

Models for areal data

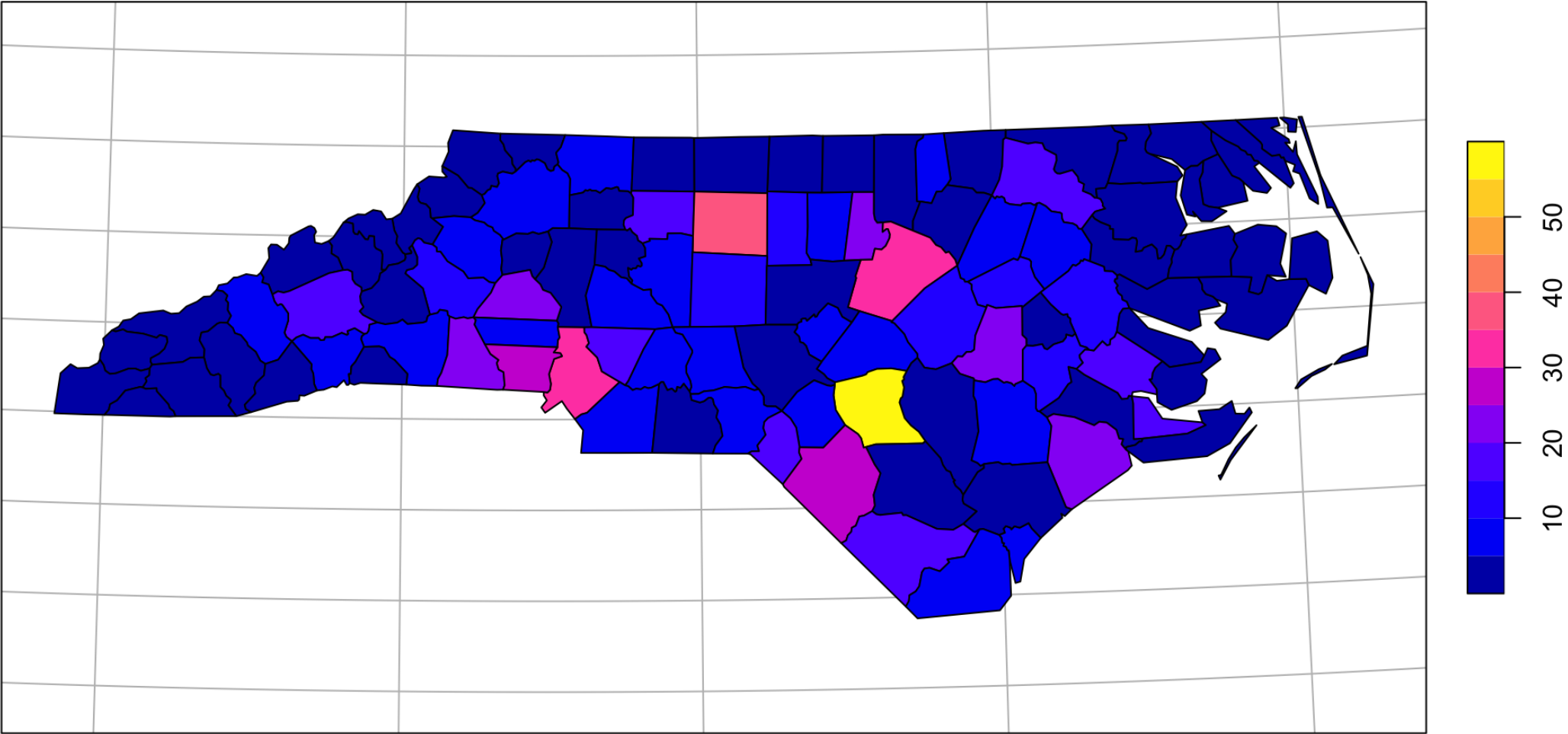
Lecture 18

Dr. Colin Rundel

areal / lattice data

Example - NC SIDS

SID79



Adjacency Matrix

```
1 1*st_touches(nc[1:12,], sparse=FALSE)
```

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]	[,11]	[,12]
[1,]	0	1	0	0	0	0	0	0	0	0	0	0
[2,]	1	0	1	0	0	0	0	0	0	0	0	0
[3,]	0	1	0	0	0	0	0	0	0	1	0	0
[4,]	0	0	0	0	0	0	1	0	0	0	0	0
[5,]	0	0	0	0	0	1	0	0	1	0	0	0
[6,]	0	0	0	0	1	0	0	1	0	0	0	0
[7,]	0	0	0	1	0	0	0	1	0	0	0	0
[8,]	0	0	0	0	0	1	1	0	0	0	0	0
[9,]	0	0	0	0	1	0	0	0	0	0	0	0
[10,]	0	0	1	0	0	0	0	0	0	0	0	1
[11,]	0	0	0	0	0	0	0	0	0	0	0	1
[12,]	0	0	0	0	0	0	0	0	0	1	1	0

Normalized spatial weight matrix

```
1 dukestm::normalize_weights( st_touches(nc[1:12,], sparse=FALSE) )
```

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]	[,11]	[,12]
[1,]	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[2,]	0.5	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[3,]	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0
[4,]	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
[5,]	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.5	0.0	0.0	0.0
[6,]	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.5	0.0	0.0	0.0	0.0
[7,]	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0
[8,]	0.0	0.0	0.0	0.0	0.0	0.5	0.5	0.0	0.0	0.0	0.0	0.0
[9,]	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[10,]	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5
[11,]	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
[12,]	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.5	0.0

EDA - Moran's I

If we have observations at n spatial locations (s_1, \dots, s_n)

$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y(s_i) - \bar{y}) (y(s_j) - \bar{y})}{\sum_{i=1}^n (y(s_i) - \bar{y})^2}$$

where \mathbf{w} is a normalized spatial weight matrix.

