



UiO • Matematisk institutt

Det matematisk-naturvitenskapelige fakultet

STK-4051/9051 Computational Statistics Spring 2020
Exercise 15

Instructor: Odd Kolbjørnsen, oddkol@math.uio.no



7.7. Consider a hierarchical nested model

$$Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ijk}, \quad (7.29)$$

where $i = 1, \dots, I$, $j = 1, \dots, J_i$, and $k = 1, \dots, K$. After averaging over k for each i and j , we can rewrite the model (7.29) as

$$Y_{ij} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ij}, \quad i = 1, \dots, I, \quad j = 1, \dots, J_i, \quad (7.30)$$

where $Y_{ij} = \sum_{k=1}^K Y_{ijk}/K$. Assume that $\alpha_i \sim N(0, \sigma_\alpha^2)$, $\beta_{j(i)} \sim N(0, \sigma_\beta^2)$, and $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$, where each set of parameters is independent a priori. Assume that σ_α^2 , σ_β^2 , and σ_ϵ^2 are known. To carry out Bayesian inference for this model, assume an improper flat prior for μ , so $f(\mu) \propto 1$. We consider two forms of the Gibbs sampler for this problem [546]:

$$Y_{ij} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ij}, \quad i = 1, \dots, I, \quad j = 1, \dots, J_i, \quad (7.30)$$

where $Y_{ij} = \sum_{k=1}^K Y_{ijk}/K$. Assume that $\alpha_i \sim N(0, \sigma_\alpha^2)$, $\beta_{j(i)} \sim N(0, \sigma_\beta^2)$, and $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$, where each set of parameters is independent a priori. Assume that σ_α^2 , σ_β^2 , and σ_ϵ^2 are known. To carry out Bayesian inference for this model, assume an improper flat prior for μ , so $f(\mu) \propto 1$. We consider two forms of the Gibbs sampler for this problem [546]:

- a. Let $n = \sum_i J_i$, $y_{..} = \sum_{ij} y_{ij}/n$, and $y_{i.} = \sum_j y_{ij}/J_i$ hereafter. Show that at iteration t , the conditional distributions necessary to carry out Gibbs sampling for this

$$\mu^{(t+1)} | (\boldsymbol{\alpha}^{(t)}, \boldsymbol{\beta}^{(t)}, \mathbf{y}) \sim N\left(y_{..} - \frac{1}{n} \sum_i J_i \alpha_i^{(t)} - \frac{1}{n} \sum_{j(i)} \beta_{j(i)}^{(t)}, \frac{\sigma_\epsilon^2}{n}\right),$$

$$\alpha_i^{(t+1)} | (\mu^{(t+1)}, \boldsymbol{\beta}^{(t)}, \mathbf{y}) \sim N\left(\frac{J_i V_1}{\sigma_\epsilon^2} \left(y_{i.} - \mu^{(t+1)} - \frac{1}{J_i} \sum_j \beta_{j(i)}^{(t)}\right), V_1\right),$$

$$\beta_{j(i)}^{(t+1)} | (\mu^{(t+1)}, \boldsymbol{\alpha}^{(t+1)}, \mathbf{y}) \sim N\left(\frac{V_2}{\sigma_\epsilon^2} (y_{ij} - \mu^{(t+1)} - \alpha_i^{(t+1)}), V_2\right),$$

where

$$V_1 = \left(\frac{J_i}{\sigma_\epsilon^2} + \frac{1}{\sigma_\alpha^2}\right)^{-1} \quad \text{and} \quad V_2 = \left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_\beta^2}\right)^{-1}.$$

$$p(\mu | \alpha, \beta, \mathbf{y})$$

$$\propto p(\mu, \alpha, \beta, \mathbf{y})$$

$$= p(\mu)p(\alpha)p(\beta)p(\mathbf{y} | \mu, \alpha, \beta)$$

$$\propto p(\mu)p(\mathbf{y} | \mu, \alpha, \beta)$$

$$\propto \prod_{i=1}^I \prod_{j=1}^{J_i} \exp\left[-\frac{1}{2\sigma_\varepsilon^2} (y_{ij} - \mu - \alpha_i - \beta_{j(i)})^2\right]$$

