Bulletin of Applied Mathematics and Mathematics Education
Volume 4, Number 1, 2024 pp. 1- 8
https://dx.doi.org/10.12928/bamme.v4i1.9351
e-ISSN 2776-1029, p-ISSN 2776-1002

The Zero-Sum Discrete-Time Feedback Linear Quadratic Dynamic Game: From Two-Player Case to Multi-Player Case

Muhammad Wakhid Musthofa[1*], Jacob Engwerda[2]

[1]UIN Sunan Kalijaga Yogyakarta, Jl. Marsda Adisucipto no. 1, Sleman, DIY 55281, Indonesia
[2]Tilburg University, Warandelaan 2, 5037 AB Tilburg, Netherlands

*Corresponding E-mail: muhammad.musthofa@uin-suka.ac.id

ARTICLE INFO

Article History
Received 15 November 2023
Revised 7 June 2024
Accepted 7 June 2024

Keywords
Discrete-time
Feedback information structure
Linear quadratic dynamic game
Multi-player

How to cite this article:
Musthofa, M. W., & Engwerda, J., (2024). The Zero-Sum Discrete-Time Feedback Linear Quadratic Dynamic Game: from Two-Player Case to Multi-Player Case. *Bulletin of Applied Mathematics and Mathematics Education*, 4(1), 1-8.

ABSTRACT

The two-person zero-sum discrete-time feedback linear quadratic dynamic game is considered. A method that provides a saddle-point for the zero-sum discrete dynamic game is developed to derive a necessary and sufficient condition under which the game has a feedback saddle-point solution. Existence solutions, which are described in terms of a sequence of nonnegative definite algebraic Riccati matrices, are constructed. Next, a generalization of such a game to a multi-player case is studied. Using the results in a two-person case, the characterization of a feedback saddle-point solution for the multi-player game is derived.

Introduction

A linear quadratic dynamic game is a type of game in which two or more players interact with each other over time, and each player's objective is to minimize a quadratic cost function that depends on their own actions and the actions of the other players (Lukes & Russell, 1971; Engwerda, 2005). This game has various applications in different fields, and their usefulness is still being explored, such as signal processing (Pachter & Pham, 2010), reinforcement learning (uz Zaman et al., 2023; Mazumdar et al., 2019), pricing mechanism (Ratliff et al., 2012), identification goal (Köpf, 2017), and partial state observations (Yu et al., 2022).

Besides that, in the area of control sciences, the linear quadratic dynamic game is a fundamental concept in optimal control (Jank & Kun, 2002) and can be utilized as a risk-sensitive dynamic game formulation (Moon & Başar, 2017; Xu & Wu, 2023) or robust control (Musthofa et al, 2016; Van Den Broek, 2023), or it can be used as a standard for evaluating the performance of multi-agent reinforcement learning algorithms involving two agents competing in continuous state-control environments (Zhang, 2021).

The linear quadratic dynamic game for continuous time has been studied by Engwerda (2005), Delfour (2007), and Garcia et al. (2020). For the discrete-time horizon, the game has been studied by Xu & Mukaidani (2003), Pachter & Pham (2010), and Kebriaei & Iannelli (2017). The generalization of the game from nonsingular systems into singular (descriptor) systems has been reported by Musthofa et al. (2013) for open continuous-time games and by Musthofa et al. (2021) for discrete-time games. All of the mentioned works above use the type of two-person game. Unfortunately, the generalization of such games from two-person into multi-player games has been lacking until now.

This paper aims to generalize the discrete-time linear quadratic dynamic game from a two-person case into a multi-player case. The multi-player games can be used to study the behavior of agents in complex environments where the actions of one agent can affect the actions of other agents (Mahajan et al., 2022; Couto & Pal, 2023). The games can also be used to study the emergence of cooperation and competition among agents (Omidshafiei et al., 2020; Gokhale & Traulsen, 2014). In this paper, we will find the necessary and sufficient condition under which the generalized (multi-player) game has a feedback saddle-point equilibrium, using the results of the two-person game case.

