Adversarial Regression with Multiple Learners
Supplementary Material

Liang Tong   Sixie Yu   Scott Alfeld   Yevgeniy Vorobeychik

A. Proofs

A.1. Proof of Lemma 1

Proof. We derive the best response of the attacker by using the first order condition. Let $\nabla_{X'} c_a(\{\boldsymbol{\theta}_i\}_{i=1}^n, \mathbf{X}')$ denote the gradient of c_a with respect to $\mathbf{X}'$. Then

$$\nabla_{X'} c_a = 2\sum_{i=1}^{n}(\mathbf{X}'\boldsymbol{\theta}_i - \mathbf{z})\boldsymbol{\theta}_i^\top + 2\lambda(\mathbf{X}' - X).$$

Due to convexity of c_a, let $\nabla_{X'} c_a = \mathbf{0}$, we have

$$\mathbf{X}^* = (\lambda\mathbf{X} + \mathbf{z}\sum_{i=1}^{n}\boldsymbol{\theta}_i^\top)(\lambda\mathbf{I} + \sum_{i=1}^{n}\boldsymbol{\theta}_i\boldsymbol{\theta}_i^\top)^{-1}.$$

□

A.2. Proof of Lemma 2

Proof. 1. First, we prove that $\mathbf{A}_n = \lambda\mathbf{I} + \sum_{i=1}^n \boldsymbol{\theta}_i\boldsymbol{\theta}_i^\top$ is invertible, and its inverse matrix, $\mathbf{A}_n^{-1}$, is positive definite by using mathematical induction.

When $n = 1$, $\mathbf{A}_1 = \lambda\mathbf{I} + \boldsymbol{\theta}_1\boldsymbol{\theta}_1^\top$. As $\lambda\mathbf{I}$ is an invertible square matrix and $\boldsymbol{\theta}_1$ is a column vector, by using *Sherman-Morrison formula*, $\mathbf{A}_1$ is invertible.

$$\mathbf{A}_1^{-1} = \frac{1}{\lambda}(\mathbf{I} - \frac{\boldsymbol{\theta}_1\boldsymbol{\theta}_1^\top}{\lambda + \boldsymbol{\theta}_1^\top\boldsymbol{\theta}_1}).$$

For any non-zero column vector $\mathbf{u}$, we have

$$\mathbf{u}^\top\mathbf{A}_1^{-1}\mathbf{u} = \frac{\lambda\mathbf{u}^\top\mathbf{u} + \mathbf{u}^\top\mathbf{u}\boldsymbol{\theta}_1^\top\boldsymbol{\theta}_1 - \mathbf{u}^\top\boldsymbol{\theta}_1\boldsymbol{\theta}_1^\top\mathbf{u}}{\lambda(\lambda + \boldsymbol{\theta}_1^\top\boldsymbol{\theta}_1)}.$$

As $\mathbf{u}^\top\mathbf{u} > 0$ and $\lambda > 0$, according to *Cauchy-Schwarz inequality*,

$$\mathbf{u}^\top\mathbf{u}\boldsymbol{\theta}_1^\top\boldsymbol{\theta}_1 \geq \mathbf{u}^\top\boldsymbol{\theta}_1\boldsymbol{\theta}_1^\top\mathbf{u},$$

Then, $\mathbf{u}^\top\mathbf{A}_1^{-1}\mathbf{u} > 0$. Thus, $\mathbf{A}_1^{-1}$ is a positive definite matrix.

We then assume that when $n = k(k \geq 1)$, $\mathbf{A}_k$ is invertible and $\mathbf{A}_k^{-1}$ is positive definite. Then, when $n = k + 1$,

$$\mathbf{A}_{k+1} = \mathbf{A}_k + \boldsymbol{\theta}_{k+1}\boldsymbol{\theta}_{k+1}^\top.$$

As $\mathbf{A}_k$ is invertible, $\boldsymbol{\theta}_{k+1}$ is a column vector. By using *Sherman-Morrison formula*, we have that $\mathbf{A}_{k+1}$ is invertible, and

$$\mathbf{A}_{k+1}^{-1} = \mathbf{A}_k^{-1} - \frac{\mathbf{A}_k^{-1}\boldsymbol{\theta}_{k+1}\boldsymbol{\theta}_{k+1}^\top\mathbf{A}_k^{-1}}{1 + \boldsymbol{\theta}_{k+1}^\top\mathbf{A}_k^{-1}\boldsymbol{\theta}_{k+1}}.$$