$$\propto \prod_{i=1}^I \prod_{j=1}^{J_i} \exp\left[-\frac{1}{2\sigma_\varepsilon^2} [(y_{ij} - \alpha_i - \beta_{j(i)})^2 - 2(y_{ij} - \alpha_i - \beta_{j(i)})\mu + \mu^2]\right]$$

$$\propto \exp\left[-\frac{1}{2\sigma_\varepsilon^2} \left[n\mu^2 - 2 \sum_{i=1}^I \sum_{j=1}^{J_i} (y_{ij} - \alpha_i - \beta_{j(i)})\mu\right]\right]$$

$$\propto \exp\left[-\frac{n}{2\sigma_\varepsilon^2} \left[\mu - \frac{1}{n} \sum_{i=1}^I \sum_{j=1}^{J_i} (y_{ij} - \alpha_i - \beta_{j(i)})\right]^2\right]$$

$$\propto \exp\left[-\frac{n}{2\sigma_\varepsilon^2} \left[\mu - \left(y_{..} - \frac{1}{n} \sum_{i=1}^I J_i \alpha_i - \frac{1}{n} \sum_{i=1}^I \sum_{j=1}^{J_i} \beta_{j(i)}\right)\right]^2\right]$$

$$\phi(z; \text{Mean, Var}) \propto \exp\left\{-\frac{1}{2} \cdot \frac{1}{\text{Var}} (z - \text{Mean})^2\right\}$$

$$n \left(\mu - \frac{1}{n} x\right)^2 = (n\mu^2 - 2\mu x + \dots)$$

$$\begin{aligned}
 & p(\alpha_i | \mu, \boldsymbol{\alpha}_{-i}, \boldsymbol{\beta}, \mathbf{y}) \\
 & \propto p(\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{y}) \\
 & = p(\mu) p(\boldsymbol{\alpha}) p(\boldsymbol{\beta}) p(\mathbf{y} | \mu, \boldsymbol{\alpha}, \boldsymbol{\beta}) \\
 & \propto p(\alpha_i) p(\mathbf{y} | \mu, \boldsymbol{\alpha}, \boldsymbol{\beta})
 \end{aligned}$$

$$c_1 \left(\alpha - \frac{1}{c_1} x \right)^2 = (c_1 \alpha^2 - 2\alpha x + \dots)$$

$$\propto \exp\left(-\frac{1}{2\sigma_\alpha^2} \alpha_i^2\right) \prod_{j=1}^{J_i} \exp\left[-\frac{1}{2\sigma_\varepsilon^2} (y_{ij} - \mu - \alpha_i - \beta_{j(i)})^2\right]$$

$$\propto \exp\left(-\frac{1}{2\sigma_\alpha^2} \alpha_i^2\right) \prod_{j=1}^{J_i} \exp\left[-\frac{1}{2\sigma_\varepsilon^2} \left[(y_{ij} - \mu - \beta_{j(i)})^2 - 2(y_{ij} - \mu - \beta_{j(i)})\alpha_i + \alpha_i^2 \right]\right]$$

$$\propto \exp\left[-\frac{1}{2} \left[\frac{1}{\sigma_\alpha^2} \alpha_i^2 + \frac{1}{\sigma_\varepsilon^2} J_i \alpha_i^2 - 2 \frac{1}{\sigma_\varepsilon^2} \sum_{j=1}^{J_i} (y_{ij} - \mu - \beta_{j(i)}) \alpha_i \right]\right]$$

$$\propto \exp\left[-\frac{1}{2} \left[\frac{1}{\sigma_\alpha^2} + \frac{J_i}{\sigma_\varepsilon^2} \right] \left[\alpha_i - \frac{1}{\sigma_\varepsilon^2} \frac{\sum_{j=1}^{J_i} (y_{ij} - \mu - \beta_{j(i)})}{\frac{1}{\sigma_\alpha^2} + \frac{J_i}{\sigma_\varepsilon^2}} \right]^2 \right]$$