Some properties of Moran's I when there is no spatial autocorrelation / dependence:

- $E(I) = -1/(n - 1)$
- $Var(I) = (\text{Something ugly but closed form}) - E(I)^2$
- $\lim_{n \rightarrow \infty} \frac{I - E(I)}{\sqrt{Var(I)}} \sim N(0, 1)$ (via the CLT)

NC SIDS & Moran's I

Lets start by using a normalized spatial weight matrix for w (basedd on shared county borders).

```
1 morans_I = function(y, w) {  
2   # w here can be either the adjacency matrix or the normalized weight r  
3   w = normalize_weights(w)  
4   n = length(y)  
5   num = sum(w * (y-mean(y)) %*% t(y-mean(y)))  
6   denom = sum( (y-mean(y))^2 )  
7   (n/sum(w)) * (num/denom)  
8 }  
9  
10 A = 1*st_touches(nc, sparse=FALSE)  
11 morans_I(y = nc$SID74, A)
```

```
[1] 0.1477405
```

```
1 ape::Moran.I(nc$SID74, weight = A)
```

\$observed

[1] 0.1477405

\$expected

[1] -0.01010101

\$sd

[1] 0.06265435

\$p.value

[1] 0.01176074

EDA - Geary's C

Like Moran's I, if we have observations at n spatial locations (s_1, \dots, s_n)

$$C = \frac{n-1}{2 \sum_{i=1}^n \sum_{j=1}^n w_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y(s_i) - y(s_j))^2}{\sum_{i=1}^n (y(s_i) - \bar{y})^2}$$

where \mathbf{w} is a normalized spatial weights matrix.

Some properties of Geary's C:

- $0 < C < 2$
 - If $C \approx 1$ then no spatial autocorrelation
 - If $C > 1$ then negative spatial autocorrelation
 - If $C < 1$ then positive spatial autocorrelation
- Geary's C is inversely related to Moran's I

NC SIDS & Geary's C

Again using an adjacency matrix for w (shared county borders).

```
1 gearys_C = function(y, w) {  
2   # w here can be either the adjacency matrix or the normalized weight matrix  
3   w = normalize_weights(w)  
4  
5   n = length(y)  
6   y_i = y %*% t(rep(1,n))  
7   y_j = t(y_i)  
8   ((n-1)/(2*sum(w))) * (sum(w * (y_i-y_j)^2) / sum( (y - mean(y))^2 ))  
9 }  
10  
11 A = 1*st_touches(nc, sparse=FALSE)  
12 gearys_C(y = nc$SID74, w = A)
```

```
[1] 0.8438767
```

Spatial Correlogram

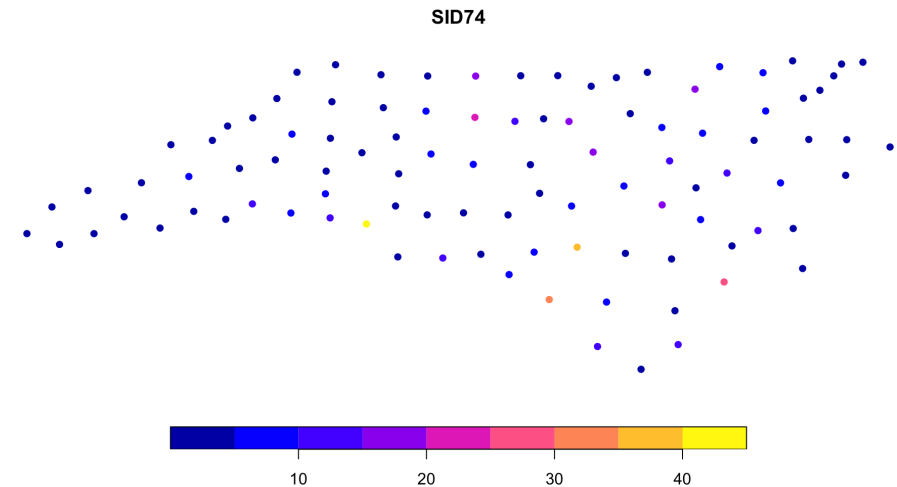
Rather than using the touches predicate to determine adjacency, we can alternatively use a distance based approach.

