

Assignment 16

③ Evaluate the score fn for $\pi(s, a; \vec{\theta}) = \frac{e^{\vec{\phi}(s, a)^T \vec{\theta}}}{\sum_b e^{\vec{\phi}(s, b)^T \vec{\theta}}}$

$$\log \pi = \vec{\phi}(s, a)^T \vec{\theta} - \log(\sum_b e^{\vec{\phi}(s, b)^T \vec{\theta}})$$

$$\vec{\nabla}_{\theta} \log \pi = \vec{\phi}(s, a) - \frac{1}{\sum_b e^{\vec{\phi}(s, b)^T \vec{\theta}}} \cdot \sum_b e^{\vec{\phi}(s, b)^T \vec{\theta}} \cdot \vec{\phi}(s, b)$$

$$= \boxed{\vec{\phi}(s, a) - \sum_b \pi(s, b; \vec{\theta}) \vec{\phi}(s, b)}$$

Construct $Q(s, a; \vec{w})$

Consider: $\boxed{Q(s, a; \vec{w}) = \vec{\phi}(s, a)^T \vec{w} - \sum_b \pi(s, b; \vec{\theta}) \vec{\phi}(s, b)^T \vec{w}}$

We can check that $\vec{\nabla}_{\vec{w}} Q = \vec{\phi}(s, a) - \sum_b \pi(s, b; \vec{\theta}) \vec{\phi}(s, b) = \vec{\nabla}_{\theta} \log \pi$

Check that Q is zero-mean

$$\mathbb{E}_{\pi} [Q(s, a; \vec{w})] = \sum_{a \in \mathcal{A}} \pi(s, a; \vec{\theta}) \cdot Q(s, a; \vec{w}) = 0$$

$$= \sum_{a \in \mathcal{A}} \pi(s, a; \vec{\theta}) \cdot (\vec{\phi}(s, a)^T \vec{w} - \sum_b \pi(s, b; \vec{\theta}) \vec{\phi}(s, b)^T \vec{w})$$

$$= \underbrace{\sum_{a \in \mathcal{A}} \pi(s, a; \vec{\theta}) \vec{\phi}(s, a)^T \vec{w}}_{=1} - \underbrace{\left(\sum_{a \in \mathcal{A}} \pi(s, a; \vec{\theta}) \right) \left(\sum_b \pi(s, b; \vec{\theta}) \vec{\phi}(s, b)^T \vec{w} \right)}_{\text{equal}} = 0$$

$$= 0$$