Lecture 1

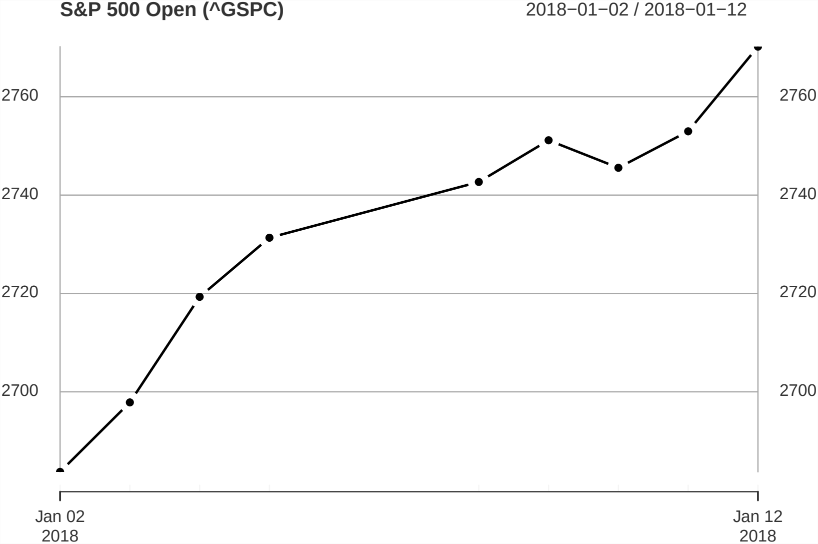
Spatio-temporal data & Linear Models

Colin Rundel

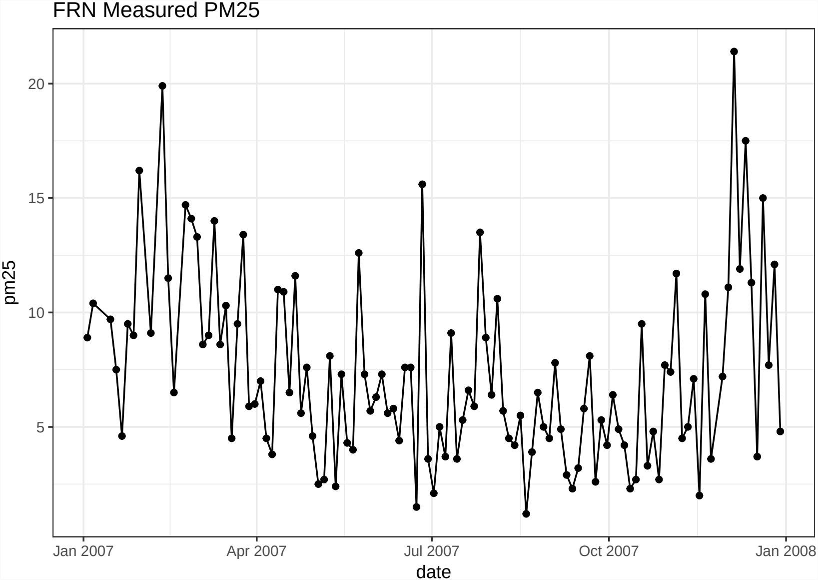
1/16/2018

Spatio-temporal data

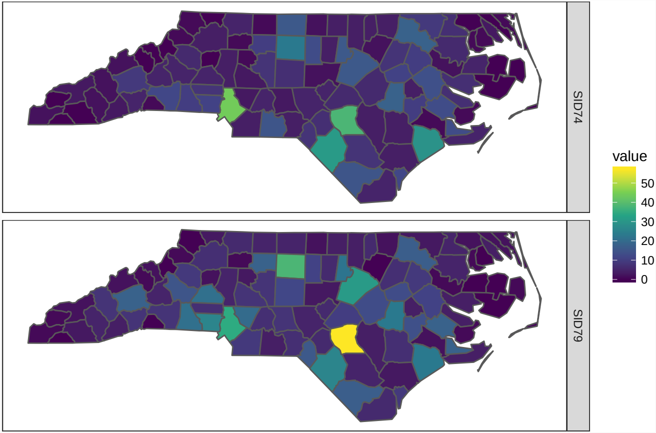
Time Series Data - Discrete

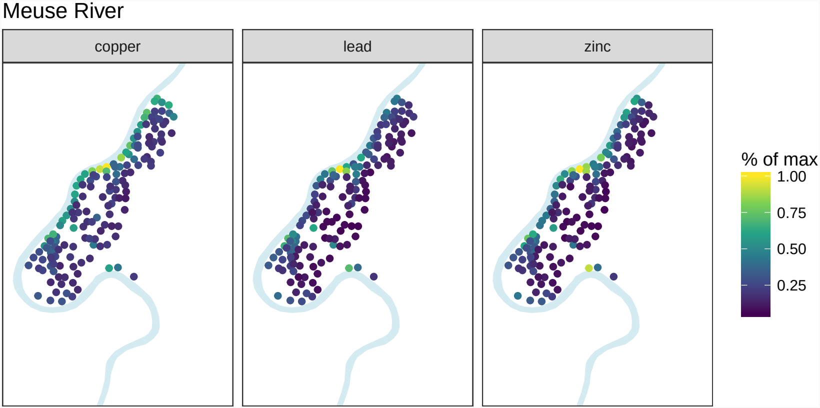


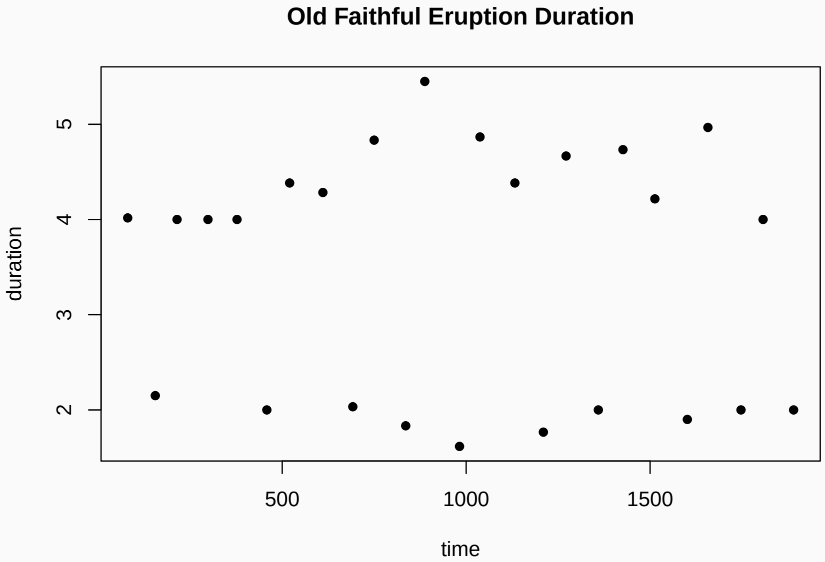
Time Series Data - Continuous



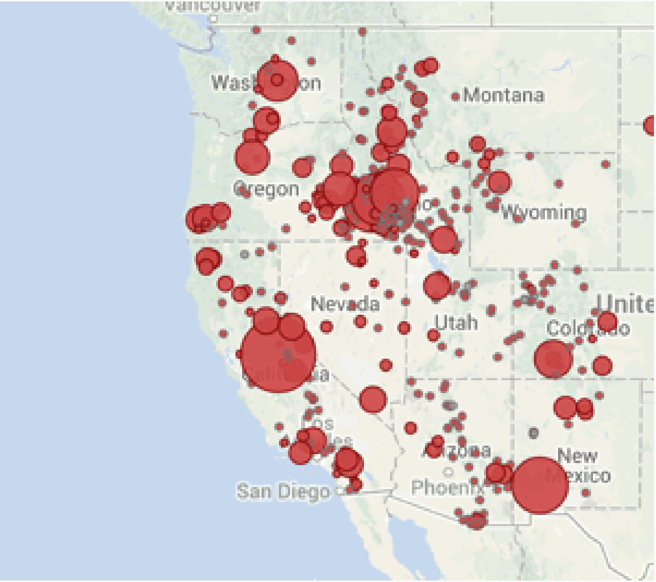
Spatial Data - Areal



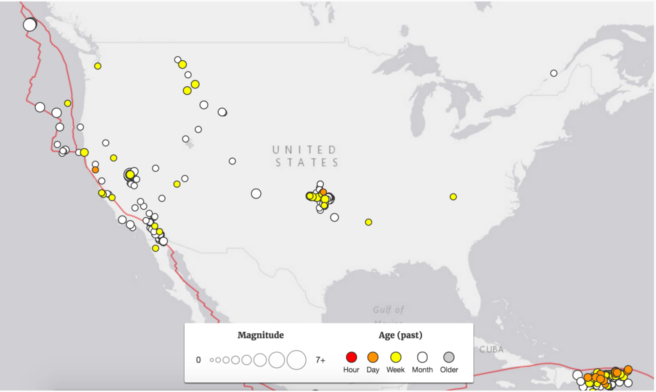

Meuse River
copper
lead
zinc
% of max
1.00
0.75
0.50
0.25


Old Faithful Eruption Duration
duration
time
2
3
4
5
500
1000
1500

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

$$
\begin{aligned}
\underset{n\times 1}{\mathbf{Y}} = \underset{n\times p}{\mathbf{X}}\,\underset{p\times 1}{\boldsymbol{\beta}} + \underset{n\times 1}{\boldsymbol{\epsilon}} \\
\boldsymbol{\epsilon} \sim N(\underset{n\times 1}{\mathbf{0}},\ \sigma^2\,\underset{n\times n}{\mathbb{1}_n})
\end{aligned}
$$

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

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

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

Multivariate Normal Distribution - Density

