ADVANCED BAYESIAN MODELING

Noninformative Prior Analysis

Assume an ordinary normal-theory regression:

$$y \mid \theta, X \quad \sim \quad \mathrm{N}(X\beta,\, \sigma^2 I)$$

i.e.,

$$y_i \mid \theta, X_i \quad \sim \quad \text{indep. } \mathrm{N}(X_i\beta,\, \sigma^2) \qquad\qquad i = 1, \ldots, n$$

with

$$\theta \quad = \quad (\beta, \sigma^2)$$

As before:

y is $n \times 1$ $\qquad$ X is $n \times k$ $\qquad$ β is $k \times 1$

X_i is row i of X $\qquad$ I is the $n \times n$ identity matrix

Noninformative Prior

The "standard" noninformative prior has density

$$p(\theta \mid X) \;=\; p(\beta, \sigma^2 \mid X) \;\propto\; (\sigma^2)^{-1} \qquad\qquad \sigma^2 > 0$$

which is improper.

Remark: Equivalent to an improper flat prior on $(\beta, \log \sigma^2)$

Note: In general, a prior may depend on X, since X is treated as constant.

Posterior

Posterior is proper, provided

- The columns of X are linearly independent, and
- $n > k$

In this case, the classical *ordinary least squares* estimates exist:

$$\hat{\beta} \;=\; (X^T X)^{-1} X^T y \qquad\qquad s^2 \;=\; \frac{1}{n-k}(y - X\hat{\beta})^T (y - X\hat{\beta})$$

The posterior is related to these ...

The full posterior has an explicit form:

$$\beta \mid \sigma^2, y, X \;\sim\; \mathrm{N}\big(\hat{\beta},\, \sigma^2 V_\beta\big) \qquad \text{where} \quad V_\beta \;=\; (X^T X)^{-1}$$

$$\sigma^2 \mid y, X \;\sim\; \text{Inv-}\chi^2\big(n-k, s^2\big)$$

When $n - k > 2$, explicit posterior inferences include

$$\mathrm{E}(\beta \mid y, X) \;=\; \hat{\beta} \qquad \mathrm{var}(\beta \mid y, X) \;=\; \frac{n-k}{n-k-2} s^2 V_\beta$$

$$\mathrm{E}(\sigma^2 \mid y, X) \;=\; \frac{n-k}{n-k-2} s^2$$

Direct posterior simulation is easy:

1. Simulate σ^2 from its posterior marginal.
2. Simulate β from its posterior conditional (given simulated σ^2).

Alternatively, the actual posterior marginals for elements of β are t-distributions that have been shifted and scaled.

Prediction

For unobserved $\tilde{y}$ satisfying

$$\tilde{y} \mid \theta, \tilde{X} \quad \sim \quad \mathrm{N}(\tilde{X}\beta,\, \sigma^2 I) \qquad\qquad \tilde{X} \ \text{observed}$$

can easily simulate from posterior predictive distribution:

1. Simulate $\theta = (\beta, \sigma^2)$ from posterior as before.
2. Simulate $\tilde{y}$ given the simulated θ.

Alternatively, there is an explicit form (BDA3, Sec. 14.2).