Method

We solve the feedback saddle-point (FSP) equilibrium in the following step for the zero-sum discrete-time feedback linear quadratic dynamic game. First, we consider the game as a two-player case. We start our work by defining a cost function for both players. Next, we assume the actions given by both players are in a linear feedback control form. From this, we define the FSP equilibrium and follow by stating a theorem that characterizes the FSP equilibrium for the two-person dynamic game. Next, we generalize the two-person dynamic game by considering the game that involves n players. We named it the multi-player dynamic game. For this game, we also find the existence of an FSP solution.

Results and Discussion

In this paper, we consider a zero-sum discrete-time feedback linear quadratic dynamic game defined by the dynamical system

$$x(k+1) = Ax(k) + B_1u(k) + B_2w(k), x(1) = \bar{x}, k \in [1, K], \quad (1)$$

where x is the vector state of the system, and the matrices $A \in \mathbb{R}^{n\times n}$, $B_i \in \mathbb{R}^{n\times m_i}, i = 1,2$. An $m_1 \times 1$ vector $u \in \mathcal{U} \subset \mathbb{R}^{n\times m_1}$ expresses the control variable for the system, where $\mathcal{U}$ represents the collection of control functions that are locally square summable and result in a stable system. An $m_2 \times 1$ vector $w \in \mathcal{L}_2[1, K]$ expressed the disturbance for the systems, where $\mathcal{L}_2[1, K]$ defines the set of all square summable functions on $[1, K]$.

Relating to the dynamical system (1), we define a cost function for the first player (minimizer)

$$J_\gamma(u(k), w(k)) = \sum_{k=1}^{K} (x^T(k)Qx(k) + u^T(k)R_1u(k) - \gamma^2 w^T(k)R_2w(k))$$
$$+x^T(K+1)Q_Kx(K+1) \quad (2)$$

where $Q, Q_K \geq 0$, $R_i > 0$, $i = 1,2$. The parameter $\gamma \in \mathbb{R}$ determines the weight or significance of the second player's action (maximizer/disturbance), whose cost function is $-J(u(k), w(k))$.

 10.12928/bamme.v4i1.9351

Furthermore, in this paper, we set a feedback information structure as a type of information for the dynamic game (1,2). So, in this game framework, both players give their actions in a linear feedback control form

$$u(k) = F_u(k)x(k) \in \mathcal{U} \quad \text{and} \quad w(k) = F_w(k)x(k) \in \mathcal{L}_2[1, K]. \tag{3}$$

Finally, we define this paper's main object of study, the feedback saddle-point (FSP) equilibrium (Engwerda, 2005).

Definition 1. The pair $(u^*(k), w^*(k))$ is an FSP equilibrium for the differential game (1,2) if for every $((u^*(k), w(k)), u(k), w^*(k)))$, the following inequality holds

$$J_\gamma(u^*(k), w(k)) \leq J_\gamma\big(u^*(k), w^*(k)\big) \leq J_\gamma\big(u(k), w^*(k)\big). \tag{4}$$

Then, the addressed problem in this paper is to find the FSP solution for the dynamic game (1,2) under the feedback information structure.

Two Players Case

To find the FSP equilibrium for the dynamic game (1,2), we need to state Theorem 2.4 (Basar & Bernhard, 2008), which states the FSP solution for the discrete-time dynamic game. The theorem is as follows.

Theorem 1. *Consider the zero-sum discrete dynamic game*

$$x(k+1) = f\big(x(k), u(k), w(k)\big), k \in [1, K], \tag{5}$$

with the finite-horizon cost function is given by

$$J(u, w) = \sum_{k=1}^{K} g(x(k+1), u(k), w(k), x(k)), \tag{6}$$

and the Isaacs equation in discrete-time

$$\begin{aligned} V_k(x) &= \min_{u \in U_s} \max_{w \in L_2[1,K]} \big[g\big(x(k+1), u(k), w(k), x(k)\big) + V_{k+1} f\big(x(k), u(k), w(k)\big)\big] \\ &= \max_{w \in L_2[1,K]} \min_{u \in U_s} \big[g\big(x(k+1), u(k), w(k), x(k)\big) + V_{k+1} f\big(x(k), u(k), w(k)\big)\big] \\ &= g\big(f\big(x(k), u(k), w(k)\big), u^*(k), w^*(k), x(k)\big) + V_{k+1} f\big(x(k), u^*(k), w^*(k)\big), \\ V_{K+1}(x) &= 0. \end{aligned} \tag{7}$$

