Bayesian Basics

We are interested in obtaining the posterior distribution for making inferences.

$$posterior = \frac{likelihood * prior}{\int likelihood * prior}$$

It is often easier to work without the denominator, and recognize the posterior density.

$$posterior \propto likelihood * prior$$

Want to know characteristics about the posterior (such as posterior mean), and sample from the posterior.

Review of Marginal and Conditional Distributions

$$p(y|x) = \frac{p(x,y)}{p(y)} \text{ is conditional of y given x}$$
$$p(x) = \int p(x,y)\delta y = \int p(y|x)p(x)\delta y \text{ is marginal of x}$$
$$\text{Under independence, } p(x,y) = p(y|x)p(x)$$

Basic Hierarchical Model Setup

Normal-Gamma Gaussian Linear Model:

$$(y_i|\theta, \sigma^2) \sim N\left(\theta, \sigma^2\right), i = 1, ..., n$$
$$(\theta|\sigma^2) \sim N\left(\mu, t\tau^2\sigma^2\right)$$
$$\sigma^2 \sim Inv-Gamma(\frac{d}{2}, \frac{\eta}{2})$$

Reframed to work with precision $\omega = 1/\sigma^2$, called the Normal-Gamma Prior for (θ, ω):

$$(y_i|\theta, \omega) \sim N\left(\theta, (\omega)^{-1}\right), i = 1, ..., n$$
$$(\theta|\omega) \sim N\left(\mu, (k\omega)^{-1}\right)$$
$$\omega \sim Gamma(\frac{d}{2}, \frac{\eta}{2})$$

Hyperparameter interpretation:

- $\mu =$ Prior guess for θ.
- $\tau =$ Prior signal-to-noise ratio for σ^2.
- $d =$ Prior sample size for σ^2.
- $\eta =$ Prior sum of squares for σ^2; $\frac{\eta}{d} \approx$ prior guess for σ^2.
- $k = \frac{1}{\tau^2}$ is prior sample size for θ.

Prior

Joint prior:

$$p(\theta, \omega) = p(\theta|\omega)p(\omega)$$

Marginal priors:

$$\begin{aligned} p(\omega) &= p(\omega) \\ p(\theta|\omega) &= \int p(\theta, \omega)\delta\omega = \int p(\theta|\omega)p(\omega)\delta\omega \end{aligned}$$

Likelihood (Sampling Model)

$$p(y|\theta, \omega) = \prod_{i=1}^{n} p(y_i|\theta, \omega)$$

Posterior

Joint posterior (up to a constant of proportionality):

$$\begin{aligned} p(\theta, \omega|y) &\propto p(y|\theta, \omega)p(\theta, \omega) \\ &= p(y|\theta, \omega)p(\theta|\omega)p(\omega) \end{aligned}$$

Marginal posteriors.

$$\begin{aligned} p(\theta|y) &= \int p(\theta, \omega|y)\delta\omega \\ p(\omega|y) &= \int p(\theta, \omega|y)\delta\theta \end{aligned}$$

Conditional posteriors. (Full conditionals.)

$$\begin{aligned} p(\theta|y, \omega) &\propto p(y|\theta, \omega)p(\theta) \\ p(\omega|y, \theta) &\propto p(y|\theta, \omega)p(\omega) \end{aligned}$$

- Obtain the joint posterior, and read off forms of conditionals by considering all other variables constant.
- Need all full conditionals to sample from joint posterior using Gibbs Sampler or similar. (Is easier than sampling from joint posterior.)

More Complex Hierarchical Model Setup

Normal-Gamma Gaussian Multivariate Linear Model (Bayesian Regression):

$$\begin{aligned} (y|\beta,\omega,\Lambda) &\sim N\left(X\beta, (\omega\Lambda)^{-1}\right) \\ \lambda_i \overset{iid}{\sim} Gamma\left(\frac{h}{2}, \frac{h}{2}\right) \\ (\beta|\omega) &\sim N\left(m, (\omega K)^{-1}\right) \\ \omega &\sim Gamma\left(\frac{d}{2}, \frac{\eta}{2}\right) \end{aligned}$$

Prior

Joint prior:

$$p(\beta,\omega,\Lambda) = p(\beta|\omega)p(\omega)p(\Lambda)$$

Likelihood

$$p(y|\beta,\omega,\Lambda) \propto N\left(X\beta, (\omega\Lambda)^{-1}\right)$$

Posterior

Joint posterior:

$$p(\beta,\omega,\Lambda|y) \propto p(y|\beta,\omega,\Lambda)p(\beta|\omega)p(\omega)p(\Lambda)$$

Full posterior conditionals.

$$\begin{aligned} p(\beta|\omega,\Lambda,y) &\propto p(y|\beta,\omega,\Lambda)p(\beta|\omega)p(\omega)p(\Lambda) \\ &\propto p(y|\beta,\omega,\Lambda)p(\beta|\omega) \end{aligned}$$

$$\begin{aligned} p(\omega|\beta,\Lambda,y) &\propto p(y|\beta,\omega,\Lambda)p(\beta|\omega)p(\omega)p(\Lambda) \\ &\propto p(y|\beta,\omega,\Lambda)p(\beta|\omega)p(\omega) \end{aligned}$$

$$\begin{aligned} p(\lambda_i|y,\beta,\omega) &\propto p(y|\lambda_i,\beta,\omega)p(\lambda_i|\beta,\omega)p(\beta|\omega)p(\omega) \\ &\propto p(y|\lambda_i,\beta,\omega)p(\lambda_i) \end{aligned}$$

- For full conditionals, read these off of the joint posterior, by considering all variables on the right side of the $|$ to be constant.