Supplement to Submission IEEE-TAC 23-2094 Titled "Data-Driven Learning Control for Discrete-Time Non-Zero-Sum Game Output Regulation by Internal Model"

Kedi Xie, Maobin Lu, Xiao Yu, Weiyao Lan

Abstract

This note is a supplementary material for the submitted manuscript IEEE-TAC 23-2094 titled "Data-driven learning control for discrete-time non-zero-sum game output regulation by internal model" by the same authors. It gives important details to the stability and convergence analyses, and is referred to at appropriate places in the manuscript.

I. Proof of Theorem 1

For the iteration $k=0$ and the player $j=1$, since $\mathrm{col}(\mathcal{K}_1,\mathcal{K}_2,\dots,\mathcal{K}_N)=\mathrm{col}(K_1^0,K_2^0,\dots,K_N^0)$ is a stabilizing control gain, the Lyapunov equation

$$\left(\bar{A}_c^0\right)^{\mathrm{T}} P_1^0\left(\bar{A}_c^0\right)-P_1^0+\bar{Q}_1+\sum\nolimits_{i=1}^{N}\left(K_i^0\right)^{\mathrm{T}} R_{i1}K_i^0=\mathbf{0} \tag{1}$$

has a unique positive definite solution P_1^0. By $K_1^1=\left(R_{11}+\bar{B}_1^{\mathrm{T}}P_1^0\bar{B}_1\right)^{-1}\bar{B}_1^{\mathrm{T}}P_1^0\left(\bar{A}-\sum_{i\neq 1}^{N}\bar{B}_i\mathcal{K}_i\right)$, we have

$$\begin{aligned}
&\left(\bar{A}-\sum_{i=0}^{N}\bar{B}_iK_i^0\right)^{\mathrm{T}}P_1^0\left(\bar{A}-\sum_{i=0}^{N}\bar{B}_iK_i^0\right)+\left(K_1^0\right)^{\mathrm{T}}R_{11}K_1^0\\
=&\left(\bar{A}-\bar{B}_1K_1^1-\sum_{i=2}^{N}\bar{B}_iK_i^0\right)^{\mathrm{T}}P_1^0\left(\bar{A}-\bar{B}_1K_1^1-\sum_{i=2}^{N}\bar{B}_iK_i^0\right)+\left(K_1^1\right)^{\mathrm{T}}R_{11}K_1^1\\
&+(K_1^1-K_1^0)^{\mathrm{T}}\left(R_{11}+\bar{B}_1^{\mathrm{T}}P_1^0\bar{B}_1\right)^{-1}(K_1^1-K_1^0).
\end{aligned} \tag{2}$$

Combining (1) and the definition of $\bar{A}_c^0=\bar{A}-\sum_{i=0}^{N}\bar{B}_i\mathcal{K}_i=\bar{A}-\sum_{i=0}^{N}\bar{B}_iK_i^0$, (2) can be rewritten as

$$\left(\bar{A}-\bar{B}_1K_1^1-\sum_{i=2}^{N}\bar{B}_iK_i^0\right)^{\mathrm{T}}P_1^0\left(\bar{A}-\bar{B}_1K_1^1-\sum_{i=2}^{N}\bar{B}_iK_i^0\right)-P_1^0+Q'=0 \tag{3}$$

where $Q'=\bar{Q}_1+\left(K_1^1\right)^{\mathrm{T}}R_{11}K_1^1+(K_1^1-K_1^0)^{\mathrm{T}}\left(R_{11}+\bar{B}_1^{\mathrm{T}}P_1^0\bar{B}_1\right)^{-1}(K_1^1-K_1^0)+\sum_{i=2}^{N}\left(K_i^0\right)^{\mathrm{T}}R_{i1}K_i^0$. Since $P_1^0>0$ and $Q'<0$, the matrix $\bar{A}-\bar{B}_1K_1^1-\sum_{i=2}^{N}\bar{B}_iK_i^0=\bar{A}-\sum_{i=1}^{N}\bar{B}_i\mathcal{K}_i$ is Schur for the player $j=2$ at the iteration $k=0$. Along the same derivation for every player $j\in\mathcal{N}$, and all iteration $k=0,1,2,\dots$, it is not hard to have that the estimated K^k is a stabilizing control gain at every iteration k. This completes the proof.

II. Proof of Theorem 2

To begin with, we prove that the sequence $\{J_j^0,J_j^1,J_j^2,\dots\},\forall j\in\mathcal{N}$, is a nonincreasing function series as follows. For the player $j=1$, according to (47) and (48), one can obtain

$$\begin{aligned}
&P_1^k-P_1^{k+1}\\
=&(\bar{A}_c^{k+1})^{\mathrm{T}}(P_1^k-P_1^{k+1})\bar{A}_c^{k+1}+\sum_{i=1}^{N}(K_i^k)^{\mathrm{T}}R_{i1}K_i^k-\sum_{i=1}^{N}(K_i^{k+1})^{\mathrm{T}}R_{i1}K_i^{k+1}\\
&+\sum_{i=1}^{N}(\bar{B}_iK_i^k)^{\mathrm{T}}P_1^k\sum_{i=1}^{N}\bar{B}_iK_i^k-\sum_{i=1}^{N}(\bar{B}_iK_i^{k+1})^{\mathrm{T}}P_1^k\sum_{i=1}^{N}\bar{B}_iK_i^{k+1}-2\bar{A}^{\mathrm{T}}P_1^k\sum_{i=1}^{N}\bar{B}_iK_i^k+2\bar{A}^{\mathrm{T}}P_1^k\sum_{i=1}^{N}\bar{B}_iK_i^{k+1}.
\end{aligned} \tag{4}$$

