ps2-3-solve

May 13, 2021

1 (a)

$$
\begin{aligned}
\theta_{\text{MAP}} &= \arg\max_\theta p(\theta|x,y) \\
&= \arg\max_\theta \frac{p(\theta,x,y)}{p(x,y)} \\
&= \arg\max_\theta p(\theta,x,y) \\
&= \arg\max_\theta \frac{p(\theta,x,y)}{p(x)} \\
&= \arg\max_\theta \frac{p(\theta,x,y)}{p(\theta,x)} \frac{p(\theta,x)}{p(x)} \\
&= \arg\max_\theta p(y|x,\theta)p(\theta|x)
\end{aligned}
$$

because we assume that $p(\theta) = p(\theta|x)$, so

$$
\theta_{\text{MAP}} = \arg\max_\theta p(y|x,\theta)p(\theta)
$$

2 (b)

$$
\begin{aligned}
\theta_{\text{MAP}} &= \arg\max_\theta p(\theta|x,y) \\
&= \arg\max_\theta p(y|x,\theta)p(\theta) \\
&= \arg\min_\theta -p(y|x,\theta)p(\theta) \\
&= \arg\min_\theta -\log p(y|x,\theta)p(\theta) \\
&= \arg\min_\theta -\log p(y|x,\theta) - \log p(\theta)
\end{aligned}
$$

for θ, we have $\theta \sim \mathcal{N}(0, \eta^2 I)$, so

$$
p(\theta_i) = \frac{1}{\eta\sqrt{2\pi}} \exp\big(-\frac{\theta_i^2}{2\eta^2}\big)
$$

then,

$$
\begin{aligned}
p(\theta) &= \prod_{i=1}^{n} p(\theta_i) \\
\log p(\theta) &= \log \prod_{i=1}^{n} p(\theta_i) \\
&= \sum_{i=1}^{n} \log p(\theta_i) \\
&= \sum_{i=1}^{n} \log \frac{1}{\eta\sqrt{2\pi}} \exp\big(-\frac{\theta_i^2}{2\eta^2}\big) \\
&= \sum_{i=1}^{n} \log \frac{1}{\eta\sqrt{2\pi}} - \frac{\theta_i^2}{2\eta^2} \\
&= n \log \frac{1}{\eta\sqrt{2\pi}} - \frac{1}{2\eta^2} \sum_{i=1}^{n} \theta_i^2 \\
&= n \log \frac{1}{\eta\sqrt{2\pi}} - \frac{1}{2\eta^2} \|\theta\|_2^2
\end{aligned}
$$

back to θ_{MAP}, we have

$$
\begin{aligned}
\theta_{\text{MAP}} &= \arg\min_{\theta} -\log p(y|x,\theta) - \log p(\theta) \\
&= \arg\min_{\theta} -\log p(y|x,\theta) - n \log \frac{1}{\eta\sqrt{2\pi}} + \frac{1}{2\eta^2} \|\theta\|_2^2 \\
&= \arg\min_{\theta} -\log p(y|x,\theta) + \frac{1}{2\eta^2} \|\theta\|_2^2 \\
&= \arg\min_{\theta} -\log p(y|x,\theta) + \lambda \|\theta\|_2^2
\end{aligned}
$$

and,

$$
\lambda = \frac{1}{2\eta^2}
$$

3 (c)

we have $\epsilon \sim \mathcal{N}(0, \sigma^2)$, it means $y - x^T\theta \sim \mathcal{N}(0, \sigma^2)$, so

$$
\begin{aligned}
p(y^{(i)}|x^{(i)},\theta) &= \frac{1}{\sigma\sqrt{2\pi}}\exp\big(-\frac{(y^{(i)}-(x^{(i)})^T\theta)^2}{2\sigma^2}\big)\\
p(y|x,\theta) &= \prod_{i=1}^{m} p(y^{(i)}|x^{(i)},\theta)\\
\log p(y|x,\theta) &= \sum_{i=1}^{m}\log\frac{1}{\sigma\sqrt{2\pi}}\exp\big(-\frac{(y^{(i)}-(x^{(i)})^T\theta)^2}{2\sigma^2}\big)\\
&= m\log\frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2\sigma^2}\sum_{i=1}^{m}(y^{(i)}-(x^{(i)})^T\theta)^2\\
&= m\log\frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2\sigma^2}\|\vec{y}-X\theta\|_2^2
\end{aligned}
$$

