Dynamic Games with Dynamic Uncertainty

Renren Zhang*

Abstract

Uncertainty has been studied a lot in both control theory and game theory. Especially in control theory, how to use feedback to deal with uncertainty is the core problem. The paper initially explores dynamic games with dynamic uncertainty from the perspective of control theory. As a starting point toward investigating the problem, the easy-to-understand boolean dynamic game described by a boolean mapping f is selected and the dynamic uncertainty is represented by a set of boolean mapping $\mathcal{F}$. All game participants know the dynamics of the system f is in $\mathcal{F}$ but do not know which one it is. To get some initial insight to the problem, it is formulated as a Bayesian game. Some results of the existence and computation of the Nash equilibrium are given in the paper.

1 Introduction

Incomplete information has been widely studied in both control theory and game theory, because incomplete information appears in a large number in practice. In the classical dynamic games with incomplete information [1], the rational agents have the uncertainty about the preferences of the other players which fundamentally affect the agents' strategic behaviours. In these game, the nature first choose the types or preferences of the players and then the agents choose their strategies to maximize their own payoff functions (More details can be seen, e.g., [2, 3, 4] and references therein). In control theory, the primary objective of feedback a basic concept is to reduce the effects of the plant uncertainty. And the relationship between structure uncertainty and feedback mechanism is a longstanding fundamental issue in control theory [5]. There are some fundament results about this relationship in the literature (see, e.g., [6, 5, 7, 8]).

The concepts of strategy in game theory and feedback in control theory are essentially the same, which are the mappings from information set to action set. Therefore, the methods in these two fields may be used for reference by the other party. In this paper, we deal with the Nash equilibrium of the dynamic games with dynamic uncertainty from the perspective of control theory. We assume that the dynamic of the system is described by a function f from the state and the control of the rational agents spaces to the state space. The dynamic uncertainty is described by a functions set $\mathcal{F}$. All game participants know the dynamics of the system f is in $\mathcal{F}$ but do not know which one it is. As a starting point toward investigating the problem, the easy-to-understand boolean dynamic game described by a boolean mapping f is selected and the dynamic uncertainty is represented by a set of boolean mapping $\mathcal{F}$. First we will consider the optimal control problem of this uncertain dynamic systems, and then the corresponding results will be used to analysis the Nash equilibrium of the dynamic games with dynamic uncertainty.

The rest of the paper is organized as follows. In Section 2, the optimal control with dynamic uncertainty is formulated and the optimal control of the finite-horizon and infinite-horizon of the problem are given. In Section

*Research Center for Mathematics and Interdisciplinary Sciences, Shandong University, Qingdao, Shandong Province, 266237, China(e-mails: rrz@sdu.edu.cn). This work was supported in part by the Strategic Research Program of Chinese Academy of Sciences under Grant No. XDA27000000.

3, the dynamic game with dynamic uncertainty is formulated and the Nash equilibrium of the finite-horizon case is given. Section 4 will conclude the paper with some remarks.

2 Optimal control with dynamic uncertainty

2.1 Problem formulation

Consider the following optimal control problem. Let $X = \{1, \cdots, n\}$, $U = \{1, \cdots, m\}$, $F \subseteq AF = X^{X \times U}$ be a non-empty subset, and p_0 be a probability distribution on F. The dynamic of the system is described by

$$\begin{cases} x(t+1) = f(x(t), u(t)), \\ \quad x(0) = x_0 \in X, \ t = 0, \cdots, T-1, \end{cases} \tag{1}$$

where f is an unknown function, but the controller knows that $f \in F$ and its probability is $p_0(f)$. The payoff function is

$$J(u(\cdot)) = K(x(T)) + \sum_{t=0}^{T-1} S(x(t), u(t)). \tag{2}$$

The controller wants to maximize the expectation of the payoff function.

We assume that the controller can observe the system state at any time t. Define the history at time t as $H_t = \{(x_0, u_0), \cdots, (x_{t-1}, u_{t-1}), x_t\}$ and the history as $H = \cup_{t \in [T-1]} H_t$. Hence, the admissible controls set is

$$\mathcal{U} = U^H = \{u = (u(t))_{t \in [T-1]} : H \to U\}. \tag{3}$$

Here we only consider the deterministic control strategies.

2.2 Results

The section we will give the characterization of the value function and optimal control.

First, we consider the finite-horizon problem, i.e., $T < +\infty$. In this case we know that the optimal control problem is well-defined, because the admissible control strategies are finite and the value of any admissible control is finite.

