# CS285 hw4

oliver song

October 2023

## 1

### 1.1

$$
\begin{aligned}
Q^\pi - \hat{Q}^\pi &= (I - \gamma P^\pi)^{-1} r - \hat{Q}^\pi \\
&= (I - \gamma P^\pi)^{-1} r - (I - \gamma P^\pi)^{-1}(I - \gamma P^\pi)\hat{Q}^\pi \\
&= (I - \gamma P^\pi)^{-1}[\hat{Q}^\pi (I - \gamma \hat{P}^\pi) - \hat{Q}^\pi (I - \gamma P^\pi)] \\
&= \gamma (I - \gamma P^\pi)^{-1}(P^\pi - \hat{P}^\pi)\hat{Q}^\pi \\
&= \gamma (I - \gamma P^\pi)^{-1}(P - \hat{P})\hat{V}^\pi
\end{aligned} \tag{1}
$$

### 1.2

We can acquire from the previous question that

$$(P - \hat{P})\hat{V}^\pi = \frac{1}{\gamma}(Q^\pi - \hat{Q}^\pi)(I - \gamma P^\pi)$$

, and we can further have

$$
\begin{aligned}
||(P - \hat{P})\hat{V}^\pi||_\infty &= ||\frac{1}{\gamma}(Q^\pi - \hat{Q}^\pi)(I - \gamma P^\pi)||_\infty \\
&\leq \frac{1-\gamma}{\gamma}||\frac{\gamma}{(1-\gamma)^2} c\sqrt{\frac{|S|\log 1/\delta}{N}}||_\infty \\
&= \frac{1}{1-\gamma}\sqrt{\frac{2|S|\log(2|S||A|/\delta)}{N}}
\end{aligned} \tag{2}
$$

So term 1 holds and term 2 does not hold. Similarly, term 3 holds and term 4 doesn't hold.