back to θ_{MAP}, we have

$$
\begin{aligned}
\theta_{\text{MAP}} &= \arg\min_{\theta} -\log p(y|x,\theta) + \lambda\|\theta\|_2^2\\
&= \arg\min_{\theta} -m\log\frac{1}{\epsilon\sqrt{2\pi}} + \frac{1}{2\sigma^2}\|\vec{y}-X\theta\|_2^2 + \frac{1}{2\eta^2}\|\theta\|_2^2\\
&= \arg\min_{\theta} \|\vec{y}-X\theta\|_2^2 + \frac{\sigma^2}{\eta^2}\|\theta\|_2^2
\end{aligned}
$$

let $J(\theta) = \|\vec{y}-X\theta\|_2^2 + \frac{\sigma^2}{\eta^2}\|\theta\|_2^2$, and solve $\triangledown_\theta J(\theta) = 0$

for $\triangledown_\theta J(\theta)$,

$$
\begin{aligned}
\triangledown_\theta J(\theta) &= \triangledown_\theta\big(\|\vec{y}-X\theta\|_2^2 + \frac{\sigma^2}{\eta^2}\|\theta\|_2^2\big)\\
&= \triangledown_\theta\big((\vec{y}-X\theta)^T(\vec{y}-X\theta) + \frac{\sigma^2}{\eta^2}\theta^T\theta\big)\\
&= 2(-X)^T(\vec{y}-X\theta) + 2\frac{\sigma^2}{\eta^2}\theta
\end{aligned}
$$

then

$$
\begin{aligned}
\triangledown_\theta J(\theta) &= 0\\
2(-X)^T(\vec{y}-X\theta) + 2\frac{\sigma^2}{\eta^2}\theta &= 0\\
X^T X\theta + \frac{\sigma^2}{\eta^2}\theta &= X^T\vec{y}\\
\theta_{\text{MAP}} &= (X^T X + \frac{\sigma^2}{\eta^2}I)^{-1}X^T\vec{y}
\end{aligned}
$$

4 (d)

now, $\theta \sim \mathcal{L}(0, bI)$

for $\log p(\theta)$,

$$\begin{aligned}
p(\theta_i) &= \frac{1}{2b} \exp\big(-\frac{|\theta_i|}{b}\big) \\
p(\theta) &= \prod_{i=1}^{n} p(\theta_i) \\
\log p(\theta) &= \sum_{i=1}^{n} \log \frac{1}{2b} \exp\big(-\frac{|\theta_i|}{b}\big) \\
&= n \log \frac{1}{2b} - \frac{1}{b} \sum_{i=1}^{n} |\theta_i| \\
&= n \log \frac{1}{2b} - \frac{1}{b} \|\theta\|_1
\end{aligned}$$

back to θ_{MAP},

$$\begin{aligned}
\theta_{\text{MAP}} &= \arg\min_{\theta} - \log p(y|x, \theta) - \log p(\theta) \\
&= \arg\min_{\theta} -m \log \frac{1}{\sigma\sqrt{2\pi}} + \frac{1}{2\sigma^2} \|\vec{y} - X\theta\|_2^2 - n \log \frac{1}{2b} + \frac{1}{b} \|\theta\|_1 \\
&= \arg\min_{\theta} \|\vec{y} - X\theta\|_2^2 + \frac{2\sigma^2}{b} \|\theta\|_1
\end{aligned}$$

so,

$$J(\theta) = \|\vec{y} - X\theta\|_2^2 + \gamma \|\theta\|_1 \gamma = \frac{2\sigma^2}{b}$$

[ ]: