(2.a)

$$\begin{aligned}
\mathbb{E}_{\substack{s\sim p(s)\\ a\sim\pi_0(s,a)}} \frac{\pi_1(s,a)}{\pi_0(s,a)}R(s,a) &= \sum_{s,a} p(s)\pi_0(s,a)\frac{\pi_1(s,a)}{\pi_0(s,a)}R(s,a) \\
&= \sum_{s,a} p(s)\pi_1(s,a)R(s,a) \\
&= \mathbb{E}_{\substack{s\sim p(s)\\ a\sim\pi_0(s,a)}} R(s,a)
\end{aligned}$$

(2.b)

The denominator is equal to one.

(2.c)

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

if $\pi_0 \neq \pi_1$., then the sampling is biased.

(2.d)

1.

$$\begin{aligned}
equation &= \mathbb{E}_{s\sim p(s)}[\mathbb{E}_{a\sim\pi_1(s,a)}\hat{R(s,a)} + \mathbb{E}_{a\sim\pi_0(s,a)}[\frac{\pi_1(s,a)}{\hat{\pi_0}(s,a)}(R(s,a) - \hat{R(s,a)})]] \\
&= \mathbb{E}_{s\sim p(s)}[\mathbb{E}_{a\sim\pi_1(s,a)}\hat{R(s,a)} + \mathbb{E}_{a\sim\pi_1(s,a)}R(s,a) - \mathbb{E}_{a\sim\pi_1(s,a)}\hat{R(s,a)}] \\
&= \mathbb{E}_{\substack{s\sim p(s)\\ a\sim\pi_1(s,a)}} R(s,a)
\end{aligned}$$

2.

$$\begin{aligned}
equation &= \mathbb{E}_{s\sim p(s)}[\mathbb{E}_{a\sim\pi_1(s,a)}R(s,a) + \mathbb{E}_{a\sim\pi_0(s,a)}[\frac{\pi_1(s,a)}{\hat{\pi_0}(s,a)}(R(s,a) - R(s,a))]] \\
&= \mathbb{E}_{\substack{s\sim p(s)\\ a\sim\pi_1(s,a)}} R(s,a)
\end{aligned}$$

(2.e)

We should estimate what is easier to estimate.

So in these two situations, we estimate $\hat{\pi_0}(s,a)$ and $\hat{R(s,a)}$ respectively.

(3.)

$$
\begin{aligned}
\sum_{i=1}^{m} \|x^{(i)} - f_u(x^{(i)})\|_2^2 &= \sum_{i=1}^{m} \|x^{(i)} - uu^T x^{(i)}\|_2^2 \\
&= \sum_{i=1}^{m} (x^{(i)} - uu^T x^{(i)})^T (x^{(i)} - uu^T x^{(i)}) \\
&= \sum_{i=1}^{m} (x^{(i)T} x^{(i)} - 2x^{(i)T} uu^T x^{(i)} + x^{(i)T} uu^T x^{(i)}) \\
&= \sum_{i=1}^{m} (x^{(i)T} x^{(i)}) - \sum_{i=1}^{m} u^T x^{(i)} x^{(i)T} u
\end{aligned}
$$

Minimizing this is just maximizing

$$
\sum_{i=1}^{m} u^T x^{(i)} x^{(i)T} u = u^T (\sum_{i=1}^{m} x^{(i)} x^{(i)^T}) u
$$

This is the same objective for maximizing variance. The solution is given by the first eigenvector of the empirical covariance matrix.

(4.a)

In this case, $g'(w_j^T x^{(i)}) = \frac{1}{2\sqrt{\pi}} \exp(-\frac{1}{2}(w_j^T x^{(i)})^2)$. And thus

$$
\begin{aligned}
\nabla_W l(W) &= \sum_{i=1}^{m} \left(W^{-T} + \sum_{j=1}^{n} \nabla_W (-\frac{1}{2} \log 2\pi - \frac{1}{2} (w_j^T x^{(i)})^2 \right) \\
&= \sum_{i=1}^{m} \left(W^{-T} - \begin{bmatrix} \left(w_1^T x^{(i)}\right) \\ \left(w_2^T x^{(i)}\right) \\ \vdots \\ \left(w_n^T x^{(i)}\right) \end{bmatrix} * x^{(i)^T} \right) \\
&= \sum_{i=1}^{m} \left(W^{-T} - W x^{(i)} x^{(i)^T} \right) \\
&= mW^{-T} - WX^T X
\end{aligned}
$$

Setting this to zero, we get

$$
W^T W = \frac{1}{m} (X^T X)^{-1}
$$

We can easily see that, if W is a solution, let R be an arbitrary orthogonal, since $(RW)^T (RW) = W^T R^T R W = W^T W$

(4.b)

$$l(W) = \sum_{i=1}^{m} (\log |W| + \sum_{j=1}^{n} (-\log 2 - |w_j^T x^{(i)}|))$$

and

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

Where $l^{(i)}$ is an indicator vector:

$$l_j^{(i)} = \begin{cases} 1, & w_j^T x^{(i)} > 0 \\ -1, & w_j^T x^{(i)} \leq 0 \end{cases}$$

And we can simplify this to

$$\nabla_W l(W) = mW^{-T} - LX$$

where the i-th column of L is $l^{(i)}$. And the update rule for a single example:

$$W := W + \alpha(W^{-T} - lx^T)$$

where l is as defined above.

(5.a)

$$\begin{aligned} |B(V_1)(s) - B(V_2)(s)| &= \gamma | \max_{a \in A} \sum_{s' \in S} P_{sa}(s') V_1(s') - \max_{a \in A} \sum_{s' \in S} P_{sa}(s') V_2(s') | \\ &\leq \gamma \max_{a \in A} | \sum_{s' \in S} P_{sa}(s') V_1(s') - \sum_{s' \in S} P_{sa}(s') V_2(s') | \\ &= \gamma \max_{a \in A} | \sum_{s' \in S} P_{sa}(s')(V_1(s') - V_2(s')) | \\ &\leq \gamma \max_{a \in A} \sum_{s' \in S} P_{sa(s')} |V_1 - V_2| \\ &= \gamma |V_1 - V_2| \max_{a \in A} \sum_{s' \in S} P_{sa}(s') \\ &= \gamma |V_1 - V_2| \end{aligned}$$

Since this is true for any s, we get

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

(5.b)

Suppose that V_1 and V_2 are two fixed points, then

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

This is true only when $\|V_1 - V_2\|_\infty = 0$, or $V_1 = V_2$. So fixed point is unique.