Kedi Xie and Maobin Lu are with the School of Automation, Beijing Institute of Technology, Beijing 100081, China, and are also with the Beijing Institute of Technology Chongqing Innovation Center, Chongqing 401120, China. (e-mail: kedixie@bit.edu.cn; lumaobin@bit.edu.cn).

Xiao Yu and Weiyao Lan are with the Department of Automation, Xiamen University, Xiamen, China. (e-mail: xiaoyu@xmu.edu.cn; wylan@xmu.edu.cn).

Then, one can obtain

$$\begin{aligned}
&\bar{x}_z(t)^{\mathsf T}(P_1^k - P_1^{k+1})\bar{x}_z(t) - \bar{x}_z(t+1)^{\mathsf T}(P_1^k - P_1^{k+1})\bar{x}_z(t+1)\\
=&\bar{x}_z^{\mathsf T}\left((P_1^k - P_1^{k+1}) - (\bar{A}_c^{k+1})^{\mathsf T}(P_1^k - P_1^{k+1})\bar{A}_c^{k+1}\right)\bar{x}_z\\
=&\sum_{i=1}^N(\bar{u}_i^k)^{\mathsf T}R_{i1}\bar{u}_i^k - \sum_{i=1}^N(\bar{u}_i^{k+1})^{\mathsf T}R_{i1}\bar{u}_i^{k+1} + \sum_{i=1}^N(\bar{B}_i\bar{u}_i^k)^{\mathsf T}P_1^k\sum_{i=1}^N\bar{B}_i\bar{u}_i^k - \sum_{i=1}^N(\bar{B}_i\bar{u}_i^{k+1})^{\mathsf T}P_1^k\sum_{i=1}^N\bar{B}_i\bar{u}_i^{k+1}\\
&- 2\bar{x}_z^{\mathsf T}\bar{A}^{\mathsf T}P_1^k\sum_{i=1}^N\bar{B}_i\bar{u}_i^k + 2\bar{x}_z^{\mathsf T}\bar{A}^{\mathsf T}P_1^k\sum_{i=1}^N\bar{B}_i\bar{u}_i^{k+1}\\
=&\sum_{i=1}^N(\bar{u}_i^k - \bar{u}_i^{k+1})^{\mathsf T}R_{i1}(\bar{u}_i^k - \bar{u}_i^{k+1}) + 2\sum_{i=1}^N(\bar{u}_i^{k+1})^{\mathsf T}R_{i1}(\bar{u}_i^k - \bar{u}_i^{k+1})\\
&+ \sum_{i=1}^N(\bar{u}_i^k - \bar{u}_i^{k+1})^{\mathsf T}\bar{B}_i^{\mathsf T}P_1^k\bar{B}_i(\bar{u}_i^k - \bar{u}_i^{k+1}) + 2\sum_{i=1}^N(\bar{u}_i^{k+1})^{\mathsf T}\bar{B}_i^{\mathsf T}P_1^k\bar{B}_i(\bar{u}_i^k - \bar{u}_i^{k+1})\\
&+ \sum_{i=1}^N\sum_{h\neq i}^N(\bar{B}_h\bar{u}_h^k)^{\mathsf T}P_1^k\bar{B}_i\bar{u}_i^k - \sum_{i=1}^N\sum_{h\neq i}^N(\bar{B}_h\bar{u}_h^{k+1})^{\mathsf T}P_1^k\bar{B}_i\bar{u}_i^{k+1} - 2\bar{x}_z^{\mathsf T}\bar{A}^{\mathsf T}P_1^k\sum_{i=1}^N\bar{B}_i\bar{u}_i^k + 2\bar{x}_z^{\mathsf T}\bar{A}^{\mathsf T}P_1^k\sum_{i=1}^N\bar{B}_i\bar{u}_i^{k+1}\\
=&\sum_{i=1}^N(\bar{u}_i^k - \bar{u}_i^{k+1})^{\mathsf T}(R_{i1} + \bar{B}_i^{\mathsf T}P_1^k\bar{B}_i)(\bar{u}_i^k - \bar{u}_i^{k+1}) + 2\sum_{i=1}^N(\bar{u}_i^{k+1})^{\mathsf T}(R_{i1} + \bar{B}_i^{\mathsf T}P_1^k\bar{B}_i)(\bar{u}_i^k - \bar{u}_i^{k+1})\\
&- \sum_{i=1}^N\bar{x}_z^{\mathsf T}(\bar{A} - \sum_{h\neq i}^N\bar{B}_hK_h^k)^{\mathsf T}P_1^k\bar{B}_i(\bar{u}_i^k - \bar{u}_i^{k+1}) - \sum_{i=1}^N\bar{x}_z^{\mathsf T}(\bar{A} - \sum_{h\neq i}^N\bar{B}_hK_h^{k+1})^{\mathsf T}P_1^k\bar{B}_i(\bar{u}_i^k - \bar{u}_i^{k+1})\\
=&\sum_{i=1}^N\bar{x}_z^{\mathsf T}\Big(\left(K_i^k - K_i^{k+1}\right)^{\mathsf T}(R_{i1} + \bar{B}_i^{\mathsf T}P_1^k\bar{B}_i)(K_i^k - K_i^{k+1})\\
&\quad + 2(\bar{A} - \sum_{h\neq i}^N\bar{B}_hK_h^k)^{\mathsf T}P_i^k\bar{B}_i(R_{ii} + \bar{B}_i^{\mathsf T}P_i^k\bar{B}_i)^{-1}(R_{i1} + \bar{B}_i^{\mathsf T}P_1^k\bar{B}_i)(K_i^k - K_i^{k+1})\\
&\quad - (\bar{A} - \sum_{h\neq i}^N\bar{B}_hK_h^k)^{\mathsf T}P_1^k\bar{B}_i(K_i^k - K_i^{k+1}) - (\bar{A} - \sum_{h\neq i}^N\bar{B}_hK_h^{k+1})^{\mathsf T}P_1^k\bar{B}_i\;(K_i^k - K_i^{k+1})\Big)\bar{x}_z\\
:=&\sum_{i=1}^N\bar{x}_z^{\mathsf T}\Omega_i\bar{x}_z
\end{aligned}$$