Then,

$$\mathbf{u}^\top \mathbf{A}_{k+1}^{-1}\mathbf{u} = \frac{\mathbf{u}^\top \mathbf{A}_k^{-1}\mathbf{u} + \mathbf{u}^\top \mathbf{A}_k^{-1}\mathbf{u} \cdot \boldsymbol{\theta}_{k+1}^\top \mathbf{A}_k^{-1}\boldsymbol{\theta}_{k+1} - \mathbf{u}^\top \mathbf{A}_k^{-1}\boldsymbol{\theta}_{k+1} \cdot \boldsymbol{\theta}_{k+1}^\top \mathbf{A}_k^{-1}\mathbf{u}}{1 + \boldsymbol{\theta}_{k+1}^\top \mathbf{A}_k^{-1}\boldsymbol{\theta}_{k+1}}$$

As $\mathbf{A}_k^{-1}$ is a positive definite matrix, we have $\mathbf{u}^\top \mathbf{A}_k^{-1}\mathbf{u} > 0$ and $\boldsymbol{\theta}_{k+1}^\top \mathbf{A}_k^{-1}\boldsymbol{\theta}_{k+1} > 0$. By using *Extended Cauchy-Schwarz inequality*, we have

$$\mathbf{u}^\top \mathbf{A}_k^{-1}\mathbf{u}\boldsymbol{\theta}_{k+1}^\top \mathbf{A}_k^{-1}\boldsymbol{\theta}_{k+1} > \mathbf{u}^\top \mathbf{A}_k^{-1}\boldsymbol{\theta}_{k+1}\boldsymbol{\theta}_{k+1}^\top \mathbf{A}_k^{-1}\mathbf{u}.$$

Then, $\mathbf{A}_{k+1}^{-1}$ is positive definite. Hence, $\mathbf{A}_n = \lambda \mathbf{I} + \sum_{i=1}^n \boldsymbol{\theta}_i\boldsymbol{\theta}_i^\top$ is invertible, and $\mathbf{A}_n^{-1}$ is positive definite. Similarly, we can prove that $\mathbf{A}_{-i}$ is invertible, and $\mathbf{A}_{-i}^{-1}$ is positive definite.

2. We have proved that $\mathbf{A}_n$ and $\mathbf{A}_{-i}$ are invertible. Then, the result can be obtained by using *Sherman-Morrison formula*.

3. Let $\mathbf{A}_{-i,-j} = \mathbf{A}_{-i} - \boldsymbol{\theta}_j\boldsymbol{\theta}_j^\top$. As $\mathbf{A}_{-i,-j}$ is a symmetric matrix, its inverse, $\mathbf{A}_{-i,-j}^{-1}$ is also symmetric. Using a similar approach to the one above, we can prove that $\mathbf{A}_{-i,-j}$ is invertible and $\mathbf{A}_{-i,-j}^{-1}$ is positive definite. By using *Sherman-Morrison formula*, we have

$$\mathbf{A}_{-i}^{-1} = \mathbf{A}_{-i,-j}^{-1} - \frac{\mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_j\boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}}{1 + \boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_j}.$$

Then,

$$\begin{aligned}\boldsymbol{\theta}_i^\top \mathbf{A}_{-i}^{-1}\boldsymbol{\theta}_i &= \boldsymbol{\theta}_i^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i - \frac{\boldsymbol{\theta}_i^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_j \cdot \boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i}{1 + \boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_j} \\ &= \boldsymbol{\theta}_i^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i - \frac{(\boldsymbol{\theta}_i^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_j)^2}{1 + \boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_j} \\ &\le \boldsymbol{\theta}_i^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i.\end{aligned}$$

We then iteratively apply *Sherman-Morrison formula* and get

$$\begin{aligned}\boldsymbol{\theta}_i^\top \mathbf{A}_{-i}^{-1}\boldsymbol{\theta}_i &\le \boldsymbol{\theta}_i^\top \mathbf{A}_0^{-1}\boldsymbol{\theta}_i \\ &= \frac{1}{\lambda}\boldsymbol{\theta}_i^\top \boldsymbol{\theta}_i.\end{aligned}$$