Let there exist a function $V_k(\cdot), k \geq 1$, and two policies $u \in \mathcal{U}$, $w \in \mathcal{L}_2[1, K]$ generated by (7)*. Then, the pair $(u^*(k), w^*(k))$ provides a saddle-point for the zero-sum discrete dynamic game* (5,6)*, with the value of the saddle-point is given by $V_1(x(1))$.*

Using the above theorem, we arrive at a theorem that characterizes the FSP equilibrium for the dynamic game (1,2).

Theorem 2. Consider a fixed $\gamma > 0$ and the two-person zero-sum dynamic game with closed-loop perfect-state (CLPS) information pattern (1,2). Then,

1. There exists a unique feedback saddle-point (FSP) solution if

$$\Theta := \gamma^2 R_2 - B_2^T M(k+1) B_2 > 0, k \in [1, K], \tag{8}$$

where the sequence of nonnegative definite matrices $M(k+1), k \in [1, K]$, is generated by

$$M(k) = Q + A^T M(k+1) \Lambda^{-1} A, M(k+1) = Q_K, \tag{9}$$

where

$$\Lambda(k) := I + (B_1 R_1^{-1} B_1^T - \gamma^{-2} B_2 R_2^{-1} B_2^T) M(k+1). \tag{10}$$

2. Under condition (8), the matrices $\Lambda(k), k \in [1, K]$ are invertible, and the unique feedback saddle-point (FSP) equilibrium is

$$u^*(k) = -R_1^{-1} B_1^T M(k+1) \Lambda^{-1}(k) A x(k), \tag{11}$$

$$w^*(k) = \gamma^{-2} R_2^{-1} B_2^T M(k+1) \Lambda^{-1}(k) A x(k). \tag{12}$$

3. If the matrix (8) possesses a negative eigenvalue for some $k \in [1, K]$, the game does not possess a saddle-point under the (CLPS) information structure, and its upper value becomes unbounded.

Proof of Theorem 2. Parts 1. and 2. of Theorem 2 follow from Theorem 1 by showing that the unique feedback saddle-point (FSP) equilibrium (11,12) uniquely solves the equation (7) for the dynamic game (5,6). In this case, the corresponding value function being

$$V_k(x) = x^T M(k+1) x, \quad k \in [1, K].$$

Using this result, it follows that for each $k \in [1, K]$ in the two-person zero-sum dynamic game (1,2), the existence of unique (FSP) equilibrium for the corresponding static game is ensured by the positive definiteness of the matrix in (8). Furthermore, using some matrix manipulations (see, e.g., Basar & Olsder, 1998) shows that the positive definiteness of (8) implies the invertibility of (10). Then, the consequence of this is f there exists a value $k \in [1, K]$ such that equation (8) has a negative eigenvalue, then the related static game does not possess a saddle-point and its upper value becomes unbounded. This proves part 3. □

Multi-Player Case

In this section, we generalize the dynamic game (1,2) by considering that there exist n players in this game. In this game setting, we consider the following model

$$x(k+1) = Ax(k) + \sum_{i=1}^{N} B_{ui} u_i(k) + B_w w(k), \; x(1) = \bar{x}, \; k \in [1, K], \tag{13}$$

with cost functions

$$J_{\gamma_i}(u_i(k), w(k)) := \sum_{k=1}^{K} \left(x^T(k) Q_i x(k) + \sum_{j=1}^{K} u_j^T(k) R_{ij} u_j - \gamma_i^2 w^T(k) R_w w(k) \right) \tag{14}$$

$$+ x^T(K+1) Q_{iK} x(K+1),$$

where $Q_i \geq 0, Q_{iK} \geq 0, R_{ij} \geq 0, R_{ii} > 0$, and $R_w > 0$.

The following theorem states the existence and characterization of the multi-player game.

