

Lecture 1

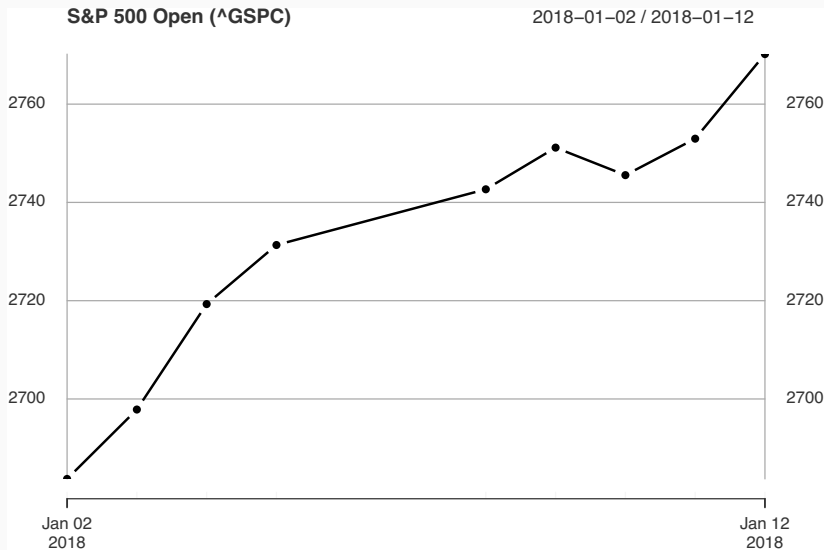
Spatio-temporal data & Linear Models

Colin Rundel

1/16/2018

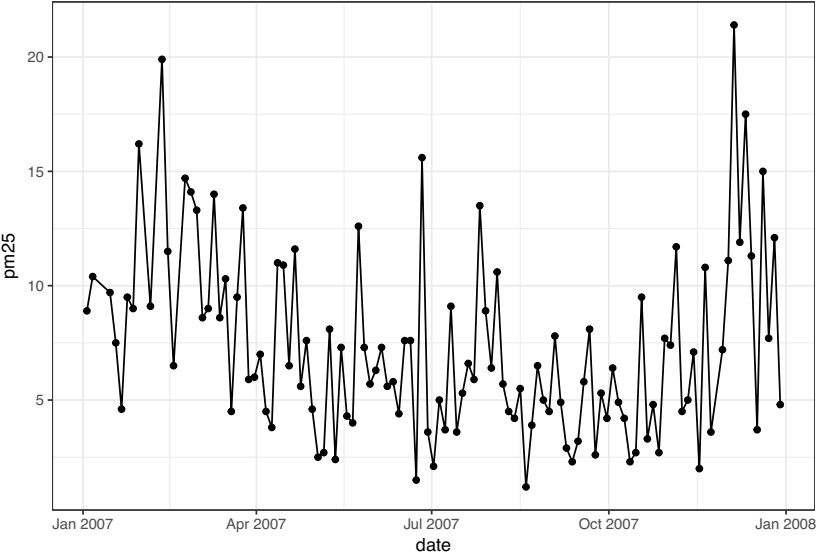
Spatio-temporal data

Time Series Data - Discrete

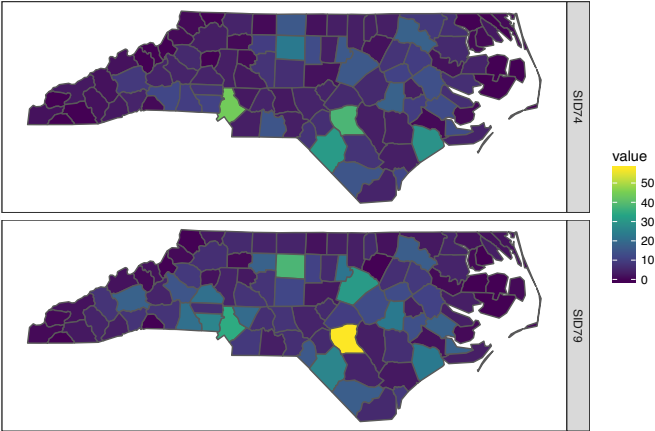


Time Series Data - Continuous

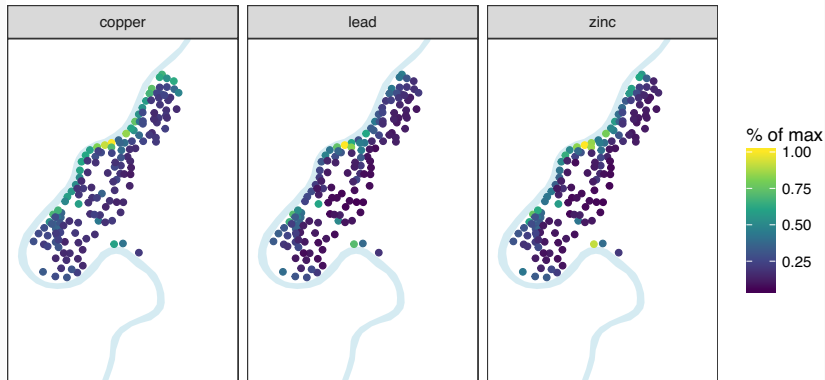
FRN Measured PM25



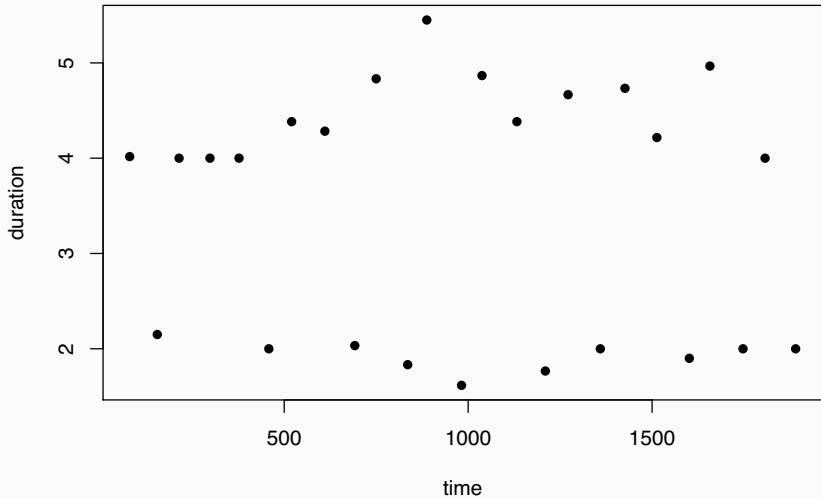
Spatial Data - Areal



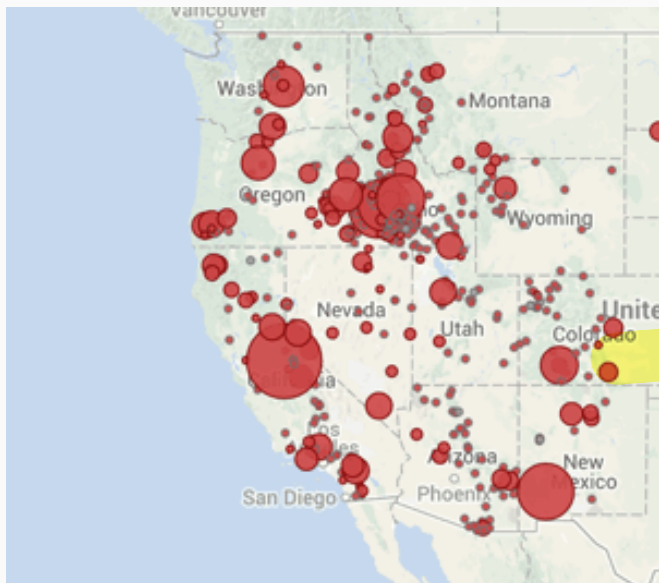
Meuse River



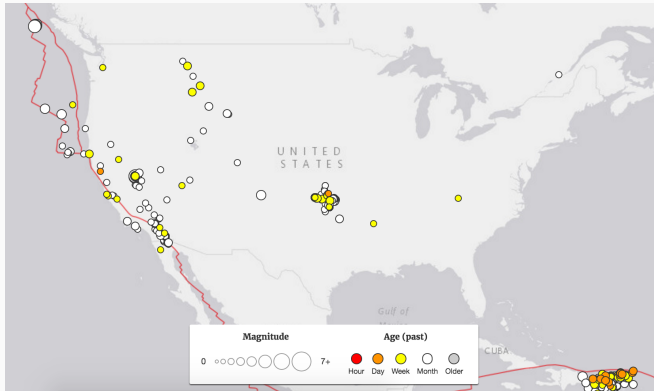
Old Faithful Eruption Duration



Point Pattern Data - Space



Point Pattern Data - Space + Time



(Bayesian) Linear Models

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

$$Y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \epsilon_i$$
$$\epsilon_i \sim N(0, \sigma^2)$$

which we can also express using matrix notation as

$$\mathbf{Y} = \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\epsilon}$$
$$\epsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{1}_n)$$