□

A.3. Proof of Theorem 2

Proof. As presented in Lemma 3, we have

$$\ell(\mathbf{X}^*\boldsymbol{\theta}_i, \mathbf{y}) \le \ell(\mathbf{B}_{-i}\mathbf{A}_{-i}^{-1}\boldsymbol{\theta}_i, \mathbf{y}) + \frac{1}{\lambda^2}||\mathbf{z} - \mathbf{y}||_2^2(\boldsymbol{\theta}_i^\top \boldsymbol{\theta}_i)^2.$$

By using *Sherman-Morrison formula*,

$$\begin{aligned}\ell(\mathbf{B}_{-i}\mathbf{A}_{-i}^{-1}\boldsymbol{\theta}_i, \mathbf{y}) &= ||\mathbf{B}_{-i}(\mathbf{A}_{-i,-j}^{-1} - \frac{\mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_j\boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}}{1 + \boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_j})\boldsymbol{\theta}_i - \mathbf{y}||_2^2 \\ &\le ||\frac{\mathbf{B}_{-i}\mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i}{1 + \boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_j} - \mathbf{y}||_2^2 + \triangle_1(\boldsymbol{\theta})\end{aligned}$$

where $j \neq i$, and $\triangle_1(\boldsymbol{\theta})$ is a continuous function of $\boldsymbol{\theta} = \{\boldsymbol{\theta}_i\}_{i=1}^n$. As the action space $\boldsymbol{\Theta}$ is bounded, then $0 \leq \triangle_1(\boldsymbol{\theta}) < \infty$. Hence, we have

$$\begin{aligned}
\ell(\mathbf{B}_{-i}\mathbf{A}_{-i}^{-1}\boldsymbol{\theta}_i, \mathbf{y}) &\leq || \frac{\mathbf{B}_{-i}\mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i}{1 + \boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_j} - \mathbf{y}||_2^2 + \triangle_1(\boldsymbol{\theta}) \\
&= || \frac{\mathbf{B}_{-i}\mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i - \mathbf{y} - \boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_j \mathbf{y}}{1 + \boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_j} ||_2^2 + \triangle_1(\boldsymbol{\theta}) \\
&\leq ||\mathbf{B}_{-i}\mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i - \mathbf{y} - \boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_j \mathbf{y}||_2^2 + \triangle_1(\boldsymbol{\theta}) \\
&= ||(\mathbf{B}_{-i,-j} + \mathbf{z}\boldsymbol{\theta}_j^\top)\mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i - \mathbf{y} - \boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_j \mathbf{y}||_2^2 + \triangle_1(\boldsymbol{\theta}) \\
&= ||(\mathbf{B}_{-i,-j}\mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i - \mathbf{y}) + (\mathbf{z} - \mathbf{y})\boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i + \boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^\top(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j)\mathbf{y}||_2^2 + \triangle_1(\boldsymbol{\theta}) \\
&\leq \ell(\mathbf{B}_{-i,-j}\mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i, \mathbf{y}) + ||(\mathbf{z} - \mathbf{y})||_2^2(\boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i)^2 + \triangle_2(\boldsymbol{\theta})
\end{aligned}$$

where $\triangle_2(\boldsymbol{\theta})$ is a continuous function of $\boldsymbol{\theta}$ and $0 \leq \triangle_2(\boldsymbol{\theta}) < \infty$. Let $\mathbf{A}_{-i,-j,-k} = \mathbf{A}_{-i,-j} - \boldsymbol{\theta}_k\boldsymbol{\theta}_k^\top$, then, similarly, $(\boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i)^2$ can be further relaxed as follows.

