Variational Approx. for Linear Models

$$y = X\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n)$$

$$\pi(\delta, \beta) = \pi(\delta)\pi(\beta \mid \delta) \qquad \delta = (\delta_1, \dots, \delta_p) \in \{0,1\}^p$$

$$\pi(\delta) \propto \left(\frac{q}{1-q}\right)^{\|\delta\|}, \quad \beta_j \mid \delta_j \overset{ind}{\sim} \begin{cases} N(0, 1/g) & \delta_j = 1 \\ N(0, \gamma) & \delta_j = 0 \end{cases}$$

Choose $q = 1/p^{\mu}$ $\mu > 1$ enforces sparsity.

$$\pi(\delta, \beta \mid y) \propto \exp\left\{ -\frac{1}{2\sigma^2} \|Y - X\beta\|_2^2 - \frac{1}{2}\beta^T D^{-1} \beta \right\} \prod_{j=1}^{p} \left(\sqrt{\frac{g}{2\pi}}\, q\right)^{\delta_j} \left(\frac{1-q}{\sqrt{2\pi\gamma}}\right)^{1-\delta_j}$$

$$D_j = \begin{cases} \gamma & \delta_j = 0 \\ 1/g & \delta_j = 1 \end{cases}$$

<u>Gibbs</u>: $\delta_j \sim \text{Bern}(\alpha_j)$ $\qquad \alpha_j = \dfrac{1}{1 + \frac{1-q}{q}\left(\frac{1}{\gamma g}\right)^{1/2} e^{-\theta_j^2/2\,(1/\gamma - g)}}$

$$\theta \mid \delta \sim MVN(\mu, \Sigma) \qquad \mu = (X^T X + \sigma^2 D^{-1})^{-1} X^T Y$$

$$\Sigma = \sigma^2 (X^T X + \sigma^2 D^{-1})^{-1}$$

Message: Choose initial distribution more carefully

<u>e.g.</u> use Lasso solution to start.

Variational Approximation

$$q(\delta, \beta) = \prod_{j=1}^{p} \text{Bern}(q_j)\, N(\mu_j, v_j^2)$$

Conjugate model $\Longrightarrow$ CAVI

z_j update: $$z_j = \frac{1}{1 + \frac{1-z}{z}\left(\frac{1}{r_j}\right)^{1/2} e^{-1/2(\mu_j^2 + v_j^2)(1/\gamma - \rho)}}$$

$$v_j = \frac{1}{\sigma^2 z_j \rho + \frac{(1-z_j)}{\gamma} + \frac{\|x_j\|_2^2}{\sigma^2}}$$

$$\mu_j = \frac{\langle X_j, y - X_{-j}\mu_{-j} \rangle}{\sigma^2 z_j \rho + \frac{(1-z_j)}{\gamma} + \|x_j\|_2^2}$$

Algo: Choose $(z^{(0)}, \mu^{(0)}, v^{(0)})$

Iterate: Update $z^{(j)}, \mu^{(j)}, v^{2(j)}$

$v^{2\,(0)}$ should be small.