MCMC Examples

Rat Population growth data (Gelfand et al. 1990, JASA 1990)

	Weights Y_{ij} of rat i on day x_{ij}				
	$x_1 = 8$	$x_2 = 15$	$x3 = 22$	$x_4 = 29$	$x_5 = 36$
rat 1	151	199	246	283	320
rat 2	145	199	249	293	354
...					
rat 30	153	200	244	286	324

- Y_{ij} weight of the ith rat at measurement point j
- x_{ij} denotes the rat's age in days at time point j
- **Stage I: Sampling Distribution**

$$Y_{ij} \sim N(\alpha_i + \beta_i x_{ij}, \sigma^2),\ i = 1, \ldots, k = 30\ j = 1, \ldots, n = 5$$

- **Stage II: Prior**

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} \sim N\left(\begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix}, \Sigma\right),\ i = 1, \ldots, k$$

- **Stage III: Hyperprior**

$$\sigma^2 \sim IG(a, b)$$

$$\begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} \sim N\left(\begin{bmatrix} \eta_0 \\ \eta_1 \end{bmatrix}, C\right)$$

$$\Sigma^{-1} \sim W((\rho R)^{-1}, \rho),\ E(\Sigma^{-1}) = R^{-1},\ var(\Sigma) \propto \rho^{-1}$$

- we seek the marginal posterior distribution for α_0, β_0 given the observed data and predictive intervals for the individual future growth given the first-week measurement
- the number of unknown parameters is 66: (30 α_is + 30 β_is + α_0 + β_0 + σ^2 + 3 unique component of Σ).
- let's re-write the sampling distribution as:

$$\boldsymbol{y}_i \sim N(X_i \boldsymbol{\theta}_i, \sigma^2 I_n),\ i = 1, \ldots, k = 30\ j = 1, \ldots, n = 5$$

where:

$$\boldsymbol{y}_i^t = (y_{i1}, \ldots, y_{in_i}),\ X_i = \begin{pmatrix} 1 & x_{i1} \\ \vdots & \vdots \\ 1 & x_{in_i} \end{pmatrix},\ \boldsymbol{\theta}_i^t = (\alpha_i, \beta_i)$$

- find the full conditional distributions

$$\begin{aligned}
\boldsymbol{\theta}_i \mid \boldsymbol{y}, \boldsymbol{\theta}_0, \Sigma^{-1}, \sigma^2 &\sim N\left(D_i\left[\sigma^{-2} X_i^t \boldsymbol{y}_i + \Sigma^{-1}\boldsymbol{\theta}_0\right], D_i\right) \\
&\quad \text{ind } i = 1, \ldots, k \\
\boldsymbol{\theta}_0 \mid \boldsymbol{y}, \{\boldsymbol{\theta}_i\}, \Sigma^{-1}, \sigma^2 &\sim N\left(V\left[k\Sigma^{-1}\bar{\boldsymbol{\theta}} + C^{-1}\boldsymbol{\eta}\right], V\right) \\
\Sigma^{-1} \mid \boldsymbol{y}, \{\boldsymbol{\theta}_i\}, \boldsymbol{\theta}_0, \sigma^2 &\sim W\left(\left[\sum_{i=1}^k (\boldsymbol{\theta}_i - \boldsymbol{\theta}_0)^t(\boldsymbol{\theta}_i - \boldsymbol{\theta}_0) + \rho R\right]^{-1}, k + \rho\right) \\
\sigma^2 \mid \boldsymbol{y}, \{\boldsymbol{\theta}_i\}, \boldsymbol{\theta}_0, \Sigma &\sim IG\left(\frac{kn}{2} + a, \left[\frac{1}{2}\sum_{i=1}^k (\boldsymbol{y}_i - X_i\boldsymbol{\theta}_i)^t(\boldsymbol{y}_i - X_i\boldsymbol{\theta}_i) + b^{-1}\right]^{-1}\right)
\end{aligned}$$

where

- $D_i^{-1} = \sigma^{-2} X_i^t X_i + \Sigma^{-1},\ \boldsymbol{\theta}_0^t = (\alpha_0, \beta_0)^t$
- $V = (k\Sigma^{-1} + C^{-1})^{-1},\ \bar{\boldsymbol{\theta}} = \frac{1}{k}\sum_{i=1}^k \boldsymbol{\theta}_i$

Prior Specification

- so that it is now reasonable to think of α_i and β_i as independent a priori. Thus we set $\Sigma = diag(\sigma_\alpha^2, \sigma_\beta^2)$ and replace the Wishart prior by a product of Inverse Gamma distributions.
- prior specification $C^{-1} = 0,\ a = b = 1/0.0001,\ \rho = 2,\ R = \begin{pmatrix} 100 & 0 \\ 0 & 0.1 \end{pmatrix}$

Reparametrization

- each rat was weighted one a week for five consecutive weeks. As a result, we may simplify our computations by rewriting the likelihood as
- Stage I:

$$Y_{ij} \sim N(\alpha_i + \beta_i(x_{ij} - \bar{x}), \sigma^2),\ i = 1, \ldots, k,\ j = 1, \ldots, n$$

- so that it is now reasonable to think of α_i and β_i as independent a priori. Thus we set $\Sigma = diag(\sigma_\alpha^2, \sigma_\beta^2)$ and replace the Wishart prior by a product of Inverse Gamma distributions.

