B Refinement on Lemma 3.1 and Lemma 3.2 in the manuscript

We will provide the proof of our refined versions of Lemma 3.1 and Lemma 3.2 in the original manuscript. We have colored the part where the statement of the lemma and proof differs from that of the original manuscript.

Recall that the modified projected Bellman equation, equation (9) in the original manuscript is

$$(X^\top DX + \eta I)\theta - \gamma X^\top DP\Pi_{X\theta}X\theta = X^\top DR. \tag{1}$$

To proceed, let us define the following matrix:

$$\Gamma_\eta := X(X^T DX + \eta I)^{-1} X^T D.$$

The above matrix can be viewed as a modified weighted Euclidean projection matrix, and will be used to simplify the bounds in Lemma 3.1 and Lemma 3.2 in the original manuscript.

Now, we provide the proof of refined versions of Lemma 3.1 and Lemma 3.2 in the original manuscript as follows:

Lemma 3.1. *Suppose $\gamma||\Gamma_\eta||_\infty < 1$ holds. Then the solution of modified projected Bellman equation in (1) exists and is unique.*

Proof. To show the existence and uniqueness of the solution of (1), we use Banach fixed-point theorem. First, consider (1), it is enough show the existence and uniqueness of the solution of the following equation:

$$y = \Gamma_\eta(R + \gamma P\Pi_y y), \quad y \in \mathbb{R}^h. \tag{2}$$

This is because y satisfying the above equation is in the image of X, and X is a full-column rank matrix. To this end, we will apply Banach fixed point theorem:

$$\begin{aligned}
||y_1 - y_2||_\infty &= ||X(X^T DX + \eta I)^{-1}(\gamma X^T DP\Pi_{y_1}y_1 - \gamma X^T DP\Pi_{y_2}y_2)||_\infty \\
&\leq \gamma||X(X^T DX + \eta I)^{-1}X^\top D||_\infty ||\Pi_{y_1}y_1 - \Pi_{y_2}y_2||_\infty \\
&\leq \gamma||X(X^T DX + \eta I)^{-1}X^\top D||_\infty ||y_1 - y_2||_\infty \\
&\leq \gamma||y_1 - y_2||_\infty.
\end{aligned}$$

The second inequality follows from the non-expansiveness property of the max-operator. Now we can use Banach fixed-point theorem to conclude existence and uniqueness of (2). This completes the proof.

□

Now, we list the cases when the condition $\gamma||\Gamma_\eta||_\infty < 1$ should hold:

(A) The condition $\gamma||X^\top||_\infty||X||_\infty + ||X^\top DX||_\infty < \eta$ implies $\gamma||\Gamma_\eta||_\infty < 1$. The proof is in Lemma C.3.

(B) Suppose $\gamma||\Gamma||_\infty < 1$ so that the solution point of original projected Bellman equation exists and is unique. Then, consider the condition $0 \leq \eta < \frac{(1-\gamma||\Gamma||_\infty)||(X^T DX)^{-1}||_\infty^{-1}}{\gamma||(X^T DX)^{-1}||_\infty||X||_\infty||X^T D||_\infty + (1-\gamma||\Gamma||_\infty)}$ implies $\gamma||\Gamma_\eta||_\infty < 1$. The proof is in Lemma C.4.

(C) There are examples when $\gamma||\Gamma_\eta||_\infty < 1$ for all $\eta > 0$.

Now, we give remarks on the case (A), (B) and (C).

From (A), we can guarantee the existence and uniqueness of the solution with appropriate scaling of the feature matrix X. For example, if $\max(||X||_\infty, ||X^\top||_\infty) < 1$, it is enough to choose $\eta > 2$ to meet the condition in Lemma 3.1.. It is worth noting that scaling the values of the feature matrix is a commonly employed technique in the both theoretical literature or in practice. In sum, the refined bound tells us that η does not need to be too large, and one can easily select η in practice.

For (B), we note that the condition, $\gamma||\Gamma||_\infty < 1$ of (B) is used to derive the known error bounds for the solution of original projected Bellman equation, which is provided in [1]. Therefore, without any special conditions, the error bound in the above becomes small for neighborhood of $\eta = 0$, which will be discussed along with Lemma 3.2.

