

# ST740 HW2

Chenyin Gao

2023-10-09

**1**

**(a)**

First, we know the posterior distribution of  $\theta_i$  given the rest:

$$\theta_i \mid \text{rest} \sim \mathcal{N}\left(\frac{\sum_{j=1}^n Y_{ij}}{m+n}, \frac{\sigma^2}{m+n}\right),$$

where  $\hat{\theta}_i = \sum_{j=1}^n Y_{ij}/(m+n)$ ,  $E(\hat{\theta}_i) = n\theta_i^*/(m+n)$  and  $\text{var}(\hat{\theta}_i) = n\sigma^2/(m+n)^2$ . Thus,  $E[D(\hat{\theta}, \theta^*)]$  can be computed by

$$E[D(\hat{\theta}, \theta^*)] = E\left[\sum_{i=1}^p (\hat{\theta}_i - \theta_i)^2\right] = \frac{m \sum_{i=1}^n \theta_i^2}{(m+n)^2} + \frac{np\sigma^2}{(m+n)^2}.$$

**(b)**

If  $p$  is fixed and  $n \rightarrow \infty$ , we have

$$\frac{m \sum_{i=1}^n \theta_i^2}{(m+n)^2} + \frac{np\sigma^2}{(m+n)^2} \asymp \frac{\|\theta^*\|^2}{n} + \frac{1}{n},$$

which converges to zero as long as  $\|\theta^*\|^2$  is bounded.

**(c)**

If  $n = p^k$ , plug it back to the formula and obtain

$$\frac{m \sum_{i=1}^n \theta_i^2}{(m+n)^2} + \frac{np\sigma^2}{(m+n)^2} = \frac{m\|\theta^*\|^2}{(m+p^k)^2} + \frac{p^{k+1}\sigma^2}{(m+p^k)^2}.$$

Obviously, we need  $k > 1$  for  $p^{2k}$  to dominate  $p^{k+1}$  and make the second term negligible. Besides, we need

$$\left(\frac{\|\theta^*\|^2}{m} \vee \frac{m\|\theta^*\|^2}{p^{2k}}\right) \rightarrow 0,$$

which ensure the first term converge to zero.

## 2

### (a)

Derive the full conditional distribution of  $\theta_k$  given the other parameters

$$\begin{aligned}
\pi(\theta_k | \theta_{-k}, \sigma^2) &\propto (\sigma^2)^{-n/2} \exp \left[ -\frac{1}{2\sigma^2} \sum_{i,j} A_{ij} (\theta_i - \theta_j)^2 \right] \\
&\propto \exp \left[ -\frac{1}{2\sigma^2} \left\{ \sum_i A_{ik} (\theta_i^2 + \theta_k^2 - 2\theta_i \theta_k) + \sum_j A_{kj} (\theta_k^2 + \theta_j^2 - 2\theta_k \theta_j) \right\} \right] \\
&\propto \exp \left[ -\frac{1}{2\sigma^2} \left\{ \left( \sum_j A_{kj} + \sum_i A_{ik} \right) \theta_k^2 - 2\theta_k \left( \sum_i A_{ik} \theta_i + \sum_j A_{kj} \theta_j \right) \right\} \right] \\
&\propto \exp \left[ -\frac{1}{2\sigma^2} \{ 2M_k \theta_k^2 - 4\theta_k M_k \bar{\theta}_k \} \right], \\
&\propto \exp \left[ -\frac{1}{2(\sigma^2/2M_k)} (\theta_k - \bar{\theta}_k)^2 \right],
\end{aligned}$$

where  $A_{kk} = 0$  for any  $k = 1, \dots, n$ ,  $A_{ij} = A_{ji}$  by definition, and  $M_k = \sum_{j=1}^n A_{kj}$  and  $\bar{\theta}_k = M_k^{-1} \sum_{j=1}^n A_{kj} \theta_j$ . Thus, we have

$$\theta_k | \theta_{-k}, \sigma^2 \sim \mathcal{N} \left( \bar{\theta}_k, \frac{\sigma^2}{2M_k} \right),$$

This prior is reasonable in a sense it assumes a conjugate Gaussian distribution with mean and variances based on the average of its neighbors.

