Controllability of Non-cooperative Dynamical Games

ZHANG Renren¹, GUO Lei²

1. Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190
E-mail: zhangrenren13@mails.ucas.ac.cn

2. Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190
E-mail: lguo@amss.ac.cn

Abstract: In this paper, we consider a new class of control systems based on non-cooperative dynamical games. It involves a hierarchal decision making structure: one leader and multiple followers. Given any strategy of the leader, the followers form a non-cooperative dynamical game which may reach a Nash equilibrium. We will study the leader’s controllability of such an equilibrium which has not been investigated before. It seems to be a new direction of dynamical games from the perspective of control and beyond the frameworks of both the traditional control theory and game theory. First, we will give the definition of the controllability, and then give some necessary and sufficient conditions of the controllability for non-cooperative linear-quadratic dynamical games.

Key Words: Non-cooperative dynamical games, hierarchical structure, controllability

1 Introduction

Multiagent systems, as a starting point for investigating complex systems, have received a great deal of research attention in recent years from various fields including control systems. These systems consist of a number of dynamic and heterogeneous agents, interacting and/or competing with each other, and may be the most complicated ones to handle [1]. The agents may act selfishly so as to improve their own payoffs which may be different from one to another. So the theory of games appears to be a natural tool in modeling and analyzing these kind of dynamic systems. In the framework of the current control theory, the systems to be controlled are completely “passive” in an essential way, in the sense that they have no intention to complete with controllers to achieve their own objectives or “payoffs” [8], [9]. By investigating the control systems based on dynamical games, we may understand more about multiagent systems. The strategy of each agent can be viewed as its respective control. In the real world, complex systems usually involve hierarchical decision making. In dealing with these systems, the Stackelberg differential game [2] appears to be a useful tool.

In this paper, we study the non-cooperative dynamical games. The dynamics are governed by a first order ordinary differential equation. Here we only investigate the system which has one leader and multiple followers. In practical systems, the leader can be the decision maker of any human organization. In the existing literature, most of the workers mainly concern with the Stackelberg games in the case the payoff functions of leader and followers are given in some sense [10], [11], [12], [13]. Here we change our perspective on the system a bit. We temporarily ignore the payoff function of the leader and consider the non-cooperative dynamical games of the followers for any given strategy of the leader. So we can investigate the leader’s ability in controlling the system, we will call it controllability which will be clearly studied in the Section 2. It reflects the leader’s influence of the non-cooperative dynamical system to some extent. As a starting point, we will consider the non-cooperative linear-quadratic payoffs dynamical games.

The remainder of this paper is organized as follows. In Section 2, we give the problem formulation and the main results. Section 3 gives the proof of some results. Section 4 will conclude the paper.

2 Problem Formulation and Main Results

2.1 Problem Formulation

We consider the non-cooperative linear-quadratic dynamical game with one leader and multiple followers. The system is described by the following differential equation

\[
\dot{X}(t) = A(t)X(t) + \sum_{i=1}^{l} B_i(t)u_i(t) + B(t)u(t),
\]

where \(X(t) \in \mathbb{R}^n\), \(A(t) \in \mathbb{R}^{n \times n}\), \(B_i(t) \in \mathbb{R}^{n \times m_i}\), \(B(t) \in \mathbb{R}^{n \times m}\), \(l\) is the number of the followers. We define \(A(t)\) as system matrix, \(B_i(t)(i = 1,2,\cdots,l)\) as action matrix for follower \(i\), \(B(t)\) as action matrix for the leader, they are all piecewise continuous matrix functions of time. \(u_i(t) \in \mathbb{R}^{m_i}\) is the strategy of the follower \(i\), here we only consider the open-loop strategy, \(u(t) \in \mathbb{R}^m\) is the strategy of the leader. \(X(t)\) is the state of the system. It is assume that the data \(A(t), B_i(t), B(t)\) is the common knowledge.