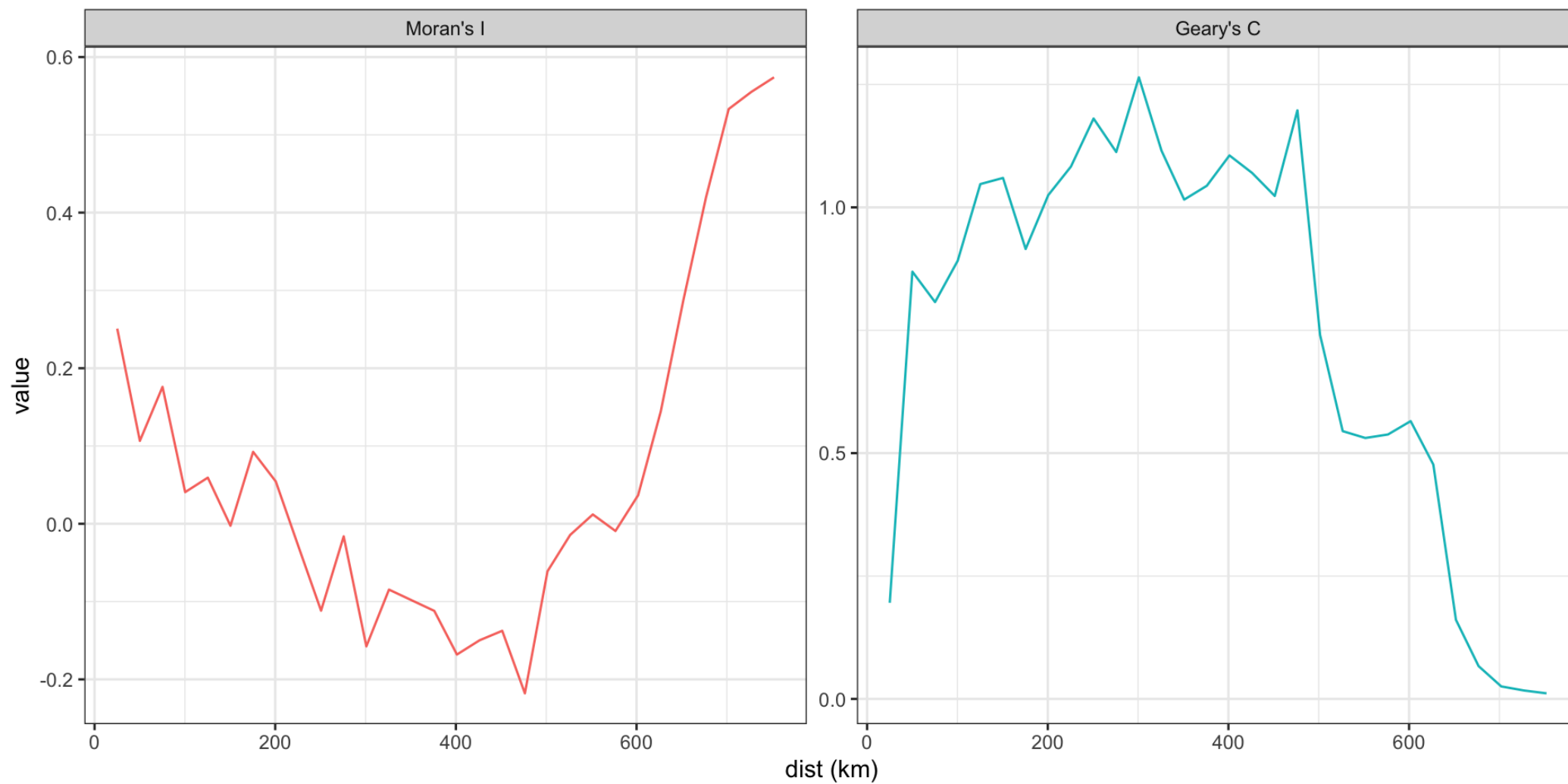
In this case we will define the elements of the adjacency matrix as

$$\{\mathbf{A}\}_{ij} = \mathbb{1}_{\{\text{distance between } s_i \text{ and } s_j \text{ is less than } d\}}$$

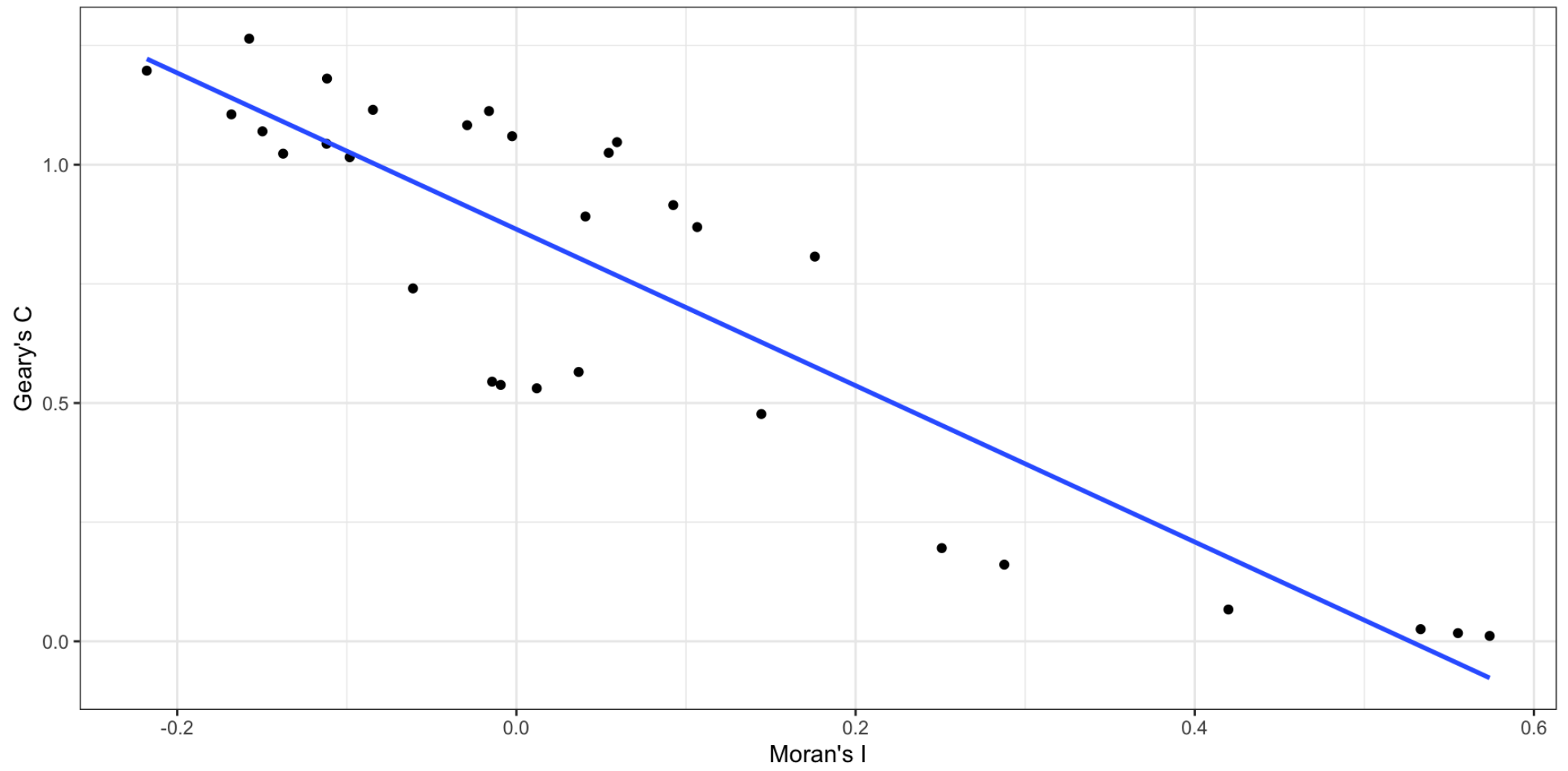
We can then construct a correlogram by varying d and then calculating Moran's I and Geary's C for each \mathbf{A} .

```
1 nc_pt = st_centroid(nc)
2 plot(nc_pt[, "SID74"], pch=16)
```





