Homework 2 - Part 1

Xiyang Dai

April 1, 2014

QUESTION 1

$$
\begin{aligned}
P(\theta_g|y_{g1},\dots,y_{gm}) &\propto P(y_{g1},\dots,y_{gm}|\theta_g)P(\theta_g)\\
&\propto \left(\prod_i P(y_{gi}|\theta_g)\right)P(\theta_g)\\
&\propto \left(\prod_i \frac{1}{\sqrt{2\pi\sigma_g^2}}\exp\left\{-\frac{1}{2}\frac{(y_{gi}-\theta_g)^2}{\sigma_g^2}\right\}\right)\frac{1}{\sqrt{2\pi\tau^2}}\exp\left\{-\frac{1}{2}\frac{\theta_g^2}{\tau^2}\right\}\\
&\propto \exp\left\{-\frac{1}{2}\left(\sum_i \frac{(y_{gi}-\theta_g)^2}{\sigma_g^2}+\frac{\theta_g^2}{\tau^2}\right)\right\}\\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{\tau^2\sum_i(y_{gi}-\theta_g)^2+\theta_g^2\sigma_g^2}{\sigma_g^2\tau^2}\right)\right\}\\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{\tau^2\sum_i y_{gi}^2-\tau^2 2\sum_i y_{gi}\theta_g+\tau^2\sum_i\theta_g^2+\theta_g^2\sigma_g^2}{\sigma_g^2\tau^2}\right)\right\}
\end{aligned}
$$

Any term that dosen't have a θ can be seen as a constant, so I drop them:

$$
\begin{aligned}
&\propto \exp\left\{-\frac{1}{2}\left(\frac{-2m\tau^2\bar{y}\theta_g+m\tau^2\theta_g^2+\theta_g^2\sigma_g^2}{\sigma_g^2\tau^2}\right)\right\}\\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{-2m\tau^2\bar{y}\theta_g+\theta_g^2(m\tau^2+\sigma_g^2)}{\sigma_g^2\tau^2}\right)\right\}\\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{\theta_g^2-2\frac{m\tau^2\bar{y}\theta_g}{m\tau^2+\sigma_g^2}}{\frac{\sigma_g^2\tau^2}{m\tau^2+\sigma_g^2}}\right)\right\}\\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{(\theta_g-\frac{m\tau^2\bar{y}}{m\tau^2+\sigma_g^2})^2}{\frac{\sigma_g^2\tau^2}{m\tau^2+\sigma_g^2}}\right)\right\}
\end{aligned}
$$

Let $\lambda_g=\frac{m\tau^2}{m\tau^2+\sigma_g^2}$, then we have:

$$\propto \exp\left\{-\frac{1}{2}\left(\frac{(\theta_g-\lambda_g\bar{y})^2}{\lambda_g\frac{\sigma_g^2}{m}}\right)\right\}$$

Hence, I show that

$$P(\theta_g|y_{g1},\dots,y_{gm})\sim N(\lambda_g\bar{y}_g,\lambda_g\frac{\sigma_g^2}{m})$$

As the $m \to \infty$, the posterior distribution will $\lambda_g = \frac{m\tau^2}{\sigma_g^2 + m\tau^2} \to 1$, $\lambda_g \bar{y}_g \to \bar{y}_g$ and $\lambda_g \frac{\sigma_g^2}{m} = \frac{\tau^2\sigma^2}{\sigma_g^2 + m\tau^2} \to 0$. So, we can conclude that as the m increases, the posterior distribution tends to be $N(\bar{y}_g, 0)$.