$$\begin{aligned}
(\boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i)^2 &= (\boldsymbol{\theta}_j^\top (\mathbf{A}_{-i,-j,-k}^{-1} - \frac{\mathbf{A}_{-i,-j,-k}^{-1}\boldsymbol{\theta}_k\boldsymbol{\theta}_k^\top \mathbf{A}_{-i,-j,-k}^{-1}}{1 + \boldsymbol{\theta}_k^\top \mathbf{A}_{-i,-j,-k}^{-1}\boldsymbol{\theta}_k})\boldsymbol{\theta}_i)^2 \\
&\leq (\boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j,-k}^{-1}\boldsymbol{\theta}_i)^2 + \triangle_3(\boldsymbol{\theta})
\end{aligned}$$

where $0 \leq \triangle_3(\boldsymbol{\theta}) < \infty$, using the same approach,$(\boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i)^2$ can be further and iteratively relaxed as follows,

$$\begin{aligned}
(\boldsymbol{\theta}_j^\top \mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i)^2 &\leq (\boldsymbol{\theta}_j^\top \mathbf{A}_0^{-1}\boldsymbol{\theta}_i)^2 + \triangle_4(\boldsymbol{\theta}) \\
&= \frac{1}{\lambda^2}(\boldsymbol{\theta}_j^\top \boldsymbol{\theta}_i)^2 + \triangle_4(\boldsymbol{\theta})
\end{aligned}$$

where $0 \leq \triangle_4(\boldsymbol{\theta}) < \infty$. Combining the results above, we can iteratively relax $\ell(\mathbf{B}_{-i}\mathbf{A}_{-i}^{-1}\boldsymbol{\theta}_i, \mathbf{y})$ as follows,

$$\begin{aligned}
\ell(\mathbf{B}_{-i}\mathbf{A}_{-i}^{-1}\boldsymbol{\theta}_i, \mathbf{y}) &\leq \ell(\mathbf{B}_{-i,-j}\mathbf{A}_{-i,-j}^{-1}\boldsymbol{\theta}_i, \mathbf{y}) + \frac{1}{\lambda^2}||\mathbf{z} - \mathbf{y}||_2^2(\boldsymbol{\theta}_j^\top \boldsymbol{\theta}_i)^2 + \triangle_5(\boldsymbol{\theta}) \\
&\leq \ell(\mathbf{X}\boldsymbol{\theta}_i, \mathbf{y}) + \frac{1}{\lambda^2}||\mathbf{z} - \mathbf{y}||_2^2 \sum_{j \neq i}(\boldsymbol{\theta}_j^\top \boldsymbol{\theta}_i)^2 + \triangle(\boldsymbol{\theta})
\end{aligned}$$

where $0 \leq \triangle_5(\boldsymbol{\theta}) < \infty$ and $0 \leq \triangle(\boldsymbol{\theta}) < \infty$. Then,

$$\begin{aligned}
\ell(\mathbf{X}^*\boldsymbol{\theta}_i, \mathbf{y}) &\leq \ell(\mathbf{B}_{-i}\mathbf{A}_{-i}^{-1}\boldsymbol{\theta}_i, \mathbf{y}) + \frac{1}{\lambda^2}||\mathbf{z} - \mathbf{y}||_2^2(\boldsymbol{\theta}_i^\top \boldsymbol{\theta}_i)^2 \\
&\leq \ell(\mathbf{X}\boldsymbol{\theta}_i, \mathbf{y}) + \frac{1}{\lambda^2}||\mathbf{z} - \mathbf{y}||_2^2 \sum_{j=1}^{n}(\boldsymbol{\theta}_j^\top \boldsymbol{\theta}_i)^2 + \triangle(\boldsymbol{\theta}).
\end{aligned}$$

Hence,

$$\begin{aligned}
c_i(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{-i}) &= \beta\ell(\mathbf{X}^*\boldsymbol{\theta}_i, \mathbf{y}) + (1 - \beta)\ell(\mathbf{X}\boldsymbol{\theta}_i, \mathbf{y}) \\
&\leq \ell(\mathbf{X}\boldsymbol{\theta}_i, \mathbf{y}) + \frac{\beta}{\lambda^2}||\mathbf{z} - \mathbf{y}||_2^2 \sum_{j=1}^{n}(\boldsymbol{\theta}_j^\top \boldsymbol{\theta}_i)^2 + \epsilon
\end{aligned}$$

where ϵ is a constant such that $\epsilon = \beta * \max_{\boldsymbol{\theta}}\{\triangle(\boldsymbol{\theta})\} < \infty$. □

