David Dahl

Illustrating of the Gibbs Sampler

"mean" "precision" = 1/"VARIANCE"

"Model for the data" $\left\{ \quad X_1, X_2, \ldots, X_n \mid \mu, \lambda \overset{i.i.d.}{\sim} N(\mu, \lambda) \right.$

"prior distribution" $\left\{ \begin{array}{ll} \mu \sim N(\mu_0, \lambda_0) \\ \lambda \sim Ga(\alpha_0, \beta_0) \end{array} \right.$ independence is implied by the notation, where we choose the parametrization of the gamma distribution where $E(\lambda) = \alpha/\beta$

We are interested in making inference about μ and λ from the observed data. That is, we make inference from the posterior distribution of μ and λ given the data:

likelihood prior

$$\underbrace{p(\mu, \lambda \mid x_1, \ldots, x_m)}_{\text{posterior}} = \frac{p(x_1, \ldots, x_m \mid \mu, \lambda)\, p(\mu, \lambda)}{p(x_1, \ldots, x_m)}$$

marginal distribution of data which is constant with respect to μ and λ, so

$$\propto p(x_1, \ldots, x_m \mid \mu, \lambda)\, p(\mu, \lambda)$$

$$= p(x_1, \ldots, x_m \mid \mu, \lambda)\, p(\mu)\, p(\lambda) \quad \text{by independence}$$

$$= \left[\prod_{i=1}^{m} \frac{\lambda}{\sqrt{2\pi}} \exp\left(-\frac{\lambda}{2}(x_i - \mu)^2\right) \right] \frac{\lambda_0}{\sqrt{2\pi}} \exp\left(-\frac{\lambda_0}{2}(\mu - \mu_0)^2\right) \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)} \lambda^{\alpha_0 - 1} \exp(-\beta_0 \lambda)\, I\{\lambda > 0\}$$

$$\propto \exp\left(-\frac{\lambda}{2} \sum (x_i - \mu)^2 - \frac{\lambda_0}{2}(\mu - \mu_0)^2 - \beta_0 \lambda\right) \lambda^{\alpha_0 - 1}\, I\{\lambda > 0\}$$

Unfortunately, this kernel is not from a known family. Therefore, estimating μ with $\hat{\mu} = E(\mu \mid x_1, \ldots, x_m)$ AND λ with $\hat{\lambda} = E(\lambda \mid x_1, \ldots, x_m)$ will be difficult to do analytically.

Instead, we can use Monte Carlo integration, specifically we can use Markov chain Monte Carlo. Even more specifically, we can use the Gibbs sampler since the full conditionals are available.

David Dahl

Full conditional of μ:

$$(*) \quad \begin{cases} p(\mu \mid \lambda, x_1, \dots, x_m) = \dfrac{p(\mu, \lambda \mid x_1, \dots, x_m)}{p(\lambda \mid x_1, \dots, x_m)} & \text{by the definition of conditional probability} \\[2ex] \quad \propto \; p(\mu, \lambda \mid x_1, \dots, x_m) & \text{which is the posterior} \end{cases}$$

$$\propto \; \exp\left(-\frac{\lambda}{2} \sum (x_i - \mu)^2 - \frac{\lambda_0}{2} (\mu - \mu_0)^2\right)$$

$\vdots \;\leftarrow$ skipping a few steps where we "complete the square"

$$\propto \; \frac{(\lambda_0 + m\lambda)}{\sqrt{2\pi}} \exp\left(-\frac{(\lambda_0 + m\lambda_0)}{2} \left(\mu - \frac{\lambda_0\mu_0 + \lambda m \bar{x}}{\lambda_0 + m\lambda}\right)^2\right)$$

"precision"

$$(1) \quad \Rightarrow \quad \mu \mid \lambda, x_1, \dots, x_m \;\sim\; N\left(\frac{\lambda_0\mu_0 + \lambda m \bar{x}}{\lambda_0 + m\lambda}, \; \lambda_0 + m\lambda\right)$$

"mean"

Full conditional of λ:

Likewise, it is straight-forward to show that

$$(2) \quad \lambda \mid \mu, x_1, \dots, x_m \;\sim\; Ga\left(\alpha_0 + \tfrac{1}{2} n, \; \beta_0 + \tfrac{1}{2} \sum (x_i - \mu)^2\right)$$

Gibbs sampler: Set the initial state $\mu^{(1)}, \lambda^{(1)}$

For $t = 2, 3, \dots, B$:

- Sample $\mu^{(t)}$ using (1) above with $\lambda = \lambda^{(t-1)}$
- Sample $\lambda^{(t)}$ using (2) above with $\mu = \mu^{(t)}$

Metropolis-Hastings Justification:

For updating μ from $\mu^{(t-1)}$ to $\mu^{(t)}$:

$$r_{MH} = \frac{p(\mu^{(t)}, \lambda^{(t-1)} \mid x_1, \dots, x_m)}{p(\mu^{(t-1)}, \lambda^{(t-1)} \mid x_1, \dots, x_m)} \; \frac{p(\mu^{(t-1)} \mid \lambda^{(t-1)}, x_1, \dots, x_m)}{p(\mu^{(t)} \mid \lambda^{(t-1)}, x_1, \dots, x_m)}$$

$$\underline{\text{by } (*) \text{ above}} \quad = \frac{\cancel{p(\mu^{(t)} \mid \lambda^{(t-1)}, x_1, \dots, x_m)}}{\cancel{p(\mu^{(t-1)} \mid \lambda^{(t-1)}, x_1, \dots, x_m)}} \; \frac{\cancel{p(\mu^{(t-1)} \mid \lambda^{(t-1)}, x_1, \dots, x_m)}}{\cancel{p(\mu^{(t)} \mid \lambda^{(t-1)}, x_1, \dots, x_m)}}$$

$$= 1 \quad \Rightarrow \quad \text{All proposals are accepted!!!}$$

A similiar calculation holds for updating λ.