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
$$x'_i = (1, x_{i1}, x_{i2}, ..., x_{ip})$$
 $\beta = (\mu, \beta_1, \beta_2, ..., \beta_p)'$

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

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

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

- Likelihood

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

Likelihood (cont.)

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

Note: MLE is OLS.

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

Prior

Normal prior for reg. coefficients

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

If we assume IID, zero-mean prior we have

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

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

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

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

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

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

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

Special (most commonly used) case

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

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

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

Posterior

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

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

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

Discuss shrinkage and connection to Ridge Regression.

Outline of a Sampler

Target distribution

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

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

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

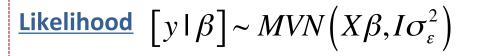
Outline of a sampler

- 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



Semi-conjugate prior