A.4. Proof of Theorem 4

Proof. We have known that $\langle \mathcal{N}, \boldsymbol{\Theta}, (\widetilde{c}_i) \rangle$ has at least NE, and each learner has an nonempty, compact and convex action space $\boldsymbol{\Theta}$. Hence, we can apply Theorem 2 and Theorem 6 of Rosen (1965). That is, for some fixed $\{r_i\}_i^n(0 < r_i <$

$1, \sum_{i=1}^{n} r_i = 1$), if the matrix in Eq. (1) is positive definite, then $\langle \mathcal{N}, \boldsymbol{\Theta}, (\widetilde{c}_i) \rangle$ has a unique NE.

$$Jr(\boldsymbol{\theta}) = \begin{bmatrix} r_1 \nabla_{\boldsymbol{\theta}_1,\boldsymbol{\theta}_1} \widetilde{c}_1(\boldsymbol{\theta}) & \dots & r_1 \nabla_{\boldsymbol{\theta}_1,\boldsymbol{\theta}_n} \widetilde{c}_1(\boldsymbol{\theta}) \\ \vdots & & \vdots \\ r_n \nabla_{\boldsymbol{\theta}_n,\boldsymbol{\theta}_1} \widetilde{c}_n(\boldsymbol{\theta}) & \dots & r_n \nabla_{\boldsymbol{\theta}_n,\boldsymbol{\theta}_n} \widetilde{c}_n(\boldsymbol{\theta}) \end{bmatrix} \tag{1}$$

By taking second-order derivatives, we have

$$\nabla_{\boldsymbol{\theta}_i,\boldsymbol{\theta}_i} \widetilde{c}_i(\boldsymbol{\theta}) = 2\mathbf{X}^\top \mathbf{X} + \frac{2\beta||\mathbf{z}-\mathbf{y}||_2^2}{\lambda^2}(4\boldsymbol{\theta}_i\boldsymbol{\theta}_i^\top + 2\boldsymbol{\theta}_i^\top\boldsymbol{\theta}_i\mathbf{I} + \sum_{j\neq i}\boldsymbol{\theta}_j\boldsymbol{\theta}_j^\top)$$

and

$$\nabla_{\boldsymbol{\theta}_i,\boldsymbol{\theta}_j} \widetilde{c}_i(\boldsymbol{\theta}) = \frac{2\beta||\mathbf{z}-\mathbf{y}||_2^2}{\lambda^2}(\boldsymbol{\theta}_i^\top\boldsymbol{\theta}_j\mathbf{I} + \boldsymbol{\theta}_j\boldsymbol{\theta}_i^\top)$$

We first let $r_1 = r_2 = ... = r_n = \frac{1}{n}$ and decompose $Jr(\boldsymbol{\theta})$ as follows,

$$Jr(\boldsymbol{\theta}) = \frac{2}{n}\mathbf{P} + \frac{2\beta||\mathbf{z}-\mathbf{y}||_2^2}{\lambda^2 n}(\mathbf{Q}+\mathbf{S}+\mathbf{T}), \tag{2}$$

where $\mathbf{P}$ and $\mathbf{Q}$ are *block diagonal matrices* such that $\mathbf{P}_{ii} = \mathbf{X}^\top\mathbf{X}$, $\mathbf{P}_{ij} = \mathbf{0}$, $\mathbf{Q}_{ii} = 4\boldsymbol{\theta}_i\boldsymbol{\theta}_i^\top + \boldsymbol{\theta}_i^\top\boldsymbol{\theta}_i\mathbf{I}$ and $\mathbf{Q}_{ij} = \mathbf{0}$, $\forall i, j \in \mathcal{N}, j \neq i$. $\mathbf{S}$ and $\mathbf{T}$ are *block symmetric matrices* such that $\mathbf{S}_{ii} = \boldsymbol{\theta}_i^\top\boldsymbol{\theta}_i\mathbf{I}$, $\mathbf{S}_{ij} = \boldsymbol{\theta}_i^\top\boldsymbol{\theta}_j\mathbf{I}$, $\mathbf{T}_{ii} = \sum_{j\neq i}\boldsymbol{\theta}_j\boldsymbol{\theta}_j^\top$ and $\mathbf{T}_{ij} = \boldsymbol{\theta}_j\boldsymbol{\theta}_i^\top, \forall i, j \in \mathcal{N}, j \neq i$.

