Lecture 1

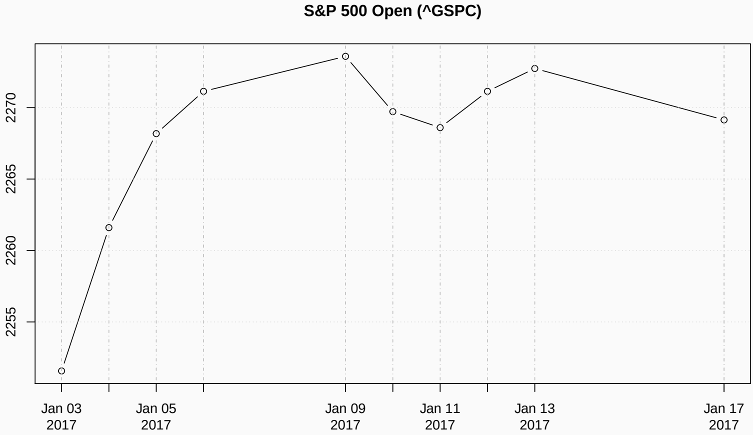
Spatio-temporal data & Linear Models

Colin Rundel

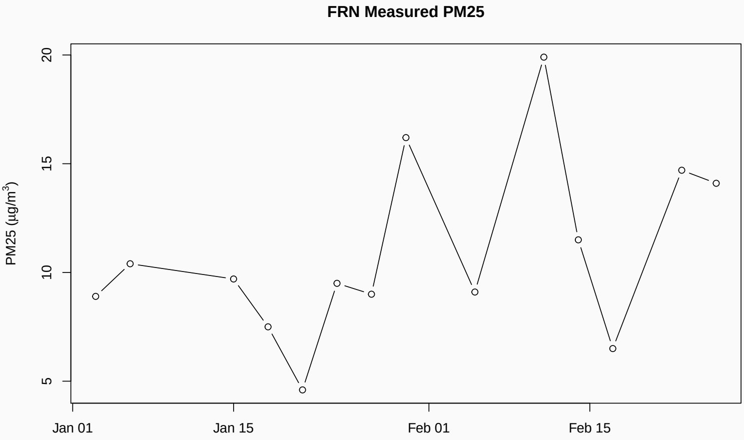
1/18/2017

Spatio-temporal data

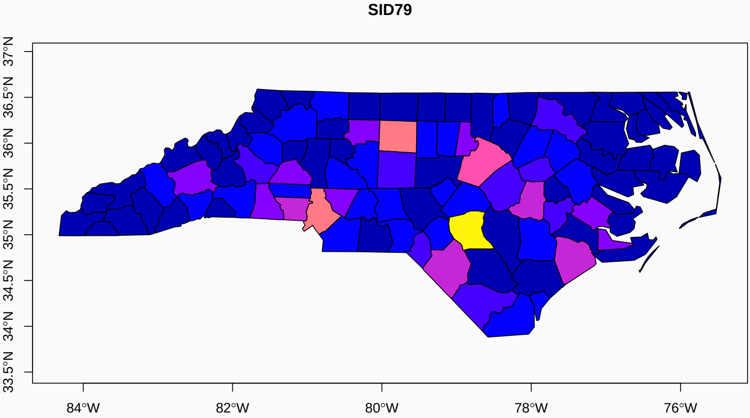
Time Series Data - Discrete

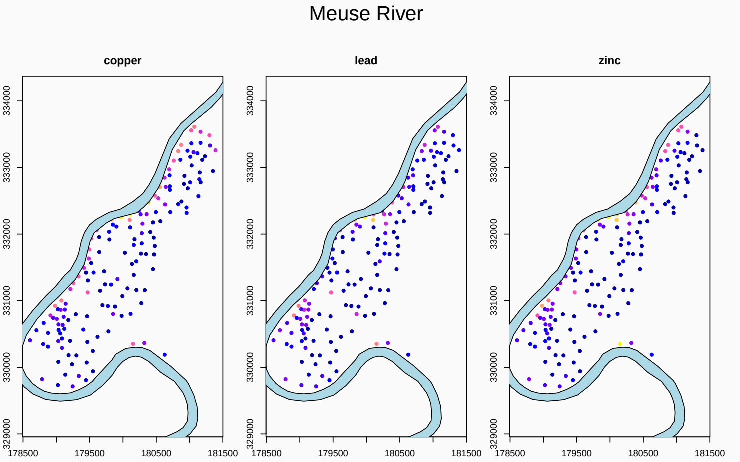


Time Series Data - Continuous

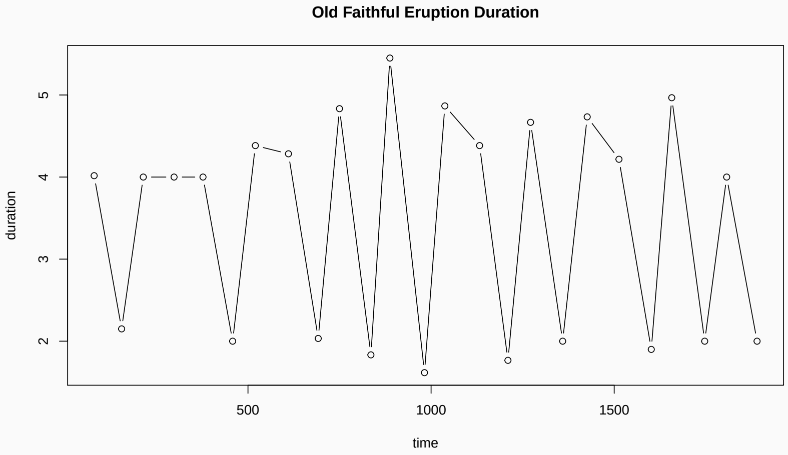


Spatial Data - Areal

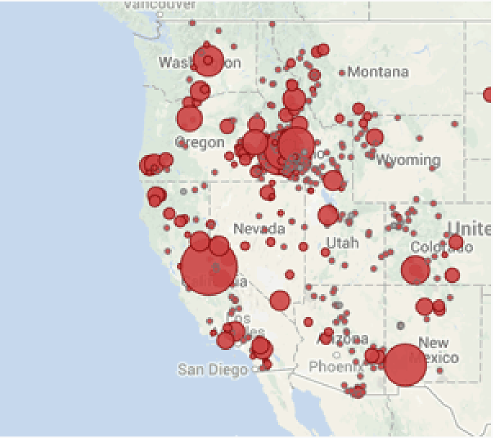



Meuse River
copper
lead
zinc
334000
333000
332000
331000
330000
329000
178500
179500
180500
181500

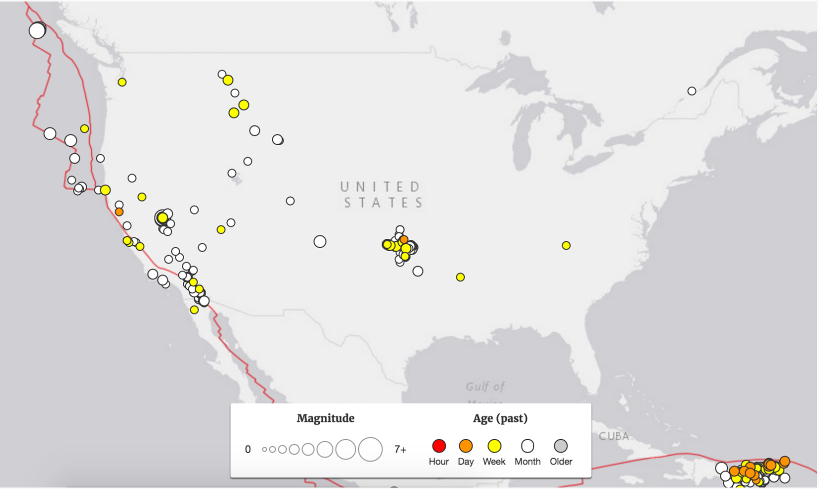
Point Pattern Data - Time



Point Pattern Data - Space



Point Pattern Data - Space + Time



(Bayesian) Linear Models

Pretty much everything we a going to see in this course will fall under the umbrella of linear or generalized linear models.

$$
\begin{aligned}
Y_i &= \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \epsilon_i \\
\epsilon_i &\sim N(0, \sigma^2)
\end{aligned}
$$

which we can also express using matrix notation as

$$
\underset{n\times 1}{\boldsymbol{Y}} = \underset{n\times p}{\boldsymbol{X}}\, \underset{p\times 1}{\boldsymbol{\beta}} + \underset{n\times 1}{\boldsymbol{\epsilon}}
$$

$$
\boldsymbol{\epsilon} \sim N(\underset{n\times 1}{\boldsymbol{0}},\ \sigma^2 \underset{n\times n}{\mathbb{I}_n})
$$