The payoff function to be minimized by each \(u_i\) is:

\[
J_i(u_i) = \frac{1}{2} \int_0^T \left[ X^T(t)Q(t)X(t) + u_i^T(t)R_i(t)u_i(t) \right] dt,
\]

i.e.,

\[
u_i = \arg \min_{u_i \in \mathbb{R}^m} J_i(u_i)
\]
where \( Q_i^T(t) = Q_i(t) \geq 0 \), \( R_i^T(t) = R_i(t) > 0 \), are also common knowledge. Given any strategy of the leader, the followers will minimize the payoff function of their own. Throughout this paper, \( \mathbb{R}^{n \times m} \) denotes the set of all \( n \times m \) real matrices, \( \text{Im}(A) (A \in \mathbb{R}^{n \times m}) \) denotes the \( n \) dimensional vector space spanned by the column of the matrix \( A \), and \( \text{rank}(A) (A \in \mathbb{R}^{n \times m}) \) denotes the rank of the matrix \( A \). For any given matrix or vector \( X, X^T \) denotes the transpose of it. \( I_{n \times n} \) is the \( n \) by \( n \) unit matrix. \( \lambda(A) \) means the eigenvalues of matrix \( A \). For any given two real vector space with the same dimension \( V_1 \) and \( V_2 \), \( V_1 + V_2 \) denotes the sum of two vector space, and if \( V_1 \cap V_2 = \{0\} \), then we will rewrite it as \( V_1 \oplus V_2 \), called the direct sum. If we take this model as a control system, of which the controller input is the strategy of the leader and the output is the state of the system, then we can study the control property of this system, such as controllability. For any given strategy of the leader, the followers will form an \( I \) person non-cooperative game. We will consider the Nash equilibrium of this non-cooperative game. That is, for those set of strategies of the followers which are such that if one follower deviates from his strategy his/her payoff will increase, the definition can be found in [5]. The followers have an open-loop information structure \( \eta_t = (x_{0t}, t \in [0, T]) \). That is the followers already have to formulate their strategies at the moment the system starts to evolve and these strategies cannot changed once the system is running. We assume that the followers can get the Nash equilibrium and act with it and the Nash equilibrium always exist in this model we have constructed. We can see that the state of the system is decided by the strategy of the leader, that is to say, for any input \( u(t) \) of the leader, to some extent the state of the system \( X(t) (0 \leq t \leq T) \) is determined, so we can see it as a control system. In this paper, we study the power or influence of the leader, which can be reflected by leader’s controllability.

**Definition 2.1** The system (1), (2) is called controllable, if for any given initial state \( X_0 \) at time 0 and any terminal state \( X_T \) at time \( T \), there is a permissible strategy of the leader \( u(t) \), under which the Nash-equilibrium exists and is unique and the state of the system \( X(t) (0 \leq t \leq T) \) satisfy the condition: \( X(0) = X_0 \) and \( X(T) = X_T \).

**Remark 2.1** According to different control purpose, the permissible control or strategy is different also, for the sake of simplicity, we call the piecewise continuous function set as permissible control set in this paper, element in the set is called permissible control or strategy. We also assume that the strategies of followers are permissible control.

**Remark 2.2** In this paper, the terminal time \( T \) is fixed and finite. In the definition, we not only requires the equilibrium exists, but also to be unique, that is because if these conditions are not satisfied, then we cannot predict the strategies of the leaders and the evolution of the system.

From the definition, we know that if the system is controllable, then the leader can drive the state of the system to any he (she) want to though the follower act selfishly so as to improve their own payoffs. In this case, the leader is able to completely control the state of the system.

### 2.2 Main Results

First, we use the maximum principle [3] to solve the model, for any given strategy of the leader \( u(t) \), the strategies of the followers \( u_i(t) (i = 1, 2, \ldots, I) \) satisfy the follow:

\[
\begin{align*}
\dot{X}(t) &= A(t)X(t) + \sum_{i=1}^{I} B_i(t)u_i(t) + B(t)u(t) \\
\dot{\phi}(t) &= Q(t)X(t) - A^T(t)\phi(t), i = 1, 2, \ldots, I \\
X(0) &= X_0, \phi(T) = 0, i = 1, 2, \ldots, I \end{align*}
\]

where the \( \phi_i(t) (i = 1, 2, \ldots, I) \) are adjoint variables, the coupled forward-backward differential equations are called adjoint equations.