$$\propto \exp\left[-\frac{1}{2} \left[\frac{1}{\sigma_\alpha^2} + \frac{J_i}{\sigma_\varepsilon^2} \right] \left[\alpha_i - J_i \frac{1}{\sigma_\varepsilon^2} \frac{y_{i\cdot} - \mu - \frac{1}{J_i} \sum_{j=1}^{J_i} \beta_{j(i)}}{\frac{1}{\sigma_\alpha^2} + \frac{J_i}{\sigma_\varepsilon^2}} \right]^2 \right]$$

$$\begin{aligned}
 & p(\beta_{j(i)} | \mu, \alpha_{-i}, \beta, \mathbf{y}) \\
 & \propto p(\mu, \alpha, \beta, \mathbf{y}) \\
 & = p(\mu) p(\alpha) p(\beta) p(\mathbf{y} | \mu, \alpha, \beta) \\
 & \propto p(\beta_{j(i)}) p(\mathbf{y} | \mu, \alpha, \beta) \\
 & \propto \exp\left(-\frac{1}{2\sigma_\beta^2} \beta_{j(i)}^2\right) \exp\left[-\frac{1}{2\sigma_\varepsilon^2} (y_{ij} - \mu - \alpha_i - \beta_{j(i)})^2\right] \\
 & \propto \exp\left(-\frac{1}{2\sigma_\beta^2} \beta_{j(i)}^2\right) \exp\left[-\frac{1}{2\sigma_\varepsilon^2} [(y_{ij} - \mu - \alpha_i)^2 - 2(y_{ij} - \mu - \alpha_i)\beta_{j(i)} + \beta_{j(i)}^2]\right] \\
 & \propto \exp\left[-\frac{1}{2} \left[\frac{1}{\sigma_\beta^2} \beta_{j(i)}^2 + \frac{1}{\sigma_\varepsilon^2} \beta_{j(i)}^2 - 2\frac{1}{\sigma_\varepsilon^2} (y_{ij} - \mu - \alpha_i)\beta_{j(i)}\right]\right] \\
 & \propto \exp\left[-\frac{1}{2} \left[\frac{1}{\sigma_\beta^2} + \frac{1}{\sigma_\varepsilon^2}\right] \left[\beta_{j(i)} - \frac{1}{\sigma_\varepsilon^2} \frac{(y_{ij} - \mu - \alpha_i)}{\frac{1}{\sigma_\beta^2} + \frac{1}{\sigma_\varepsilon^2}}\right]^2\right] \\
 & \propto \exp\left[-\frac{1}{2} \left[\frac{1}{\sigma_\beta^2} + \frac{1}{\sigma_\varepsilon^2}\right] \left[\beta_{j(i)} - \frac{1}{\sigma_\varepsilon^2} \frac{y_{ij} - \mu - \alpha_i}{\frac{1}{\sigma_\beta^2} + \frac{1}{\sigma_\varepsilon^2}}\right]^2\right]
 \end{aligned}$$

$c_1 \left(\beta - \frac{1}{c_1} x \right)^2 = (c_1 \beta^2 - 2\beta x + \dots)$

b. The convergence rate for a Gibbs sampler can sometimes be improved via reparameterization. For this model, the model can be reparameterized via hierarchical centering (Section 7.3.1.4). For example, let Y_{ij} follow (7.30), but now let $\eta_{ij} = \mu + \alpha_i + \beta_{j(i)}$ and $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$. Then let $\gamma_i = \mu + \alpha_i$ with $\eta_{ij} | \gamma_i \sim N(\gamma_i, \sigma_\beta^2)$ and $\gamma_i | \mu \sim N(\mu, \sigma_\alpha^2)$. As above, assume σ_α^2 , σ_β^2 , and σ_ϵ^2 are known, and assume a flat prior for μ . Show that the conditional distributions necessary to carry out Gibbs sampling for this model are given by

$$\mu^{(t+1)} | (\boldsymbol{y}^{(t)}, \boldsymbol{\eta}^{(t)}, \mathbf{y}) \sim N\left(\frac{1}{I} \sum_i \gamma_i^{(t)}, \frac{1}{I} \sigma_\alpha^2\right),$$

