PS – 4

Page No.:
Date:

2. (a) If $\hat{\pi}_0 = \pi_0$,

$$E_{\substack{s\sim p(s)\\ a\sim\pi_0(s,a)}} \frac{\pi_1(s,a)}{\hat{\pi}_0(s,a)} R(s,a)$$

$$= E_{\substack{s\sim p(s)\\ a\sim\pi_0(s,a)}} \frac{\pi_1(s,a)}{\pi_0(s,a)} R(s,a)$$

$$= \cancel{E_{\substack{s\sim p(s)\\ a\sim\pi_0(s,a)}}} \sum_{(s,a)} R(s,a)\, p(s)\, \cancel{\pi_0(s,a)} \cdot \frac{\pi_1(s,a)}{\cancel{\pi_0(s,a)}}$$

$$= E_{\substack{s\sim p(s)\\ a\sim\pi_1(s,a)}} R(s,a)$$

(b) If $\hat{\pi}_0 = \pi_0$,

$$\frac{E_{\substack{s\sim p(s)\\ a\sim\pi_0(s,a)}} \frac{\pi_1(s,a)}{\hat{\pi}_0(s,a)} R(s,a)}{E_{\substack{s\sim p(s)\\ a\sim\pi_0(s,a)}} \frac{\pi_1(s,a)}{\hat{\pi}_0(s,a)}}$$

$$= \frac{E_{\substack{s\sim p(s)\\ a\sim\pi_1(s,a)}} R(s,a)}{1 = \left\{ \sum_{(s,a)} p(s)\, \cancel{\pi_0(s,a)} \frac{\pi_1(s,a)}{\cancel{\pi_0(s,a)}} \right.} = E_{\substack{s\sim p(s)\\ a\sim\pi_1(s,a)}} R(s,a)$$

(c) If there exists a single sample

$$\frac{E_{\substack{s\sim p(s)\\ a\sim\pi_0(s,a)}} \frac{\pi_1(s,a)}{\hat{\pi}_0(s,a)} R(s,a)}{E_{\substack{s\sim p(s)\\ a\sim\pi_0(s,a)}} \frac{\pi_1(s,a)}{\hat{\pi}_0(s,a)}}$$

$$= \frac{\cancel{p(s)} R(s,a) \cancel{\frac{\pi_1(s,a)}{\hat{\pi}_0(s,a)}}}{\cancel{p(s)} \cancel{\frac{\pi_1(s,a)}{\hat{\pi}_0(s,a)}}} = R(s,a)$$

3. $$\arg\min_{u:u^Tu=1} \sum_{i=1}^{m} \|x^{(i)} - f_u(x^{(i)})\|_2^2$$

$$= \arg\min_{u:u^Tu=1} \sum_{i=1}^{m} (x^{(i)} - uu^T x^{(i)})^T (x^{(i)} - uu^T x^{(i)})$$

$$= \arg\min_{u:u^Tu=1} \sum_{i=1}^{m} (x^{(i)T} - x^{(i)T} uu^T)(x^{(i)} - uu^T x^{(i)})$$

$$= \arg\min_{u:u^Tu=1} \sum_{i=1}^{m} x^{(i)T} x^{(i)} - x^{(i)T} uu^T x^{(i)} - x^{(i)T} uu^T + x^{(i)T} uu^T uu^T x^{(i)}$$

$$= \arg\max_{u:u^Tu=1} \sum_{i=1}^{m} x^{(i)T} uu^T (x^{(i)}) = \arg\max_{u:u^Tu=1} u^T \sum_{i=1}^{m} x^{(i)} x^{(i)T} u$$

4. (a) $$\nabla_W l = \nabla_W \sum_{i=1}^{n} \left(\log|W| + \sum_{j=1}^{d} \log \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(w_j^T x^{(i)})^2\right)\right)$$

$$= (W^{-1})^T \cdot m - \sum_{i=1}^{n} W x^{(i)} x^{(i)T}$$

$$= m(W^{-1})^T - WX^TX = 0$$

$$\Rightarrow W^TW = m(X^TX)^{-1}$$

If $Y \to$ ortho, then

$$(YW)^T(YW) = W^T Y^T Y W = (W^TW).$$

(b) $$\nabla_W l(W) = \nabla_W \left(\log|W| + \sum_{i=1}^{m} \log \frac{1}{2} \exp(-|w_j^T x^{(i)}|)\right)$$

$$= (W^{-1})^T - \text{sign}(Wx^{(i)})\, x^{(i)T} = \Delta$$

$$\Rightarrow W^+ = W + \alpha\Delta$$

(c)

5. (a) $\|B(V_1) - B(V_2)\|_\infty = \gamma \left\| \max_{a \in A} \sum_{s' \in S} P_{sa}(s')[V_1(s') - V_2(s')] \right\|_\infty$

$= \gamma \max_{s \in S} \left| \max_{a \in A} \sum_{s' \in S} P_{sa}(s')[V_1(s') - V_2(s')] \right|$

$\leq \gamma \|V_1 - V_2\|_\infty$

(b) $\|V_1 - V_2\|_\infty = \|B(V_1) - B(V_2)\|_\infty \leq \gamma \|V_1 - V_2\|_\infty$

$\|V_1 - V_2\|_\infty = 0 \Rightarrow V_1 = V_2$