Theorem 3. *Consider the set of coupled equations*

 10.12928/bamme.v4i1.9351

$$M_i(k) = Q_i + \sum_{j\neq i}^{N} F_j^T(k) R_{ij} F_j(k) + \left(A + \sum_{j\neq i}^{N} B_j F_J\ \ (k)\right)^T M_i(k+1)\Lambda_i^{-1}(k)\left(A + \sum_{j\neq i}^{N} B_j F_J\ \ (k)\right)$$

$$M_i(K+1) = Q_{iK};\, i = 1,2,\dots,N, \tag{15}$$

where

$$\Lambda_i(k) \coloneqq I + \left(B_i R_{ii}^{-1} B_i^T - \gamma_i^{-2} B_w R_w^{-1} B_w^T\right) M_i(k+1), \tag{16}$$

and

$$\begin{bmatrix} F_1(k) \\ \vdots \\ F_N(k) \end{bmatrix} := -G^{-1}(k) \begin{bmatrix} Z_1(k+1) \\ \vdots \\ Z_N(k+1) \end{bmatrix} A. \tag{17}$$

Here, with $Z_i(k+1) \coloneqq R_{ii}^{-1} B_i^T M_i(k+1)\Lambda_i^{-1}(k)$,

$$G(k) \coloneqq \begin{bmatrix} I & Z_1(k+1)B_2 & \cdots & Z_1(k+1)B_N \\ \vdots & \vdots & \ddots & \vdots \\ Z_N(k+1)B_1 & \dots & Z_N(k+1)B_{N-1} & I \end{bmatrix}.$$

Then,

1. *For all players, there exists a unique feedback saddle-point solution if*

   a. $\Theta_i(k) \coloneqq \gamma_i^2 R_w - B_w^T M_i(k+1) B_w > 0,,\ k \in [1, K], i = 1, \dots, N;$ (18)

   and

   b. $G(k)$ *is invertible,,* $k \in [1, K]$*;* (19)

   where the sequence of nonnegative definite matrices $M_i(k+1)$*,* $k \in [1, K]$*, is generated by* (15).

2. *Under conditions (18,19), the matrices* $\Lambda_i(k)$*,* $k \in [1, K]$*, are invertible, and the unique feedback saddle-point policies are*

$$u_i^*(k) = -R_{ii}^{-1} B_i^T M_i(k+1)\Lambda_i^{-1}(k)\left(A + \sum_{j\neq i}^{N} B_j F_j(k)\right) x(k) \tag{20}$$

   where the worst-case control for player i that can occur is

$$w_i^*(k) = -\gamma_i^{-2} R_w^{-1} B_w^T M_i(k+1)\Lambda_i^{-1}(k)\left(A + \sum_{j\neq i}^{N} B_j F_j(k)\right) x(k). \tag{21}$$

3. *If for some i, the matrix* (18) *has a negative eigenvalue for some* $k \in [1, K]$*, then the game does not admit a saddle-point for player i under the (CLPS) information structure, and its upper value becomes unbounded.*

Proof of Theorem 3. Assuming all players use a state feedback control $u_i(k) = F_i(k)x(k)$. Player *i* is confronted with the optimization problem to minimize

$$J_{\gamma_i}\big(u_i(k),w(k)\big) := \sum_{k=1}^{K}\left(x^T(k)\left[Q_i+\sum_{j\neq i}^{N}F_j^T(k)R_{ij}F_j\right]x(k)+u_i^T(k)R_{ii}u_i(k)-\gamma_i^2 w^T(k)R_w w(k)\right)$$
$$+x^T(K+1)Q_{iK}x(K+1),$$

subject to

$$x(k+1)=\left[A+\sum_{j\neq i}^{N}B_jF_j(k)\right]x(k)+B_iu_i(k)+B_ww(k), x(1)=\bar{x}, k\in[1,K].$$

By Theorem 2 this problem has a solution

$$u_i(k)=-R_{ii}^{-1}B_i^TM_i(k+1)\Lambda_i^{-1}(k)\left(A+\sum_{j\neq i}^{N}B_jF_j(k)\right)x(k),$$

where

$$M_i(k)=Q_i+\sum_{j\neq i}^{N}F_j^T(k)R_{ij}F_j(k)+\left(A+\sum_{j\neq i}^{N}B_jF_J\ \ (k)\right)^T M_i(k+1)\Lambda_i^{-1}(k)\left(A+\sum_{j\neq i}^{N}B_jF_J\ \ (k)\right)$$