$$\gamma_i^{(t+1)} | (\mu^{(t+1)}, \boldsymbol{\eta}^{(t)}, \mathbf{y}) \sim N\left(V_3 \left(\frac{1}{\sigma_\beta^2} \sum_j \eta_{ij}^{(t)} + \frac{\mu^{(t+1)}}{\sigma_\alpha^2}\right), V_3\right),$$

$$\eta_{ij}^{(t+1)} | (\mu^{(t+1)}, \boldsymbol{y}^{(t+1)}, \mathbf{y}) \sim N\left(V_2 \left(\frac{y_{ij}}{\sigma_\epsilon^2} + \frac{\gamma_i^{(t+1)}}{\sigma_\beta^2}\right), V_2\right),$$

where

$$V_3 = \left(\frac{J_i}{\sigma_\beta^2} + \frac{1}{\sigma_\alpha^2}\right)^{-1}.$$

$$Y_{ij} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ij}, \quad i = 1, \dots, I, \quad j = 1, \dots, J_i,$$

$$Y_{ij} = \eta_{ij} + \epsilon_{ij}$$

$$\eta_{ij} \sim N(\gamma_i, \sigma_\beta^2)$$

$$\gamma_i \sim N(\mu, \sigma_\alpha^2)$$

$$f(\mu) \propto 1$$

$$p(\mu | \boldsymbol{\gamma}, \boldsymbol{\eta}, \mathbf{y}) \propto p(\mu, \boldsymbol{\gamma}, \boldsymbol{\eta}, \mathbf{y}) = p(\mu)p(\boldsymbol{\gamma} | \mu)p(\boldsymbol{\eta} | \boldsymbol{\gamma})p(\mathbf{y} | \boldsymbol{\eta})$$

$$\propto p(\mu)p(\boldsymbol{\gamma} | \mu)$$

$$\propto \prod_{i=1}^I \exp\left[-\frac{1}{2\sigma_\alpha^2}(\gamma_i - \mu)^2\right] = \exp\left[-\frac{1}{2\sigma_\alpha^2}\left[I\mu^2 - 2\sum_{i=1}^I \gamma_i \mu\right]\right]$$

$$\propto \exp\left[-\frac{I}{2\sigma_\alpha^2}\left[\mu - \frac{1}{I}\sum_{i=1}^I J_i \gamma_i\right]^2\right]$$

$$\begin{aligned}
 Y_{ij} &= \eta_{ij} + \varepsilon_{ij} \\
 \eta_{ij} &\sim N(\gamma_i, \sigma_\beta^2) \\
 \gamma_i &\sim N(\mu, \sigma_\alpha^2) \\
 f(\mu) &\propto 1
 \end{aligned}$$

$$\begin{aligned}
 p(\gamma_i | \mu, \gamma_{-i}, \boldsymbol{\beta}, \mathbf{y}) &\propto p(\mu, \boldsymbol{\gamma}, \boldsymbol{\eta}, \mathbf{y}) \\
 &= p(\mu) p(\boldsymbol{\gamma} | \mu) p(\boldsymbol{\eta} | \boldsymbol{\gamma}) p(\mathbf{y} | \boldsymbol{\eta}) \propto p(\gamma_i | \mu) p(\boldsymbol{\eta}_i | \gamma_i) \\
 &\propto \exp\left(-\frac{1}{2\sigma_\alpha^2}(\gamma_i - \mu)^2\right) \prod_{j=1}^{J_i} \exp\left[-\frac{1}{2\sigma_\beta^2}(\eta_{ij} - \gamma_i)^2\right] \\
 &\propto \exp\left[-\frac{1}{2}\left[\left(\frac{1}{\sigma_\alpha^2} + \frac{J_i}{\sigma_\beta^2}\right)\gamma_i^2 - 2\left(\frac{1}{\sigma_\alpha^2}\mu + \frac{1}{\sigma_\beta^2}\sum_{j=1}^{J_i}\eta_{ij}\right)\gamma_i\right]\right] \\
 &\propto \exp\left[-\frac{1}{2}\left[\frac{1}{\sigma_\alpha^2} + \frac{J_i}{\sigma_\beta^2}\right]\left[\gamma_i - \frac{\frac{1}{\sigma_\alpha^2}\mu + \frac{1}{\sigma_\beta^2}\sum_{j=1}^{J_i}\eta_{ij}}{\frac{1}{\sigma_\alpha^2} + \frac{1}{\sigma_\beta^2}}\right]^2\right]
 \end{aligned}$$