Inverse correlation of Moran's I and Geary's C



Autoregressive Models

AR Models - Time

Lets return to the simplest case, an $AR(1)$ process

$$y_t = \delta + \phi y_{t-1} + w_t$$

where $w_t \sim N(0, \sigma_w^2)$ and $|\phi| < 1$, then

$$E(y_t) = \frac{\delta}{1 - \phi}$$

$$Var(y_t) = \frac{\sigma^2}{1 - \phi}$$

$$\gamma(h) = \phi^h \frac{\sigma^2}{1 - \phi}$$

$$\rho(h) = \phi^h$$

AR Models - Time - Joint Distribution

We also saw that this $AR(1)$ model can be represented using a multivariate normal distribution,

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \sim N \left(\frac{\delta}{1-\phi} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \frac{\sigma^2}{1-\phi} \begin{pmatrix} 1 & \phi & \cdots & \phi^{n-1} \\ \phi & 1 & \cdots & \phi^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi^{n-1} & \phi^{n-2} & \cdots & 1 \end{pmatrix} \right)$$

In writing down the likelihood we also saw that an $AR(1)$ is 1st order Markovian,

$$\begin{aligned} f(y_1, \dots, y_n) &= f(y_1) f(y_2 \mid y_1) f(y_3 \mid y_2, y_1) \cdots f(y_n \mid y_{n-1}, y_{n-2}, \dots, y_1) \\ &= f(y_1) f(y_2 \mid y_1) f(y_3 \mid y_2) \cdots f(y_n \mid y_{n-1}) \end{aligned}$$

Alternative Definitions for y_t

$$y_t = \delta + \phi y_{t-1} + w_t$$

vs.

$$y_t \mid y_{t-1} \sim N(\delta + \phi y_{t-1}, \sigma^2)$$

In the case of time, both of these definitions result in the same multivariate distribution for \mathbf{y} given on the previous slide.

AR in Space

• s1	• s2	• s3	• s4	• s5	• s6	• s7	• s8	• s9	• s10
---------	---------	---------	---------	---------	---------	---------	---------	---------	----------

Even in the simplest spatial case there is no clear / unique ordering,

$$\begin{aligned} f(y(s_1), \dots, y(s_{10})) &= f(y(s_1)) f(y(s_2) \mid y(s_1)) \cdots f(y(s_{10}) \mid y(s_9), y(s_8), \dots, y(s_1)) \\ &= f(y(s_{10})) f(y(s_9) \mid y(s_{10})) \cdots f(y(s_1) \mid y(s_2), y(s_3), \dots, y(s_{10})) \\ &= ? \end{aligned}$$

Instead we need to think about things in terms of their neighbors / neighborhoods. We define $N(s_i)$ to be the set of neighbors of location s_i .

Defining the Spatial AR model

Here we will consider a simple average of neighboring observations, just like with the temporal AR model we have two options in terms of defining the autoregressive process,

- Simultaneous Autoregressive (SAR)

$$y(s) = \delta + \phi \frac{1}{|N(s)|} \sum_{s' \in N(s)} y(s') + N(0, \sigma^2)$$

- Conditional Autoregressive (CAR)

$$y(s) \mid y(-s) \sim N \left(\delta + \phi \frac{1}{|N(s)|} \sum_{s' \in N(s)} y(s'), \sigma^2 \right)$$

SAR

Simultaneous Autoregressive (SAR)

Using

$$y(s) = \phi \frac{1}{|N(s)|} \sum_{s' \in N(s)} y(s') + N(0, \sigma^2)$$

we want to find the distribution of $\mathbf{y} = \left(y(s_1), y(s_2), \dots, y(s_n) \right)^t$.