Next, we prove that $\mathbf{P}$ is *positive definite*, and $\mathbf{Q}$, $\mathbf{S}$ and $\mathbf{T}$ are *positive semi-definite*. Let $\mathbf{u} = [\mathbf{u}_1^\top, ..., \mathbf{u}_n^\top]^\top$ be an $nd \times 1$ vector, where $\mathbf{u}_i \in \mathbb{R}^{d\times 1} (i \in \mathcal{N})$ are not all zero vectors.

1. $\mathbf{u}^\top\mathbf{P}\mathbf{u} = \sum_{i=1}^{n} \mathbf{u}_i^\top\mathbf{X}^\top\mathbf{X}\mathbf{u}_i = \sum_{i=1}^{n} ||\mathbf{X}\mathbf{u}_i||_2^2$. As the columns of $\mathbf{X}$ are linearly independent and $\mathbf{u}_i$ are not all zero vectors, there exists at least one $\mathbf{u}_i$ such that $\mathbf{X}\mathbf{u}_i \neq \mathbf{0}$. Hence, $\mathbf{u}^\top\mathbf{P}\mathbf{u} > 0$ which indicates that $\mathbf{P}$ is positive definite.
2. Similarly, $\mathbf{u}^\top\mathbf{Q}\mathbf{u} \geq 0$ which indicates that $\mathbf{Q}$ is a positive semi-definite matrix.
3. Let's $\mathbf{S}^* \in \mathbb{R}^{n\times n}$ be a symmetric matrix such that $\mathbf{S}^*_{ii} = \boldsymbol{\theta}_i^\top\boldsymbol{\theta}_i$ and $\mathbf{S}^*_{ij} = \boldsymbol{\theta}_i^\top\boldsymbol{\theta}_j, \forall i, j \in \mathcal{N}, j \neq i$. Hence, $\mathbf{S}_{ij} = \mathbf{S}^*_{ij}\mathbf{I}$, $\forall i, j \in \mathcal{N}$. Note that $\mathbf{S}^* = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, ..., \boldsymbol{\theta}_n]^\top[\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, ..., \boldsymbol{\theta}_n]$ is a positive semi-definite matrix, as it is also symmetric, there exists at least one lower triangular matrix $\mathbf{L}^* \in \mathbb{R}^{n\times n}$ with non-negative diagonal elements (Higham, 1990) such that

   $$\mathbf{S}^* = \mathbf{L}^*{\mathbf{L}^*}^\top \text{(Cholesky Decomposition)}$$

   Let $\mathbf{L}$ be a block matrix such that $\mathbf{L}_{ij} = \mathbf{L}^*_{ij}\mathbf{I}$, $\forall i, j \in \mathcal{N}$. Therefore, $(\mathbf{L}\mathbf{L}^\top)_{ij} = (\mathbf{L}^*{\mathbf{L}^*}^\top)_{ij}\mathbf{I} = \mathbf{S}^*_{ij}\mathbf{I} = \mathbf{S}_{ij}$ which indicates that $\mathbf{S} = \mathbf{L}\mathbf{L}^\top$ is a positive semi-definite matrix.
4. Since

   $$\begin{aligned} \mathbf{u}^\top\mathbf{T}\mathbf{u} &= \sum_{i=1}^{n}\sum_{j\neq i}(\mathbf{u}_i^\top\boldsymbol{\theta}_j)^2 + \sum_{i=1}^{n}\sum_{j\neq i}(\mathbf{u}_i^\top\boldsymbol{\theta}_j)(\mathbf{u}_j^\top\boldsymbol{\theta}_i) \\ &= \sum_{i=1}^{n}\sum_{j\neq i}[\frac{1}{2}(\mathbf{u}_i^\top\boldsymbol{\theta}_j)^2 + \frac{1}{2}(\mathbf{u}_j^\top\boldsymbol{\theta}_i)^2 + (\mathbf{u}_i^\top\boldsymbol{\theta}_j)(\mathbf{u}_j^\top\boldsymbol{\theta}_i)] \\ &= \frac{1}{2}\sum_{i=1}^{n}\sum_{j\neq i}(\mathbf{u}_i^\top\boldsymbol{\theta}_j + \mathbf{u}_j^\top\boldsymbol{\theta}_i)^2 \\ &\geq 0, \end{aligned}$$

   $\mathbf{T}$ is a positive semi-definite matrix.

