Multiple regression

- $$y = f(x_1, x_2, \ldots x_p; \theta) + \varepsilon$$
- Data:

$$(y_1, x_{11,}, x_{21}, \ldots x_{p1}),$$

$$(y_2, x_{12}, x_{22}, \ldots, x_{p2}),$$

$$\ldots\ldots$$

$$(y_n, x_{1n}, x_{2n}, \ldots, x_{pn})$$

- Multiple linear regression

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + \varepsilon_i,$$

Variable selection

- Goal
  - Find the “best” subset of predictors
- Problem
  - 2^p possible models
  - How to search in the model space?
  - What criterion to evaluate different models?

Stochastic search variable selection (SSVS)

- Use a set of binary indicator variables $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_p)$ to represent sub-models (George & McCulloch 1993)
- Bayesian: find the posterior distribution of γ_i
- Ranking predictors by $p_i = \Pr(\gamma_i = 1)$

- Now, it is an estimation problem!
  - Bayesian: posterior = prior x likelihood (data)

A Bayesian hierarchical model

$$\mathbf{Y}|\beta, \sigma^2 \sim N_n(\mathbf{X}\beta, \sigma^2\mathbf{I})$$

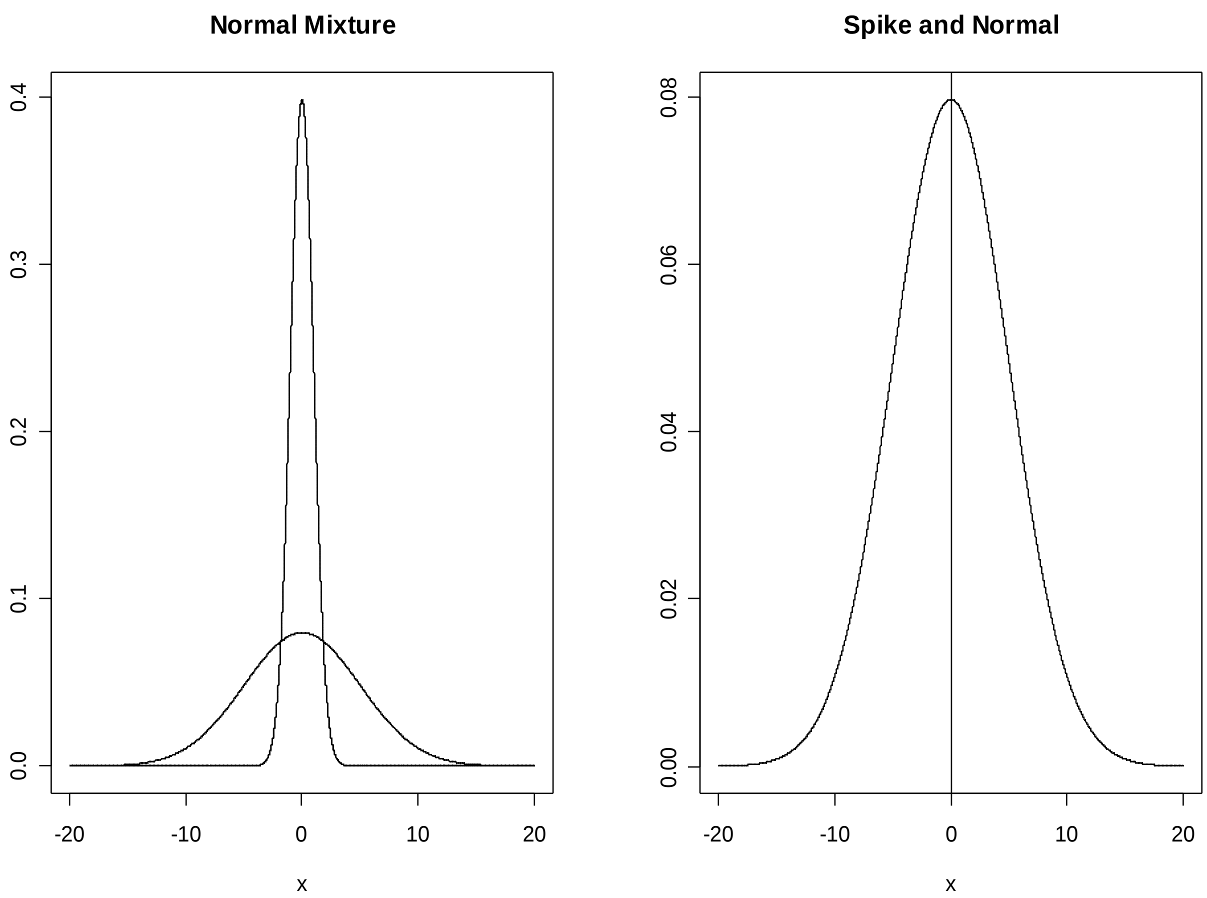
$$\beta_i|\gamma_i \sim \gamma_i N(0, \sigma_i^2) + (1 - \gamma_i)I_0,\ i = 1, 2, \ldots, p$$

$$\sigma^2|\gamma \sim IG(\nu_{\boldsymbol{\gamma}}/2, \nu_{\boldsymbol{\gamma}}\lambda_{\boldsymbol{\gamma}}/2)$$

$$\gamma_i \sim Bernoulli(p_i),\ i = 1, 2, \ldots, p$$

$$f(\beta, \sigma^2, \gamma|\mathbf{Y}) \propto f(Y|\beta, \sigma^2) f(\beta|\gamma) f(\sigma^2|\gamma) f(\gamma).$$

Prior distribution of $\beta_i \mid \gamma_i$



Gibbs sampler for SSVS

$$\boldsymbol{\beta}_{\boldsymbol{\gamma}}|\sigma^{-2}, \boldsymbol{\gamma}, \mathbf{Y} \sim N(\sigma^2 \mathbf{A}_{\gamma} \mathbf{X}'_{\gamma} \mathbf{Y}, \mathbf{A}_{\gamma}),$$

where $\mathbf{A}_{\gamma} = (\sigma^{-2}\mathbf{X}'_{\gamma}\mathbf{X}_{\gamma} + (\mathbf{D}_{\gamma}\mathbf{I}_{\gamma}\mathbf{D}_{\gamma})^{-1})^{-1}$, and $\mathbf{D}_{\boldsymbol{\gamma}} = diag(\sigma_1, \ldots, \sigma_p)$

$$\sigma^2|\boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{Y} \sim IG\left(\frac{n + \nu_{\boldsymbol{\gamma}}}{2}, \frac{(\mathbf{y} - \mathbf{x}_{\gamma}\boldsymbol{\beta}_{\boldsymbol{\gamma}})'(\mathbf{y} - \mathbf{x}_{\gamma}\boldsymbol{\beta}_{\boldsymbol{\gamma}}) + \nu_{\boldsymbol{\gamma}}\lambda_{\boldsymbol{\gamma}}}{2}\right)$$

$$\gamma_i|\boldsymbol{\beta}, \sigma^2, \boldsymbol{\gamma}_{(i)}, \quad \text{Bernoulli } B(1, \tilde{p}_i)$$