STATS370: Final Project

Erich Trieschman

2022 Fall Quarter

1 Introduction

Priors

$$
\begin{aligned}
\theta &= (\sigma^2, \tau, \mu_1, \mu_2, \gamma_1, \gamma_2) \\
p(\sigma^2) &\propto \frac{1}{\sigma^2} \\
p(\tau) &\sim Unif[0,1] \\
p(\mu) &= p(\mu_1, \mu_2) \propto 1 \text{ (improper uniform)} \\
p(\gamma) &= p(\gamma_1, \gamma_2) \propto 1 \text{ (improper uniform)}
\end{aligned}
$$

Gene expression distributions

$$
\begin{aligned}
(y_i|g_i = 1) &\sim N(\mu, \sigma^2 I) \\
(y_i|g_i = 2) &\sim N(\gamma, \sigma^2 I) \\
(y_i|g_i = 3) &\sim N(\frac{1}{2}(\mu + \gamma), \sigma^2 I) \\
(y_i|g_i = 4) &\sim N(\tau\mu + (1-\tau)\gamma, \sigma^2 I)
\end{aligned}
$$

Likelihood

$$
\begin{aligned}
L(Y|\theta) &= \prod_{i=1}^{n} p(y_i|\theta), \text{ for } Y = (y_i, \dots, y_n) \\
&= \prod_{i\in g_1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-\frac{1}{2\sigma^2}(y_i-\mu)^T(y_i-\mu)] \times \prod_{i\in g_2} \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-\frac{1}{2\sigma^2}(y_i-\gamma)^T(y_i-\gamma)] \\
&\quad \times \prod_{i\in g_3} \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-\frac{1}{2\sigma^2}(y_i-\frac{1}{2}(\mu+\gamma))^T(y_i-\frac{1}{2}(\mu+\gamma))] \\
&\quad \times \prod_{i\in g_4} \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-\frac{1}{2\sigma^2}(y_i-(\tau\mu+(1-\tau)\gamma))^T(y_i-(\tau\mu+(1-\tau)\gamma))] \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp[-\frac{1}{2\sigma^2}(\sum_{i\in g_1}(y_i-\mu)^T(y_i-\mu) + \sum_{i\in g_2}(y_i-\gamma)^T(y_i-\gamma) \\
&\quad + \sum_{i\in g_3}(y_i-\frac{1}{2}(\mu+\gamma))^T(y_i-\frac{1}{2}(\mu+\gamma)) + \sum_{i\in g_4}(y_i-(\tau\mu+(1-\tau)\gamma))^T(y_i-(\tau\mu+(1-\tau)\gamma)))]
\end{aligned}
$$

Posterior

$$p(\theta|Y) \propto p(\theta)L(Y|\theta)$$

2 Metropolis Hasting

3 Gibbs sampling

Posterior conditioned on each parameter

$$p(\theta[i]|Y,\theta[-i]) = \frac{p(\theta[i],\theta[-i]|Y)}{p(\theta[-i]|Y)} = f(\theta[i]) \propto p(\theta|Y) \text{ with fixed } \theta[-i], Y$$

Posterior conditional probability of τ

$$\begin{aligned}
p(\tau|Y,\theta[-\tau]) &\propto p(\theta)L(Y|\theta) \\
&\propto p(\tau)\prod_{i\in g_4}\frac{1}{\sqrt{2\pi\sigma^2}}\exp[-\frac{1}{2\sigma^2}(y_i-(\tau\mu+(1-\tau)\gamma))^T(y_i-(\tau\mu+(1-\tau)\gamma))] \\
&\propto \exp[-\frac{1}{2\sigma^2}\sum_{i\in g_4}(y_i-(\tau\mu+(1-\tau)\gamma))^T(y_i-(\tau\mu+(1-\tau)\gamma))], \text{ for } \tau\in[0,1] \\
&\propto \exp[-\frac{1}{2\sigma^2}\sum_{i\in g_4}(y_i^T y_i-2y_i^T(\tau\mu+(1-\tau)\gamma)+(\tau\mu+(1-\tau)\gamma)^T(\tau\mu+(1-\tau)\gamma))], \text{ for } \tau\in[0,1] \\
&\propto \exp[-\frac{1}{2\sigma^2}\sum_{i\in g_4}\tau^2(\mu^T\mu-2\mu^T\gamma+\gamma^T\gamma)-2\tau(y_i^T\mu-y_i^T\gamma-\mu^T\gamma+\gamma^T\gamma)], \text{ for } \tau\in[0,1] \\
&\propto \exp[-\frac{1}{2\sigma^2}(n_4\tau^2(\mu-\gamma)^T(\mu-\gamma)-2\tau(\mu-\gamma)^T[\sum_{i\in g_4}(y_i-\gamma), \text{ for } \tau\in[0,1] \\
&\propto \exp\left[-\frac{1}{2}*\frac{n_4(\mu-\gamma)^T(\mu-\gamma)}{\sigma^2}\left(\tau-\frac{(\mu-\gamma)^T(\sum_{i\in g_4}(y_i-\gamma))}{(\mu-\gamma)^T(\mu-\gamma)}\right)^2\right], \text{ for } \tau\in[0,1] \\
&\sim Norm\left(\mu=\frac{(\mu-\gamma)^T(\sum_{i\in g_4}(y_i-\gamma))}{(\mu-\gamma)^T(\mu-\gamma)}, \sigma^2=\frac{\sigma^2}{n_4(\mu-\gamma)^T(\mu-\gamma)}\right), \text{ truncated to } [0,1]
\end{aligned}$$

