

**PROBLEM SET 2**  
**STAT 221**

WON I. LEE

1. POSTERIOR AND LIKELIHOOD

The joint posterior of the model is given by the following:

$$p(\mu, \sigma^2, \log \theta | Y, w) \propto p(Y | \mu, \sigma^2, \log \theta, w) p(\mu, \sigma^2, \log \theta | w)$$

where we treat  $w$  as constant and fixed (thus conditioned on) throughout the calculation. We have:

$$p(Y | \mu, \sigma^2, \log \theta, w) = \prod_{j=1}^J \prod_{n=1}^N p(Y_{jn} | \mu, \sigma^2, \log \theta_j, w_j) = \prod_{j=1}^J \prod_{n=1}^N \text{Pois}(Y_{jn} | w_j \theta_j)$$

and:

$$p(\mu, \sigma^2, \log \theta | w) = p(\log \theta | \mu, \sigma^2, w) p(\mu, \sigma^2 | w) \propto \frac{1}{\sigma^2} \prod_{j=1}^J \mathcal{N}(\log \theta_j | \mu, \sigma^2)$$

Putting these together, we have:

$$p(\mu, \sigma^2, \log \theta | Y, w) \propto \frac{1}{\sigma^2} \prod_{j=1}^J \mathcal{N}(\log \theta_j | \mu, \sigma^2) \prod_{n=1}^N \text{Pois}(Y_{jn} | w_j \theta_j)$$

The conditional posterior is given by:

$$p(\log \theta_j | Y, w, \mu, \sigma^2) \propto p(Y_{j\cdot} | \log \theta_j, Y_{-j\cdot}, w, \mu, \sigma^2) p(\log \theta_j | Y_{-j\cdot}, w, \mu, \sigma^2)$$

where  $Y_{j\cdot}$  indicates  $Y_{j1}, \dots, Y_{jN}$  and  $Y_{-j\cdot}$  indicates all other values of  $Y$  (i.e. all other columns). We note that by the presumed i.i.d. assumptions:

$$p(Y_{j\cdot} | \log \theta_j, Y_{-j\cdot}, w, \mu, \sigma^2) = p(Y_{j\cdot} | \log \theta_j, w) = \prod_{n=1}^N \text{Pois}(Y_{jn} | w_j \theta_j)$$

and:

$$p(\log \theta_j | Y_{-j\cdot}, w, \mu, \sigma^2) = p(\log \theta_j | \mu, \sigma^2) = \mathcal{N}(\log \theta_j | \mu, \sigma^2)$$

so putting this together:

$$p(\log \theta_j | Y, w, \mu, \sigma^2) \propto \mathcal{N}(\log \theta_j | \mu, \sigma^2) \prod_{n=1}^N \text{Pois}(Y_{jn} | w_j \theta_j)$$

The conditional posterior is unimodal, as verified by both analytical observation and numerical visualization. That is, we substituted reasonable values for the parameters and computed the posterior numerically at grid-points of  $\log \theta_j$  to verify

that the posterior is unimodal. It is also log-concave, since taking derivatives twice of the log-posterior according to  $\log \theta_j$  yields:

$$\frac{\partial^2}{\partial \theta_j^2} \log p(\log \theta_j | Y_{-j, \cdot}, w, \mu, \sigma^2) = -\sigma^{-2} - nw_j \theta_j < 0$$

## 2. SIMULATE DATA

Please see `wonlee_ps2_functions.R` to see the desired function.

## 3. EVALUATE COVERAGE (SIMPLE)

We first started by designing the simulation as requested, with 25 values of  $Y$  for each  $\theta$  value to start. We found that with the requested parameters, i.e.  $J = 1000$ ,  $N = 2$ , and using 25 data points, we required an average system time of ... to perform a simulation for a fixed parameter setting  $\mu, \sigma^2$ .

## 4. EVALUATE COVERAGE WITH EXPOSURE WEIGHTS

## 5. EVALUATE COVERAGE WITH MISSPECIFICATION

## 6. INTERPRETATION AND RESULTS