Combining the results above, $Jr(\boldsymbol{\theta})$ is a positive definite matrix which indicates that $\langle \mathcal{N}, \boldsymbol{\Theta}, (\widetilde{c}_i) \rangle$ has a unique NE. As Theorem 3 points out, the game has at least one symmetric NE. Therefore, the NE is unique and must be symmetric. □

B. Experiment Results

B.1. Supplementary results for the redwine dataset

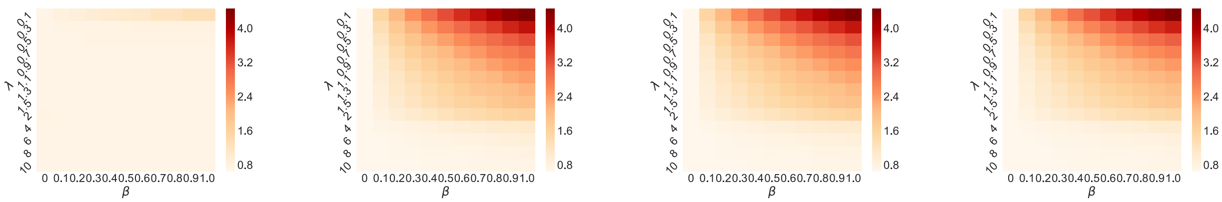

Figure 1. Overestimated $\mathbf{z}$, $\hat{\lambda} = 0.5$, $\hat{\beta} = 0.8$.The average RMSE across different values of actual λ and β on redwine dataset. From left to right: *MLSG*, *Lasso*, *Ridge*, *OLS*.

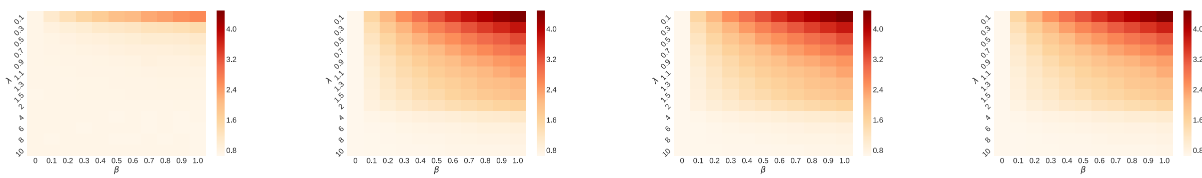

Figure 2. Overestimated $\mathbf{z}$, $\hat{\lambda} = 1.5$, $\hat{\beta} = 0.8$. The average RMSE across different values of actual λ and β on redwine dataset. From left to right: *MLSG*, *Lasso*, *Ridge*, *OLS*.

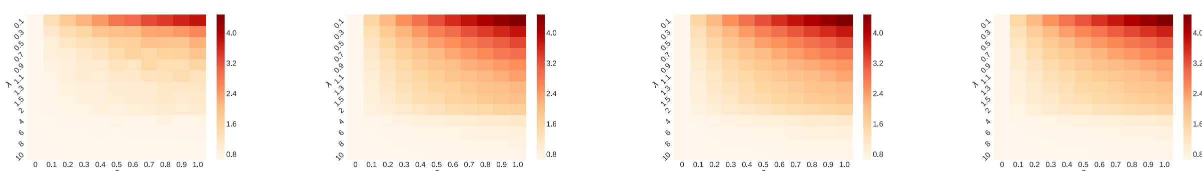

Figure 3. Underestimated $\mathbf{z}$, $\hat{\lambda} = 1.5$, $\hat{\beta} = 0.8$. The average RMSE across different values of actual λ and β on redwine dataset. From left to right: *MLSG*, *Lasso*, *Ridge*, *OLS*.