$$\begin{aligned}
 Y_{ij} &= \eta_{ij} + \varepsilon_{ij} \\
 \eta_{ij} &\sim N(\gamma_i, \sigma_\beta^2) \\
 \gamma_i &\sim N(\mu, \sigma_\alpha^2) \\
 f(\mu) &\propto 1
 \end{aligned}$$

$$\begin{aligned}
 p(\eta_{ij} | \mu, \gamma, \boldsymbol{\eta}_{-ij}, \mathbf{y}) &\propto p(\mu, \gamma, \boldsymbol{\eta}, \mathbf{y}) \\
 &= p(\mu) p(\gamma | \mu) p(\boldsymbol{\eta} | \gamma) p(\mathbf{y} | \boldsymbol{\eta}) \propto p(\eta_{ij}) p(y_{ij} | \eta_{ij}) \\
 &\propto \exp\left(-\frac{1}{2\sigma_\beta^2} (\eta_{ij} - \gamma_i)^2\right) \exp\left[-\frac{1}{2\sigma_\varepsilon^2} (y_{ij} - \eta_{ij})^2\right] \\
 &\propto \exp\left[-\frac{1}{2} \left[\left(\frac{1}{\sigma_\beta^2} + \frac{1}{\sigma_\varepsilon^2}\right) \eta_{ij}^2 - 2\left(\frac{1}{\sigma_\beta^2} \gamma_i + \frac{1}{\sigma_\varepsilon^2} y_{ij}\right) \eta_{ij}\right]\right] \\
 &\propto \exp\left[-\frac{1}{2} \left[\frac{1}{\sigma_\beta^2} + \frac{1}{\sigma_\varepsilon^2}\right] \left[\eta_{ij} - \frac{\frac{1}{\sigma_\beta^2} \gamma_i + \frac{1}{\sigma_\varepsilon^2} y_{ij}}{\frac{1}{\sigma_\beta^2} + \frac{1}{\sigma_\varepsilon^2}}\right]^2\right] \\
 &\propto \exp\left[-\frac{1}{2} \left[\frac{1}{\sigma_\beta^2} + \frac{1}{\sigma_\varepsilon^2}\right] \left[\beta_{j(i)} - \frac{1}{\sigma_\varepsilon^2} \frac{y_{ij} - \mu - \alpha_i}{\frac{1}{\sigma_\alpha^2} + \frac{1}{\sigma_\varepsilon^2}}\right]^2\right]
 \end{aligned}$$

7.8 a Gibbs for $p(\mu, \alpha, \beta, y)$

- Initialization

```

sig2.alpha = 86
sig2.beta = 58
sig2.eps = 1
n = nrow(d)
y.bar = mean(d$Moisture)
y.dot <-with(d, tapply(Moisture, Batch, mean))
J = 2
I = 15
D=1000
L = 1000
N = D+L
V1 = 1/(J/sig2.eps + 1/sig2.alpha)
V2 = 1/(1/sig2.eps+1/sig2.beta)
V3 = 1/(J/sig2.beta+1/sig2.alpha)
y = matrix(d$Moisture,ncol=J,byrow=T)

```

} Given in text

$$V_1 = \left(\frac{J_i}{\sigma_\epsilon^2} + \frac{1}{\sigma_\alpha^2} \right)^{-1} \quad \text{and} \quad V_2 = \left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_\beta^2} \right)^{-1}$$

$$V_3 = \left(\frac{J_i}{\sigma_\beta^2} + \frac{1}{\sigma_\alpha^2} \right)^{-1}$$

