

# LECTURE 14: SPATIAL GLMS

# CLASS INTRO

# INTRO QUESTIONS

- Describe geometric anisotropy. Discuss this in the context of a data set.
- For Today:
  - Finish geometric anisotropy
  - Spatial Generalized Linear Models

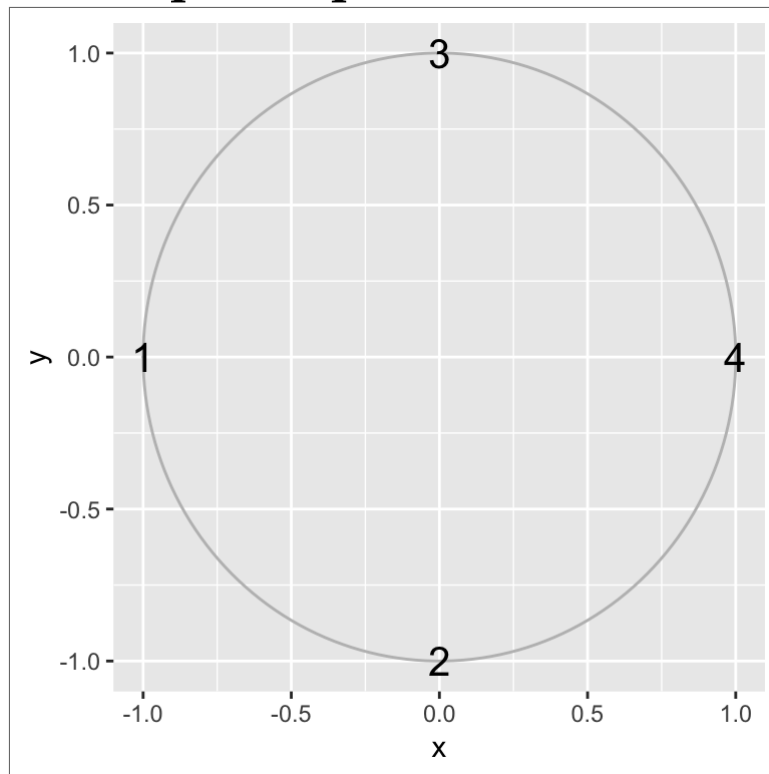
# GEOMETRIC ANISOTROPY

## GEOMETRIC ANISOTROPY MODEL

- Let  $Y(s) = \mu(s) + w(s) + \epsilon(s)$ , thus  $Y(s) \sim N(\mu(s), \Sigma(\tau^2, \sigma^2, \phi, B))$ , where  $B = L^T L$ .
- The covariance matrix is defined as  $\Sigma(\tau^2, \sigma^2, \phi, B) = \tau^2 I + \sigma^2 H((\mathbf{h}^T B \mathbf{h}^T)^{\frac{1}{2}})$ , where  $H((\mathbf{h}^T B \mathbf{h}^T)^{\frac{1}{2}})$  has entries of  $\rho((\mathbf{h}_{ij}^T B \mathbf{h}_{ij}^T)^{\frac{1}{2}})$  with  $\rho()$  being a valid correlation function, typically including  $\phi$  and  $\mathbf{h}_{ij} = s_i - s_j$ .

# GEOMETRIC ANISOTROPY VISUAL

- Consider four points positioned on a unit circle.



- How far apart are each set of points?

## GEOMETRIC ANISOTROPY EXERCISE 1

Now consider a set of correlation functions. For each, calculate the correlation matrix and discuss the impact of  $B$  on the correlation. Furthermore, how does  $B$  change the geometry of the correlation?