Define $V(t, x, p)$ is the optimal value when the system starts from time t, state x, and the probability distribution on F is p. By the fact above we know that the value function exists. If at time t, the system state is x, probability distribution on F is p, and the state at time $t+1$ is y when the control input is u, then we can get that the posterior probability distribution on F at time $t+1$ is $p(x, y; p, u)$ which can be calculated by the following:

1. If $p(f) = 0$, then $p(x, y; p, u)(f) = 0$;

2. If $p(f) \neq 0$ but $f(x, u) \neq y$, then $p(x, y; p, u)(f) = 0$;

3. Denote all the function which do not satisfy 1 and 2 as set $F(t+1)$. For any $f \in F(t+1)$,

$$p(x, y; p, u)(f) = \frac{p(f)}{\sum_{\overline{f} \in F(t+1)} p(\overline{f})}.$$

Because $f \in F$, we know that $F(t+1) \neq \emptyset$, $p(x, y; p, u)$ is well-defined.

Now we can get the principle of optimality theorem as follows.

Theorem 1 *The value function $V(t,x,p)$ satisfies the equation*

$$V(t,x(t),p) = \max_{u\in U}\Big\{S(x(t),u) + \sum_{y\in X}\mu(x(t),u,y;p)V(t+1,y,p(x(t),y;p,u))\Big\}, \tag{4}$$

and $V(T,x,p) = K(x)$, where $\mu(x,u,y;p) = \sum_{f\in F, f(x,u)=y} p(f)$. The optimal control u^* can be get by the following:

$$u^*(t,x(t),p(t)) = arg\max_{u\in U}\Big\{S(x(t),u) + \sum_{y\in X}\mu(x(t),u,y;p)V(t+1,y,p(x(t),y;p,u))\Big\}, \tag{5}$$

where $p(t) = p(x(t-1),x(t);p(t-1),u(t-1))$ when $t\geq 1$ and $p(0) = p_0$.

Proof: This can be proved by the dynamic programming method and we omit here.

Remark 2.1 *From the above theorem, we can calculate the value function $V(t,x,p)$ going from backward. Hence, we can get the optimal value of the original problem and the optimal control.*

Remark 2.2 *We should note that the control strategy is not stationary, the control $u(t)$ at time t depends on all the inputs and states before time t. This because the controller have to identities the system dynamic f. This is different form the classical optimal control.*

If the system degenerates to complete information, i.e., the prior probability p_0 belongs to the set $\{e_1,\cdots,e_{|F|}\}\in R^{|F|}\}$, then we can verify that Theorem 1 degenerates to the classical principle of optimality.

Corollary 2.1 *If $p_0(f) = 1$ for some $f\in F$, then Equation (4) in Theorem 1 can be simplified to the following form:*

$$V(t,x(t)) = \max_{u\in U}\Big\{S(x(t),u) + V(t+1,f(x(t),u))\Big\}, \tag{6}$$

which is the classical principle of optimality.

Proof: Because $p_0(f) = 1$, we get that $\mu(x,u,y;p) = 1$, $p(x(t),y;p,u)(f) = 1$ and $p(x(t),y;p,u)(f') = 0$ for all other function $f'\in F$.

If the time is infinite, what is the optimal control?

Consider the infinite-horizon optimal control problem and assume that the system dynamic model is the same as (1), but the payoff function is

$$J(u(\cdot)) = \sum_{t=0}^{\infty}\alpha^t S(x(t),u(t)). \tag{7}$$

where $\alpha\in(0,1)$ is the discount factor.

First we know that the payoff is finite for any control strategy, because X and U are finite set and

$$|\sum_{t=0}^{\infty}\alpha^t S(x(t),u(t))| \leq \sum_{t=0}^{\infty}\alpha^t|S(x(t),u(t))| \leq \sum_{t=0}^{\infty}\alpha^t K \leq \infty.$$

Now we give a similar principle of optimality of the infinite-horizon case.

Theorem 2 *The following equation has a unique solution $V^*(x,p)$:*

$$V(x,p) = \max_{u\in U}\Big\{S(x,u) + \alpha\sum_{y\in X}\mu(x,u,y;p)V(y,p(x,y;p,u))\Big\}, \tag{8}$$

where the probability distribution μ is calculated as follows:

$$\mu(x,u,y;p) = \sum_{f\in F, f(x,u)=y} p(f). \tag{9}$$

The optimal control problem (1) and (8) has a solution and its value function is $V^*(x,p)$.

Proof: First, we proof Equation (9) has a unique solution.

Let $B(X \times \Delta(AF))$ be the set of all bounded functions on compact set $X \times \Delta(AF)$ Define the following operation:

$$\begin{aligned} &T : B(X \times \Delta(AF)) \to B(X \times \Delta(AF)) \\ &T(v(x,p)) = \max_{u \in U}\Big\{S(x,u) + \alpha \sum_{y \in X} \mu(x,u,y;p)v(y,p(x,y;p,u))\Big\}. \end{aligned} \tag{10}$$

It is obvious that the operator is well-defined.