First we can define a weight matrix \mathbf{W} where

$$\{\mathbf{W}\}_{ij} = \begin{cases} 1/|N(s_i)| & \text{if } j \in N(s_i) \\ 0 & \text{otherwise} \end{cases}$$

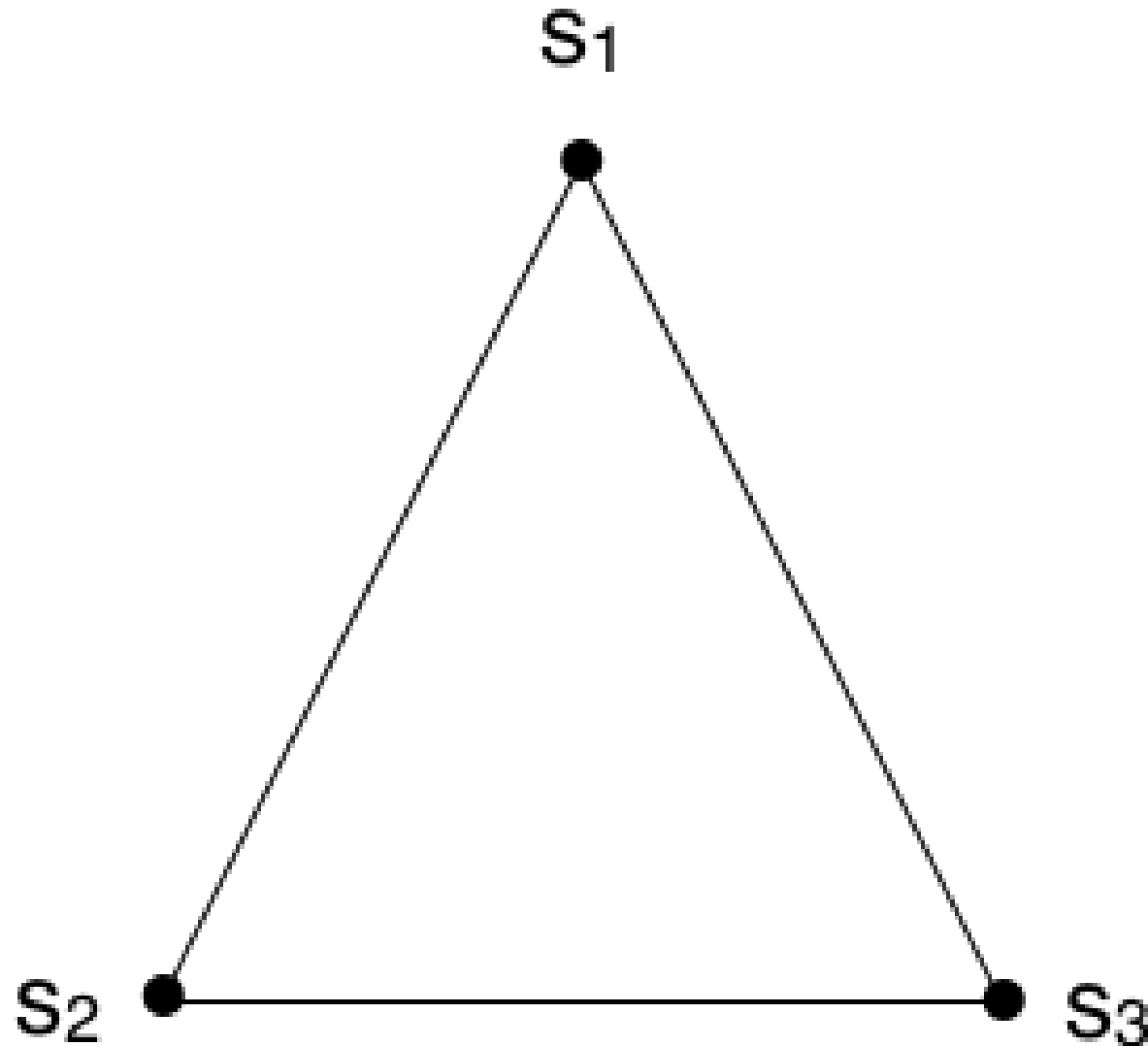
then we can write \mathbf{y} as follows,

$$\mathbf{y} = \phi \mathbf{W} \mathbf{y} + \boldsymbol{\epsilon}$$

where

$$\boldsymbol{\epsilon} \sim N(0, \sigma^2 \mathbf{I})$$

A toy example



$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

$$\mathbf{W} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

Back to SAR

$$\mathbf{y} = \phi \mathbf{W} \mathbf{y} + \epsilon$$

CAR

Conditional Autogressive (CAR)

This is a bit trickier, in the case of the temporal AR process we actually went from joint distribution \rightarrow conditional distributions (which we were then able to simplify).

Since we don't have a natural ordering we can't get away with this (at least not easily).

Going the other way, conditional distributions \rightarrow joint distribution is difficult because it is possible to specify conditional distributions that lead to an improper joint distribution.