It follows from Theorem 1 that $\bar{A}_c^{k+1}$ is Schur for all $k = 0, 1, 2, \ldots$. If $\Omega_i \geq 0$, the following Lyapunov equation

$$(P_1^k - P_1^{k+1}) - (\bar{A}_c^{k+1})^{\mathsf T}(P_1^k - P_1^{k+1})\bar{A}_c^{k+1} = \Omega_i \tag{5}$$

has a unique positive semi-definite solution $(P_1^k - P_1^{k+1}) \geq 0$ since $\bar{A}_c^{k+1}$ is Schur. We next give a sufficient condition for $\Omega_i \geq 0$ as follows

$$\underline{\sigma}(\underline{R}_{i1})\|\Delta K\| \geq 2\bar{\sigma}(\bar{R}_{i1})\|\bar{A}_h^k\|\|P_i^k\|\|\bar{B}_i\| + (\|\bar{A}_h^k\| + \|\bar{A}_h^{k+1}\|)\|P_1^k\|\|\bar{B}_1\| \tag{6}$$

where $\underline{\sigma}(\underline{R}_{i1})$ is the minimum singular value of $R_{i1}+\bar{B}_i^{\mathsf T}P_1^k\bar{B}_i$, $\bar{\sigma}(\bar{R}_{i1})$ is the maximum singular value of $(R_{ii}+\bar{B}_i^{\mathsf T}P_i^k\bar{B}_i)^{-1}(R_{i1}+\bar{B}_i^{\mathsf T}P_1^k\bar{B}_i)$, $\Delta K = K_i^k - K_i^{k+1}$, and $\bar{A}_h^k = \bar{A} - \sum_{h\neq i}^N\bar{B}_hK_h^k$. Using the same derivation for all player $j \in \mathcal{N}$, we have that for small norm $\|B_i\|$, the value function J_j^k is a nonincreasing function. According to the principle of optimality [1], we have

$$J_j^k \geq \sum_{t=t_0}^{\infty}\left(\bar{x}_z^{\mathsf T}\bar{Q}_j\bar{x}_z + (\bar{u}_j^*)^{\mathsf T}R_{jj}\bar{u}_j^* + \sum\nolimits_{i\neq j}^N(\bar{u}_i^*)^{\mathsf T}R_{ij}\bar{u}_i^*\right) = J_j^*, \qquad \forall j \in \mathcal{N} \tag{7}$$

where u_j^* is the unique solution to the non-zero-sum game which can minimize the cost functions $J_j^* = J_j(x_z, u_1^*, \ldots, u_N^*, t_0), \forall j \in \mathcal{N}$ in the sense of Nash equilibrium. Since the value function $J^k = \{J_1^k, J_2^k, \ldots, J_N^k\}$ is a nonincreasing function, and there exists a unique solution J^* to the non-zero-sum game problem, it follows from (7) that $J^k = \{J_1^k, J_2^k, \ldots, J_N^k\}$ converge to $J^* = \{J_1^*, J_2^*, \ldots, J_N^*\}$ as $k \to \infty$. Thus, the estimated control gain K^k converges to K^*. This completes the proof.

References

[1] F. L. Lewis, D. L. Vrabie, and V. L. Syrmos, *Reinforcement Learning and Optimal Adaptive Control*. John Wiley & Sons, Inc., 2012.