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

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

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. *When $\eta > \gamma||X^\top||_\infty||X||_\infty + ||X^\top DX||_\infty$, a 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, we define the operator $\mathcal{T}_\eta$ as follows:

$$\mathcal{T}_\eta(\theta) := (X^T DX + \eta I)^{-1}(X^T DR + \gamma X^T DP\Pi_{X\theta}X\theta)$$

We show that $\mathcal{T}_\eta$ is contraction mapping. The existence and uniqueness of (1) follows from the Banach fixed-point theorem.

$$\begin{aligned}
||\theta_1 - \theta_2||_\infty &= ||(X^T DX + \eta I)^{-1}(\gamma X^T DP\Pi_{X\theta_1}X\theta_1 - \gamma X^T DP\Pi_{X\theta_2}X\theta_2)||_\infty \\
&\leq \gamma||(X^T DX + \eta I)^{-1}||_\infty||X^T||_\infty||\Pi_{X\theta_1}X\theta_1 - \Pi_{X\theta_2}X\theta_2||_\infty \\
&\leq \gamma||(X^T DX + \eta I)^{-1}||_\infty||X^T||_\infty||X\theta_1 - X\theta_2||_\infty \\
&\leq \gamma||(X^T DX + \eta I)^{-1}||_\infty||X^T||_\infty||X||_\infty||\theta_1 - \theta_2||_\infty \\
&\leq \gamma\frac{1}{\eta - \|X^\top DX\|_\infty}||X^T||_\infty||X||_\infty||\theta_1 - \theta_2||_\infty.
\end{aligned}$$

The first inequality follows from the sub-multiplicativity of matrix norm and $||DP||_\infty \leq 1$. The second inequality follows from the non-expansiveness property of the max-operator. We have $\gamma\frac{1}{\eta-\|X^\top DX\|_\infty}||X^T||_\infty||X||_\infty < 1$ thanks to the condition $\gamma\left\|X\right\|_\infty\left\|X^\top\right\|_\infty + \left\|X^\top DX\right\|_\infty < \eta$. The last inequality follows from Lemma C.3. Therefore, $\mathcal{T}_\eta$ is contraction mapping. Now we can use Banach fixed-point theorem to conclude existence and uniqueness of (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 3.2 in the original manuscript.

Lemma 3.2. *(a) Suppose that*

$$\gamma||X^\top||_\infty||X||_\infty + ||X^\top DX||_\infty < \eta.$$

Then, the solution of (1) exists and is unique, and the following conditions holds:

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

(b) Suppose that the solution exists, and $\gamma||\Gamma_\eta||_\infty < 1$ for all $\eta > 0$.

For both (a) and (b), we have

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

Proof. Let us first prove for the case (a). The error bound of the solution can be obtained using simple algebraic inequalities.

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

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.

Note that from Lemma C.2, we have $\gamma \left\|\Gamma_\eta\right\|_\infty < 1$. Therefore, we get

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

This finishes the proof for (a). The argument for the case (b) also holds since $\gamma \left\|\Gamma_\eta\right\|_\infty < 1$ for all η. This completes the proof. □

C Technical lemmas

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

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

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

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

Proof. We have

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

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

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

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. □

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. *Suppose $\eta > \gamma||X^\top||_\infty||X||_\infty + ||X^\top DX||_\infty$. Then, we have*

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

Proof. From (1), we have

$$\begin{aligned}
\left\|\theta^*_\eta\right\|_\infty =& \left\|(X^\top DX + \eta I)^{-1}(X^\top DR + \gamma X^\top DP\Pi_{X\theta^*_\eta} X\theta^*_\eta)\right\|_\infty \\
\leq& \frac{1}{\eta - \left\|X^\top DX\right\|_\infty} \left\|X^\top DR + \gamma X^\top DP\Pi_{X\theta^*_\eta} X\theta^*_\eta\right\|_\infty \\
\leq& \frac{1}{\eta - \left\|X^\top DX\right\|_\infty} \left\|X^\top DR\right\|_\infty + \frac{1}{\eta - \left\|X^\top DX\right\|_\infty} \left\|X^\top D\right\|_\infty \left\|X\right\|_\infty \left\|\theta^*_\eta\right\|_\infty .
\end{aligned}$$

The first inequality follows from Lemma C.3. Therefore, considering that $\eta > \left\|X^\top DX\right\|_\infty + \gamma \left\|X^\top D\right\|_\infty \left\|X\right\|_\infty$, we have

$$\frac{\eta - \left\|X^\top DX\right\|_\infty - \gamma \left\|X^\top D\right\|_\infty \left\|X\right\|_\infty}{\eta - ||X^\top DX||_\infty} \left\|\theta^*_\eta\right\|_\infty < \frac{1}{\eta - \left\|X^\top DX\right\|_\infty} \left\|X^\top DR\right\|_\infty ,$$

which leads to

$$\left\|\theta^*_\eta\right\|_\infty \leq \frac{1}{\eta - \left\|X^\top DX\right\|_\infty - \gamma \left\|X^\top D\right\|_\infty \left\|X\right\|_\infty} \left\|X^\top DR\right\|_\infty .$$

As $\eta \to 0$, the right-hand side of the above equation goes to zero, i.e., $\theta^*_\eta \to 0$ as $\eta \to 0$. $\square$