ps3-3-solve

May 28, 2021

1 (a)

$$\begin{aligned}
\mathbb{E}_{y\sim p(y;\theta)}[\triangledown_{\theta'} \log p(y;\theta')|_{\theta'=\theta}] &= \int_{-\infty}^{\infty} p(y;\theta)\ \triangledown_{\theta'} \log p(y;\theta')|_{\theta'=\theta} dy \\
&= \int_{-\infty}^{\infty} p(y;\theta) \frac{1}{p(y;\theta')}\ \triangledown_{\theta'}\ p(y;\theta')|_{\theta'=\theta} dy \\
&= \int_{-\infty}^{\infty} \triangledown_{\theta'} p(y;\theta')|_{\theta'=\theta} dy \\
&= \triangledown_{\theta'} \int_{-\infty}^{\infty} p(y;\theta')|_{\theta'=\theta} dy \\
&= \triangledown_{\theta'} 1 \\
&= 0
\end{aligned}$$

2 (b)

$$\begin{aligned}
\mathcal{I}(\theta) &= \mathrm{Cov}_{y\sim p(y;\theta)}[\triangledown_{\theta'} \log p(y;\theta')|_{\theta'=\theta}] \\
&= \mathbb{E}_{y\sim p(y;\theta)}[\triangledown_{\theta'} \log p(y;\theta')\ \triangledown_{\theta'} \log p(y;\theta')^T|_{\theta'=\theta}] \\
&\quad - \mathbb{E}_{y\sim p(y;\theta)}[\triangledown_{\theta'} \log p(y;\theta')|_{\theta'=\theta}]\mathbb{E}_{y\sim p(y;\theta)}[\triangledown_{\theta'} \log p(y;\theta')|_{\theta'=\theta}] \\
&= \mathbb{E}_{y\sim p(y;\theta)}[\triangledown_{\theta'} \log p(y;\theta')\ \triangledown_{\theta'} \log p(y;\theta')^T|_{\theta'=\theta}]
\end{aligned}$$

3 (c)

$$
\begin{aligned}
\mathbb{E}_{y\sim p(y;\theta)}[-\bigtriangledown^2_{\theta'}\log p(y;\theta')|_{\theta'=\theta}] &= \mathbb{E}_{y\sim p(y;\theta)}[-\bigtriangledown_{\theta'}\big(\frac{1}{p(y;\theta')}\bigtriangledown_{\theta'} p(y;\theta')\big)|_{\theta'=\theta}] \\
&= \mathbb{E}_{y\sim p(y;\theta)}[\frac{1}{p(y;\theta')^2}\bigtriangledown_{\theta'} p(y;\theta')\bigtriangledown_{\theta'} p(y;\theta')^T - \frac{1}{p(y;\theta')}\bigtriangledown^2_{\theta'} p(y;\theta')|_{\theta'=\theta}] \\
&= \mathbb{E}_{y\sim p(y;\theta)}[\frac{1}{p(y;\theta')^2}\bigtriangledown_{\theta'} p(y;\theta')\bigtriangledown_{\theta'} p(y;\theta')^T|_{\theta'=\theta}] \\
&\quad - \mathbb{E}_{y\sim p(y;\theta)}[\frac{1}{p(y;\theta')}\bigtriangledown^2_{\theta'} p(y;\theta')|_{\theta'=\theta}] \\
&= \mathbb{E}_{y\sim p(y;\theta)}[\frac{1}{p(y;\theta')^2}\bigtriangledown_{\theta'} p(y;\theta')\bigtriangledown_{\theta'} p(y;\theta')^T|_{\theta'=\theta}] \\
&\quad - \int_{-\infty}^{\infty} p(y;\theta)\frac{1}{p(y;\theta')}\bigtriangledown^2_{\theta'} p(y;\theta')|_{\theta'=\theta} dy \\
&= \mathbb{E}_{y\sim p(y;\theta)}[\frac{1}{p(y;\theta')^2}\bigtriangledown_{\theta'} p(y;\theta')\bigtriangledown_{\theta'} p(y;\theta')^T|_{\theta'=\theta}] \\
&= \mathbb{E}_{y\sim p(y;\theta)}[\bigtriangledown_{\theta'}\log p(y;\theta')\bigtriangledown_{\theta'}\log p(y;\theta')^T|_{\theta'=\theta}] \\
&= \mathcal{I}(\theta)
\end{aligned}
$$