Brooks' Lemma

For sets of observations \mathbf{x} and \mathbf{y} where $p(x) > 0 \ \forall x \in \mathbf{x}$ and $p(y) > 0 \ \forall y \in \mathbf{y}$ then

$$\begin{aligned} \frac{p(\mathbf{y})}{p(\mathbf{x})} &= \prod_{i=1}^n \frac{p(y_i \mid y_1, \dots, y_{i-1}, x_{i+1}, \dots, x_n)}{p(x_i \mid y_1, \dots, y_{i-1}, x_{i+1}, \dots, x_n)} \\ &= \prod_{i=1}^n \frac{p(y_i \mid x_1, \dots, x_{i-1}, y_{i+1}, \dots, y_n)}{p(x_i \mid x_1, \dots, x_{i-1}, y_{i+1}, \dots, y_n)} \end{aligned}$$

Derivation for $n = 2$

Let $\mathbf{y} = (y_1, y_2)$ and $\mathbf{x} = (x_1, x_2)$ then we can derive Brook's Lemma for this case,

$$\begin{aligned} p(y_1, y_2) &= p(y_1 \mid y_2) p(y_2) \\ &= p(y_1 \mid y_2) \frac{p(y_2 \mid x_1)}{p(x_1 \mid y_2)} p(x_1) \\ &= \frac{p(y_1 \mid y_2)}{p(x_1 \mid y_2)} p(y_2 \mid x_1) p(x_1) \\ &= \frac{p(y_1 \mid y_2)}{p(x_1 \mid y_2)} p(y_2 \mid x_1) p(x_1) \left(\frac{p(x_2 \mid x_1)}{p(x_2 \mid x_1)} \right) \\ &= \frac{p(y_1 \mid y_2)}{p(x_1 \mid y_2)} \frac{p(y_2 \mid x_1)}{p(x_2 \mid x_1)} p(x_1, x_2) \end{aligned}$$

$$p(y_1, y_2) = \frac{p(y_1 \mid y_2)}{p(x_1 \mid y_2)} \frac{p(y_2 \mid x_1)}{p(x_2 \mid x_1)} p(x_1, x_2)$$

$$\frac{p(y_1, y_2)}{p(x_1, x_2)} = \frac{p(y_1 \mid y_2)}{p(x_1 \mid y_2)} \frac{p(y_2 \mid x_1)}{p(x_2 \mid x_1)}$$

From which we can generalize for $n = 3$,

$$\frac{p(y_1, y_2, y_3)}{p(x_1, x_2, x_3)} = \frac{p(y_1 \mid y_2, y_3)}{p(x_1 \mid y_2, y_3)} \frac{p(y_2 \mid x_1, y_3)}{p(x_2 \mid x_1, y_3)} \frac{p(y_3 \mid x_1, x_2)}{p(x_3 \mid x_1, x_2)}$$

and so on.

Usefulness?

Lets repeat that last example but consider the case where $\mathbf{y} = (y_1, y_2)$ but now we let $\mathbf{x} = (y_1 = 0, y_2 = 0)$

$$\frac{p(y_1, y_2)}{p(x_1, x_2)} = \frac{p(y_1, y_2)}{p(y_1 = 0, y_2 = 0)}$$

$$p(y_1, y_2) = \frac{p(y_1 \mid y_2)}{p(y_1 = 0 \mid y_2)} \frac{p(y_2 \mid y_1 = 0)}{p(y_2 = 0 \mid y_1 = 0)} p(y_1 = 0, y_2 = 0)$$

$$p(y_1, y_2) \propto \frac{p(y_1 \mid y_2) p(y_2 \mid y_1 = 0)}{p(y_1 = 0 \mid y_2)}$$

As applied to a simple CAR model



$$y(s_1) \mid y(s_2) \sim N(\phi W_{12} y(s_2), \sigma^2)$$

$$y(s_2) \mid y(s_1) \sim N(\phi W_{21} y(s_1), \sigma^2)$$

$$\begin{aligned}
p(y(s_1), y(s_2)) &\propto \frac{p(y(s_1) \mid y(s_2)) p(y(s_2) \mid y(s_1) = 0)}{p(y(s_1) = 0 \mid y(s_2))} \\
&\propto \frac{\exp\left(-\frac{1}{2\sigma^2} (y(s_1) - \phi W_{12} y(s_2))^2\right) \exp\left(-\frac{1}{2\sigma^2} (y(s_2) - \phi W_{21} 0)^2\right)}{\exp\left(-\frac{1}{2\sigma^2} (0 - \phi W_{12} y(s_2))^2\right)} \\
&\propto \exp\left(-\frac{1}{2\sigma^2} \left((y(s_1) - \phi W_{12} y(s_2))^2 + y(s_2)^2 - (\phi W_{21} y(s_2))^2\right)\right) \\
&\propto \exp\left(-\frac{1}{2\sigma^2} \left(y(s_1)^2 - \phi W_{12} y(s_1) y(s_2) - \phi W_{21} y(s_1) y(s_2) + y(s_2)^2\right)\right) \\
&\propto \exp\left(-\frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{0}) \begin{pmatrix} 1 & -\phi W_{12} \\ -\phi W_{21} & 1 \end{pmatrix} (\mathbf{y} - \mathbf{0})^t\right)
\end{aligned}$$

Implications for \mathbf{y}

$$\boldsymbol{\mu} = \mathbf{0}$$

$$\begin{aligned}\boldsymbol{\Sigma}^{-1} &= \frac{1}{\sigma^2} \begin{pmatrix} 1 & -\phi W_{12} \\ -\phi W_{21} & 1 \end{pmatrix} \\ &= \frac{1}{\sigma^2} (\mathbf{I} - \phi \mathbf{W})\end{aligned}$$

$$\boldsymbol{\Sigma} = \sigma^2 (\mathbf{I} - \phi \mathbf{W})^{-1}$$

we can then conclude that for $\mathbf{y} = (y(s_1), y(s_2))^t$,

$$\mathbf{y} \sim N(\mathbf{0}, \sigma^2 (\mathbf{I} - \phi \mathbf{W})^{-1})$$

which generalizes for all mean 0 CAR models.

