July 04, 2019

1 Variable Selection

Start with spline model

$$Y = b(X)\beta + \epsilon$$

Select variable? select basis?

- $Y = b(\Gamma X)\beta + \epsilon$
- $Y = b(X)\Gamma\beta + \epsilon$

2 Basis Selection Model

$$Y = b(X)\Gamma\beta + \epsilon$$

Let $Z = b(X)$

2.1 prior

$$\begin{aligned}
Y|\beta, \sigma^2, \Gamma &\sim N(Z\Gamma\beta, \sigma^2 \cdot I) \\
\sigma^2 &\sim Inverse - Gamma(a, b) \\
\beta_j|\sigma^2_{\beta_j} &\overset{\text{ind}}{\sim} N(0, \sigma^2_{\beta_j}) \\
\sigma^2_{\beta_j} &\overset{\text{iid}}{\sim} Inverse - Gamma(c, d) \\
\gamma_j &\overset{\text{iid}}{\sim} Bernoulli(\rho)
\end{aligned}$$

- Z is n x p design matrix
- $\Gamma = diag(\gamma_j)$ for $j = 1, \ldots, p$
- a, b, c, d is flat prior

2.2 posterior

$$
\begin{aligned}
p(\beta, \sigma^2, \sigma_\beta^2, \Gamma | Y) \propto & p(\beta, \sigma^2, \sigma_\beta^2, \Gamma, Y) \\
\propto & p(Y|\beta, \sigma^2, \Gamma) p(\beta|\sigma_\beta^2) p(\Gamma) p(\sigma^2) p(\sigma_\beta^2) \\
\propto & \left(\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\left(Y - Z\Gamma\beta\right)'\left(Y - Z\Gamma\beta\right)\right) \\
& \times \prod_{j=1}^{p} \left(\sigma_{\beta_j}^2\right)^{-1/2} \exp\left(-\frac{1}{2}\sum_{j=1}^{p} \frac{\beta_j^2}{\sigma_{\beta_j}^2}\right) \\
& \times \prod_{j=1}^{p} \rho^{\gamma_j} \left(1-\rho\right)^{-\gamma_j} \\
& \times \left(\sigma^2\right)^{-a-1} \exp\left(-\frac{b}{\sigma^2}\right) \\
& \times \prod_{j=1}^{p} \left(\sigma_{\beta_j}^2\right)^{-c-1} \exp\left(-\sum_{j=1}^{p} \frac{d}{\sigma_{\beta_j}^2}\right)
\end{aligned}
$$

2.3 Variational Inference

$$
p(\beta, \sigma^2, \sigma_\beta^2, \Gamma | Y) \approx q(\beta, \sigma^2, \sigma_\beta^2, \Gamma) = \prod_{j=1}^{p} q_1^*(\beta_j) q_2^*(\sigma^2) \prod_{j=1}^{p} q_3^*(\sigma_{\beta_j}^2) \prod_{j=1}^{p} q_4^*(\gamma_j)
$$

Use coordinate ascent algorithm, q density of β is

$$
\begin{aligned}
q_1^*(\beta) &\propto E_{-q_1}\left[p(\beta, \sigma^2, \sigma_\beta^2, \Gamma, Y)\right] \\
&\propto E_{-q_1}\left[\exp\left(-\frac{1}{2\sigma^2}\left(Y - Z\Gamma\beta\right)'\left(Y - Z\Gamma\beta\right) - \frac{1}{2}\sum_{j=1}^{p} \frac{\beta_j^2}{\sigma_{\beta_j}^2}\right)\right] \\
&\propto \exp\left(-\frac{1}{2}\beta' \langle D \rangle - \frac{1}{2}\left\langle \frac{1}{\sigma^2} \right\rangle \left(\beta' \langle \Gamma \rangle' Z' Z \langle \Gamma \rangle \beta - 2\beta' \langle \Gamma \rangle' Z' Y\right)\right) \\
&\propto \exp\left(-\frac{1}{2}\left[\beta'\left(\langle D \rangle + \left\langle \frac{1}{\sigma^2} \right\rangle \langle \Gamma \rangle' Z' Z \langle \Gamma \rangle\right)\beta - 2\left\langle \frac{1}{\sigma^2} \right\rangle \beta' \langle \Gamma \rangle' Z' Y\right]\right) \\
&\sim N(\mu, \Sigma)
\end{aligned}
$$

Where $D = diag(\frac{1}{\sigma_{\beta_j}^2})$, ¡¿ means expectation under q functions and

$$
\Sigma = \left(\langle D \rangle + \beta' \left\langle \frac{1}{\sigma^2} \right\rangle \langle \Gamma \rangle' Z' Z \langle \Gamma \rangle\right)^{-1}, \quad \mu = \left\langle \frac{1}{\sigma^2} \right\rangle \Sigma \langle \Gamma \rangle' Z' Y
$$

q density of σ^2 is

$$
\begin{aligned}
q_2^*(\sigma^2) &\propto E_{-q_2}\left[p(\beta, \sigma^2, \sigma_\beta^2, \Gamma, Y)\right] \\
&\propto E_{-q_2}\left[\left(\sigma^2\right)^{-\left(\frac{n}{2}+a\right)-1} \exp\left(-\frac{1}{\sigma^2}\left(b + (Y - Z\Gamma\beta)'\,(Y - Z\Gamma\beta)\right)\right)\right] \\
&\propto Inverse - Gamma\left(a + \frac{n}{2}, b + \frac{1}{2}\left(Y'Y - 2\left\langle\beta\right\rangle' \left\langle\Gamma\right\rangle' Z'Y + tr\left(\left(Z'Z \odot \Omega\right)\left(\mu\mu' + \Sigma\right)\right)\right)\right)
\end{aligned}
$$

Where *odot* is hadamard product and

- $\gamma = (\gamma_1, \dots, \gamma_p)$
- $\Omega = \left\langle\gamma\right\rangle \left\langle\gamma\right\rangle' + \left\langle\Gamma\right\rangle \odot (I - \left\langle\Gamma\right\rangle)$

q density of $\sigma_{\beta_j}^2$ is

$$
\begin{aligned}
q_3^*(\sigma_\beta^2) &\propto E_{-q_3}\left[p(\beta, \sigma^2, \sigma_\beta^2, \Gamma, Y)\right] \\
&\propto \prod_{j=1}^{p}\left[\left(\sigma_{\beta_j}^2\right)^{-\left(\frac{1}{2}+c\right)-1} \exp\left(-\frac{1}{\sigma_{\beta_j}^2}\left(d + \frac{1}{2}\left\langle\beta_j^2\right\rangle\right)\right)\right] \\
q_3^*(\sigma_{\beta_j}^2) &\sim Inverse - Gamma\left(c + \frac{1}{2}, d + \frac{1}{2}\left\langle\beta_j^2\right\rangle\right)
\end{aligned}
$$

q density of γ is

$$
\begin{aligned}
q_4^*(\gamma) &\propto E_{-q_4}\left[\propto \prod_{j=1}^{p} \rho^{\gamma_j}(1-\rho)^{-\gamma_j} \exp\left(-\frac{1}{2}\frac{1}{\sigma^2}\left(\beta'\Gamma'Z'Z\Gamma\beta - 2\beta'\Gamma'Z'y\right)\right)\right] \\
&\propto \exp\left[\sum_{j=1}^{p}\gamma_j\left(\log\left(\rho/(1-\rho)\right) - \frac{1}{2}\left\langle\frac{1}{\sigma^2}\right\rangle\left\langle\beta_j^2\right\rangle Z_j'Z_j + \left\langle\frac{1}{\sigma^2}\right\rangle Z_j'\left[Y\mu_j - X_{-j}\left\langle\Gamma_{-j}\right\rangle\left(\mu_{-j}\mu_j + \Sigma_{-j,j}\right)\right]\right)\right]
\end{aligned}
$$

$q_4^*(\gamma_j) \sim Bernoulli(w_j)$

Where

$$
w_j = expit\left[\left(\log\left(\rho/(1-\rho)\right) - \frac{1}{2}\left\langle\frac{1}{\sigma^2}\right\rangle\left\langle\beta_j^2\right\rangle Z_j'Z_j + \left\langle\frac{1}{\sigma^2}\right\rangle Z_j'\left[Y\mu_j - X_{-j}\left\langle\Gamma_{-j}\right\rangle\left(\mu_{-j}\mu_j + \Sigma_{-j,j}\right)\right]\right)\right]
$$

and

- X_j means jth coloumn of X
- X_{-j} means without jth column
- $X_{-i,j}$ means jth column without ith component
- μ_j is jth component of vector and mu_j means without jth component