QUETSTION 2A

$$
\begin{aligned}
P(\theta_g|y_g) &\propto P(y_g|\theta_g)P(\theta_g) \\
&\propto \frac{1}{\sqrt{2\pi\theta_g^2}} \exp\left\{-\frac{1}{2}\frac{(y_g - \theta_g)^2}{\sigma_g^2}\right\} \frac{1}{\sqrt{2\pi\tau^2}} \exp\left\{-\frac{1}{2}\frac{(\theta_g - \mu)^2}{\tau^2}\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{(y_{gi} - \theta_g)^2}{\sigma_g^2} + \frac{(\theta_g - \mu)^2}{\tau^2}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{\tau^2(y_g - \theta_g)^2 + \sigma_g^2(\theta_g - \mu)^2}{\sigma_g^2\tau^2}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{\tau^2 y_g^2 - \tau^2 2 y_g\theta_g + \tau^2\theta_g^2 + \sigma^2\theta_g^2 - \sigma^2 2\theta_g\mu + \sigma^2\mu^2}{\sigma_g^2\tau^2}\right)\right\}
\end{aligned}
$$

Any term that dosen't have a θ can be seen as a constant, so I drop them:

$$
\begin{aligned}
&\propto \exp\left\{-\frac{1}{2}\left(\frac{-\tau^2 2 y_g\theta_g + \tau^2\theta_g^2 + \sigma^2\theta_g^2 - \sigma^2 2\theta_g\mu}{\sigma_g^2\tau^2}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{\theta_g^2(\tau^2 + \sigma^2) - 2\theta_g(\sigma^2\mu + \tau^2 y_g)}{\sigma_g^2\tau^2}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{\theta_g^2 - 2\theta_g\frac{\sigma^2\mu + \tau^2 y_g}{(\tau^2 + \sigma^2)}}{\frac{\sigma_g^2\tau^2}{(\tau^2 + \sigma^2)}}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{(\theta_g - \frac{\sigma^2\mu + \tau^2 y_g}{(\tau^2 + \sigma^2)})^2}{\frac{\sigma_g^2\tau^2}{(\tau^2 + \sigma^2)}}\right)\right\}
\end{aligned}
$$

Let $\lambda_g = \frac{\tau^2}{\tau^2 + \sigma_g^2}$, then we have:

$$
\propto \exp\left\{-\frac{1}{2}\left(\frac{(\theta_g - (\lambda_g y_g + (1 - \lambda_g)\mu))^2}{\lambda_g\sigma_g^2}\right)\right\}
$$

Hence, I show that

$$
P(\theta_g|y_g) \sim N(\lambda_g y_g + (1 - \lambda_g)\mu, \lambda_g\sigma_g^2)
$$

QUETSTION 2B

$$
\begin{aligned}
P(\theta_g|y_{g1},\dots,y_{gm}) &\propto P(y_{g1},\dots,y_{gm}|\theta_g)P(\theta_g) \\
&\propto \left(\prod_i P(y_{gi}|\theta_g)\right)P(\theta_g) \\
&\propto \left(\prod_i \frac{1}{\sqrt{2\pi\theta_g^2}}\exp\left\{-\frac{1}{2}\frac{\sum_i(y_{ig}-\theta_g)^2}{\sigma_g^2}\right\}\right)\frac{1}{\sqrt{2\pi\tau^2}}\exp\left\{-\frac{1}{2}\frac{(\theta_g-\mu)^2}{\tau^2}\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{\sum_i(y_{ig}-\theta_g)^2}{\sigma_g^2}+\frac{(\theta_g-\mu)^2}{\tau^2}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{\tau^2\sum_i(y_{ig}-\theta_g)^2+\sigma_g^2(\theta_g-\mu)^2}{\sigma_g^2\tau^2}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{\tau^2\sum_i y_{ig}^2-\tau^2 2\sum_i y_{ig}\theta_g+\tau^2\sum_i\theta_g^2+\sigma^2\theta_g^2-\sigma^2 2\theta_g\mu+\sigma^2\mu^2}{\sigma_g^2\tau^2}\right)\right\}
\end{aligned}
$$

