CAR: Conditional Auto-Regression Models

Sudipto Banerjee

University of California, Los Angeles, USA

Simple linear regression: continuous outcomes

- Units: $i = 1, 2, \ldots, n$; Outcome: y_i; Predictor: x_i
- Recall simple linear regression:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \; ; \quad i = 1, 2, \ldots n$$

- Multiple linear regression:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + \epsilon_i \; ; \quad i = 1, 2, \ldots n$$

- The errors: $\epsilon_i \stackrel{iid}{\sim} N(0, \sigma^2)$:

$$\begin{aligned} y_i \,|\, \mu_i,\, \sigma^2 &\stackrel{ind}{\sim} N(\mu_i, \sigma^2) \; ; \\ \mu_i &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} \; . \end{aligned}$$

- How do we estimate the regression coefficients and the residual variance σ^2.

$$
\begin{aligned}
y_i \,|\, \mu_i,\, \sigma^2 &\stackrel{ind}{\sim} N(\mu_i, \sigma^2)\;; \quad i = 1, 2, \ldots, n\;;\\
\mu_i &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi}\;; \quad i = 1, 2, \ldots, n\;;\\
\beta_j &\stackrel{ind}{\sim} N(0, \sigma_\beta^2)\;; \quad j = 0, 1, \ldots, p\;; 1/\sigma^2 \sim \text{Gamma}(a, b)\;.
\end{aligned}
$$

- Binary data:

$$
\begin{aligned}
y_i &\sim Ber(p_i) \; ; \quad i = 1, 2, \ldots, n \; ; \\
\log \frac{p_i}{1 - p_i} &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} \; ; \quad i = 1, 2, \ldots, n \; ; \\
& \qquad \beta_j \stackrel{ind}{\sim} N(0, \sigma_\beta^2) \; ; \quad j = 0, 1, \ldots, p \; .
\end{aligned}
$$

- Count data:

$$
\begin{aligned}
y_i &\sim Poi(\lambda_i) \; ; \quad i = 1, 2, \ldots, n \; ; \\
\log \lambda_i &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} \; ; \quad i = 1, 2, \ldots, n \; ; \\
& \qquad \beta_j \stackrel{ind}{\sim} N(0, \sigma_\beta^2) \; ; \quad j = 0, 1, \ldots, p \; .
\end{aligned}
$$

- Continuous data:

$$
\begin{aligned}
y_i \,|\, \mu_i,\, \sigma^2 &\overset{ind}{\sim} N(\mu_i, \sigma^2)\ ; \quad i = 1, 2, \ldots, n\ ;\\
\mu_i &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + w_i\ ; \quad i = 1, 2, \ldots, n\ ;\\
\beta_j &\overset{ind}{\sim} N(0, \sigma_\beta^2)\ ; \quad j = 0, 1, \ldots, p\ ;\\
w = (w_1, w_2, \ldots, w_n)^\top &\sim N(0, V_w)\ ; \quad 1/\sigma^2 \sim \text{Gamma}(a, b)\ .
\end{aligned}
$$

- Count data:

$$
\begin{aligned}
y_i &\sim Poi(\lambda_i)\ ; \quad i = 1, 2, \ldots, n\ ;\\
\log \lambda_i &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + w_i\ ; \quad i = 1, 2, \ldots, n\ ;\\
\beta_j &\overset{ind}{\sim} N(0, \sigma_\beta^2)\ ; \quad j = 0, 1, \ldots, p\ ; \quad w \sim N(0, V_w)\ .
\end{aligned}
$$

- V_w introduce dependence among w_i's—and among y_i's.

Toward spatial dependence: Adjacency or proximity matrices

- A, entries a_{ij}, ($a_{ii} = 0$); choices for a_{ij}:
  - $a_{ij} = 1$ if i, j share a common boundary (possibly a common vertex)
  - a_{ij} is an *inverse* distance between units
  - $a_{ij} = 1$ if distance between units is $\leq K$
  - $a_{ij} = 1$ for m nearest neighbors.
- A need not be symmetric.
- $\widetilde{A}$: standardize row i by $a_{i+} = \sum_j a_{ij}$ (row stochastic but need not be symmetric).
- A elements often called "weights"; perhaps nicer interpretation?

CAR models

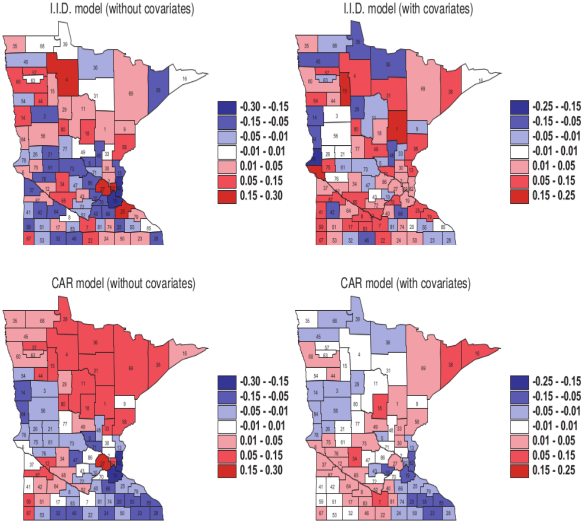
- CAR model:

$$w_i \,|\, w_{-i} \sim N\left(\frac{\rho \sum_j a_{ij} w_j}{\sum_j a_{ij}}, \frac{\sigma_w^2}{\sum_j a_{ij}}\right)$$

- Joint density for w: $w = (w_1, w_2, \ldots, w_n)^\top \sim N(0, V_w)$
- $V_w^{-1} = \left(1/\sigma_w^2\right)(D - \rho A)$; $D = \text{diag}\left(\sum_j a_{1j}, \sum_j a_{2j}, \ldots, \sum_j a_{nj}\right)$.
- Usual specification: $a_{ii} = 0$ and $a_{ij} = I(i \leftrightarrow j)$.
- $\rho = 1 \Rightarrow$ Improper distribution as $(D - A)1 = 0$ (ICAR)
  - Can be 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

Disease Mapping: Mapping Random Effects

[Infant Mortality Rates] $=$ [Intercept] $+$ [Fixed Effects] $+$ [County-wise Random Effects]



Other versions of CAR models

- Besag-York-Mollie (BYM):

$$\mu_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + w_i + v_i \ ;$$

$$w_i \,|\, w_{-i} \sim N\left(\frac{\sum_j a_{ij} w_j}{\sum_j a_{ij}}, \frac{\sigma_w^2}{\sum_j a_{ij}}\right) \quad ; \quad v_i \sim N(0, \sigma_v^2) \ .$$

- Leroux et al. (2000) proposed the following alternative CAR prior for modeling varying strengths of spatial autocorrelation using only a single set of random effects:

$$w_i \,|\, w_{-i} \sim N\left(\frac{\rho \sum_j a_{ij} w_j}{\rho \sum_j a_{ij} + 1 - \rho}, \frac{\sigma^2}{\rho \sum_j a_{ij} + 1 - \rho}\right)$$