We introduce the following variables:

\[
\begin{bmatrix}
Q(t) \\
Q(t) \\
\vdots \\
Q(t)
\end{bmatrix}
\begin{bmatrix}
X(t) \\
\phi(t) \\
0 \\
\phi(t)
\end{bmatrix} = \begin{bmatrix}
A(t) & B(t)R_i^T(t)E_i(t) & \cdots & B_i(t)R_i^T(t)E_i(t) \\
0 & -A^T(t) \\
\vdots & \ddots & \ddots & \ddots \\
0 & \cdots & 0 & -A^T(t)
\end{bmatrix}
\]

By (4), the system (3), is equivalent to the following system:

\[
\dot{X}(t) = A(t)X(t) + B(t)u(t)
\]

From the representation (5) of the system, we know that we can deal with this system using the control theory of linear system.

**Assumption 2.1** In this paper, we always assume that: For any strategy \( u(t) \) of the leader and any initial state \( X_0 \), the Nash equilibrium of the followers exists;

If the equilibrium is also unique, it will proved correct, then we know that for any strategy \( u(t) \) of the leader, the dynamic of the system is governed by (3) [4].

**Remark 2.3** Because of the assumption 2.1, the controllability of (1), (2) can be converted to study the system (3). We can restate the definition of controllability as follow: The system (1), (2) is called controllable, if for any given initial state \( X_0 \) at time 0 and any terminal state \( X_T \) at time \( T \), there is a permissible strategy of the leader \( u(t) \), under which the solution of (3) \( X(t), \phi(t) \) satisfy the condition: \( X(0) = X_0, X(T) = X_T \) and \( \phi(T) = 0 \) and {\( u_i(t) (i = 1, 2, \ldots, I) \)} is the Nash equilibrium.
We define the state transition matrix $\Phi(t,s)$:

$$\frac{\partial \Phi(t,s)}{\partial t} = -A(t)\Phi(t,s), \quad \Phi(s,s) = I_{(\tau+1)n\times(\tau+1)n}$$

(6)

and the controllable matrix of the system $W(T,0)$:

$$W(T,0) = \int_0^T \Phi(t,s)B(t)\bar{B}^T(t)\Phi^T(t,s)ds$$

(7)

This matrix is like the controllable matrix of the linear system in control theory.

**Theorem 2.1** The system (1), (2) is controllable, if and only if the follow matrix is full rank:

$$\begin{bmatrix}
\Phi(T,0) & 0 \\
I_{nd} & W(T,0)
\end{bmatrix}$$

(8)

**Remark 2.4** For any two matrixes $A$ and $B$, $[A, B]$ is the matrix based on $A$ and $B$ column as the column of it. The matrix (8) is a constant, so the rank is clear.

**Remark 2.5** We can know that if the action matrixes of the followers are all 0, then the rank condition degenerates to the common criterion in linear control theory, it is easy to test the intuition.

**Remark 2.6** If the preference matrixes of the followers $Q_i(t)$ degenerate to 0, that means the followers don’t care about the state of the system, then intuitively the followers will not impose any control, so $u_i(t)(i = 1, 2, \cdots, l)$ are 0, from the system (1), (2), we know that in this case the system degenerate to the common linear system. So the criterion of Theorem 2.1 should be equivalent to the common criterion in linear control theory, it is easy to test the intuition is right.

When the system (1), (2) is linear time-invariant, that is to say, the data $A(t)$, $B_i(t)$, $B(t)$, $Q_i(t)$, $R_i(t)$ are independent of the time $t$, respectively denoted by $A$, $B_i$, $B$, $Q_i$, $R_i$, we can get more convenient criterion.

