

STT 465

Summary of Bayesian Multiple Linear Regression:

- Regression equation
 - Likelihood
 - Prior
- Posterior distribution

Bayesian Multiple Linear Regression

- Gaussian Linear Regression Model

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

- Matrix representation

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

Then $y_i = x'_i \beta + \varepsilon_i$

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

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

- Likelihood

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

Likelihood (cont.)

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

Note: MLE is OLS.

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

Prior

Normal prior for reg. coefficients

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

If we assume IID, zero-mean prior we have

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

Posterior Density

Likelihood

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

Prior

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

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

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

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

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

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

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

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

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

Discuss shrinkage and
connection to Ridge
Regression.