General Proof

Let $\mathbf{y} = (y(s_1), \dots, y(s_n))$ and $\mathbf{0} = (y(s_1) = 0, \dots, y(s_n) = 0)$ then by Brook's lemma,

$$\begin{aligned}
 \frac{p(\mathbf{y})}{p(\mathbf{0})} &= \prod_{i=1}^n \frac{p(y_i \mid y_1, \dots, y_{i-1}, 0_{i+1}, \dots, 0_n)}{p(0_i \mid y_1, \dots, y_{i-1}, 0_{i+1}, \dots, 0_n)} = \prod_{i=1}^n \frac{\exp \left(-\frac{1}{2\sigma^2} \left(y_i - \phi \sum_{j<i} W_{ij} y_j - \phi \sum_{j>i} 0_j \right)^2 \right)}{\exp \left(-\frac{1}{2\sigma^2} \left(0_i - \phi \sum_{j<i} W_{ij} y_j - \phi \sum_{j>i} 0_j \right)^2 \right)} \\
 &= \exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n \left(y_i - \phi \sum_{j<i} W_{ij} y_j \right)^2 + \frac{1}{2\sigma^2} \sum_{i=1}^n \left(\phi \sum_{j<i} W_{ij} y_j \right)^2 \right) \\
 &= \exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n y_i^2 - 2\phi y_i \sum_{j<i} W_{ij} y_j \right) \\
 &= \exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n y_i^2 - \phi \sum_{i=1}^n \sum_{j=1}^n y_i W_{ij} y_j \right) \quad (\text{if } W_{ij} = W_{ji}) \\
 &= \exp \left(-\frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{0})^t (\mathbf{I} - \phi \mathbf{W}) (\mathbf{y} - \mathbf{0}) \right)
 \end{aligned}$$

CAR vs SAR

- Simultaneous Autoregressive (SAR)

$$y(s) = \phi \sum_{s'} W_{s s'} y(s') + \epsilon$$

$$\mathbf{y} \sim N(0, \sigma^2 ((\mathbf{I} - \phi \mathbf{W})^{-1})(\mathbf{I} - \phi \mathbf{W})^{-1})$$

- Conditional Autoregressive (CAR)

$$y(s) \mid \mathbf{y}_{-s} \sim N \left(\sum_{s'} W_{s s'} y(s'), \sigma^2 \right)$$

$$\mathbf{y} \sim N(0, \sigma^2 (\mathbf{I} - \phi \mathbf{W})^{-1})$$

Generalizations

- Adopting different weight matrices (\mathbf{W})
 - Between SAR and CAR model we move to a generic weight matrix definition (beyond average of nearest neighbors)
 - In time we varied p in the $AR(p)$ model, in space we adjust the weight matrix.
 - In general having a symmetric \mathbf{W} is helpful, but not required
- More complex Variance (beyond $\sigma^2 I$)
 - σ^2 can be a vector (differences between areal locations)
 - i.e. since areal data tends to be aggregated - adjust variance based on sample size
 - i.e. scale based on the number of neighbors

Some specific generalizations

Generally speaking we want to work with a scaled / normalized version of the weight matrix,

$$W_{ij}/W_i.$$

When W is derived from an adjacency matrix, A , we can express this as

$$W = D^{-1} A$$

where $D^{-1} = \text{diag}(1/|N(s_i)|)$.

We can also allow σ^2 to vary between locations, we can define this as $D_{\sigma^2} = \text{diag}(\sigma_i^2)$ and most often we use

$$D_{\sigma^2}^{-1} = \text{diag} \left(\frac{\sigma^2}{|N(s_i)|} \right) = \sigma^2 D^{-1}.$$

Revised SAR Model

- Formula Model

$$y(s_i) = X_{i.}\beta + \phi \sum_{j=1}^n D_{jj}^{-1} A_{ij} (y(s_j) - X_{j.}\beta) + \epsilon_i$$

$$\boldsymbol{\epsilon} \sim N(\mathbf{0}, \mathbf{D}_{\sigma^2}^{-1}) = N(\mathbf{0}, \sigma^2 \mathbf{D}^{-1})$$

- Joint Model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \phi \mathbf{D}^{-1} \mathbf{A} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + \boldsymbol{\epsilon}$$

$$\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, (\mathbf{I} - \phi \mathbf{D}^{-1} \mathbf{A})^{-1} \sigma^2 \mathbf{D}^{-1} ((\mathbf{I} - \phi \mathbf{D}^{-1} \mathbf{A})^{-1})^t)$$

Revised CAR Model

- Conditional Model

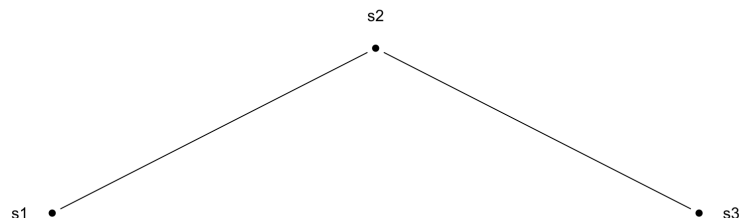
$$y(s_i) \mid y_{-s_i} \sim N \left(X_{i\cdot} \beta + \phi \sum_{j=1}^n \frac{W_{ij}}{D_{ii}} (y(s_j) - X_{j\cdot} \beta), \sigma^2 D_{ii}^{-1} \right)$$