**Theorem 2.2** Assume the system (1), (2) is time invariant. It is controllable, if and only if the follow matrix rank condition holds:

$$\text{rank}(Q_i) = n$$

(9)

where,

$$Q_i = [B, \quad AB, \quad A^2B + P_i, \quad \cdots, \quad A^{\tau+1}B + P_{i+1n-2}]$$

$$P_i = [I_e \quad 0][A^{\tau+1}B - A^{\tau+1}B]$$

(10)

**Remark 2.7** From the definition of $P_k$, we know that it is a polynomial about $A$, $B$, $Q_i$, $R_i$ and if $Q_i = 0(i = 1, 2, \cdots, l)$ or $R_i = 0(i = 1, 2, \cdots, l)$ then the $P_k = 0(k = 1, 2, \cdots, (l+1)n-2)$. In this case, this rank condition into our familiar Kalman’s controllability rank condition.

**Remark 2.8** From the theorem, in a way we could say that the more followers, the easier to achieve the controllability condition. It seems counterintuitive, but we should(156,268),(803,328) be careful to distinguish the controllable and the degree of difficulty of control. The controllability means there are some right inputs to achieve any terminal state from any initial state, but it’s not mean we can easily find the input or easily put the right input into practice. We can say that to a certain extent the controllability is the inherent attribute of the system, but how easy it is to control the system is more a problem of ability of the controller.

**Theorem 2.3** Assume the system (1), (2) is time invariant. If it is controllable, then we have the follow result: For any $s \in \lambda(\bar{A})$, we have:

$$\text{rank} \left[ \begin{bmatrix} A - sl & B, R_i^{-1}B_i, \cdots, B, R_i^{-1}B_i^T \end{bmatrix} \right] = n$$

(11)

**Remark 2.9** We should note that the condition here is necessary but not sufficient. From this we know even if the system without the followers is controllable, i.e. $(A, B)$ is controllable, the system (1), (2) maybe not controllable when introducing the followers into it.

### 3 Proof of the Theorem

This section give the proofs of the three theorems in section 2.

#### 3.1 Proof of Theorem 2.1

To proof this theorem, we need the following claims. We first define a matrix.

$$W = \int_0^T \Phi(T,s)\bar{B}(s)\Phi^T(T,s)ds \quad \text{is permissible control}$$

(12)

**Claim 3.1** $\text{Im}(W(T,0)) = W$.

Proof: first we show $\text{Im}(W(T,0)) \subset W$. Let $\{e_i, i = 1, \cdots, (l+1)n\}$ be a basis of the space $\mathbb{R}^{(l+1)n}$, then we have the following:

$$W(T,0)e_i = \int_0^T \Phi(T,s)\bar{B}(s)\Phi^T(T,s)ds e_i$$
Claim 3.2 If the assumption 2.1 holds, then:

\[
\text{rank} \left( \begin{bmatrix} I_n, 0 \end{bmatrix} \Phi^{-1}(T, 0) \begin{bmatrix} I_n \ 0 \end{bmatrix} \right) = n \tag{13}
\]

Claim 3.3 If the assumption 2.1 holds, then for any given strategy \( u(t) \) of the leader, the two-point boundary value problem (3) has a unique solution. So the unique solution is the optimal strategy trajectory and optimal strategy of the followers.

Proof: Because assumption 2.1, (3) must has solution. We only need to prove the unique. From the unique solution of ODE, we only need to proof the terminal state \( X(T) = X_T \) is unique. If there is another solution \( X'(t) \) of (3) such that \( X_T' = X'(T) \neq X_T \), then we have:

\[
v \triangleq \int_0^T \Phi(T, s)\bar{B}(s)u(s)ds = 0 \Rightarrow [I_n, 0] \Phi(T, 0)^{-1} \begin{bmatrix} I_n \\
0 \end{bmatrix} (X_T - X_T') \Rightarrow X_T = X_T' \text{ because of claim 3.2}
\]

So the claim 3.3 holds.

