HW 16 Q3

$$\pi(a|s,\theta) = \frac{e^{\theta^T \phi(s,a)}}{\sum_{b\in A} e^{\theta^T \phi(s,b)}}$$

$$\log \pi(a|s,\theta) = \theta^T \phi(s,a) - \log\left(\sum_{b\in A} e^{\theta^T \phi(s,b)}\right)$$

$$\Rightarrow \frac{\partial \log \pi(a|s,\theta)}{\partial \theta_i} = \phi_i(s,a) - \frac{\sum_b \phi_i(s,b) e^{\theta^T \phi(s,b)}}{\sum_b e^{\theta^T \phi(s,b)}}$$

$$= \phi_i(s,a) - \sum_{b\in A} \pi(b|s,\theta) . \phi_i(s,b) = \phi_i(s,a) - \mathbb{E}_{\pi}\left[\phi_i(s,\cdot)\right]$$

Then; $\nabla_\theta \log \pi(a|s,\theta) = \phi(s,a) - \mathbb{E}_{\pi}\left[\phi(s,\cdot)\right]$ $\square$.

To have $\nabla_w Q(s,a;w) = \nabla_\theta \log \pi(a|s;\theta)$

$\Rightarrow Q(s,a;w) = w^T \nabla_\theta \log \pi(a|s;\theta)$. We will show that: $\sum_{a\in A} \pi(a|s;\theta) Q(s,a;w) = 0$

$\Leftrightarrow \sum_{a\in A} \pi(a|s;\theta)\, w^T \nabla_\theta \log \pi(a|s;\theta)$

We know that $\nabla_\theta \pi(a|s;\theta) = \pi(a|s;\theta) \nabla_\theta \log \pi(a|s;\theta)$

$\Leftrightarrow \sum_{a\in A} w^T \nabla_\theta \pi(a|s;\theta) = 0 \Leftrightarrow w^T \nabla_\theta \underbrace{\sum \pi(a|s;\theta)}_{=0}$

as $\sum \pi(a|s;\theta) = 1 \Rightarrow \nabla_\theta \sum \pi(a|s;\theta) = 0$

then this holds. $\square$.