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

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

and its log density is given by

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

Maximum Likelihood - β

Maximum Likelihood - σ^2

Likelihood:

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

Bayesian Model

Likelihood:

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

Priors:

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

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

Deriving the posterior

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

Deriving the posterior

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

where,

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

Deriving the posterior

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

where,

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

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

Deriving the posterior

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

where,

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

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

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

Deriving the Gibbs sampler (σ^2 step)

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

Deriving the Gibbs sampler (β step)

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

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

```r
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)
```

Least squares fit

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

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

Least squares fit

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

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

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

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

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)

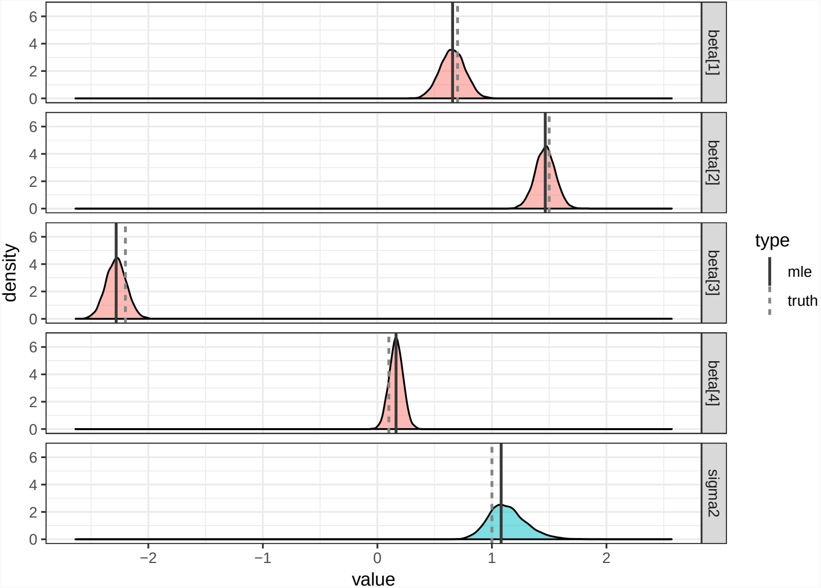
```
m = rjags::jags.model(
  textConnection(model),
  data = d
)
```

```
## 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")

samp = rjags::coda.samples(
  m, variable.names=c("beta","sigma2"),
  n.iter=5000, progress.bar="none"
)
```

Results



density
value
beta[1]
beta[2]
beta[3]
beta[4]
sigma2
type
mle
truth
−2
−1
0
1
2
0
2
4
6

Results (zoom)