③

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

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

$$\nabla_\theta \log \pi(s,a;\theta) = \phi(s,a) - \frac{1}{\sum_{b\in\mathcal{A}} e^{\phi(s,b)^T\theta}} \cdot \sum_{b\in\mathcal{A}} \phi(s,b)\, e^{\phi(s,b)^T\theta}$$

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

$$= \phi(s,a) - \sum_{b\in\mathcal{A}} \phi(s,b)\, \pi(s,b;\theta)$$

$$\boxed{\nabla_\theta \log \pi(s,a;\theta) = \phi(s,a) - \sum_{b\in\mathcal{A}} \phi(s,b)\, \pi(s,b;\theta)}$$

Compatible Function Approximation Theorem requires:

$$\frac{\partial Q(s,a;w)}{\partial w_i} = \frac{\partial \log \pi(s,a;\theta)}{\partial \theta_i} \quad \text{for all } i = 1 \ldots m$$

Therefore:

$$Q(s,a;w) = \sum_{i=1}^{m} \frac{\partial \log \pi(s,a;\theta)}{\partial \theta_i} w_i = \nabla_\theta \log \pi(s,a;\theta)^T w$$

$$\boxed{Q(s,a;w) = \phi(s,a)^T w - \left(\sum_{b \in A} \phi(s,b)\, \pi(s,b;\theta)\right)^T w}$$

Show $\mathbb{E}_\pi [Q(s,a;w)] = 0$:

$$\sum_{a \in A} \pi(s,a;\theta)\, Q(s,a;w) = \sum_{a \in A} \pi(s,a;\theta) \left(\sum_{i=1}^{m} \frac{\partial \log \pi(s,a;\theta)}{\partial \theta_i} w_i \right)$$

by construction

$$= \sum_{a \in A} \left(\sum_{i=1}^{m} \frac{\partial \pi(s,a;\theta)}{\partial \theta_i} \cdot w_i \right)$$

definition of score

$$= \sum_{i=1}^{m} \left(\sum_{a \in A} \frac{\partial \pi(s,a;\theta)}{\partial \theta_i} \right) w_i$$

$$= \sum_{i=1}^{m} \frac{\partial}{\partial \theta_i} \left(\sum_{a \in A} \pi(s,a;\theta) \right) \cdot w_i$$

$$= \sum_{i=1}^{m} \frac{\partial}{\partial \theta_i} (1) \cdot w_i$$

$$= 0$$

With our specific $Q(s,a;w)$:

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

$$= \sum_{a\in\mathcal{A}} \tilde{\pi}(s,a;\theta)\left[\phi(s,a)^T w - \left(\sum_{b\in\mathcal{A}} \phi(s,b)\,\tilde{\pi}(s,b;\theta)\right)^T w\right]$$

$$= \sum_{a\in\mathcal{A}}\left[\frac{e^{\phi(s,a)^T\theta}\,\phi(s,a)^T w}{\sum_{b\in\mathcal{A}} e^{\phi(s,b)^T\theta}} - \frac{e^{\phi(s,a)^T\theta}}{\cancel{\sum_{b\in\mathcal{A}} e^{\phi(s,b)^T\theta}}}\cdot\left(\sum_{b\in\mathcal{A}}\left[\frac{\phi(s,b)\,\cancel{e^{\phi(s,b)^T\theta}}}{\sum_{c\in\mathcal{A}} e^{\phi(s,c)^T\theta}}\right]\right)^T w\right]$$

$$= \sum_{a\in\mathcal{A}}\left[\frac{e^{\phi(s,a)^T\theta}\,\phi(s,a)^T w}{\sum_{b\in\mathcal{A}} e^{\phi(s,b)^T\theta}} - e^{\phi(s,a)^T\theta}\cdot\frac{\sum_{b\in\mathcal{A}}\phi(s,b)^T w}{\sum_{c\in\mathcal{A}} e^{\phi(s,c)^T\theta}}\right]$$

$$= \sum_{a\in\mathcal{A}}\left[\frac{e^{\phi(s,a)^T\theta}\cdot\phi(s,a)^T w - e^{\phi(s,a)^T\theta}\cdot\sum_{b\in\mathcal{A}}\phi(s,b)^T w}{\sum_{b\in\mathcal{A}} e^{\phi(s,b)^T\theta}}\right]$$

$$= \sum_{a\in\mathcal{A}}\left[\frac{0}{\sum_{b\in\mathcal{A}} e^{\phi(s,b)^T\theta}}\right]$$

$$= 0$$