$M_i(K+1)=Q_{iK}$

Provided

$$\Lambda_i(k):=I+\big(B_iR_{ii}^{-1}B_i^T-\gamma_i^{-2}B_wR_w^{-1}B_w^T\big)M_i(k+1)$$

is invertible and the set of equations

$$F_i(k)=-R_{ii}^{-1}B_i^TM_i(k+1)\Lambda_i^{-1}(k)\left(A+\sum_{j\neq i}^{N}B_jF_j(k)\right) \tag{22}$$

has a solution $(F_1,\ldots F_N)$.

Notice (22) can be rewritten with $Z_i(k+1):=R_{ii}^{-1}B_i^TM_i(k+1)\Lambda_i^{-1}(k)$ as

$G(k)\begin{bmatrix}F_1(k)\\ \vdots\\ F_N(k)\end{bmatrix}=-\begin{bmatrix}Z_1(k+1)A\\ \vdots\\ Z_N(k+1)A\end{bmatrix}$. So, $(F_1(k),\ldots F_N(k))$ is uniquely determined if $G(k)$ is invertible. Furthermore, from Theorem 2item 1, it follows that $\Lambda_i(k)$ is invertible if (19) applies. □

Remarks 1. The solution presented in Theorem 3 can be calculated recursively backward in time. By first calculating $M_i(k+1), i=1,2,..,N$ in (15), next calculating $\Lambda_i(k), i=1,2,\ldots,N$ in (16), followed by the calculation of $Z_i(k+1), i=1,2,\ldots,N$, and next $G(k)$. Finally, $F_i(k), i=1,2,\ldots,N$ can then be determined from (17), and this sequentially for $k=N,N-1,\ldots,1$.

Conclusion

This paper considered the zero-sum discrete-time feedback linear quadratic dynamic game. We have characterized the necessary and sufficient conditions under which the game has a feedback saddle-point solution. We solve the problem assuming feedback information structure. Here, we have shown the critical role played by the matrices (8) for the two-player games and (18,19) for the

 10.12928/bamme.v4i1.9351

multi-player games to guarantee the existence of feedback saddle-point solutions for games.

The information structure of the game addressed in this paper is restricted to feedback information structure. To find an FSP equilibrium in a linear quadratic dynamic game for another information structure such as open-loop, sampled-data, and delayed-state is still an open problem to be analyzed for future research.

Acknowledgment

The primary author expresses gratitude for the assistance the Institute for Research and Community Service of UIN Sunan Kalijaga Yogyakarta provided through research funding in 2022.

References

Başar, T., & Bernhard, P. (2008). *H-infinity Optimal Control and Related Minimax Design Problems: a Dynamic Game Approach.* Springer Science & Business Media.

Başar, T., & Olsder, G. J. (1998). *Dynamic noncooperative game theory*. Society for Industrial and Applied Mathematics.

Couto, M. C., & Pal, S. (2023). *Introspection dynamics in asymmetric multiplayer games*. Dynamic Games and Applications, 13(4), 1256-1285, https://doi.org/10.1007/s13235-023-00525-8

Delfour, M. C. (2007). *Linear quadratic differential games: Saddle point and Riccati differential equation*. *SIAM journal on control and optimization*, 46(2), 750-774.

Engwerda, J. (2005). *LQ dynamic optimization and differential games*. John Wiley & Sons.

Garcia, E., Casbeer, D. W., Pachter, M., Curtis, J. W., & Doucette, E. (2020, July). *A two-team linear quadratic differential game of defending a target*. In *2020 American Control Conference (ACC)* (pp. 1665-1670). IEEE.

Gokhale, C. S., & Traulsen, A. (2014). Evolutionary multi-player games. *Dynamic Games and Applications*, 4, 468-488.