Posterior conditional probability of σ^2

$$\begin{aligned}
p(\sigma^2|Y,\theta[-\sigma^2]) \propto & p(\theta)L(Y|\theta)\propto p(\sigma^2)\prod_{i=1}^{n}p(y_i|\theta) \\
\propto & \frac{1}{\sigma^2}*\left(\frac{1}{\sigma}\right)^n\exp\left[-\frac{1}{2\sigma^2}M\right], \text{ where } M= \\
& \sum_{i\in g_1}(y_i-\mu)^T(y_i-\mu)+\sum_{i\in g_2}(y_i-\gamma)^T(y_i-\gamma)+\sum_{i\in g_3}(y_i-\frac{1}{2}(\mu+\gamma))^T(y_i-\frac{1}{2}(\mu+\gamma)) \\
& +\sum_{i\in g_4}(y_i-(\tau\mu+(1-\tau)\gamma))^T(y_i-(\tau\mu+(1-\tau)\gamma)) \\
\propto & (\sigma^2)^{-\frac{n}{2}-1}\exp\left[-\frac{M}{2}\frac{1}{\sigma^2}\right] \\
\sim & InvGamma\left(\alpha=\frac{n}{2}, \beta=\frac{M}{2}\right)
\end{aligned}$$

Posterior conditional probability of μ

$$
\begin{aligned}
p(\mu|Y,\theta[-\mu]) \propto & p(\theta)L(Y|\theta) \\
\propto & \prod_{i\in g_1}\frac{1}{\sqrt{2\pi\sigma^2}}\exp[-\frac{1}{2\sigma^2}(y_i-\mu)^T(y_i-\mu)] \times \prod_{i\in g_3}\frac{1}{\sqrt{2\pi\sigma^2}}\exp[-\frac{1}{2\sigma^2}(y_i-\frac{1}{2}(\mu+\gamma))^T(y_i-\frac{1}{2}(\mu+\gamma))] \\
& \times \prod_{i\in g_4}\frac{1}{\sqrt{2\pi\sigma^2}}\exp[-\frac{1}{2\sigma^2}(y_i-(\tau\mu+(1-\tau)\gamma))^T(y_i-(\tau\mu+(1-\tau)\gamma))] \\
\propto & \exp[-\frac{1}{2\sigma^2}(n_1\mu^T\mu-2\mu^T\left(\sum_{i\in g_1}y_i\right)-\mu^T\left(\sum_{i\in g_3}y_i\right)+\frac{n_3}{4}\mu^T\mu+\frac{n_3}{2}\mu^T\gamma \\
& -2\tau\mu^T\left(\sum_{i\in g_4}y_i\right)+\tau^2n_4\mu^T\mu+2\tau(1-\tau)n_4\mu^T\gamma)] \\
\propto & \exp\left[-\frac{1}{2\sigma^2}(n_1+\frac{n_3}{4}+n_4\tau^2)\left(\mu^T\mu-2\mu^T\frac{\sum_{i\in g_1}y_i+\frac{1}{2}\sum_{i\in g_3}y_i+\tau\sum_{i\in g_4}y_i+(\frac{n_3}{4}+n_4\tau(1-\tau))\gamma}{n_1+\frac{n_3}{4}+n_4\tau^2}\right)\right] \\
\propto & \exp\left[-\frac{1}{2\phi^2}(\mu-\psi)^T(\mu-\psi)\right], \text{ where} \\
& \psi=\frac{\sum_{i\in g_1}y_i+\frac{1}{2}\sum_{i\in g_3}y_i+\tau\sum_{i\in g_4}y_i+(\frac{n_3}{4}+n_4\tau(1-\tau))\gamma}{n_1+\frac{n_3}{4}+n_4\tau^2}, \ \phi^2=\frac{\sigma^2}{n_1+\frac{n_3}{4}+n_4\tau^2} \\
\sim & N(\mu=\psi,\Sigma=\phi^2I)
\end{aligned}
$$

Posterior conditional probability of γ

By symmetry with posterior conditional probability of μ,

$$
\begin{aligned}
p(\gamma|Y,\theta[-\gamma]) \sim & N(\mu=\psi',\Sigma=\phi'^2I), \text{ where} \\
& \psi'=\frac{\sum_{i\in g_2}y_i+\frac{1}{2}\sum_{i\in g_3}y_i+(1-\tau)\sum_{i\in g_4}y_i+(\frac{n_3}{4}+n_4\tau(1-\tau))\gamma}{n_2+\frac{n_3}{4}+n_4(1-\tau)^2}, \ \phi'^2=\frac{\sigma^2}{n_2+\frac{n_3}{4}+n_4(1-\tau)^2}
\end{aligned}
$$

4 Hamiltonian Monte Carlo

5 Importance sampling