(1) T is monotonous, i.e., for any $u \le v$, we have $T(u) \le T(v)$.

Because for any $v_1 \le v_2$,

$$\begin{aligned} &S(x,u) + \alpha \sum_{y \in X} \mu(x,u,y;p)v_1(y,p(x,y;p,u)) \\ \le\ &S(x,u) + \alpha \sum_{y \in X} \mu(x,u,y;p)v_2(y,p(x,y;p,u)), \end{aligned}$$

we have $T(v_1) \le T(v_2)$.

(2) For any constant $c \in R$ and $v \in B(X \times \Delta(AF))$, it is easily to prove that $T(v + c) = T(v) + \alpha c$. Because $v_1 - \|v_1 - v_2\|_\infty \le v_2 \le v_1 + \|v_1 - v_2\|_\infty$, we have

$$\begin{aligned} T(v_1) - \alpha\|v_1 - v_2\|_\infty &= T(v_1 - \|v_1 - v_2\|_\infty) \\ &\le T(v_2) \le T(v_1 + \|v_1 - v_2\|_\infty) \\ &= T(v_1) + \alpha\|v_1 - v_2\|_\infty, \end{aligned}$$

and so we get $-\alpha\|v_1 - v_2\|_\infty \le T(v_2) - T(v_1) \le \alpha\|v_1 - v_2\|_\infty$. It means the following relation holds:

$$\|T(v_1) - T(v_2)\|_\infty \le \alpha\|v_1 - v_2\|_\infty,$$

i.e., T is a contraction operator on Banach space $B(X \times \Delta(AF))$. Hence, there is a unique bounded function $V^* \in B(X \times \Delta(AF))$ that satisfies $T(V^*) = V^*$, so Equation (9) has a unique solution.

By the dynamic programming and V^*, we can construct a control strategy u^* as follows:

$$u^*(t,x(t),p(t)) = arg \max_{u \in U}\Big\{S(x(t),u) + \sum_{y \in X} \mu(x(t),u,y;p)V^*(y,p(x(t),y;p,u))\Big\},$$

where $p(t) = p(x(t-1),x(t);p(t-1),u(t-1))$ when $t \ge 1$ and $p(0) = p_0$. It is easy to verify that u^* is an optimal control.

The proof is complete.

For the infinite time optimal control problem, we can give some properties of the optimal control strategy.

Theorem 3 *For any prior probability p_0 and initial state x_0, if $\{u_t^*(\cdot) : t \ge 0\}$ is an optimal control and $\{p_t^*(\cdot) : t \ge 0\}$ is the corresponding probability estimation, then there is a sufficient large time LT and $p \in \Delta(F)$ such that $p_t^* = p$ for any $t \ge LT$ and u_t^* is periodic variation or remain unchanged.*

Proof: If there is not uncertainty, by the finite of the state and control spaces, the optimal control u_t^* is periodic variation or constant. Because the possible dynamics f are finite and p_t changed only if some possible functions are excluded, there must exists $LT > 0$ and $p \in \Delta(F)$ such that $p_t^* = p$ for any $t \ge LT$. No possible functions are excluded after time LT and the system dynamics are deterministic on the optimal trajectory. Hence after time LT, the optimal control has the same properties as in the deterministic case.

Remark 2.3 *We should note that the controller may not know the true dynamic of the system all the time, but he will constraint his control set under which he know how the system responds.*

3 Dynamic games with dynamic uncertainty

3.1 Problem formulation

Now we consider the dynamic games with dynamic uncertainty with L players. Let $X = \{0, \cdots, n\}$, $U_i = \{0, \cdots, m_i\}(i = 1, \cdots, L)$, $U = \prod_{i=1}^{L} U_i$, $F \subset X^{X \times U}$ be a non-empty subset, and p_0 be a probability distribution on F. The dynamic of the system is

$$\begin{cases} x(t+1) = f(x(t), u_1(t), \cdots, u_L(t)), \\ \quad x(0) = x_0 \in X,\ t = 0, \cdots, T-1, \end{cases} \tag{11}$$

where f is an unknown function, but the controller knows that $f \in F$ and its probability is $p_0(f)$. The payoff function of player i is

$$J_i(u_i(\cdot), u_{-i}(\cdot)) = K_i(x(T)) + \sum_{t=0}^{T-1} S_i(x(t), u_i(t), u_{-i}(t)). \tag{12}$$

Each player wants to maximize the expectation of its own payoff functions. All the information of the game is common knowledge. Denote the game as $G(0, x_0, p_0)$.

We assume that each player can observe the system state at any time and the actions of all players before this time, i.e, the only incomplete information is system dynamic f. At each time t, all players act in a predetermined sequence. In this case, how about the Nash equilibrium and how to calculate it when it exists?