1.  $\rho() = \exp(-\mathbf{h}_{ij}^T B \mathbf{h}_{ij}^T)^{\frac{1}{2}})$ , where  $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

2.  $\rho() = \exp(-\mathbf{h}_{ij}^T B \mathbf{h}_{ij}^T)^{\frac{1}{2}})$ , where  $B = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$

3.  $\rho() = \exp(-\mathbf{h}_{ij}^T B \mathbf{h}_{ij}^T)^{\frac{1}{2}})$ , where  $B = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}$

# GEOMETRIC ANISOTROPY: SOLUTION 1

$$\rho() = \exp(-\mathbf{h}_{ij}^T I \mathbf{h}_{ij}^T)^{\frac{1}{2}})$$

Implied Distance

0.00	1.41	1.41	2.00
1.41	0.00	2.00	1.41
1.41	2.00	0.00	1.41
2.00	1.41	1.41	0.00

Correlation

1.000	0.243	0.243	0.135
0.243	1.000	0.135	0.243
0.243	0.135	1.000	0.243
0.135	0.243	0.243	1.000



## GEOMETRIC ANISOTROPY: SOLUTION 2

$$\rho() = \exp(-\mathbf{h}_{ij}^T B \mathbf{h}_{ij}^T)^{\frac{1}{2}}), \text{ where } B = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

Implied Distance

0.00	1.73	1.73	2.83
1.73	0.00	2.00	1.73
1.73	2.00	0.00	1.73
2.83	1.73	1.73	0.00

Correlation

1.000	0.177	0.177	0.059
0.177	1.000	0.135	0.177
0.177	0.135	1.000	0.177
0.059	0.177	0.177	1.000



## GEOMETRIC ANISOTROPY: SOLUTION 3

$$\rho() = \exp(-\mathbf{h}_{ij}^T B \mathbf{h}_{ij}^T)^{\frac{1}{2}}), \text{ where } B = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}$$

Implied Distance

0.00	1.41	2.45	3.46
1.41	0.00	2.00	2.45
2.45	2.00	0.00	1.41
3.46	2.45	1.41	0.00

Correlation

1.000	0.243	0.086	0.031
0.243	1.000	0.135	0.086
0.086	0.135	1.000	0.243
0.031	0.086	0.243	1.000



## MORE GEOMETRIC ANISOTROPY

- The matrix  $B$  relates to the orientation of a transformed ellipse.
- The (effective) range for any angle  $\eta$  is determined by the equation

$$\rho(r_\eta(\tilde{\mathbf{h}}_\eta^T B \tilde{\mathbf{h}}_\eta)^{\frac{1}{2}}) = .05,$$

where  $\tilde{\mathbf{h}}_\eta$  is a unit vector in the direction  $\eta$ .

# FITTING GEOMETRIC ANISOTROPY MODELS

- Okay, so if we suspect that geometric anisotropy is present, how do we fit the model? That is, what is necessary in estimating this model?
- In addition to  $\sigma^2$  and  $\tau^2$  we need to fit  $B$ .
- What about  $\phi$ ? What is  $\phi$  when  $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ?

## PRIORS FOR $B$

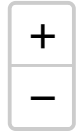
- While  $B$  is a matrix, it is just another unknown parameter.
- Hence, to fit a Bayesian model we need a prior distribution for  $B$ .
- One option for the positive definite matrix is the Wishart distribution, which is a bit like a matrix-variate gamma distribution.

# SPATIAL GLMS



# DATA VIZ MOTIVATION

## OZONE EXCEEDANCE IN COLORADO





# GENERALIZED LINEAR MODEL NOTATION

There are three components to a generalized linear model:

1. Sampling Distribution: such as Poisson or Binomial
2. Linear combination of predictors:  $\eta = X\beta$
3. A link function to map the linear combination of predictors to the support of the sampling distribution.

# LOGISTIC REGRESSION OVERVIEW

Write out the complete model specification for logistic regression.

