STT 465
Bayesian Multiple Linear Regression:

- Regression equation
  - Likelihood
  - Prior
  - Posterior distribution
  - Outline of a Sampler
  - Gibbs sampler with unknown variances
  - Gibbs sampler with single-coefficients updates.

Bayesian Multiple Linear Regression

- **Gaussian Linear Regression Model**

$$y_i = \sum_{j=1}^{p} x_{ij}\beta_j + \varepsilon_i$$

- **Matrix representation**

Let $x_i' = \left(1, x_{i1}, x_{i2}, \ldots, x_{ip}\right)$ $\beta = \left(\mu, \beta_1, \beta_2, \ldots, \beta_p\right)'$

Then $y_i = x_i'\beta + \varepsilon_i$

Stack equations 1-n to get $y = X\beta + \varepsilon$

Probability assumptions (for now iid errors) $\varepsilon \sim MVN\left(0, \mathrm{I}\sigma_\varepsilon^2\right)$

- **Likelihood**

$$\left[y \mid \beta\right] \sim MVN\left(X\beta, R\right)$$

Likelihood (cont.)

$$p\left(y \mid X, \beta, \sigma_{\varepsilon}^{2}\right) = \left(2\pi\right)^{-n/2} \left\| I\sigma_{\varepsilon}^{2} \right\|^{-1/2} Exp\left\{ -\frac{1}{2\sigma_{\varepsilon}^{2}} \left(y - X\beta\right)' \left(y - X\beta\right) \right\}$$

Note: MLE is OLS.

$$\hat{\beta}_{OLS} = \left[X'X \right]^{-1} X'y$$

Prior

Normal prior for reg. coefficients

$$\beta \sim N\left(\beta_0, \Sigma_0\right) = \left(2\pi\right)^{-n/2} \left\|\Sigma_0\right\|^{-1/2} Exp\left\{-\left(\beta - \beta_0\right)' \Sigma_0^{-1} \left(\beta - \beta_0\right)\right\}$$

If we assume IID, zero-mean prior we have

$$\beta \sim N\left(0, I\sigma_\beta^2\right) = \left(2\pi\right)^{-n/2} \left\|I\sigma_\beta^2\right\|^{-1/2} Exp\left\{-\frac{\beta'\beta}{2\sigma_\beta^2}\right\}$$

Posterior Density

Likelihood

$$p(y \mid X,\beta) = (2\pi)^{-n/2} \left\| I\sigma_\varepsilon^2 \right\|^{-1/2} Exp\left\{-\frac{1}{2}(y - X\beta)'\left[I\sigma_\varepsilon^2\right]^{-1}(y - X\beta)\right\}$$

Prior

$$\beta \sim N\left(0, I\sigma_\beta^2\right) = (2\pi)^{-n/2} \left\| I\sigma_\beta^2 \right\|^{-1/2} Exp\left\{-\frac{1}{2}\beta'\left[I\sigma_\beta^2\right]^{-1}\beta\right\}$$

General for of the posterior density (derivation presented in class)

$$p(\beta \mid y) \sim N\left(C^{-1}rhs, C^{-1}\right)$$

$$C = \left[X'R^{-1}X + \Sigma_0^{-1}\right] \qquad \left[X'R^{-1}X + \Sigma_0^{-1}\right]\hat{\beta} = X'R^{-1}y + \Sigma_0^{-1}\beta_0$$

$$rhs = X'R^{-1}y + \Sigma_0^{-1}\beta_0$$

Special (most commonly used) case

Likelihood [iid residuals] $R = I\sigma_\varepsilon^2$

Prior [iid mean-zero effects] $R = I\sigma_\varepsilon^2$

$$\beta \sim N\left(0, I\sigma_0^2\right)$$

Posterior

$$p\left(\beta \mid y\right) \sim N\left(C^{-1}rhs, C^{-1}\right)$$

$$C = \left[X'X\sigma_\varepsilon^{-2} + I\sigma_0^2\right]$$

$$rhs = X'y\sigma_\varepsilon^{-2}$$

Discuss shrinkage and connection to Ridge Regression.

Outline of a Sampler

Target distribution

$$p(\beta \mid y) \sim N\left(C^{-1}rhs, C^{-1}\right)$$

$$C = \left[X'X\sigma_{\varepsilon}^{-2} + I\sigma_{0}^{2}\right]$$

$$rhs = X'y\sigma_{\varepsilon}^{-2}$$

Outline of a sampler

- Compute C and it's inverse C^{-1}
- Compute rhs
- Compute the posterior mean $C^{-1}rhs$
- To sample from MVN
    - draw iid standard normal
    - pre-multiply the iid standard normal with the (upper-triangular) Cholesky of the inverse of C.
    - add the solution.

Sampler with unknown variances

Likelihood $\left[y \mid \beta\right] \sim MVN\left(X\beta, I\sigma_{\varepsilon}^{2}\right)$

Semi-conjugate prior