Jank, G. and Kun, G. Optimal Control of Disturbed Linear-Quadratic Differential Games, *European Journal of Control*, Volume 8, Issue 2, 2002, Pages 152-162

Kebriaei, H., & Iannelli, L. (2017). *Discrete-time robust hierarchical linear-quadratic dynamic games*. *IEEE Transactions on Automatic Control*, 63(3), 902-909.

Köpf, F., Inga, J., Rothfuß, S., Flad, M., & Hohmann, S. (2017). *Inverse reinforcement learning for identification in linear-quadratic dynamic games*. *IFAC-PapersOnLine*, 50(1), 14902-14908.

Lukes, D. L., & Russell, D. L. (1971). *A global theory for linear-quadratic differential games. Journal of Mathematical Analysis and Applications*, 33(1), 96–123.

Mahajan, A., Samvelyan, M., Gupta, T., Ellis, B., Sun, M., Rocktäschel, T., & Whiteson, S. (2022). *Generalization in cooperative multi-agent systems. arXiv preprint* arXiv:2202.00104.

Mazumdar, E., Ratliff, L. J., Jordan, M. I., & Sastry, S. S. (2019). *Policy-gradient algorithms have no guarantees of convergence in linear quadratic games—arXiv preprint* arXiv:1907.03712.

Moon, J., & Başar, T. (2016). Linear quadratic risk-sensitive and robust mean field games. *IEEE Transactions on Automatic Control*, 62(3), 1062-1077.

Musthofa, M. W., Salmah, , Engwerda, J. C., & Suparwanto, A. (2013). *Feedback saddle point equilibria for soft-constrained zero-sum linear quadratic descriptor differential game*. *Archives of Control Sciences*, 23(4), 473-493.

Musthofa, M. W., Engwerda, J., & Suparwanto, A. (2016). *Robust Optimal Control Design Using a Differential Game Approach for Open-Loop Linear Quadratic Descriptor Systems*. *Journal of Optimization Theory and Applications*, 168(3), 1046-1064.

Musthofa, M. W. (2021). *The Open-Loop Zero-Sum Linear Quadratic Index One Discrete-Time Soft-*

Constrained Descriptor Dynamic Games. *Journal of Mathematical Control Science and Applications*, 7(1), 1 – 10.

Omidshafiei, S., Tuyls, K., Czarnecki, W. M., Santos, F. C., Rowland, M., Connor, J., & Munos, R. (2020). *Navigating the landscape of multi-player games*. *Nature communications*, 11(1), 5603.

Pachter, M., & Pham, K. D. (2010). *Discrete-time linear-quadratic dynamic games*. *Journal of Optimization Theory and Applications*, 146, 151-179.

Ratliff, L. J., Coogan, S., Calderone, D., & Sastry, S. S. (2012, October). *Pricing in linear-quadratic dynamic games*. In *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)* (pp. 1798-1805). IEEE.

Van Den Broek, W. A., Engwerda, J. C., & Schumacher, J. M. (2003). *Robust equilibria in indefinite linear-quadratic differential games*. *Journal of Optimization Theory and Applications*, pp. 119, 565–595.

Wu, J. (2023). *Learning Zero-Sum Linear Quadratic Games with Improved Sample Complexity and Last-Iterate Convergence* (Master's thesis).

Xu, H., & Mukaidani, H. (2003). *The linear quadratic dynamic game for discrete-time descriptor systems*. *International Game Theory Review*, 5(04), 361-374.

Xu, R. and Wu, T., *Risk-sensitive large-population linear-quadratic-Gaussian games with major and minor agents*, *Asian J Control* (2023), 1–13, DOI 10.1002/asjc.3106.

Yu, C., Li, Y., Li, S., & Chen, J. (2022). *Inverse linear quadratic dynamic games using partial state observations*. *Automatica*, 145, 110534.

uz Zaman, M. A., Miehling, E., & Başar, T. (2023). *Reinforcement learning for non-stationary discrete-time linear–quadratic mean-field games in multiple populations*. *Dynamic Games and Applications*, 13(1), 118-164.

Zhang, K., Yang, Z., & Başar, T. (2021). *Multi-agent reinforcement learning: A selective overview of theories and algorithms*. Handbook of reinforcement learning and control, 321-384.

 10.12928/bamme.v4i1.9351