Now we can proof the theorem 2.1. First we proof the rank condition is sufficient. Let

\[
W_1 = \Phi(T, 0) \begin{bmatrix} 0 \\
I_{(t+1)n} \end{bmatrix}, W_2 = W(T, 0), W = (W_1, W_2).
\]

the rank condition is equivalent to W is a matrix of full row rank, so W has right invertible element, denoted by

\[
W^{-1} = \begin{bmatrix} W_1^{-1} \\
W_2^{-1} \end{bmatrix},
\]

i.e. \( WW^{-1} = I_{(t+1)n} \) or \( W_1 W_1^{-1} + W_2 W_2^{-1} = I_{(t+1)n} \), for any initial state \( X_0 \) and any terminal state \( X_T \), let

\[
u(t) = \bar{B}(t)\Phi(T, t)LW_2^{-1}(\Phi(T, 0)[-X_0 \ 0] + [X_T \ 0]) \in \mathbb{R}^n
\]

\[
\Phi(0, T) \begin{bmatrix} X_0 \\
0 \end{bmatrix} + \int_0^T \Phi(T, s)\bar{B}(s)u(s)ds = \Phi(0, T) \begin{bmatrix} X_0 \\
0 \end{bmatrix} + \Phi(T, 0) \begin{bmatrix} X_T \\
0 \end{bmatrix} + \int_0^T \Phi(T, s)\bar{B}(s)u(s)ds
\]

From the remark 2.3 and claim 3.3, we know that the system (1), (2) is controllable.

To proof the necessary of the rank condition, let \( X_T = 0 \), because of the controllability of the system, for any initial state \( X_0 \in \mathbb{R}^n \), \( \exists \varphi_0 \in \mathbb{R}^n \) such that

\[
\Phi(T, 0) \begin{bmatrix} X_0 \\
0 \end{bmatrix} + \int_0^T \Phi(T, s)\bar{B}(s)u(s)ds = \begin{bmatrix} 0 \\
0 \end{bmatrix}
\]

\[
\forall X_0 \in \mathbb{R}^n,
\]

\[
\Phi(T, 0) \begin{bmatrix} X_0 \\
0 \end{bmatrix} = \Phi(T, 0) \begin{bmatrix} 0 \\
-\varphi_0 \end{bmatrix} + \int_0^T \Phi(T, s)\bar{B}(s)u(s)ds
\]
\[ \Rightarrow \text{Im}(\Phi(T,0) \begin{bmatrix} I_n \\ 0 \end{bmatrix}) \subset \text{Im}(\Phi(T,0) \begin{bmatrix} 0 \\ I_{nl} \end{bmatrix}, W) \]

Because of the invertibility of $\Phi(T,0)$, we have
\[ \text{Im} \left( \Phi(T,0) \begin{bmatrix} I_n \\ 0 \end{bmatrix} \right) \supset \text{Im} \left( \Phi(T,0) \begin{bmatrix} 0 \\ I_{nl} \end{bmatrix} \right) = \mathbb{R}^{(l+1)n} \]
\[ \Rightarrow \left[ \Phi(T,0) \begin{bmatrix} 0 \\ I_{nl} \end{bmatrix}, W \right] \text{ is full rank.} \]

By the claim 3.1, we get the result:
\[ \text{Im}(\Phi(T,0) \begin{bmatrix} 0 \\ I_{nl} \end{bmatrix}, W(T,0)) \text{ is full rank} \]

Hence theorem 2.1 is true.

### 3.2 Proof of Theorem 2.2

We need the following results on linear time-invariant system, which can be found [6].