7.8 a Gibbs for $p(\mu, \alpha, \beta, \mathbf{y})$

- Gibbs sampler

```
#Gibbs sampling
for(i in 1:N)
{
  #Sample mu
  mu = rnorm(1, y.bar - sum(J*alpha)/n - sum(beta)/n, sqrt(sig2.eps)/n)
  #Sample alpha
  alpha = rnorm(I, J*V1*(y.dot - mu - rowSums(beta))/J / sig2.eps, sqrt(V1))
  #Sample beta
  beta[,1] = rnorm(I, V2*(y[,1] - mu - alpha), sqrt(V2))
  beta[,2] = rnorm(I, V2*(y[,2] - mu - alpha), sqrt(V2))
  #Store simulations
  muM1[k,i] = mu
  alphaM[k,i,] = alpha
  betaM[k,i,,] = beta
}
```

$$\mu^{(t+1)} | (\alpha^{(t)}, \beta^{(t)}, \mathbf{y}) \sim N\left(y_{..} - \frac{1}{n} \sum_i J_i \alpha_i^{(t)} - \frac{1}{n} \sum_{j(i)} \beta_{j(i)}^{(t)}, \frac{\sigma_\epsilon^2}{n}\right),$$

$$\alpha_i^{(t+1)} | (\mu^{(t+1)}, \beta^{(t)}, \mathbf{y}) \sim N\left(\frac{J_i V_1}{\sigma_\epsilon^2} \left(y_{i.} - \mu^{(t+1)} - \frac{1}{J_i} \sum_j \beta_{j(i)}^{(t)}\right), V_1\right)$$

$$\beta_{j(i)}^{(t+1)} | (\mu^{(t+1)}, \alpha^{(t+1)}, \mathbf{y}) \sim N\left(\frac{V_2}{\sigma_\epsilon^2} (y_{ij} - \mu^{(t+1)} - \alpha_i^{(t+1)}), V_2\right),$$

```
sig2.eps = 1
```

7.8 b) Gibbs for $p(\mu, \gamma, \eta, \mathbf{y})$

```

#Initialization
mu = y.bar
gamma = y.dot-y.bar
eta = matrix(0,nrow=I,ncol=J)
y = matrix(d$Moisture,ncol=J,byrow=T)
#Gibbs sampling
for(i in 1:N)
{
  #Sample mu
  mu = rnorm(1,mean(gamma),sqrt(sig2.alpha)/I)
  #Sample gamma
  gamma = rnorm(I,V3*(rowSums(eta)/sig2.beta + mu/sig2.alpha),sqrt(V3))
  #Sample eta
  eta[,1] = rnorm(I,V2*(y[,1]/sig2.eps+gamma/sig2.beta),sqrt(V2))
  eta[,2] = rnorm(I,V2*(y[,2]/sig2.eps+gamma/sig2.beta),sqrt(V2))

  muM2[k,i] = mu
  gammaM[k,i,] = gamma
  etaM[k,i,,] = eta
}

```

$$\mu^{(t+1)} | (\boldsymbol{\gamma}^{(t)}, \boldsymbol{\eta}^{(t)}, \mathbf{y}) \sim N\left(\frac{1}{I} \sum \gamma_i^{(t)}, \frac{1}{I} \sigma_\alpha^2\right),$$

$$\gamma_i^{(t+1)} | (\mu^{(t+1)}, \boldsymbol{\eta}^{(t)}, \mathbf{y}) \sim N\left(V_3 \left(\frac{1}{\sigma_\beta^2} \sum_j \eta_{ij}^{(t)} + \frac{\mu^{(t+1)}}{\sigma_\alpha^2} \right), V_3\right),$$

$$\eta_{ij}^{(t+1)} | (\mu^{(t+1)}, \boldsymbol{\gamma}^{(t+1)}, \mathbf{y}) \sim N\left(V_2 \left(\frac{y_{ij}}{\sigma_\epsilon^2} + \frac{\gamma_i^{(t+1)}}{\sigma_\beta^2} \right), V_2\right),$$

«7d» STAN model