Flour beetle mortality data Prentice Biometrics 1976

Dosage w_i	# killed y_i	# exposed w_i
1.6907	6	59
1.7242	13	60
1.7552	18	62
1.7842	28	56
1.8113	52	63
1.8369	53	59
1.8610	61	62
1.8839	60	60

- these data record the number of adult flour beetles killed after five hours of exposure to various levels of gaseous carbon disulphide (CS_2)
- we use a generalized logit model (Prentice (1976))

$$P(death \mid w) = h(w) = \{\exp(x)/(1 + \exp(x))\}^{m_1}$$

where

- $x = \frac{w - \mu}{\sigma},\ \mu \in \mathcal{R}$ and $\sigma^2,\ m_1 > 0$
- μ, σ^2, m_1 are unknown parameters
- This model is a generalization of the logit and probit models currently used in biossay experiments
- $m_1 = 1$ logistic model
- $m_1 < 1$ then $P(death \mid w)$ is negatively skewed

- $m_1 > 1$ then $P(death \mid w)$ is positively skewed
- the form of $P(death \mid w)$ approaches the distribution function of an extreme minimum value distribution for $m_1 \to \infty$ and a reflected exponential distribution for $m_1 \to 0$
- MLE with a Newton-Raphson algorithm give $\hat{\mu} = 1.818, \ \hat{\sigma} = 0.16, \ \hat{m_1} = .279$
- our goal is to approximate $p(\mu, \sigma, m_1 \mid \text{data})$ by implementing a Metropolis Algorithm
- Prior distributions

$$\begin{aligned} m_1 &\sim \text{Gamma}(a_0, b_0) \\ \mu &\sim N(c_0, d_0) \\ \sigma^2 &\sim IG(e_0, f_0) \end{aligned}$$

- Likelihood-prior

$$\begin{aligned} p(\mu, \sigma, m_1 \mid \text{data}) &\propto f(\text{data} \mid \mu, \sigma, m_1)\pi(\mu, \sigma, m_1) \\ &\propto \left\{\prod_{i=1}^k [h(w_i)]^{y_i}[1-h(w_i)]^{n_i-y_i}\right\} \frac{{m_1}^{a_0-1}}{\sigma^{2(e0+1)}} \\ &\times \ \exp\left[-\frac{1}{2}\left(\frac{\mu-c_0}{d_0}\right)^2 - \frac{m_1}{b_0} - \frac{1}{f_0\sigma^2}\right] \end{aligned}$$

- we make a change of variable to $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3) = (\mu, \frac{1}{2}\log\sigma^2, \log r$ this transform the parameter space in $\mathcal{R}^3$ and also helps to symmetrize the posterior distribution, our target density is now

$$\begin{aligned} p(\boldsymbol{\theta} \mid \text{data}) &\propto \left\{\prod_{i=1}^k [h(w_i)]^{y_i}[1-h(w_i)]^{n_i-y_i}\right\} \exp(a_0\theta_3 - 2e_0\theta_2 \\ &\times \ \exp\left[-\frac{1}{2}\left(\frac{\theta_0-c_0}{d_0}\right)^2 - \frac{\exp(\theta_3)}{b_0} - \frac{exp(-2\theta_2)}{f_0}\right] \end{aligned}$$

- prior specifications
- m_1: $a_0 = .25, \ b_0 = 4$ so that $E[m_1] = 1, \ std[m_1] = 2$
- μ and σ^2: $c_0 = 2, \ d_0 = 10, e_0 = 2, \ f_0 = 1000$ that is $E[\sigma^2] = 0.001. \ std[\sigma^2] = .5$
- proposal distribution $N_3(\boldsymbol{\theta}^{(t-1)}, \tilde{\Sigma})$ where

$$\tilde{\Sigma} = diag(.00012, .033, .10)$$

- to accelerate convergence, form the output of the first algorithm get a new estimate of $\tilde{\Sigma}$ as $\frac{1}{N}\sum_{j=1}^N(\boldsymbol{\theta}_j - \bar{\boldsymbol{\theta}})(\boldsymbol{\theta}_j - \bar{\boldsymbol{\theta}})^t$ where j indexes Monte Carlo samples.