribution as $(D - A)1 = 0$ (ICAR)
  - Can be still used as a prior for random effects
  - Cannot be used directly as a data generating model

CAR models

- CAR model:

$$w_i \,|\, w_{-i} \sim N\left(\frac{\rho}{n_i}\sum_{j\,|\,i\sim j} w_j, \tau_w n_i\right)$$

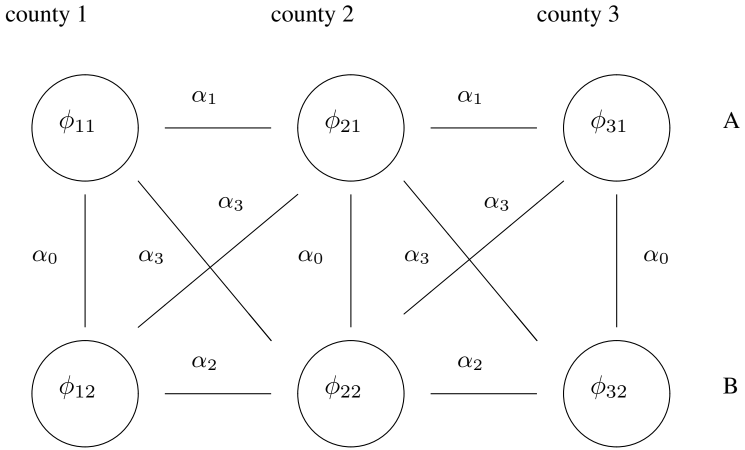
- Apply Brook's Lemma to obtain joint density for w
- $w = (w_1, w_2, \ldots, w_k)^\top \sim N(0, \tau_w(D - \rho A))$ where $D = \mathrm{diag}(n_1, n_2, \ldots, n_k)$
- $\rho = 1 \Rightarrow$ Improper distribution as $(D - A)1 = 0$ (ICAR)
  - Can be still used as a prior for random effects
  - Cannot be used directly as a data generating model
- $\rho < 1 \Rightarrow$ Proper distribution with added parameter flexibility

- y_{ij} is the disease count for disease j in county i.
- x_{ij} are explanatory, region-level spatial covariates for disease j.

$$p_2(\beta, \theta) \times N(\phi \,|\, 0, \Sigma_\theta) \times \prod_{i=1}^{n} \prod_{j=1}^{p} p_1(y_{ij} \,|\, x_{ij}^\top \beta_j + \phi_{ij})$$

- How do we model the ϕ_{ij}'s?

Multivariate disease mapping using graphs



Conditional bivariate CAR: Jin et al. (2005)

- Build a hierarchical bivariate spatial model for $p = 2$ outcomes:

$$N(\phi_1 \,|\, 0, \tau_1(D - \rho_1 A)) \times N(\phi_2 \,|\, C\phi_1, \tau_2(D - \rho_2 A))$$

- $\mathrm{E}[\phi_2 \,|\, \phi_1] = C\phi_1$. Assume that the elements of C are

$$c_{ij} = \begin{cases} \eta_0 & \text{if } j = i \\ \eta_1 & \text{if } j \sim i \text{ (i.e., if region } j \text{ is a neighbor of region } i\text{)} \\ 0 & \text{otherwise} \end{cases} \quad .$$

- $C = \eta_0 I + \eta_1 A$, where η_0 and η_1 control spatial smoothing for cross-covariances.
- Call this CBCAR$(\rho_1, \rho_2, \eta_0, \eta_1, \tau_1, \tau_2)$. Some special cases emerge:
  - Separable MCAR
  - Kim, Sun and Tsutakawa (2001): Two-fold CAR model with smoothing of cross-correlations
- Generalizations: $C = \sum_j \eta_j A^j$