For an n-dimension multivate normal distribution with covariance $\mathbf{\Sigma}$ (positive semidefinite) can be written as

$$\underset{n\times 1}{\boldsymbol{Y}} \sim N(\underset{n\times 1}{\boldsymbol{\mu}}, \underset{n\times n}{\mathbf{\Sigma}}) \text{ where } \{\mathbf{\Sigma}\}_{ij} = \rho_{ij}\sigma_i\sigma_j$$

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}, \begin{pmatrix} \rho_{11}\sigma_1\sigma_1 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \vdots & \ddots & \vdots \\ \rho_{n1}\sigma_n\sigma_1 & \cdots & \rho_{nn}\sigma_n\sigma_n \end{pmatrix}\right)$$

Multivariate Normal Distribution - Density

For the n dimensional multivate normal given on the last slide, its density is given by

$$(2\pi)^{-n/2} \;\det(\boldsymbol{\Sigma})^{-1/2} \;\exp\left(-\frac{1}{2}\underset{1\times n}{(Y-\boldsymbol{\mu})'}\underset{n\times n}{\boldsymbol{\Sigma}^{-1}}\underset{n\times 1}{(Y-\boldsymbol{\mu})}\right)$$

and its log density is given by

$$-\frac{n}{2}\log 2\pi - \frac{1}{2}\log\det(\boldsymbol{\Sigma}) - -\frac{1}{2}\underset{1\times n}{(Y-\boldsymbol{\mu})'}\underset{n\times n}{\boldsymbol{\Sigma}^{-1}}\underset{n\times 1}{(Y-\boldsymbol{\mu})}$$

A Simple Linear Regression Example

Lets generate some simulated data where the underlying model is known and see how various regression preceedures function.

$$\beta_0 = 0.7, \quad \beta_1 = 1.5, \quad \beta_2 = -2.2, \quad \beta_3 = 0.1$$

$$n = 100, \quad \epsilon_i \sim N(0, 1)$$

Generating the data

```r
set.seed(01172017)
n = 100
beta = c(0.7,1.5,-2.2,0.1)
eps = rnorm(n)

X0 = rep(1, n)
X1 = rt(n,df=5)
X2 = rt(n,df=5)
X3 = rt(n,df=5)

X = cbind(X0, X1, X2, X3)
Y = X %*% beta + eps
d = data.frame(Y,X[,-1])
```

Least squares fit

Let $\hat{Y}$ be our estimate for Y based on our estimate of $\boldsymbol{\beta}$,

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 = X\,\hat{\boldsymbol{\beta}}$$

Let $\hat{Y}$ be our estimate for Y based on our estimate of $\boldsymbol{\beta}$,

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 = X\hat{\boldsymbol{\beta}}$$

The least squares estimate, $\hat{\boldsymbol{\beta}}_{ls}$, is given by

$$\underset{\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - X_{i\cdot}\boldsymbol{\beta}\right)^2$$

Let $\hat{Y}$ be our estimate for Y based on our estimate of $\boldsymbol{\beta}$,

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 = X\,\hat{\boldsymbol{\beta}}$$

The least squares estimate, $\hat{\boldsymbol{\beta}}_{ls}$, is given by

$$\underset{\boldsymbol{\beta}}{\arg\min} \sum_{i=1}^{n} \left(Y_i - X_{i\cdot}\boldsymbol{\beta}\right)^2$$

With a bit of calculus and algebra we can derive

$$\hat{\boldsymbol{\beta}}_{ls} = (X^t X)^{-1} X^t\, Y$$

Maximum Likelihood

Frequentist Fit

```r
lm(Y ~ ., data=d)$coefficients
```

```
## (Intercept)          X1          X2          X3
##  0.73726738  1.65321096 -2.16499958  0.07996257
```

```r
(beta_hat = solve(t(X) %*% X, t(X)) %*% Y)
```

```
##           [,1]
## X0  0.73726738
## X1  1.65321096
## X2 -2.16499958
## X3  0.07996257
```

$$Y_1, \ldots, Y_{100} \mid \boldsymbol{\beta}, \sigma^2 \sim N(\mathbf{X}_{i\cdot}\boldsymbol{\beta}, \sigma^2)$$

$$\beta_0, \beta_1, \beta_2, \beta_3 \sim N(0, \sigma^2_\beta = 100)$$

$$\tau^2 = 1/\sigma^2 \sim \text{Gamma}(a = 1, b = 1)$$

Deriving the posterior

$$
\begin{aligned}
[\beta_0, \beta_1, \beta_2, \beta_3, \sigma^2 \mid Y] &= \frac{[Y \mid \boldsymbol{\beta}, \sigma^2]}{[Y]}[\boldsymbol{\beta}, \sigma^2] \\
&\propto [Y \mid \boldsymbol{\beta}, \sigma^2][\boldsymbol{\beta}][\sigma^2]
\end{aligned}
$$