B.2. Supplementary results for the boston dataset

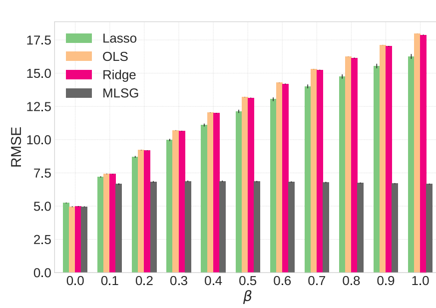

Figure 4. The defender knows λ, β, and $\mathbf{z}$. RMSE of $\mathbf{y}'$ and $\mathbf{y}$ on boston dataset. The defender knows λ, β, and $\mathbf{z}$.

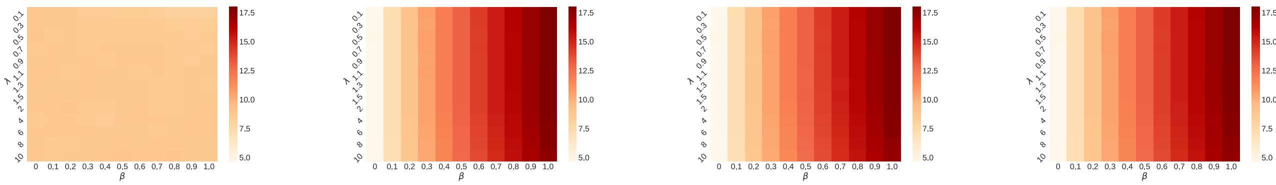

Figure 5. Overestimated **z**, $\hat{\lambda} = 0.3$, $\hat{\beta} = 0.8$. The average RMSE across different values of actual λ and β on boston dataset. From left to right: *MLSG*, *Lasso*, *Ridge*, *OLS*.

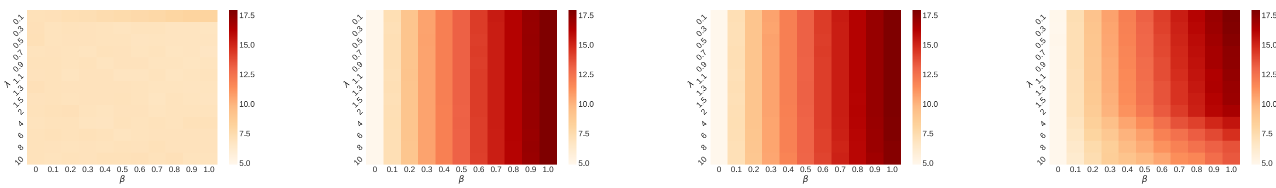

Figure 6. Underestimated **z**, $\hat{\lambda} = 0.3$, $\hat{\beta} = 0.8$. The average RMSE across different values of actual λ and β on boston dataset. From left to right: *MLSG*, *Lasso*, *Ridge*, *OLS*.

B.3. Supplementary results for the PDF dataset

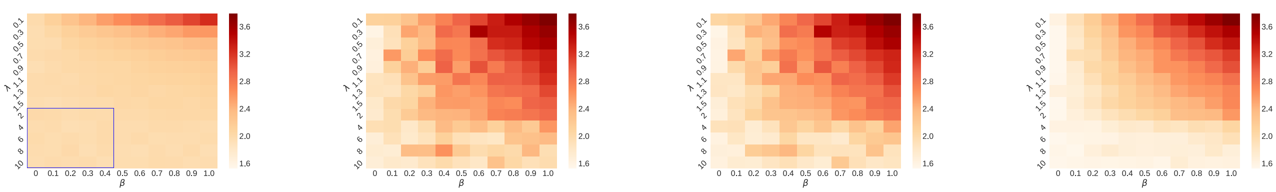

Figure 7. Overestimated **z**, $\hat{\lambda} = 1.5$, $\hat{\beta} = 0.5$. The average RMSE across different values of actual λ and β on PDF dataset. From left to right: *MLSG*, *Lasso*, *Ridge*, *OLS*.

References

Higham, N. J. Analysis of the cholesky decomposition of a semi-definite matrix. In *Reliable Numerical Computation*, pp. 161–185. University Press, 1990.

Rosen, J. B. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*, pp. 520–534, 1965.