**Claim 3.4**
\[ \text{Im}(\begin{bmatrix} B, AB, A^2B, \ldots, A^{(l+1)n-1}B \end{bmatrix}) = W, \text{ W is the invariant subspace of } A. \]

**Proof of the theorem 2.2.**

By the linear control theory, we get $\Phi(T,0) = e^{A_T}$. Because $W$ is invariant subspace of $A$, $W$ is also invariant subspace of $A^n (m \in N)$. We have a finite representation of matrix exponential function as follow, which can be found in advanced linear algebra tutorial or [7]:
\[ \exists \alpha(T) \in C^1(\mathbb{R}), \text{s.t. } e^{A_T} = \sum_{k=0}^{(l+1)n-1} \alpha_k(T)A^k \quad (14) \]

From this representation, we know $W$ is also invariant subspace of $\Phi(T,0)$, and because of the invertibility of $\Phi(T,0), \Phi(T,0)W = W$ and $\Phi^{-1}(T,0)W = W$.

From theorem 2.1 and the claim 3.4 the controllability equivalent to the full rank condition:
\[ \text{rank } \left( \Phi(T,0) \begin{bmatrix} 0 \\ I_{nl} \end{bmatrix}, W \right) = (l + 1)n \]
\[ \Rightarrow \text{rank } \left( \Phi(T,0) \begin{bmatrix} 0 \\ I_{nl} \end{bmatrix}, \Phi(T,0)W \right) = (l + 1)n \]
\[ \Rightarrow \text{rank } \left( \begin{bmatrix} 0 \\ I_{nl} \end{bmatrix}, \Phi^{-1}(T,0)W \right) = (l + 1)n \]
\[ \Rightarrow \text{rank } \left( \begin{bmatrix} 0 \\ I_{nl} \end{bmatrix}, W \right) = (l + 1)n \]
\[ \Rightarrow \text{rank } (I_{nl}, 0)W = n \]

It is easy to calculate the matrix $[I_{nl}, 0]W$ is just the matrix in theorem 2.2.

Hence theorem 2.2 holds.

### 3.3 Proof of Theorem 2.3

Assume that the system is controllable.

If there is a $s_0 \in \lambda(A)$ which does not satisfy the condition in the theorem 2.3, i.e.
\[ \text{rank } (A - s_0I_n, B_1R_1^{-1}B_1^T, \ldots, B_lR_l^{-1}B_l^T, B)) < n \]
\[ \Rightarrow \exists 0 \neq z \in \mathbb{C}^n, \text{s.t.} \]
\[ z^T(A - s_0I_n, B_1R_1^{-1}B_1^T, \ldots, B_lR_l^{-1}B_l^T, B) = 0 \]

By the definition of $A$

\[ \Rightarrow z^T[I_n, 0](A - s_0I_{(l+1)n}) = 0 & z^TB = 0 \]
\[ \Rightarrow [Z]^T \left( A - s_0I_{(l+1)n} \right) = 0 & [Z]^TB = 0 \]
\[ \Rightarrow [Z]^T \left( \begin{bmatrix} A^k & 0 \\ 0 & 1 \end{bmatrix}B \right) = 0, k = 0, 1, \ldots, (l + 1)n - 1 \]
\[ \Rightarrow [Z]^T \left( B, AB, A^2B, \ldots, A^{(l+1)n-1}B \right) = 0 \]
\[ \Rightarrow z^T[I_n, 0]B, AB, A^2B, \ldots, A^{(l+1)n-1}B = 0 \]

By the definition of $Q_c$ in theorem 2.2

\[ \Rightarrow z^TQ_c = 0 & 0 \neq z \in \mathbb{C}^n \]
\[ \Rightarrow \text{rank } (Q_c) < n \]

This is contradiction against the theorem 2.2.

So the theorem 2.3 holds.

### 4 Conclusion

In this paper, we have defined and analyzed the controllability of a class of dynamical games. As a starting point, we have investigated the non-cooperative linear-quadratic dynamical games with perfect information. Some necessary and sufficient conditions are given in this paper. This research direction gives a new inherent attribute of the kind of dynamical games which can be used to model and to analyze other complex systems. But a number of problems still remain open in this direction. For example, what is the controllability to meet the Assumption 2.1? How to find or to solve the strategy of the leader? How to deal with the case where the dynamics of the game is stochastic differential equation? How about the corresponding results in nonlinear systems? These problems need to be investigated in the future.

### References


