$$Y = X\beta + \epsilon \qquad \epsilon \sim N(0, \sigma^2 I_n)$$

$$\beta_0 \in l_0(s, p) \qquad s << p \qquad p >> n$$

$$\boxed{\beta_j \sim \pi_0 \delta_0 + (1 - \pi_0)\, g(\cdot)}$$

$$\pi_0 \sim U(0,1) \qquad \pi_0 \sim \text{Beta}(\kappa p, 1)$$

$$(\gamma_1, \ldots, \gamma_p) \qquad \gamma_j = I(\beta_j \neq 0) \qquad p_\gamma = \sum_{j=1}^{p} \gamma_j$$

$$\pi(\gamma) = \frac{1}{p+1} \cdot \frac{1}{\binom{p}{p_\gamma}} \approx e^{-p_\gamma \log p}$$

$$\beta_j \mid \gamma_j = \gamma_j\, g(\cdot) + (1 - \gamma_j)\, \delta_0$$

$$\pi_0 \sim \text{Beta}(\kappa p, 1)$$

$$\pi(\gamma) \propto e^{-a p_\gamma \log\left(\frac{b}{p_\gamma}\right)} \frac{1}{\binom{p}{p_\gamma}}$$

(Verify!)

Yang, Wainwright, Jordon (2016), AoS

$$\pi(\gamma) \propto \left(\frac{1}{p}\right)^{\kappa p_\gamma} I(p_\gamma \le s_0)$$

$s_0 = n$

$\gamma = (\gamma_1, \ldots, \gamma_p)$

$\{2^p\}$

$$= e^{-\kappa p_\gamma \log p}$$

$\beta_j \mid \beta_{-j}, y, X$

(β, γ)

$p(\gamma \mid y, X)$

Collapsed Gibbs sampler for variable selection

$S = \{ j : \gamma_j = 1 \}$ $\qquad X_S = (X_j : j \in S)$

$\beta_S = (\beta_j : j \in S)$ $\qquad \beta = (\beta_S, \beta_{S^c})$ $\nearrow 0$

represents a "model"

$\pi(\gamma \mid y, X) = \pi(S \mid y, X)$

$$\underline{Y} \mid \beta_S, \underline{S} \sim N\left(X_S \boxed{\beta_S}, \tfrac{1}{\phi} I\right)$$

$$\left\{ \begin{array}{l} \underline{\beta_S \mid S} \sim \prod_{j \in S} g(\beta_j) \\ \underline{S} \sim \left(\frac{1}{p}\right)^{k\, p_\gamma} I(p_\gamma \le s_0) \\ \quad\ \propto \end{array} \right.$$

$Y \mid S$

S

$\gg$ $p(S \mid Y)$

$= p(\gamma \mid Y)$

Zellner's g-prior:

$g = p^2$

$$\beta_s \mid S \sim N\left(0, \frac{1}{\phi} g \, (X_s' X_s)^{-1}\right)$$

$$\boxed{\pi(\phi) = 1/\phi}$$

- $y \mid \beta_s, S, \phi \sim N\left(X_s \beta_s, \frac{1}{\phi} I_n\right)$
- $\beta_s \mid S, \phi \sim N\left(0, g \, \phi^{-1} (X_s' X_s)^{-1}\right)$
- $S \sim \left(\frac{1}{p}\right)^{k \, p 7} I(p_1 \leq s_0)$
- $\phi \sim \pi(\phi) = 1/\phi$

$y \mid S$

$$p(y \mid S) = \int p(y \mid \beta_S, \phi, S)\, \pi(\beta_S \mid \phi)\, \pi(\phi)\, d\beta_S\, d\phi$$

$$= \boxed{\frac{\Gamma\left(\frac{n}{2}\right)(1+g)^{n/2}}{\pi^{n/2}\, \|y\|_2^n}} \frac{(1+g)^{-p_\gamma/2}}{\left(1+g\left(1-R_\gamma^2\right)\right)^{n/2}}$$

$$R_\gamma^2 = \frac{y^T \Phi_\gamma y}{\|y\|_2^2} \quad : \text{coefficient of determination}$$

$$\Phi_\gamma = \underline{X_\gamma \left(X_\gamma^T X_\gamma\right)^{-1} X_\gamma^T}$$