3.2 Results

Just like the finite game, there may not be Nash equilibrium when we only consider the pure strategies. Hence, we use mixed strategies whose definition is obvious. This is a dynamic game with dynamic uncertainty and we use the concept of subgame perfect Nash equilibrium.

Define the subgame $G(t, x, p)$ as the game whose payoff functions are the same as (13) and the dynamic is the same as (12) but the start time is t, start state is x and the probability distribution on F is p. Denote $V_i(t, x, p)$ as the value of agent i if the game $G(t, x, p)$ has a unique subgame perfect Nash equilibrium.

Theorem 4 *The dynamic game with dynamic uncertainty of dynamic has at least one subgame perfect Nash equilibrium.*

Proof: We will find one Nash equilibrium by backward induction on the dynamic game.
First, we consider the last stage subgame $G(T-1, x, p)$. This is a classical incomplete static game where the nature first randomly choose a function f from set F according to probability p and then the players choose their own control. Hence, there is at least one Nash equilibrium of game $G(T-1, x, p)$. We choose any one of them and denote $V_i(T-1, x, p)$ as the value of agent i of this equilibrium. At time $T-2$, the players face a dynamic game $G(T-2, x, p)$. According to the backward induction method, we can transform the dynamic game to a classical incomplete static game $\overline{G}(T-2, x, p)$. The same as game $G(T-1, x, p)$, the nature first randomly choose a function f from set F according to probability p at game $\overline{G}(T-2, x, p)$. Player i's payoff function of game $\overline{G}(T-2, x, p)$ is constructed as follows.

$$\overline{J}_i(u_i, u_{-i}; T-2, x, p) = S_i(x, u_i, u_{-i}) + \sum_{y \in X} \mu(x, u, y; p) V_i(T-1, y, p(x, y; p, u)) \tag{13}$$

where $u = (u_1, \cdots, u_L)$, μ is calculated as

$$\mu(x, u, y; p) = \sum_{f \in F : f(x,u) = y} p(f) \tag{14}$$

and $p(x, y; p, u)$ can be calculated by the following;

1. If $p(f) = 0$, then $p(x, y; p, u)(f) = 0$;
2. If $p(f) \neq 0$ but $f(x, u) \neq y$, then $p(x, y; p, u)(f) = 0$;
3. Denote all the function which do not satisfy 1 and 2 as set $F(t+1)$. For any $f \in F(t+1)$,

$$p(x, y; p, u)(f) = \frac{p(f)}{\sum_{\overline{f} \in F(t+1)} p(f)}.$$

The the actions are observable for any players, so the payoff functions is well-defined. Game $\overline{G}(T-2, x, p)$ and so game $G(T-2, x, p)$ has at least one Nash equilibrium. We choose any one of them and denote $V_i(T-2, x, p)$ as the value of agent i of this equilibrium. Using the same argument, we know that subgame $G(t, x, p)$ has at least one Nash equilibrium for any $t \in \{0, 1, \cdots, T-1\}$. Hence, the proof is complete.

4 Conclusion

In this paper, we have investigated the optimal control problem and dynamic games with dynamic uncertainty. The motivation for studying the games comes from rich situations in the real world, where the rational agents or controller do not have all information about the system dynamics. The study hope to shed some light on the relationship between the Nash equilibrium and the dynamic uncertainty. In this paper, we have first study the optimal control problem with dynamic uncertainty, and then using the corresponding results to investigate the dynamic games with dynamic uncertainty. The existence and calculation of the Nash equilibrium have been given by using dynamic programming. It is obvious that there are many interesting questions to be studied such as the continuous state and control spaces.

References

[1] Weiying Zhang. *Game and information economics.* Truth and Wisdom Press, 1996.

[2] Eric Rasmusen. *Game and information: An introduction to game theory.* Blackwell Publishing, 4 edition, 2006.

[3] Fudenberg D. and Jean Tirole. Perfect Bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory*, 53:236–260, 1991.

[4] Leland H. and D. Pyle. Informational asymmetries, financial structure, and financial intermediation. *Journal of Financial*, 32:371–387, 1977.

[5] Liangliang Xie and Lei Guo. How much undertainty can be dealt with by feedback. *IEEEE Transactions on Automatic Control*, 45(12):2203–2217, December 2000.

[6] Lei Guo. On critical stability of discrete-time adaptive nonlinear control. *IEEEE Transactions on Automatic Control*, 42(11):1488–1499, 1997.

[7] Chanyin Li, Liangliang Xie, and Lei Guo. A polynomial criterion for adaptive stabilizability of discrete-time nonlinear systems. *Communications in Information and Systems*, 6(4):273–298, 2006.

[8] Lei Guo. Feedback and uncertainty: Some basic problems and results. *Anuual Reviews in Control*, 49:27–36, 2020.