4 (d)

$$
\begin{aligned}
D_{KL}(p_\theta \| p_{\theta+d}) &= \mathbb{E}[\log p_\theta] - \mathbb{E}[\log p_{\theta+d}] \\
&\approx \mathbb{E}[\log p_\theta] - \mathbb{E}[\log p_\theta + d^T \bigtriangledown_{\theta'}\log p_{\theta'}|_{\theta'=\theta} + \frac{1}{2}d^T(\bigtriangledown^2_{\theta'}\log p_{\theta'}|_{\theta'=\theta})d] \\
&= \mathbb{E}[\log p_\theta] - \mathbb{E}[\log p_\theta] - \mathbb{E}[d^T \bigtriangledown_{\theta'}\log p_{\theta'}|_{\theta'=\theta}] - \mathbb{E}[\frac{1}{2}d^T \bigtriangledown^2_{\theta'}\log p_{\theta'} d|_{\theta'=\theta}] \\
&= -d^T\mathbb{E}[\bigtriangledown_{\theta'}\log p_{\theta'}|_{\theta'=\theta}] + \frac{1}{2}d^T\mathbb{E}[-\bigtriangledown^2_{\theta'}\log p_{\theta'}|_{\theta'=\theta}]d \\
&= \frac{1}{2}d^T\mathcal{I}(\theta)d
\end{aligned}
$$

5 (e)

Lagrangian:

$$
\begin{aligned}
\mathcal{L}(d,\lambda) &= \ell(\theta+d) - \lambda(D_{KL}(p_\theta\|p_{\theta+d}) - c) \\
&= \log p_{\theta+d} - \lambda(D_{KL}(p_\theta\|p_{\theta+d}) - c) \\
&\approx \log p_\theta + d^T \bigtriangledown_{\theta'}\log p_{\theta'}|_{\theta'=\theta} - \lambda(\frac{1}{2}d^T\mathcal{I}(\theta)d - c)
\end{aligned}
$$

for $\bigtriangledown_d\mathcal{L}(d,\lambda) = 0$:

$$
\begin{aligned}
\triangledown_d \mathcal{L}(d,\lambda) &= 0 \\
\triangledown_d \big(d^T \triangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta} - \lambda(\frac{1}{2} d^T \mathcal{I}(\theta) d - c)\big) &\approx 0 \\
\triangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta} &= \lambda \mathcal{I}(\theta) d \\
\widetilde{d} &= \frac{1}{\lambda} \mathcal{I}(\theta)^{-1} \triangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta}
\end{aligned}
$$

for $\triangledown_\lambda \mathcal{L}(d,\lambda) = 0$:

$$
\begin{aligned}
\triangledown_\lambda \mathcal{L}(d,\lambda) &= 0 \\
\triangledown_\lambda \big(- \lambda(\frac{1}{2} d^T \mathcal{I}(\theta) d - c)\big) &= 0 \\
\frac{1}{2} d^T \mathcal{I}(\theta) d &= c \\
\frac{1}{2} \big(\frac{1}{\lambda} \mathcal{I}(\theta)^{-1} \triangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta}\big)^T \mathcal{I}(\theta) \big(\frac{1}{\lambda} \mathcal{I}(\theta)^{-1} \triangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta}\big) &= c \\
\frac{1}{2} \frac{1}{\lambda^2} \big(\triangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta}\big)^T \mathcal{I}(\theta)^{-1} \triangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta} &= c \\
\frac{1}{2} \frac{1}{c} \big(\triangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta}\big)^T \mathcal{I}(\theta)^{-1} \triangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta} &= \lambda^2
\end{aligned}
$$

Because $\lambda \in \mathbb{R}_+$, so:

$$
\lambda = \Big(\frac{1}{2} \frac{1}{c} \big(\triangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta}\big)^T \mathcal{I}(\theta)^{-1} \triangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta} \Big)^{\frac{1}{2}}
$$

Plug λ into $\widetilde{d}$:

$$
d^* = \Big(\frac{1}{2} \frac{1}{c} \big(\triangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta}\big)^T \mathcal{I}(\theta)^{-1} \triangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta} \Big)^{-\frac{1}{2}} \mathcal{I}(\theta)^{-1} \triangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta}
$$

6 (f)

Newton's Method:

$$
\theta := \theta - H^{-1} \triangledown_\theta \ell(\theta)
$$

Natural Gradient:

$$
\begin{aligned}
\theta &:= \theta + \widetilde{d} \\
&= \theta + \frac{1}{\lambda}\mathcal{I}(\theta)^{-1} \bigtriangledown_{\theta'} \log p_{\theta'}|_{\theta'=\theta} \\
&= \theta + \frac{1}{\lambda}\mathbb{E}_{y\sim p(y;\theta)}[- \bigtriangledown^2_{\theta'} \log p(y;\theta')|_{\theta'=\theta}]^{-1} \bigtriangledown_{\theta'} \ell(\theta) \\
&= \theta - \frac{1}{\lambda}\mathbb{E}_{y\sim p(y;\theta)}[\bigtriangledown^2_{\theta'}\ell(\theta)]^{-1} \bigtriangledown_{\theta'} \ell(\theta) \\
&= \theta - \frac{1}{\lambda}\mathbb{E}_{y\sim p(y;\theta)}[H^{-1}] \bigtriangledown_{\theta'} \ell(\theta)
\end{aligned}
$$

```
[ ]:
```