$p(y \mid S)$ is large if X_S is a "good fit" for the linear.

$$\pi(s \mid y) \propto \frac{p(y \mid s)\, p(s)}{\sum_{s} p(y \mid s)\, p(s)}$$

$S \subset \{0, 1\}^p$

<u>Goal:</u> Develop a MCMC which has $\pi(s \mid y)$ as the stationary distribution

Current at γ. $\underline{N(\gamma)}$: neighborhood of γ

<u>Proposal</u>: $S(\gamma, \cdot)$ is a prob. dist. on $N(\gamma)$ which depends on γ.

<u>Step I</u>: Choose $\gamma' \in N(\gamma)$ according to $S(\gamma, \cdot)$

Step 2: Move to γ' with prob. $R(\gamma, \gamma')$ and stay at γ w.p

$1 - R(\gamma, \gamma')$.

$$R(\gamma, \gamma') = \min \left\{ 1, \frac{\pi(\gamma' \mid y)\, S(\gamma', \gamma)}{\pi(\gamma \mid y)\, S(\gamma, \gamma')} \right\}$$

Currently at γ

With prob $1/2$, you do

- Single flip update $j \in \{1, 2, \dots, p\}$ at random $\gamma_j' = 1 - \gamma_j$

- Double flip update

$S(\gamma) = \{ j \in \{1, 2, \dots, p\} : \gamma_j = 1 \}$

$S^c(\gamma)$

$(k, l) \in S(\gamma) \times S^c(\gamma)$ uniformly and create γ' by γ_k from 1 to 0 and γ_l from 0 to 1

Prove that

$N_1(\gamma) = \{ \gamma' : d_H(\gamma', \gamma) = 1 \}$

$N_2(\gamma) = \{ \gamma' : d_H(\gamma', \gamma) = 2 \}$

$$P_{MH}(\gamma, \gamma')$$

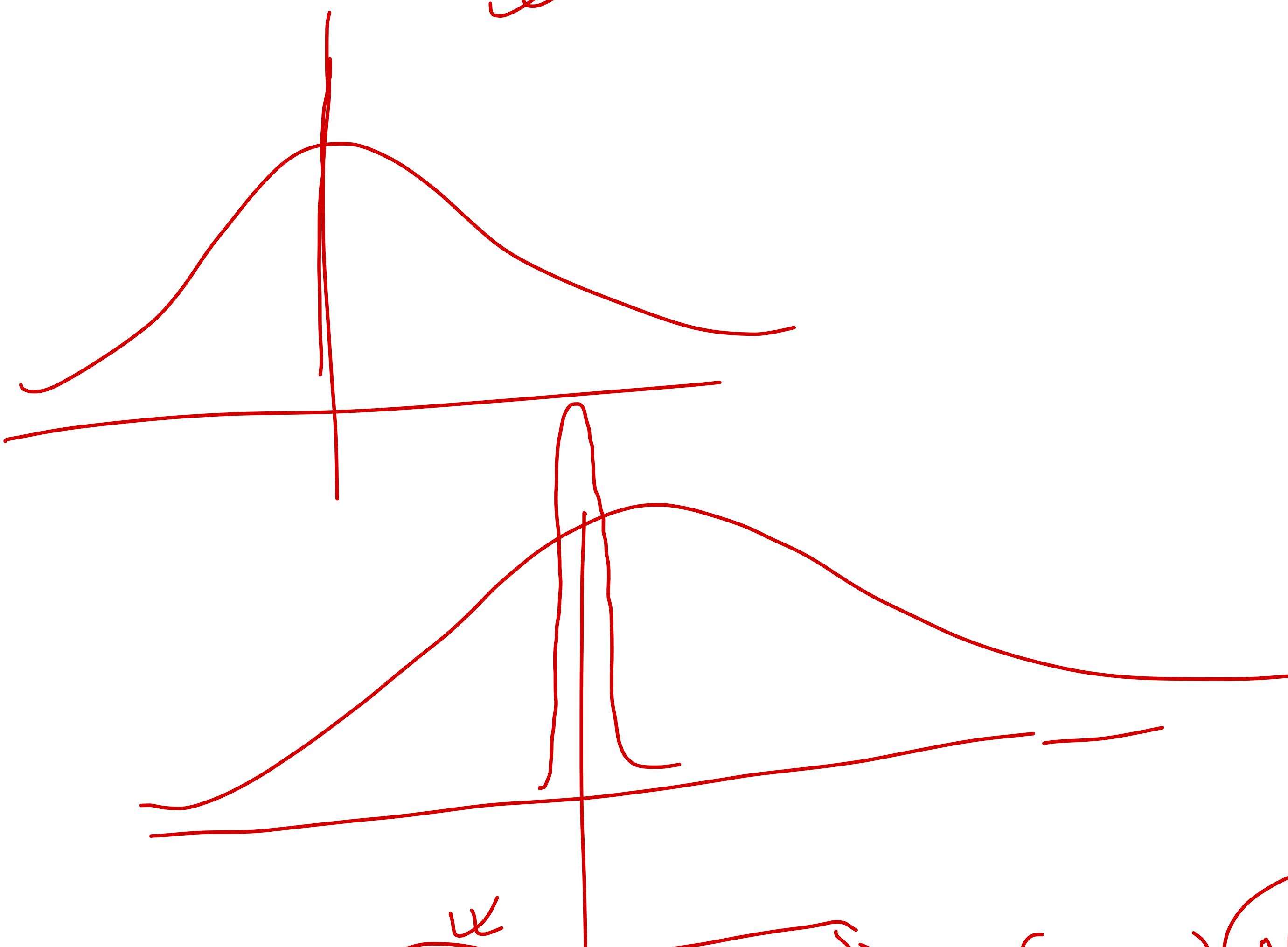
$$= \begin{cases} \frac{1}{2} \cdot \frac{1}{p} \min\left\{ 1, \frac{\pi(\gamma' \mid y)}{\pi(\gamma \mid y)} \right\}, & \gamma' \in N_1(\gamma) \\ \frac{1}{2\,|S(\gamma)|\,|S^c(\gamma)|} \min\left\{ 1, \frac{\pi(\gamma' \mid y)}{\pi(\gamma \mid y)} \right\}, & \gamma' \in N_2(\gamma) \\ 1 - \sum_{\tilde{\gamma} \neq \gamma} P_{MH}(\gamma, \tilde{\gamma}) & \text{if } \gamma = \gamma' \\ 0 & \text{if } d_H(\gamma, \gamma') > 2 \end{cases}$$

$g = p^2$ or p^3 , $k \in (1, 2)$

Non-collapsed Gibbs sampler (Jointly update β)

$\beta \mid Y, X$

$$\boxed{\beta_j = \pi_0 \delta_0 + (1-\pi_0) N(0, c_j^2)}$$



$$\beta_j = \pi_0 N(0, d_j^2) + (1-\pi_0) N(0, c_j^2)$$

$\left.\begin{array}{l} d_j \text{ is small} \\ c_j \text{ is large} \end{array}\right\}$

$$\underline{\beta_j \mid \beta_{-j}, y, X}$$

$$\beta \mid y, X$$

$$\underline{\gamma_j = 0/1}$$

w.p $1-\pi_0$

$$\beta_j \mid \gamma_j \sim (1-\gamma_j) N(0, d_j^2) + \gamma_j N(0, c_j^2)$$

$$[\beta, \gamma \mid y, X] \propto \underbrace{[y \mid X, \beta, \gamma][\beta \mid \gamma]}[\gamma]$$

$$\pi(\beta \mid \gamma) = N(0, D)$$

$$[\beta \mid \gamma, y, X) \sim N \underbrace{(y; X\beta, \sigma^2 I) N(0, D)}$$

$$[\gamma_j \mid \beta_j, \underline{y, X}] \propto [\beta_j \mid \gamma_j][\gamma_j]$$

$$p(\gamma_j = 1 \mid -) \propto (\beta_j \mid \gamma_j = 1)(1-\pi_0)$$