$n \times 1$ $n \times p$ $p \times 1$ $n \times 1$

$n \times 1$ $n \times n$

Multivariate Normal Distribution

For an n -dimension multivariate normal distribution with covariance Σ (positive semidefinite) can be written as

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

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

Multivariate Normal Distribution - Density

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

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

$1 \times n$ $n \times n$ $n \times 1$

and its log density is given by

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

$1 \times n$ $n \times n$ $n \times 1$

Maximum Likelihood - β

$$l(\beta, \sigma^2 | x, y) = -\frac{n}{2} \log 2\pi - \frac{1}{2} \log \det \Sigma - \frac{1}{2} (y - X\beta)' \Sigma^{-1} (y - X\beta)$$

$$\Sigma = \sigma^2 \mathbb{1}_n = \begin{pmatrix} \sigma^2 & & & 0 \\ & \sigma^2 & & \\ & & \ddots & \\ 0 & & & \sigma^2 \end{pmatrix}$$

$$\propto -\frac{n}{2} \log \sigma^2 - \frac{1}{2} \frac{1}{\sigma^2} (y - X\beta)' (y - X\beta)$$

$$\frac{\partial l(\beta, \sigma^2 | x, y)}{\partial \beta} \propto [(-X)' (y - X\beta) + (-X)' (y - X\beta)]$$

$$\propto -2 \underset{p \times n}{X}' \underset{n \times 1}{(y - X\beta)}$$

$$x' (y - x\beta) = 0$$

$$x' y - x' x \beta = 0$$

$$x' y = x' x \beta$$

$$(x' x)^{-1} x' y = \hat{\beta}_{MLE}$$

Maximum Likelihood - σ^2

$$\frac{\partial \ell(\beta, \sigma^2 | x, y)}{\partial \sigma^2} \propto \frac{-n}{2} \frac{1}{\sigma^2} + \frac{1}{2(\sigma^2)^2} (y - x\beta)' (y - x\beta)$$

$$\propto \frac{1}{2\sigma^2} \left(-n + \frac{1}{\sigma^2} (y - x\beta)' (y - x\beta) \right)$$

Assume $\sigma^2 > 0$

$$\sigma^2 = \frac{1}{n} (y - x\beta)' (y - x\beta)$$

$$= \frac{1}{n} \sum (y_i - x_i' \beta)^2 = \frac{1}{n} \sum (y_i - \hat{y}_i)^2$$

Likelihood:

$$\mathbf{Y} | \boldsymbol{\beta}, \sigma^2 \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{1}_n)$$

Likelihood:

$$\mathbf{Y} | \boldsymbol{\beta}, \sigma^2 \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{1}_n)$$

Priors:

$$\beta_i \sim N(0, \sigma_\beta^2) \text{ or } \boldsymbol{\beta} \sim N(\mathbf{0}, \sigma_\beta^2 \mathbf{1}_p)$$

$$\sigma^2 \sim \text{Inv-Gamma}(a, b)$$

$$\begin{aligned}[\beta, \sigma^2 | \mathbf{Y}, \mathbf{X}] &= \frac{[\mathbf{Y} | \mathbf{X}, \beta, \sigma^2]}{[\mathbf{Y} | \mathbf{X}]} [\beta, \sigma^2] \\ &\propto [\mathbf{Y} | \mathbf{X}, \beta, \sigma^2] [\beta, \sigma^2] \\ &\propto [\mathbf{Y} | \mathbf{X}, \beta, \sigma^2] [\beta | \sigma^2] [\sigma^2]\end{aligned}$$

$$\begin{aligned}[\boldsymbol{\beta}, \sigma^2 \mid \mathbf{Y}, \mathbf{X}] &= \frac{[\mathbf{Y} \mid \mathbf{X}, \boldsymbol{\beta}, \sigma^2]}{[\mathbf{Y} \mid \mathbf{X}]} [\boldsymbol{\beta}, \sigma^2] \\ &\propto [\mathbf{Y} \mid \mathbf{X}, \boldsymbol{\beta}, \sigma^2] [\boldsymbol{\beta}, \sigma^2] \\ &\propto [\mathbf{Y} \mid \mathbf{X}, \boldsymbol{\beta}, \sigma^2] [\boldsymbol{\beta} \mid \sigma^2] [\sigma^2]\end{aligned}$$