### (b)

The univariate full conditional posterior distributions of  $\theta_k$ ,  $\sigma^2$  and  $\tau^2$  can be derived as follows:

For,  $\theta_k$ , we have

$$\pi(\theta_k | \theta_{-k}, \sigma^2, \mathbf{Y}) \propto \exp \left[ -\frac{1}{2(\sigma^2/2M_k)} (\theta_k - \bar{\theta}_k)^2 \right] \cdot \exp \left\{ -\frac{(Y_k - \theta_k)^2}{2\tau^2} \right\},$$

which leads to

$$\theta_k | \theta_{-k}, \sigma^2, \tau^2, \mathbf{Y} \sim \mathcal{N} \left( \frac{\sigma^2 Y_k + 2\tau^2 M_k \bar{\theta}_k}{\sigma^2 + 2\tau^2 M_k}, \frac{\tau^2 \sigma^2}{\sigma^2 + 2\tau^2 M_k} \right).$$

by completing the square.

For  $\sigma^2$ , we have

$$\pi(\sigma^2 | \theta, \tau^2, \mathbf{Y}) \propto (\sigma^2)^{-n/2} \exp \left[ -\frac{1}{2\sigma^2} \sum_{i,j} A_{ij} (\theta_i - \theta_j)^2 \right] \cdot (\sigma^2)^{-(a+1)} \exp(-b/\sigma^2),$$

which leads to

$$\sigma^2 | \theta, \tau^2, \mathbf{Y} \sim \text{InvGamma} \left( a + \frac{n}{2}, b + \frac{1}{2} \sum_{i,j} A_{ij} (\theta_i - \theta_j)^2 \right).$$

For  $\tau^2$ , we have

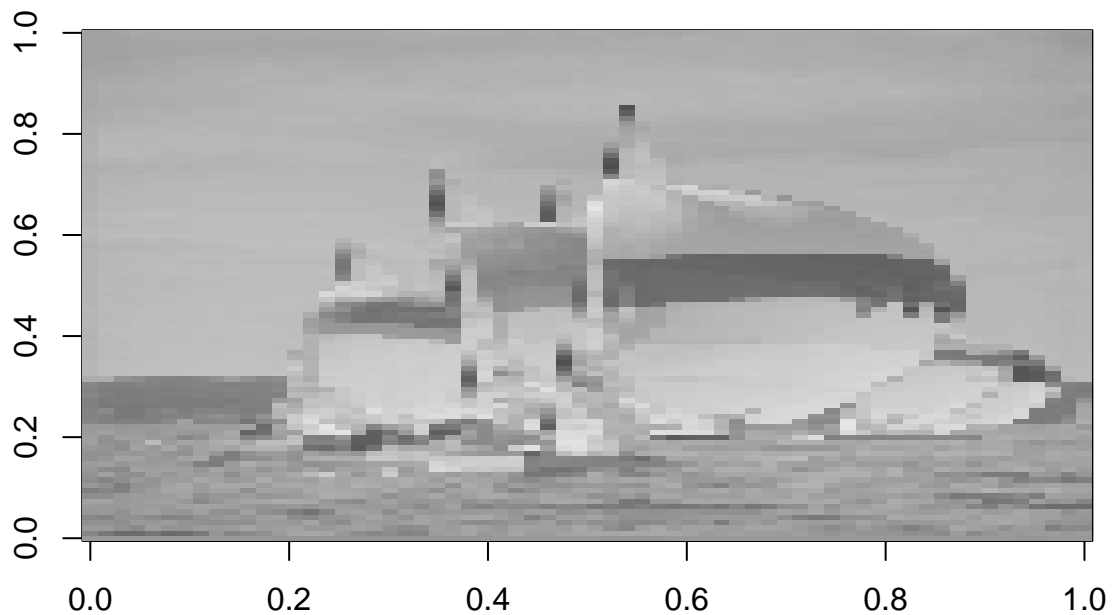
$$\pi(\tau^2 | \theta, \sigma^2, \mathbf{Y}) \propto (\tau^2)^{-n/2} \exp \left\{ -\frac{1}{2\tau^2} \sum_i (Y_i - \theta_i)^2 \right\} \cdot (\tau^2)^{-(a+1)} \exp(-b/\tau^2),$$

which leads to

$$\tau^2 \mid \theta, \sigma^2, \mathbf{Y} \sim \text{InvGamma} \left( a + \frac{n}{2}, b + \frac{1}{2} \sum_i (Y_i - \theta_i)^2 \right).$$

