PS3 - Written

1.a.

$$\frac{\nabla_\theta \pi_\theta(s_t, a_t)}{\pi_\theta(s_t, a_t)} = \nabla_\theta \ln(\pi_\theta(s_t, a_t))$$

$$= \nabla_\theta \theta^T s_t[a_t] - \frac{\nabla_\theta \sum_{a'} e^{\theta^T s_t[a_t]}}{\sum_{a'} e^{\theta^T s_t[a_t]}}$$

since $[\nabla_\theta \theta^T s_t[a_t]]_{i,j} = \begin{cases} s_t[i] & \text{iff } a_t = j \\ 0 & \text{otherwise} \end{cases}$

and

$$\left[\frac{\nabla_\theta \sum_{a'} e^{\theta^T s_t[a_t]}}{\sum_{a'} e^{\theta^T s_t[a_t]}}\right]_{i,j} = \frac{s_t[i]\, e^{\theta^T s_t[i][a_t]}}{\sum_{a'} e^{\theta^T s_t[i][a_t]}}$$

$$= s_t[i]\, \pi_\theta(s_t, j)$$

Therefore

$$\left[\frac{\nabla_\theta \pi_\theta(s_t, a_t)}{\pi_\theta(s_t, a_t)}\right]_{i,j} = \begin{cases} s_t[i][1 - \pi_\theta(s_t, j)] & \text{iff } a_t = j \\ -s_t[i]\, \pi_\theta(s_t, j) & \text{otherwise.} \end{cases}$$