where,

$$f(\mathbf{Y} \mid \mathbf{X}, \boldsymbol{\beta}, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})\right)$$

$$\begin{aligned}[\boldsymbol{\beta}, \sigma^2 \mid \mathbf{Y}, \mathbf{X}] &= \frac{[\mathbf{Y} \mid \mathbf{X}, \boldsymbol{\beta}, \sigma^2]}{[\mathbf{Y} \mid \mathbf{X}]} [\boldsymbol{\beta}, \sigma^2] \\ &\propto [\mathbf{Y} \mid \mathbf{X}, \boldsymbol{\beta}, \sigma^2] [\boldsymbol{\beta}, \sigma^2] \\ &\propto [\mathbf{Y} \mid \mathbf{X}, \boldsymbol{\beta}, \sigma^2] [\boldsymbol{\beta} \mid \sigma^2] [\sigma^2]\end{aligned}$$

where,

$$f(\mathbf{Y} \mid \mathbf{X}, \boldsymbol{\beta}, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})\right)$$

$$f(\boldsymbol{\beta} \mid \sigma_\beta^2) = (2\pi\sigma_\beta^2)^{-p/2} \exp\left(-\frac{1}{2\sigma_\beta^2}\boldsymbol{\beta}'\boldsymbol{\beta}\right)$$

$$\begin{aligned}[\boldsymbol{\beta}, \sigma^2 | \mathbf{Y}, \mathbf{X}] &= \frac{[\mathbf{Y} | \mathbf{X}, \boldsymbol{\beta}, \sigma^2]}{[\mathbf{Y} | \mathbf{X}]} [\boldsymbol{\beta}, \sigma^2] \\ &\propto [\mathbf{Y} | \mathbf{X}, \boldsymbol{\beta}, \sigma^2] [\boldsymbol{\beta}, \sigma^2] \\ &\propto [\mathbf{Y} | \mathbf{X}, \boldsymbol{\beta}, \sigma^2] [\boldsymbol{\beta} | \sigma^2] [\sigma^2]\end{aligned}$$

where,

$$f(\mathbf{Y} | \mathbf{X}, \boldsymbol{\beta}, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})\right)$$

$$f(\boldsymbol{\beta} | \sigma_\beta^2) = (2\pi\sigma_\beta^2)^{-p/2} \exp\left(-\frac{1}{2\sigma_\beta^2}\boldsymbol{\beta}'\boldsymbol{\beta}\right)$$

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

Deriving the Gibbs sampler (σ^2 step)

$$[\beta, \sigma^2 | \mathbf{Y}, \mathbf{X}] \propto (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)\right)$$

$$\cancel{(2\pi\sigma_\beta^2)^{-p/2}} \exp\left(\cancel{-\frac{1}{2\sigma_\beta^2}\beta'\beta}\right)$$

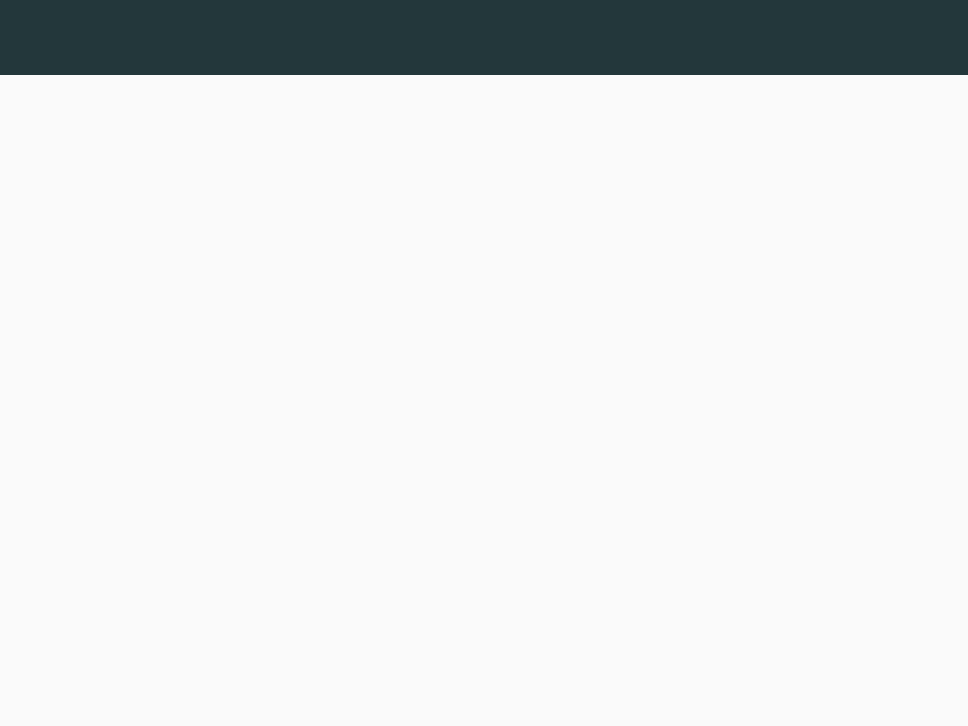
$$\cancel{\frac{b^a}{\Gamma(a)}} (\sigma^2)^{-a-1} \exp\left(-\frac{b}{\sigma^2}\right)$$