(c)

```
# plot(boats)
Y <- grayscale(boats)[,384:1,1,1]
# image(Y,col=gray.colors(100))
y <- Y
y <- Y[1:nrow(y)%4==2,] # Reduced image for speed
y <- y[,1:ncol(y)%4==2]
image(y,col=gray.colors(100))
```



```
# initialization placeholder for iterations
gibbs.obj <- list(list(theta = y,
                      sigma2 = 1e-5, tau2 = 1e-5))
difs <- matrix(0, nrow = height, ncol = width)
for(it in 2:niters){
  # update the current status
  theta <- gibbs.obj[[it-1]]$theta
  sigma2 <- gibbs.obj[[it-1]]$sigma2
  tau2 <- gibbs.obj[[it-1]]$tau2
  for(r.idx in 1:nrow(y)){
    for (c.idx in 1:ncol(y)) {
      # document the neighborhood
```

```

if(r.idx == 1 & c.idx == 1){ # on the corner
  M.ij <- c(theta[r.idx, c.idx + 1],
            theta[r.idx + 1, c.idx])
} else if(r.idx == 1 & c.idx == width){
  M.ij <- c(theta[r.idx, c.idx - 1],
            theta[r.idx + 1, c.idx])
} else if(r.idx == height & c.idx == 1){
  M.ij <- c(theta[r.idx, c.idx + 1],
            theta[r.idx - 1, c.idx])
} else if(r.idx == height & c.idx == width){
  M.ij <- c(theta[r.idx, c.idx - 1],
            theta[r.idx - 1, c.idx])
} else if(r.idx == 1){ # on the border
  M.ij <- c(theta[r.idx, c.idx - 1],
            theta[r.idx, c.idx + 1],
            theta[r.idx + 1, c.idx])
} else if(c.idx == 1){
  M.ij <- c(theta[r.idx, c.idx + 1],
            theta[r.idx - 1, c.idx],
            theta[r.idx + 1, c.idx])
} else if(r.idx == height){
  M.ij <- c(theta[r.idx, c.idx - 1],
            theta[r.idx, c.idx + 1],
            theta[r.idx - 1, c.idx])
} else if(c.idx == width){
  M.ij <- c(theta[r.idx, c.idx - 1],
            theta[r.idx - 1, c.idx],
            theta[r.idx + 1, c.idx])
} else{
  M.ij <- c(theta[r.idx, c.idx - 1],
            theta[r.idx, c.idx + 1],
            theta[r.idx - 1, c.idx],
            theta[r.idx + 1, c.idx])
}

}

# update theta_ij
theta[r.idx, c.idx] <- rnorm(1,
                            mean = (sigma2 * y[r.idx, c.idx] +
                                     2 * tau2 * length(M.ij) * mean(M.ij))/
                            (sigma2 + 2 * tau2 * length(M.ij)),
                            sd = sqrt(tau2 * sigma2/
                                       (sigma2 + 2 * tau2 * length(M.ij))))

# document the sum of squares
diffs[r.idx, c.idx] <- sum((theta[r.idx, c.idx] - M.ij)^2)
}
}

# update sigma2
sigma2 <- rinvgamma(1, shape = a + n/2,
                   scale = b + 0.5 * sum(diffs))

# update tau2