Any term that dosen't have a θ can be seen as a constant, so I drop them:

$$
\begin{aligned}
&\propto \exp\left\{-\frac{1}{2}\left(\frac{-2m\tau^2\bar{y}_g\theta_g+m\tau^2\theta_g^2+\sigma^2\theta_g^2-\sigma^2 2\theta_g\mu}{\sigma_g^2\tau^2}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{\theta_g^2(m\tau^2+\sigma^2)-2\theta_g(\sigma^2\mu+m\tau^2\bar{y}_g)}{\sigma_g^2\tau^2}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{\theta_g^2-2\theta_g\frac{\sigma^2\mu+m\tau^2\bar{y}_g}{(\tau^2+m\sigma^2)}}{\frac{\sigma_g^2\tau^2}{(\tau^2+m\sigma^2)}}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{(\theta_g-\frac{\sigma^2\mu+m\tau^2\bar{y}_g}{(\tau^2+m\sigma^2)})^2}{\frac{\sigma_g^2\tau^2}{(\tau^2+m\sigma^2)}}\right)\right\}
\end{aligned}
$$

Let $\lambda_g = \frac{m\tau^2}{\tau^2+m\sigma_g^2}$, then we have:

$$
\propto \exp\left\{-\frac{1}{2}\left(\frac{(\theta_g-(\lambda_g y_g+(1-\lambda_g)\mu))^2}{\lambda_g\frac{\sigma_g^2}{m}}\right)\right\}
$$

Hence, I show that

$$
P(\theta_g|y_{g1},\dots,y_{gm}) \sim N(\lambda_g y_g+(1-\lambda_g)\mu, \lambda_g\frac{\sigma_g^2}{m})
$$

QUETSTION 3A

$$
\begin{aligned}
E[\bar{Y}_g] &= E[\frac{1}{m}\sum_i Y_{ig}] \\
&= \frac{1}{m}\sum_i E[Y_{ig}] \\
&= \frac{1}{m} m\mu \\
&= \mu \\
Var[\bar{Y}_g] &= E[Var[\bar{Y}_g|\theta_g]] + Var[E[\bar{Y}_g|\theta_g]] \\
&= E[Var[\frac{1}{m}\sum_i Y_{ig}|\theta_g]] + Var[E[\frac{1}{m}\sum_i Y_{ig}|\theta_g]] \\
&= E[\frac{1}{m^2}\sum_i Var[Y_{ig}|\theta_g]] + Var[\frac{1}{m}\sum_i E[Y_{ig}|\theta_g]]
\end{aligned}
$$

Since I know $Y_g - \theta_g = \epsilon_g \sim N(0, \sigma^2)$ and $\theta_g \sim N(\mu, \tau^2)$, I can induct:

$$
\begin{aligned}
&= E[\frac{1}{m^2} m\sigma^2] + Var[\frac{1}{m} m\theta_g] \\
&= E[\frac{\sigma^2}{m}] + Var[\theta] \\
&= \frac{\sigma^2}{m} + \tau^2
\end{aligned}
$$

Hence, I show that

$$\bar{Y}_g|\mu \sim N(\mu, \frac{\sigma^2}{m} + \tau^2)$$

QUETSTION 3B

In order to estimate μ, we want to maximize the log likelihood of $P(\bar{Y}_1, \ldots, \bar{Y}_i|\mu)$:

$$
\begin{aligned}
\arg\max_\mu \ \log(P(\bar{Y}_1, \ldots, \bar{Y}_i|\mu)) &= \arg\max_\mu \sum_g \log(P(\bar{Y}_g|\mu)) \\
&\propto \arg\max_\mu \sum_g \log\left(\exp\left\{-\frac{1}{2}\frac{(\bar{Y}_g - \mu)^2}{\frac{\sigma_g^2}{m} + \tau^2}\right\}\right) \\
&\propto \arg\max_\mu -\frac{1}{2}\sum_g \frac{(\bar{Y}_g - \mu)^2}{\frac{\sigma_g^2}{m} + \tau^2}
\end{aligned}
$$