$$\sigma^2 | \cdot \propto (\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} (\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)\right)$$

$$(\sigma^2)^{-a-1} \exp\left(-\frac{1}{\sigma^2} b\right)$$

$$\propto (\sigma^2)^{-(a+n/2)-1} \exp\left(-\frac{1}{\sigma^2} \left(b + \frac{1}{2} (\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)\right)\right)$$

$$\sigma^2 | \cdot \propto \text{InverseGamma}\left(a + \frac{n}{2}, b + \frac{1}{2} (\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)\right)$$



Deriving the Gibbs sampler (β step)

$$\begin{aligned}
 [\beta, \sigma^2 | \mathbf{Y}, \mathbf{X}] &\propto \cancel{(2\pi\sigma^2)^{-n/2}} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)\right) \\
 &\quad \cancel{(2\pi\sigma_\beta^2)^{-p/2}} \exp\left(-\frac{1}{2\sigma_\beta^2}\beta'\beta\right) \quad S = \frac{\sigma^2}{\sigma_\beta^2} \\
 &\quad \frac{b^a}{\Gamma(a)} \cancel{(\sigma^2)^{-a-1}} \exp\left(-\frac{b}{\sigma^2}\right) \quad = \quad \exp\left(-\frac{S}{2\sigma^2}\beta'\beta\right)
 \end{aligned}$$

$$\beta | \cdot \propto \exp\left[-\frac{1}{2}\sigma^2(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta) - \frac{S}{2\sigma^2}\beta'\beta\right]$$

$$\propto \exp\left[-\frac{1}{2\sigma^2}\left(\cancel{\mathbf{Y}'\mathbf{Y}} - \beta'\mathbf{X}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}\beta + \beta'\mathbf{X}'\mathbf{X}\beta + \beta'S\mathbf{1}\right)\right]$$

$$\propto \exp\left[-\frac{1}{2\sigma^2}\left(-\beta'\mathbf{X}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}\beta + \beta'(\mathbf{X}'\mathbf{X} + S\mathbf{1})\beta\right)\right]$$

$$\beta \sim N(\mu_\beta, \Sigma_\beta)$$

$$p| \propto \exp \left[-\frac{1}{2} (\beta - \mu_\beta)' \Sigma_\beta^{-1} (\beta - \mu_\beta) \right]$$

$$\propto \exp \left[-\frac{1}{2} \left(\beta' \Sigma_\beta^{-1} \beta - \beta' \Sigma_\beta^{-1} \mu_\beta - \dots \right) \right]$$

$$\Sigma_\beta^{-1} = (X'X + S \mathbb{1}) \Rightarrow \Sigma_\beta = (X'X + S \mathbb{1})^{-1} \\ = \left(X'X + \frac{\sigma^2}{\sigma_\beta} \mathbb{1} \right)^{-1}$$

$$-\beta' \Sigma_\beta^{-1} \mu_\beta = -\beta' X' Y \Rightarrow \mu_\beta = \Sigma_\beta (X' Y)$$

$$= \left(X'X + \frac{\sigma^2}{\sigma_\beta} \mathbb{1} \right)^{-1} X' Y$$

A Quick Example

Some Fake Data

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

```
set.seed(01162018)
n = 100
beta = c(0.7, 1.5, -2.2, 0.1)
eps = rnorm(n)

d = data_frame(
  X1 = rt(n, df=5),
  X2 = rt(n, df=5),
  X3 = rt(n, df=5)
) %>%
  mutate(Y = beta[1] + beta[2]*X1 + beta[3]*X2 + beta[4]*X3 + eps)

X = cbind(1, d$X1, d$X2, d$X3)
```

