Problem 6:

$$\frac{d(\theta^t(1-\theta)^h)}{d\theta} = t\theta^{t-1}\cdot(1-\theta)^h - \theta^t\cdot h\cdot(1-\theta)^{h-1}$$

$$= \Big((1-\theta)\cdot t - h\theta\Big)\cdot\theta^{t-1}\cdot(1-\theta)^{h-1}$$

$$\frac{d^2(\theta^t(1-\theta)^h)}{d\theta^2} = \frac{d\left[\left((1-\theta)\cdot t-\theta h\right)\cdot\theta^{t-1}\cdot(1-\theta)^{h-1}\right]}{d\theta}$$

$$= (-t-h)\cdot\theta^{t-1}\cdot(1-\theta)^{h-1} + \Big((1-\theta)\cdot t-\theta h\Big)\cdot\Big[(t-1)\theta^{t-2}(1-\theta)^{h-1}$$

$$- \theta^{t-1}\cdot(h-1)\cdot(1-\theta)^{h-2}\Big]$$

$$= \Big[\big[(1-\theta)\cdot t-\theta h\big](t-1)\cdot\theta^{t-2} - (t+h)\cdot\theta^{t-1}\Big]\cdot(1-\theta)^{h-1} -$$

$$\Big((1-\theta)t-\theta h\Big)\cdot\theta^{t-1}(h-1)\cdot(1-\theta)^{h-2}$$

$\log\theta^t(1-\theta)^h$:

$$\frac{d\log\theta^t(1-\theta)^h}{d\theta} = \frac{d\left(t\log\theta + h\log(1-\theta)\right)}{d\theta}$$

$$= \frac{t}{\theta} + \frac{h}{\theta-1}$$

$$\frac{d^2\log\theta^t(1-\theta)^h}{d\theta^2} = \frac{d\left(\frac{t}{\theta}+\frac{h}{\theta-1}\right)}{d\theta} = \frac{-t}{\theta^2} - \frac{h}{(\theta-1)^2}$$

Problem 7: for $\theta_1$ is the local maximum of $\log f(\theta)$

$\log f(\theta_1) > \log f(\theta_2)$
$\log f(\theta_1) > \log f(\theta_3)$ for arbitary $\theta_2$ and $\theta_3$ $\theta_2 < \theta_1$, $\theta_3 > \theta_1$, $\theta_3$ and $\theta_2$ is close to $\theta_1$

log is monotonic function

$$\Rightarrow \quad f(\theta_1) > f(\theta_2)$$
$$f(\theta_1) > f(\theta_3)$$

so we can also say $\theta_n$ is the local maximum of $f(\theta)$

we can conclude that we can use log to compute the local minimum or maximum $\theta$, the result remains the same as using without log but the computation is smaller.

Problem 8:

Prior: $P(\theta|a,b) = Beta(a,b) \propto \theta^{a-1}(1-\theta)^{b-1}$ $\quad E(\theta|a,b) = \frac{a}{a+b}$

Likelihood: $P(X=m|N,\theta) \propto \theta^m(1-\theta)^{N-m}$

$$P(\theta|D) = \frac{P(D|\theta)\cdot P(\theta)}{P(D)} \propto P(D|\theta)\cdot P(\theta) \propto \theta^m\cdot(1-\theta)^{N-m}\cdot\theta^{a-1}(1-\theta)^{b-1}$$

$$= \theta^{m+a-1}\cdot(1-\theta)^{N-m+b-1}$$

$$\propto Beta(m+a, N-m+b)$$

$$E(\theta|D) = \frac{m+a}{m+a+N-m+b}$$

$$= \frac{m}{N+a+b} + \frac{a}{N+a+b}$$

MLE: $$\frac{d\log\theta^m(1-\theta)^{N-m}}{d\theta} = \frac{m}{\theta} - \frac{N-m}{1-\theta} = 0$$

$$m(1-\theta) = (N-m)\cdot\theta$$

$$m - m\theta = (N-m)\theta$$

$$N\theta = m$$

$$\theta_{MLE} = \frac{m}{N}$$

$$E(\theta|D) = (1-\lambda)\cdot\frac{m}{N} + \lambda\cdot\frac{a}{a+b}$$

$$E(\theta|D) = \frac{N}{N+a+b}\cdot\frac{m}{N} + \frac{a+b}{N+a+b}\cdot\frac{a}{a+b}$$

$$\lambda = \frac{a+b}{N+a+b} \qquad 1-\lambda = \frac{N}{N+a+b}$$

Problem 4:

$$\lambda_{MAP} = \underset{\lambda}{\arg\max}\ p(\lambda \mid x, a, b)$$

$$= \underset{\lambda}{\arg\max}\ \log\left(p(x \mid \lambda) \cdot p(\lambda \mid a, b)\right)$$

$$= \underset{\lambda}{\arg\max}\ \log(\lambda^x e^{-\lambda}) - \log x! + \log\frac{b^a}{\Gamma(a)} + \log \lambda^{a-1} + \log e^{-b\lambda}$$

$$= \underset{\lambda}{\arg\max}\left(x \cdot \log\lambda - \lambda + \log\left(\frac{b^a}{\Gamma(a)}\right) + (a-1)\cdot\log\lambda - b\lambda - \log x!\right)$$

$$\frac{d\left(x \cdot \log\lambda - \lambda + \log\frac{b^a}{\Gamma(a)} + (a-1)\log\lambda - b\lambda - \log(x!)\right)}{d\lambda} = 0$$

$$\frac{x}{\lambda} - 1 + (a-1)\cdot\frac{1}{\lambda} - b = 0$$

$$\frac{a-1+x}{\lambda} = b+1$$

$$\lambda_{MAP} = \frac{a-1+x}{b+1}$$