```

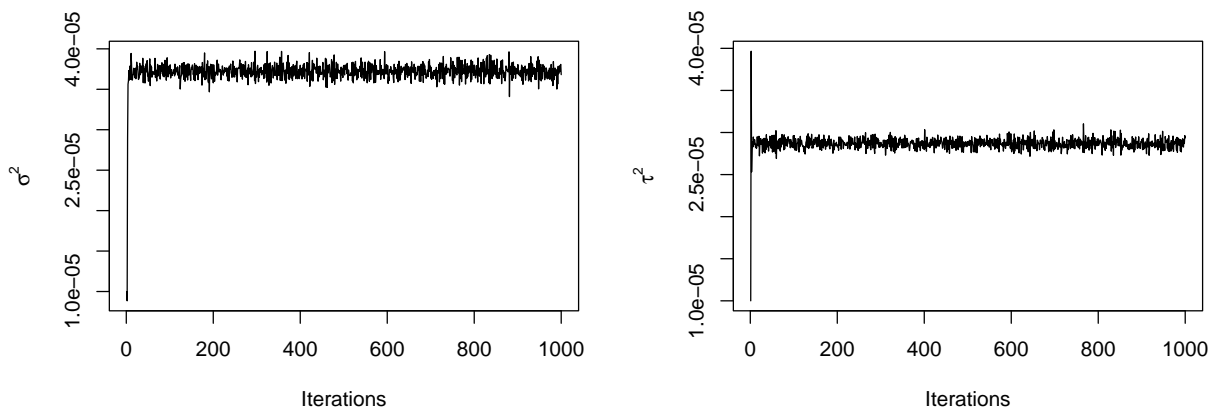
```

tau2 <- rinvgamma(1, shape = a + n/2,
                 scale = b + 0.5 * sum((y-theta)^2))

# record the updated status
gibbs.obj[[it]] <- list(theta = theta,
                      sigma2 = sigma2, tau2 = tau2)
}
save(gibbs.obj, file = 'ST740_HW2_Q2.RData')

```

Make trace plots of the variance components  $\sigma^2$  and  $\tau^2$ . One can observe that they converged pretty well (i.e., oscillate round one value).



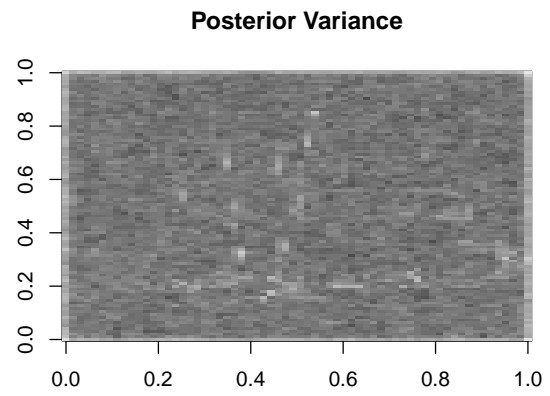
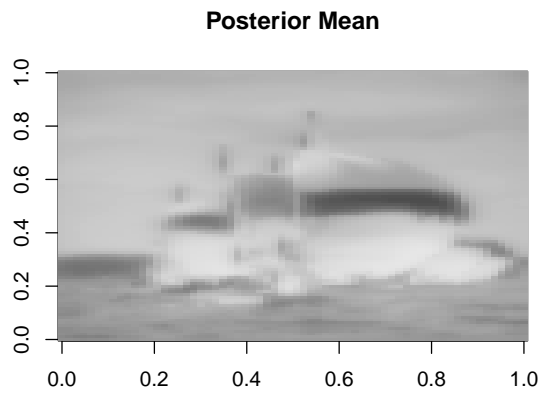
Map the posterior mean and variance of  $(\theta_1, \dots, \theta_n)$ . Let the burn-in iterations be the first 100 from 1000 samples, we can compute the posterior mean and variance of  $(\theta_1, \dots, \theta_n)$ .

```

burnin <- 100
theta.array <- abind(lapply(gibbs.obj, function(x)x$theta),
                  along = 3)

# burn-in
theta.array.post <- theta.array[, , -(1:burnin)]
# posterior mean
theta.pm <- apply(theta.array.post, 1:2, mean)
par(mfrow=c(1,2))
image(theta.pm, col=gray.colors(100), main = 'Posterior Mean')
# posterior variance
theta.pv <- apply(theta.array.post, 1:2, var)
image(theta.pv, col=gray.colors(100), main = 'Posterior Variance')

```



One can observe that

- The posterior mean image is a smoothed version of the original image.
- The posterior variance image shows that the pixels near the edges of the boats tend to have larger variances