For (C), we will provide a simple example when the condition $\gamma||\Gamma_\eta||_\infty < 1$ holds for all $\eta > 0$: Suppose the feature vector satisfies $X^\top DX = aI$ for a positive real number a such that $a|S||A| \geq 1$. Moreover, assume that $||X||_2 \leq 1$ and the sampling distribution is uniform random distribution. Then, we can easily check that $||\Gamma_\eta||_\infty \leq \frac{1}{|SA|}\frac{1}{a+\eta} < 1$ holds for all $\eta > 0$.

Lemma 3.2. *Suppose $\gamma||\Gamma_\eta||_\infty < 1$. Then, we have*

$$||X\theta_e - Q^\star||_\infty \leq \frac{1}{1-\gamma||\Gamma_\eta||_\infty}||\Gamma_\eta Q^\star - Q^\star||_\infty.$$

Proof. The error bound of the solution can be obtained using simple algebraic inequalities.

$$\begin{aligned}
||X\theta_e - Q^*||_\infty \leq & \|\Gamma_\eta \mathcal{T}(X\theta_e) - \Gamma Q^*\|_\infty + \|\Gamma_\eta Q^* - Q^*\|_\infty \\
\leq & \|\Gamma_\eta\|_\infty \|\mathcal{T}(X\theta_e) - Q^*\|_\infty + \|\Gamma_\eta Q^* - Q^*\|_\infty \\
= & \|\Gamma_\eta\|_\infty \|\mathcal{T}(X\theta_e) - \mathcal{T}(Q^*)\|_\infty + \|\Gamma_\eta Q^* - Q^*\|_\infty \\
\leq & \gamma \|\Gamma_\eta\|_\infty \|X\theta_e - Q^*\|_\infty + \|\Gamma_\eta Q^* - Q^*\|_\infty .
\end{aligned}$$

The first inequality follows from triangle inequality. The first equality follows from the fact that Q^* is the solution of optimal Bellman equation. The last inequality follows from the contraction property of the Bellman operator.

Noting that $\gamma \|\Gamma_\eta\|_\infty < 1$, we get

$$\|X\theta_e - Q^*\|_\infty \leq \frac{1}{1-\gamma\|\Gamma_\eta\|_\infty} \|\Gamma_\eta Q^* - Q^*\|_\infty .$$

This completes the proof. □

Now, we provide discussion on the bound in Lemma 3.2. We consider two extreme cases when $\eta \to 0$ and $\eta \to \infty$ indicating the tightness of our bound. Moreover, we discuss the case when the error bound goes to zero:

1) $\eta \to 0$ (valid for both (B) and (C)): Since $\Gamma_\eta \to \Gamma$, we exactly recover the error bound by fixed point of standard projected Bellman equation ($\eta = 0$), which is $\frac{||\Gamma Q^\star - Q^\star||_\infty}{1-\gamma||\Gamma||_\infty}$ provided in [1]. Thus, our bound in Lemma 3.2 is tight when $\eta \to 0$.

2) $\eta \to \infty$: As η gets larger, by Lemma 3.1 and (A), solution exists and is unique. Noting that $\Gamma_\eta \to 0$, we have $||X\theta_e - Q^\star||_\infty \leq ||Q^\star||_\infty$. Considering that $\theta_e \to 0$ as $\eta \to \infty$ (the proof is given in Lemma D.1) we should have $||X \cdot 0 - Q^\star||_\infty = ||Q^\star||_\infty$. Thus, our bound in Lemma 3.2 is tight when $\eta \to \infty$.

3) The error bound is close to zero (for (B) and (C)): An upper bound on Lemma 3.2 can be obtained by simple algebraic manipulation:

$$\begin{aligned}
||X\theta_e - Q^\star||_\infty \leq & \frac{1}{1-\gamma||\Gamma_\eta||_\infty}||\Gamma_\eta Q^\star - Q^\star||_\infty \\
\leq & \underbrace{\frac{1}{1-\gamma||\Gamma_\eta||_\infty}||\Gamma_\eta Q^\star - \Gamma Q^\star||_\infty}_{\text{(T1)}} + \underbrace{\frac{1}{1-\gamma||\Gamma_\eta||_\infty}||\Gamma Q^\star - Q^\star||_\infty}_{\text{(T2)}} .
\end{aligned}$$

Suppose that the features are well designed such that (T2) in (1) will be small. For example if, $Q^\star$ is in the range space of X, then the error term in (T2) vanishes. Moreover, we can make (T1) arbitrarily small as follows: as $\eta \to 0$, we have $||\Gamma_\eta - \Gamma|| \to 0$ while $1 - \gamma||\Gamma_\eta||_\infty > 0$.

