2.2

Since the prior given is Beta(2,2), so the PDF of prior will be

$$f(\theta) = \frac{\theta(1-\theta)}{6}$$

In MAP, we're supposed to optimize $P(x,c|\theta,\pi)f(\theta)$, and we have the likelihood function from previous question.

$$\log\left(P(x|\theta)f(\theta)\right) = \left(\sum_{i=1}^{N}\sum_{c=0}^{9}\mathbb{1}\{C^{(i)}=c\}\log\pi_c + \sum_{i=1}^{N}\sum_{j=1}^{784}\left[x_j^{(i)}\log\theta_{jc} + (1-x_j^{(i)})\log(1-\theta_{jc})\right]\right) + \log\left(\frac{\theta(1-\theta)}{6}\right)$$

$$= \left(\sum_{i=1}^{N}\sum_{c=0}^{9}\mathbb{1}\{C^{(i)}=c\}\log\pi_c + \sum_{i=1}^{N}\sum_{j=1}^{784}\left[x_j^{(i)}\log\theta_{jc} + (1-x_j^{(i)})\log(1-\theta_{jc})\right]\right) + \log(\theta) + \log(1-\theta) - \log(6)$$

Then take the derivative respect to θ, and use part of result from previous question.

$$\sum_{i=1}^{N}\left(\frac{x_j^{(i)}}{\theta_{jc}} - \frac{1-x_j^{(i)}}{1-\theta_{jc}}\right)\mathbb{1}\{C^{(i)}=c\} + \frac{1}{\theta} - \frac{1}{1-\theta} = 0$$

$$\sum_{i=1}^{N}\frac{x_j^{(i)}-\theta_{jc}}{\theta_{jc}(1-\theta_{jc})}\mathbb{1}\{C^{(i)}=c\} + \frac{1}{\theta_{jc}} - \frac{1}{1-\theta_{jc}} = 0$$

Solving for θ, we have $\hat{\theta}_{map} = \dfrac{\sum_{i=1}^{N}\left[x_j^{(i)}\mathbb{1}\{C^{(i)}=c\}\right] + 1}{\sum_{i=1}^{N}\left[\mathbb{1}\{C^{(i)}=c\}\right] + 2}$