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

$$\hat{\mathbf{Y}} = \hat{\beta}_0 + \hat{\beta}_1 \mathbf{X}_1 + \hat{\beta}_2 \mathbf{X}_2 + \hat{\beta}_3 \mathbf{X}_3 = \mathbf{X} \hat{\boldsymbol{\beta}}$$

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

$$\hat{\mathbf{Y}} = \hat{\beta}_0 + \hat{\beta}_1 \mathbf{X}_1 + \hat{\beta}_2 \mathbf{X}_2 + \hat{\beta}_3 \mathbf{X}_3 = \mathbf{X} \hat{\boldsymbol{\beta}}$$

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

$$\arg \min_{\boldsymbol{\beta}} \sum_{i=1}^n (Y_i - \mathbf{X}_i \cdot \boldsymbol{\beta})^2$$

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

$$\hat{\mathbf{Y}} = \hat{\beta}_0 + \hat{\beta}_1 \mathbf{X}_1 + \hat{\beta}_2 \mathbf{X}_2 + \hat{\beta}_3 \mathbf{X}_3 = \mathbf{X} \hat{\boldsymbol{\beta}}$$

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

$$\arg \min_{\boldsymbol{\beta}} \sum_{i=1}^n (Y_i - \mathbf{X}_i \cdot \boldsymbol{\beta})^2$$

Previously we derived,

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

Frequentist Fit

```
l = lm(Y ~ X1 + X2 + X3, data=d)
l$coefficients
```

```
## (Intercept)          X1          X2          X3
##  0.6566561    1.4657537  -2.2807109    0.1629704
```

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

```
##           [,1]
## [1,]  0.6566561
## [2,]  1.4657537
## [3,] -2.2807109
## [4,]  0.1629704
```

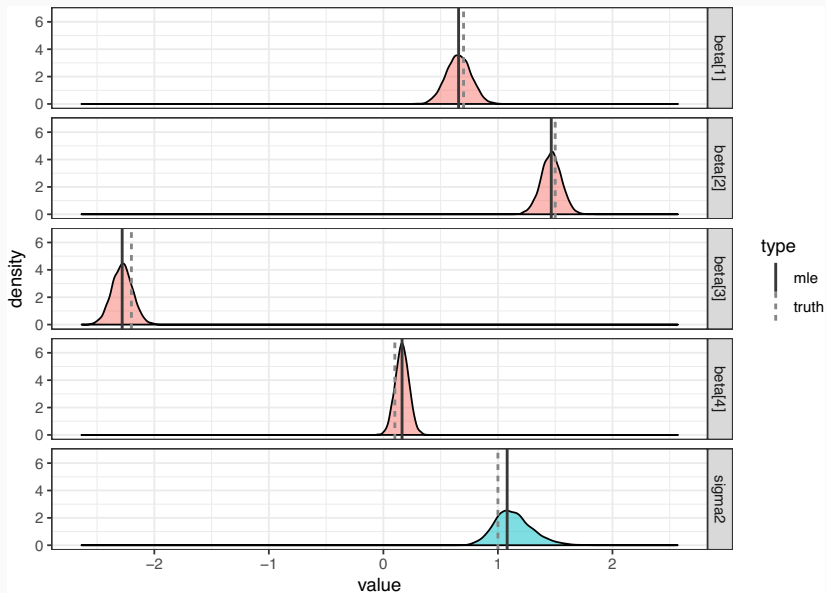
Bayesian model specification (JAGS)

```
model =  
"model{  
  # Likelihood  
  for(i in 1:length(Y)){  
    Y[i] ~ dnorm(mu[i], tau)  
    mu[i] = beta[1] + beta[2]*X1[i] + beta[3]*X2[i] + beta[4]*X3[i]  
  }  
  
  # Prior for beta  
  for(j in 1:4){  
    beta[j] ~ dnorm(0,1/100)  
  }  
  
  # Prior for sigma / tau2  
  tau ~ dgamma(1, 1)  
  sigma2 = 1/tau  
}"
```

Bayesian model fitting (JAGS)

```
m = rjags::jags.model(  
  textConnection(model),  
  data = d  
)  
  
## Compiling model graph  
##   Resolving undeclared variables  
##   Allocating nodes  
## Graph information:  
##   Observed stochastic nodes: 100  
##   Unobserved stochastic nodes: 5  
##   Total graph size: 810  
##  
## Initializing model  
  
update(m, n.iter=1000, progress.bar="none") → Burnin  
  
samp = rjags::coda.samples(  
  m, variable.names=c("beta", "sigma2"),  
  n.iter=5000, progress.bar="none"  
)
```

Results



Results (zoom)