This yields (T1) in (1) to be sufficiently small. In the end, we will have $||X\theta_e - Q^\star|| \le \epsilon$ for sufficiently small $\epsilon \ge 0$.

4) When the projected Bellman equation does not admit a fixed point around $\eta = 0$ (for (A)): In this case, we should always choose $\eta > 0$ greater than a certain number, and (T1) cannot be entirely vanished, while (T2) can be arbitrarily close to zero when Q^* is close to the range space of X. The error in (T1) cannot be overcame because it can be seen as a fundamental error caused by having a fixed point of the original projected Bellman equation that does not have a solution. However, (T1) can be still small enough in many cases when $||\Gamma - \Gamma_\eta||$ is small.

C Technical lemmas

Lemma C.1. *For $M \in \mathbb{R}^{n\times n}$, if $||M|| < 1$ for any matrix norm, we have*

$$\left\|(I-M)^{-1}\right\| \le \frac{1}{1-||M||}.$$

Lemma C.2. *Suppose that $||X^\top DX||_\infty < \eta$. Then, we have*

$$\left\|(X^\top DX + \eta I)^{-1}\right\|_\infty \le \frac{1}{\eta - ||X^\top DX||_\infty}.$$

Proof. We have

$$\begin{aligned}
\left\|(X^\top DX + \eta I)^{-1}\right\|_\infty =& \left\|\frac{1}{\eta}\left(\frac{1}{\eta}X^\top DX + I\right)^{-1}\right\|_\infty \\
=& \frac{1}{\eta}\left\|\left(\frac{1}{\eta}X^\top DX + I\right)^{-1}\right\|_\infty \\
\le& \frac{1}{\eta}\frac{1}{1-\left\|\frac{1}{\eta}X^\top DX\right\|_\infty} \\
=& \frac{1}{\eta - \|X^\top DX\|_\infty}.
\end{aligned}$$

The first inequality follows from Lemma C.1. This completes the proof. □

Lemma C.3. *For $\eta > \gamma||X^\top D||_\infty ||X||_\infty + ||X^\top DX||_\infty$, we have*

$$\gamma\left\|\Gamma_\eta\right\|_\infty < 1.$$

Proof. From the definition of Γ_η, we have

$$\begin{aligned}
\gamma\left\|\Gamma_\eta\right\|_\infty =& \gamma\left\|X^\top D(X^\top DX + \eta I)^{-1}X\right\|_\infty \\
\le& \gamma\left\|X^\top D\right\|_\infty \|X\|_\infty \frac{1}{\eta - \|X^\top DX\|_\infty} \\
<& 1.
\end{aligned}$$

The first inequality follows from Lemma C.3. The last inequality follows from the condition $\eta > \gamma||X^\top D||_\infty ||X||_\infty + ||X^\top DX||_\infty$. This completes the proof. □

Lemma C.4. *Suppose that $\gamma\|\Gamma\|_\infty < 1$ so that the unregularized PBE ($\eta = 0$) admits a unique fixed point. Then, $\gamma\|\Gamma_\eta\|_\infty < 1$ for all $0 \le \eta < \frac{(1-\gamma\|\Gamma\|_\infty)\left\|(X^T DX)^{-1}\right\|_\infty^{-1}}{\gamma\|(X^T DX)^{-1}\|_\infty \|X\|_\infty \|X^T D\|_\infty + (1-\gamma\|\Gamma\|_\infty)}$.*

