lm.bayes, summary.bayes

1 lm.bayes

Bayesian Regression - reference prior
lm.bayes(y,X,howmany)

input: vector y of responses
input: X the design matrix (n by p, this can easily be constructed with lm function, see the example)
input: howmany - a value of how many posteriors draws you want
output: posterior draws for $\beta_0, \beta_1, ..., \beta_p, \sigma^2$

This function uses a reference prior: $\pi(\beta, \sigma^2) \propto \frac{1}{\sigma^2}$. This is typical regression where X is a n by p design matrix and β is a vector of length p. y is a vector of responses of length n. $y|\beta, \sigma^2, X \sim N(X\beta, \sigma^2 I)$. Other things we need to know: $\hat{\beta} = (X'X)^{-1}X'y$ and the sample variance is $s^2 = (y - X\hat{\beta})'(y - X\hat{\beta})/(n-p)$.
Full conditional:

$$
\begin{aligned}
p(\beta, \sigma^2|X, y) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2}\pi\sigma^2} e^{\frac{-1}{2\sigma^2}(y_i - X_i\beta)'(y_i - X_i\beta)} \frac{1}{\sigma^2} \\
&\propto (\sigma^2)^{-(n+2)/2} e^{\frac{-1}{2\sigma^2}(y - X\beta)'(y - X\beta)} \\
&\propto (\sigma^2)^{-(n+2)/2} e^{\frac{-1}{2\sigma^2}(y - X\hat{\beta} + X\hat{\beta} - X\beta)'(y - X\hat{\beta} + X\hat{\beta} - X\beta)} \\
&\propto (\sigma^2)^{-(n+2)/2} e^{\frac{-1}{2\sigma^2}(y - X\hat{\beta})'(y - X\hat{\beta}) + (X\beta - X\hat{\beta})'(X\beta - X\hat{\beta})} \\
&\propto (\sigma^2)^{-(n+2)/2} e^{\frac{-(n-p)s^2}{2\sigma^2} + (\beta - \hat{\beta})'(X'X)(\beta - \hat{\beta})} \\
&\propto (\sigma^2)^{-(n+2)/2} e^{\frac{-(n-p)s^2}{2\sigma^2}} e^{\frac{-1}{2\sigma^2}(\beta - \hat{\beta})'(X'X)(\beta - \hat{\beta})}
\end{aligned}
$$

Posterior for σ^2:

$$
\begin{aligned}
\sigma^2|... &\propto \int (\sigma^2)^{-(n+2)/2} \exp \frac{-(n-p)s^2}{2\sigma^2} e^{\frac{-1}{2\sigma^2}(\beta-\hat{\beta})'(X'X)(\beta-\hat{\beta})} d\beta \\
&\propto (\sigma^2)^{-(n+2)/2} e^{\frac{-(n-p)s^2}{2\sigma^2}} \int e^{\frac{-1}{2\sigma^2}(\beta-\hat{\beta})'(X'X)(\beta-\hat{\beta})} d\beta \\
&\propto (\sigma^2)^{-(n+2)/2} e^{\frac{-(n-p)s^2}{2\sigma^2}} (2\pi)^{p/2} |\sigma^2(X'X)|^{1/2} \int (2\pi)^{-p/2} |\sigma^2(X'X)|^{-1/2} e^{\frac{-1}{2\sigma^2}(\beta-\hat{\beta})'(X'X)(\beta-\hat{\beta})} d\beta \\
&\propto (\sigma^2)^{-(n+2)/2} |\sigma^2(X'X)|^{1/2} e^{\frac{-(n-p)s^2}{2\sigma^2}} \\
&\propto (\sigma^2)^{-(n+2)/2} (\sigma^2)^{p/2} |(X'X)|^{1/2} e^{\frac{-(n-p)s^2}{2\sigma^2}} \\
&\propto (\sigma^2)^{-(\frac{n-p}{2}+1)} e^{\frac{-(n-p)s^2}{2\sigma^2}} \\
&\sim scaled - Inv - \chi^2(n-p, s^2) \\
&\sim s^2(n-k)/\chi^2(n-k)
\end{aligned}
$$

Posterior for β: ($X = QR$ and $Q'Q = I$)

$$
\begin{aligned}
\beta|... &\sim N((X'X)^{-1}X'y, \sigma^2(X'X)^{-1}) \\
&\sim [(X'X)^{-1}]^{-1/2} N(0, \sigma^2) + (X'X)^{-1}X'y \\
&\sim [(R'Q'QR)]^{1/2} N(0, \sigma^2) + \hat{\beta} \\
&\sim (R'R)^{-1/2} N(0, \sigma^2) + \hat{\beta} \\
&\sim R^{-1} N(0, \sigma^2) + \hat{\beta}
\end{aligned}
$$

2 summary.bayes

Summary for Bayesian Regression

summary.bayes(lm.bayes.object)
input: lm.bayes object
output: a table with means and 95% PIs for the parameters $\beta_0, \beta_1, ..., \beta_p, \sigma^2$ (each parameter has a row)
This function works like the summary function for lm (but now for lm.bayes)

3 fitted.values

Fitted values of a Bayesian regression

fitted.values(lm.bayes.object, X)
input: lm.bayes.object an object resulting from lm.bayes
input: X (no as.factors allowed)
output: fitted values - howmany(q) by n matrix

4 bayes.resid

Residuals of a Bayesian regression

bayes.resid(lm.bayes.object, X, y)
input: lm.bayes.object an object resulting from lm.bayes
input: X (no as.factors allowed)
input: y
output: residual values - howmany(q) by n matrix

In frequentist statistics a residual is a single value. But in Bayesian methods each residual has a distribution.

5 bayes.resid.plot

Residuals of a Bayesian regression

bayes.resid(lm.bayes.object, X, y)
input: residmat = matrix(q by n) filled with residuals
input: (optional) m is an extra parameter that controls the height of the densities, just play with it until the graphs look good. The default is m=1 if you don't provide one.
output: pretty plot

In frequentist statistics a residual is a single value. But in Bayesian methods each residual has a distribution. Here is how I like to plot them. There will be some number of draws (q="howmany" in this notation) from the posterior. y_i is the observation $y_i - fitted.function$ gives the residuals ($i = 1...n$). So, we will have an q by n matrix of residuals.