- Joint Model

$$\mathbf{y} \sim N(\mathbf{X}\beta, \Sigma_{CAR})$$

$$\begin{aligned} \Sigma_{CAR} &= (\mathbf{D}_\sigma (\mathbf{I} - \phi \mathbf{D}^{-1} \mathbf{A}))^{-1} \\ &= (1/\sigma^2 \mathbf{D} (\mathbf{I} - \phi \mathbf{D}^{-1} \mathbf{A}))^{-1} \\ &= (1/\sigma^2 (\mathbf{D} - \phi \mathbf{A}))^{-1} \\ &= \sigma^2 (\mathbf{D} - \phi \mathbf{A})^{-1} \end{aligned}$$

Toy CAR Example



$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{\Sigma} = \sigma^2 (\mathbf{D} - \phi \mathbf{W}) = \sigma^2 \begin{pmatrix} 1 & -\phi & 0 \\ -\phi & 2 & -\phi \\ 0 & -\phi & 1 \end{pmatrix}^{-1}$$

When does Σ exist?

```
1 check_sigma = function(phi) {  
2   Sigma_inv = matrix(c(1,-phi,0,-phi,2,-phi,0,-phi,1), ncol=3, byrow=TRUE)  
3   solve(Sigma_inv)  
4 }  
5  
6 check_sigma(phi=0)
```

	[,1]	[,2]	[,3]
[1,]	1	0.0	0
[2,]	0	0.5	0
[3,]	0	0.0	1

```
1 check_sigma(phi=0.5)
```

	[,1]	[,2]	[,3]
[1,]	1.1666667	0.3333333	0.1666667
[2,]	0.3333333	0.6666667	0.3333333
[3,]	0.1666667	0.3333333	1.1666667

```
1 check_sigma(phi=-0.6)
```

	[,1]	[,2]	[,3]
[1,]	1.28125	-0.46875	0.28125
[2,]	-0.46875	0.78125	-0.46875
[3,]	0.28125	-0.46875	1.28125

```
1 check_sigma(phi=1)
```

Error in solve.default(Sigma_inv): Lapack routine dgesv: system is exactly singular: U[3,3] = 0

```
1 check_sigma(phi=-1)
```

Error in solve.default(Sigma_inv): Lapack routine dgesv: system is exactly singular: U[3,3] = 0

```
1 check_sigma(phi=1.2)
```

	[,1]	[,2]	[,3]
[1,]	-0.6363636	-1.363636	-1.6363636
[2,]	-1.3636364	-1.136364	-1.3636364
[3,]	-1.6363636	-1.363636	-0.6363636

```
1 check_sigma(phi=-1.2)
```

	[,1]	[,2]	[,3]
[1,]	-0.6363636	1.363636	-1.6363636
[2,]	1.3636364	-1.136364	1.3636364
[3,]	-1.6363636	1.363636	-0.6363636

When is Σ positive semidefinite?

```
1 check_sigma_pd = function(phi) {  
2   Sigma_inv = matrix(c(1,-phi,0,-phi,2,-phi,0,-phi,1), ncol=3, byrow=TRUE)  
3   chol(solve(Sigma_inv))  
4 }  
5  
6 check_sigma_pd(phi=0)
```

```
      [,1]      [,2] [,3]  
[1,]      1 0.0000000      0  
[2,]      0 0.7071068      0  
[3,]      0 0.0000000      1
```

```
1 check_sigma_pd(phi=0.5)
```

```
      [,1]      [,2]      [,3]  
[1,] 1.080123 0.3086067 0.1543033  
[2,] 0.000000 0.7559289 0.3779645  
[3,] 0.000000 0.0000000 1.0000000
```

```
1 check_sigma_pd(phi=-0.6)
```

```
      [,1]      [,2]      [,3]  
[1,] 1.131923 -0.4141182 0.2484709  
[2,] 0.000000 0.7808688 -0.4685213  
[3,] 0.000000 0.0000000 1.0000000
```

```
1 check_sigma_pd(phi=1)
```

Error in solve.default(Sigma_inv): Lapack routine dgesv: system is exactly singular: U[3,3] = 0

```
1 check_sigma_pd(phi=-1)
```

Error in solve.default(Sigma_inv): Lapack routine dgesv: system is exactly singular: U[3,3] = 0

```
1 check_sigma_pd(phi=1.2)
```

Error in chol.default(solve(Sigma_inv)): the leading minor of order 1 is not positive

```
1 check_sigma_pd(phi=-1.2)
```

Error in chol.default(solve(Sigma_inv)): the leading minor of order 1 is not positive

Conclusions

Generally speaking just like the AR(1) model for time series we require that $|\phi| < 1$ for the CAR model to be proper.

These results for ϕ also apply in the context where σ_i^2 is constant across locations, i.e.

$$\Sigma = \left(\sigma^2 (\mathbf{I} - \phi \mathbf{D}^{-1} \mathbf{A}) \right)^{-1}$$

As a side note, the special case where $\phi = 1$ is known as an intrinsic autoregressive (IAR) model and they are popular as an *improper* prior for spatial random effects. An additional sum constraint is necessary for identifiability

$$\sum_{i=1}^n y(s_i) = 0$$