Proof. For the proof, suppose that $\gamma\|\Gamma\|_\infty < 1$. If the condition $0 \leq \eta < \frac{(1-\gamma\|\Gamma\|_\infty)\left\|(X^TDX)^{-1}\right\|_\infty^{-1}}{\gamma\|(X^TDX)^{-1}\|_\infty\|X\|_\infty\|X^TD\|_\infty+(1-\gamma\|\Gamma\|_\infty)}$ holds, it ensures $\left\|\eta(X^TDX)^{-1}\right\|_\infty < 1$ since $\frac{(1-\gamma\|\Gamma\|_\infty)}{\gamma\|(X^TDX)^{-1}\|_\infty\|X\|_\infty\|X^TD\|_\infty+(1-\gamma\|\Gamma\|_\infty)} < 1$. Then, using Gelfand's formula, we can easily prove that the spectral radius of $\eta(X^TDX)^{-1}$ is less than one. Next, note that for any two square matrices A and B, $(A-B)^{-1} = \sum_{i=0}^\infty (A^{-1}B)^i A^{-1}$ if the spectral radius of $A^{-1}B$ is less than one. Using this fact, one has

$$\begin{aligned}
\gamma\left\|\Gamma_\eta\right\|_\infty =& \left\|\gamma X(X^TDX+\eta I)^{-1}X^TD\right\|_\infty \\
=&\gamma\left\|X\sum_{i=0}^\infty (\eta(X^TDX)^{-1})^i(X^TDX)^{-1}X^TD\right\|_\infty \\
\leq&\gamma\left\|X(X^TDX)^{-1}X^TD\right\|_\infty + \gamma\left\|(X^TDX)^{-1}\sum_{i=1}^\infty \eta^i X(X^TDX)^{-i}X^TD\right\|_\infty \\
\leq&\gamma\left\|X(X^TDX)^{-1}X^TD\right\|_\infty + \gamma\eta\left\|(X^TDX)^{-1}\right\|_\infty^2\|X\|_\infty\left\|X^TD\right\|_\infty\sum_{i=0}^\infty\left\|\eta(X^TDX)^{-1}\right\|_\infty^i \\
\leq&\gamma\|\Gamma\|_\infty + \frac{\gamma\eta\left\|(X^TDX)^{-1}\right\|_\infty^2\|X\|_\infty\left\|X^TD\right\|_\infty}{1-\eta\|(X^TDX)^{-1}\|_\infty},
\end{aligned}$$

where the second line uses the matrix inverse property. Therefore, $\gamma\left\|\Gamma_\eta\right\|_\infty < 1$ holds if

$$\gamma\|\Gamma\|_\infty + \frac{\gamma\eta\left\|(X^TDX)^{-1}\right\|_\infty^2\|X\|_\infty\left\|X^TD\right\|_\infty}{1-\eta\left\|(X^TDX)^{-1}\right\|_\infty} < 1.$$

Rearranging terms, one gets the desired conclusion. $\square$

D Characterization of θ_e when $\eta \to \infty$

To clarify the dependency on η, we will use the notation θ_η^* instead of θ_e, which is the solution of (1) given η.

Lemma D.1. *We have*

$$\lim_{\eta\to\infty}\theta_\eta^* = 0.$$

Proof. From Lemma 3.1 and B, if η is large enough, solution exists and unique. Moreover, from (1), we have

$$(I-\gamma\Gamma_\eta P\Pi_{X\theta_\eta^*})X\theta_\eta^* = \Gamma_\eta R$$

Since $||\gamma\Gamma_\eta P\Pi_{X\theta_\eta^*}||_\infty < 1$, $I-\Gamma_\eta P\Pi_{X\theta_\eta^*}$ is invertible, , and

$$\begin{aligned}
||X\theta_\eta^*||_\infty =&||(I-\gamma\Gamma_\eta P\Pi_{X\theta_\eta^*})^{-1}\Gamma_\eta R||_\infty \\
\leq&\frac{1}{1-\gamma||\Gamma_\eta P\Pi_{X\theta_\eta^*}||_\infty}\left\|\Gamma_\eta R\right\|_\infty \\
\leq&\frac{1}{1-\gamma||\Gamma_\eta||_\infty}\left\|\Gamma_\eta\right\|_\infty\|R\|_\infty.
\end{aligned}$$

As $\eta \to 0$, it is easy to check that $\left\|\Gamma_\eta\right\|_\infty \to 0$, which implies $\left\|X\theta_\eta^*\right\|_\infty \to 0$. Noting that X is full-column rank matrix, we have $\theta_\eta^* \to 0$.

$\square$

References

[1] Melo, Francisco S., and M. Isabel Ribeiro. "Q-learning with linear function approximation." International Conference on Computational Learning Theory. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.