Deriving the posterior

$$
\begin{aligned}
[\beta_0, \beta_1, \beta_2, \beta_3, \sigma^2 \mid Y] &= \frac{[Y \mid \boldsymbol{\beta}, \sigma^2]}{[Y]}[\boldsymbol{\beta}, \sigma^2] \\
&\propto [Y \mid \boldsymbol{\beta}, \sigma^2][\boldsymbol{\beta}][\sigma^2]
\end{aligned}
$$

where,

$$
[Y \mid \boldsymbol{\beta}, \sigma^2] = \left(2\pi\sigma^2\right)^{-n/2} \exp\left(-\frac{\sum_{i=1}^{n}\left(Y_i - \beta_0 - \beta_1 X_{i,1} - \beta_1 X_{i,2} - \beta_3 X_{i,3}\right)^2}{2\sigma^2}\right)
$$

Deriving the posterior

$$
\begin{aligned}
[\beta_0, \beta_1, \beta_2, \beta_3, \sigma^2 \mid Y] &= \frac{[Y \mid \boldsymbol{\beta}, \sigma^2]}{[Y]}[\boldsymbol{\beta}, \sigma^2] \\
&\propto [Y \mid \boldsymbol{\beta}, \sigma^2][\boldsymbol{\beta}][\sigma^2]
\end{aligned}
$$

where,

$$
[Y \mid \boldsymbol{\beta}, \sigma^2] = \left(2\pi\sigma^2\right)^{-n/2} \exp\left(-\frac{\sum_{i=1}^{n}\left(Y_i - \beta_0 - \beta_1 X_{i,1} - \beta_1 X_{i,2} - \beta_3 X_{i,3}\right)^2}{2\sigma^2}\right)
$$

$$
[\beta_0, \beta_1, \beta_2, \beta_3 \mid \sigma^2_\beta] = \left(2\pi\sigma^2_\beta\right)^{-4/2} \exp\left(-\frac{\sum_{i=0}^{3}\beta_i^2}{2\sigma^2_\beta}\right)
$$

$$
\begin{aligned}
[\beta_0, \beta_1, \beta_2, \beta_3, \sigma^2 \mid Y] &= \frac{[Y \mid \boldsymbol{\beta}, \sigma^2]}{[Y]}[\boldsymbol{\beta}, \sigma^2] \\
&\propto [Y \mid \boldsymbol{\beta}, \sigma^2][\boldsymbol{\beta}][\sigma^2]
\end{aligned}
$$

where,

$$
[Y \mid \boldsymbol{\beta}, \sigma^2] = \left(2\pi\sigma^2\right)^{-n/2} \exp\left(-\frac{\sum_{i=1}^{n}\left(Y_i - \beta_0 - \beta_1 X_{i,1} - \beta_1 X_{i,2} - \beta_3 X_{i,3}\right)^2}{2\sigma^2}\right)
$$

$$
[\beta_0, \beta_1, \beta_2, \beta_3 \mid \sigma^2_\beta] = (2\pi\sigma^2_\beta)^{-4/2} \exp\left(-\frac{\sum_{i=0}^{3}\beta_i^2}{2\sigma^2_\beta}\right)
$$

$$
[\sigma^2 \mid a,\ b] = \frac{b^a}{\Gamma(a)}(\sigma^2)^{-a-1} \exp\left(-\frac{b}{\sigma^2}\right)
$$

Deriving the posterior (cont.)

$$
\begin{aligned}
[\beta_0, \beta_1, \beta_2, \beta_3, \sigma^2 \mid Y] \propto \\
& \left(2\pi\sigma^2\right)^{-n/2} \exp\left(-\frac{\sum_{i=1}^{n}\left(Y_i - \beta_0 - \beta_1 X_{i,1} - \beta_1 X_{i,2} - \beta_3 X_{i,3}\right)^2}{2\sigma^2}\right) \\
& (2\pi\sigma_\beta^2)^{-4/2} \exp\left(-\frac{\beta_0^2 + \beta_1^2 + \beta_2^2 + \beta_3^2}{2\sigma_\beta^2}\right) \\
& \frac{b^a}{\Gamma(a)}(\sigma^2)^{-a-1} \exp\left(-\frac{b}{\sigma^2}\right)
\end{aligned}
$$

Deriving the Gibbs sampler (σ^2 step)

Deriving the Gibbs sampler (β_i step)