Let $L(\mu) = \sum_g \frac{(\bar{Y}_g - \mu)^2}{\frac{\sigma_g^2}{m} + \tau^2}$. In order to solve this maximization problem, we need to set $\frac{\partial L}{\partial \mu} = 0$, solve μ:

$$\begin{aligned}
\frac{\partial}{\partial \mu}\left(\sum_g \frac{(\bar{Y}_g - \mu)^2}{\frac{\sigma_g^2}{m} + \tau^2}\right) &= \sum_g \frac{-2(\bar{Y}_g - \mu)}{\frac{\sigma_g^2}{m} + \tau^2} = 0 \\
\Longrightarrow \hat{\mu} &= \frac{\sum_g \frac{m}{\sigma_g^2 + m\tau^2} \bar{y}_g}{\sum_g \frac{m}{\sigma_g^2 + m\tau^2}} \\
\Longrightarrow \hat{\mu} &= \frac{\sum_g w_g \bar{y}_g}{\sum_g w_g} \quad \text{where, } w_g = \frac{m}{\sigma_g^2 + m\tau^2}.
\end{aligned}$$

Since we have $w_g \propto \frac{1}{\sigma_g^2}$, the weight will bias low variance samples. If the variance is low, we can trust the sample, so the weight will be high. If the variance is high, we cannot trust this sample, so the weight will be penalized and it will be low.

QUETSTION 3C

If we change $Var[Y_g|\theta_g] = \theta_g^2 \sigma_g^2$, $\bar{Y}_g|\mu$ will become:

$$\bar{Y}_g|\mu \sim N(\mu, \frac{\theta_g^2 \sigma_g^2}{m} + \tau^2)$$

Repeat the steps in (3b), we will get

$$\hat{\mu} = \frac{\sum_g w_g \bar{y}_g}{\sum_g w_g} \quad \text{where, } w_g = \frac{m}{\theta_g^2 \sigma_g^2 + m\tau^2}.$$

Because of $\theta_g \sim N(\mu, \tau^2)$, as $\mu \to 0$, $\theta_g \to 0$, weight w_g tends to be large. Hence, we can see this weight is actually biased towards zero.

QUETSTION 4

$$\begin{aligned}
\log \frac{P(\bar{y}_g|\mu \neq 0)}{P(\bar{y}_g|\mu = 0)} &\propto \log \frac{\exp\left\{-\frac{1}{2}\frac{(\bar{y}_g - \hat{\mu})^2}{\frac{\sigma^2}{m} + \tau^2}\right\}}{\exp\left\{-\frac{1}{2}\frac{\bar{y}_g^2}{\frac{\sigma^2}{m} + \tau^2}\right\}} \\
&\propto -\frac{1}{2}\frac{(\bar{y}_g - \hat{\mu})^2}{\frac{\sigma^2}{m} + \tau^2} + \frac{1}{2}\frac{\bar{y}_g^2}{\frac{\sigma^2}{m} + \tau^2} \\
\text{Let } w_g = \frac{m\tau^2}{\sigma^2 + m\tau^2}, & \\
&\propto 2 w_g \hat{\mu} \bar{y}_g - w_g \hat{\mu}^2
\end{aligned}$$

If we want to have more evidence that $E\bar{y}_g \neq 0$ than $E\bar{y}_g = 0$, we need to let

$$
\begin{aligned}
&\log \frac{P(\bar{y}_g|\mu \neq 0)}{P(\bar{y}_g|\mu = 0)} > 0 \\
&\Longrightarrow 2w_g\hat{\mu}\bar{y}_g - w_g\hat{\mu}^2 > 0 \\
&\Longrightarrow \bar{y}_g > \frac{\hat{\mu}}{2}
\end{aligned}
$$

So when $\bar{y}_g > \frac{\hat{\mu}}{2}$, we have more evidence that $E\bar{y}_g \neq 0$ than $E\bar{y}_g = 0$.