- Assume  $Y_i$  is the binary response for the  $i^{th}$  observation,

$$Y_i \sim \text{Bernoulli}(\pi_i)$$

$$\text{logit}(\pi_i) = X_i\beta,$$

- where  $\text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right)$

# GLM ( )

## Interpret this output

```
CO <- CO %>% mutate(north = as.numeric(Latitude > 38 ))
glm(Exceedance~north, family=binomial(link = 'logit'),data=CO) %>% summary
```

```
##
## Call:
## glm(formula = Exceedance ~ north, family = binomial(link = "logit"),
##      data = CO)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.313    0.378    0.378    0.378    1.177
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.695e-15  8.165e-01   0.000   1.0000
## north        2.603e+00  1.097e+00   2.372   0.0177 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 28.708  on 34  degrees of freedom
## Residual deviance: 22.873  on 33  degrees of freedom
## AIC: 26.873
##
## Number of Fisher Scoring iterations: 5
```



# LOGISTIC REGRESSION IN JAGS

## JAGS specification for logistic regression

```
model{
  # likelihood
  for (i in 1:N){
    y[i] ~ dbern(p[i])
    p[i] <- 1 / (1 + exp(-eta[i]))
    eta[i] <- int + x[i] * north
  }
  # prior
  int ~ dnorm(0, 1E-12)
  north ~ dnorm(0, 1E-12)
}
```

# SPATIAL LOGISTIC REGRESSION

- Assume  $Y(s_i)$  is the binary response for  $s_i$ ,  
 $Y(s_i) | \beta, w(s_i) \sim \text{Bernoulli}(\pi(s_i))$   
 $\text{logit}(\pi(s_i)) = X(s_i)\beta + w(s_i),$
- where  $W \sim N(\mathbf{0}, \sigma^2 H(\phi))$



# SPATIAL LOGISTIC REGRESSION IN JAGS

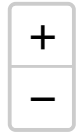
```
logistic.spatial <- "model {  
  for (i in 1:N) {  
    Y[i] ~ dbern(p[i])  
    logit(p[i]) <- mu[i] + w[i]  
    mu[i] <- beta[1]+beta[2]*x.north[i]  
    muW[i] <- 0  
  }  
  # process  
  w[1:N] ~ dmnorm(muW[],Omega[,])  
  
  # priors  
  beta[1] ~ dnorm(0.0,1E-8)  
  beta[2] ~ dnorm(0.0,1E-8)  
  spat.prec ~ dgamma(0.001, 0.001)  
  sigmasq <- 1/spat.prec  
  phi ~ dunif(1.5,2.5)  
  
  #build omega  
  for (i in 1:N){  
    for (j in 1:N){  
      H[i,j] <- (1/ spat.prec) * exp(-phi * d[i,j])  
    }  
  }  
  Omega[1:N,1:N] <- inverse(H[1:N,1:N])  
}"
```

## SPGLM( )

```
m.1 <- spGLM(y~1,
              family="binomial",
              coords=coords,
              weights=weights,
              starting=list("beta"=beta.starting, "phi"=0.06,"sigma.sq"=1, "w"=0),
              tuning=list("beta"=beta.tuning, "phi"=0.5, "sigma.sq"=0.5, "w"=0.5),
              priors=list("beta.Normal"=list(0,10), "phi.Unif"=c(0.03, 0.3), "sigma.sq"=1),
              amcmc=list("n.batch"=n.batch, "batch.length"=batch.length, "accept.rate"=0.5),
              cov.model="exponential",
              verbose=TRUE,
              n.report=10)
```

# **SPATIAL POISSON REGRESSION**

# MOTIVATION



**Leaflet**



# POISSON REGRESSION OVERVIEW

Write out the complete model specification for Poisson regression.

- Assume  $Y_i$  is the count response for the  $i^{th}$  observation,

$$Y_i \sim \text{Poisson}(\lambda_i)$$

$$\log(\lambda_i) = X_i\beta,$$

- thus  $\exp(X_i\beta) \geq 0$
- next write out a Poisson regression model with spatial random effects

# SPATIAL POISSON REGRESSION

- Assume  $Y_i$  is the count response for  $s_i$ ,

$$Y(s_i) \sim \text{Poisson}(\lambda(s_i))$$

$$\log(\lambda(s_i)) = X(s_i)\beta + w(s_i),$$

- where  $\mathbf{W} \sim N(\mathbf{0}, \sigma^2 H(\phi))$

# FITTING SPATIAL POISSON REGRESSION MODELS

1. `spGLM( )` accepts binomial and Poisson data,
2. JAGS can also be used to fit spatial GLMS models, or
3. Try writing a sampler from scratch.