Game Theory and its Applications

Part VIII: Differential Games

微分博弈论简介

清 华 大 学 数 学 科 学 系　　谢金星

办公室：理科楼1202#　电话：62787812

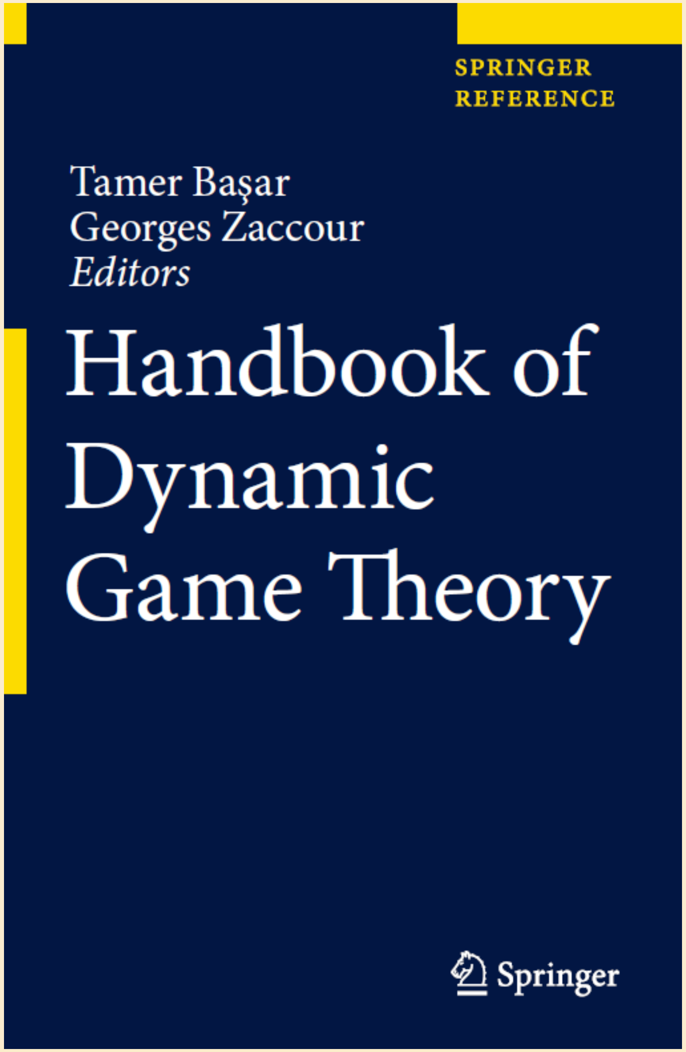
E-mail: xiejx@tsinghua.edu.cn

Outline

- Definition
- Methods

Tamer Basar • Georges Zaccour (Editors), Handbook of Dynamic Game Theory, Springer, 2018.

(over 1200 pages)



Dynamic Games

- Dynamic (state-space) games have been of considerable value to represent time, strategic behavior and interdependencies
- State variables:
  - summarize all relevant consequences of the past history of the game
  - describe the main features of a dynamic system at any instant of time
- Time can be discrete or continuous

Differential Games

- Differential games are games played by agents (*players*) who jointly control (through their actions over time, as inputs) a dynamical system described by differential state equations
  - the game evolves over a continuous-time horizon (with the length of the horizon known to all players, as common knowledge)
  - over this horizon each player is interested in optimizing a particular objective function which depends on the state variable

Example: pursuit escape game

Introduction

- Differential games are offsprings of game theory and optimal control.
- Initiated by R. Isaacs at the Rand Corporation in the late 1950s and early 1960s.
- Initial focal points: military applications and zero-sum games.
- Now, applications are found in many areas, e.g., in management science (operations management, marketing, finance), economics (industrial organization, macro, resource, environmental economics, etc.), biology, ecology, military, etc.
- Textbooks: Başar and Olsder (1982, 1995), Petrosjan (1993), Dockner et al. (2000), Jørgensen and Zaccour (2004), Engwerda (2005), Yeung and Petrosjan (2005), Haurie, Krawczyk and Zaccour (2012).

Elements of a differential game

A deterministic differential game (DG) played on a time interval $[t^0, T]$ involves the following elements:

- A set of players $M = \{1, \ldots, m\}$;
- For each player $j \in M$, a vector of controls $\mathbf{u}_j(t) \in U_j \subseteq \mathbb{R}^{m_j}$, where U_j is the set of admissible control values for Player j;
- A vector of state variables $\mathbf{x}(t) \in X \subseteq \mathbb{R}^n$, where X is the set of admissible states. The evolution of the state variables is governed by a system of differential equations, called the state equations:

$$\dot{\mathbf{x}}(t) = \frac{d\mathbf{x}}{dt}(t) = f\left(\mathbf{x}(t), \underline{\mathbf{u}}(t), t\right), \quad \mathbf{x}(t^0) = \mathbf{x}^0, \tag{1}$$

where $\underline{\mathbf{u}}(t) \triangleq (\mathbf{u}_1(t), \ldots, \mathbf{u}_m(t))$;

Elements of a differential game

- A payoff for Player $j, j \in M$,

$$J_j \triangleq \int_{t^0}^{T} g_j(\mathbf{x}(t), \underline{\mathbf{u}}(t), t)\, dt + S_j(\mathbf{x}(T)) \qquad (2)$$

  where function g_j is Player j's instantaneous payoff and function S_j is his terminal payoff;
- An information structure, i.e., information available to Player j when he selects $\mathbf{u}_j(t)$ at t;
- A strategy set Γ_j, where a strategy $\gamma_j \in \Gamma_j$ is a decision rule that defines the control $\mathbf{u}_j(t) \in U_j$ as a function of the information available at time t.

Elements of a differential game

Assumption: All feasible state trajectories remain in the interior of the set of admissible states X.

Assumption: Functions f and g are continuously differentiable in $\mathbf{x}$, $\underline{\mathbf{u}}$ and t. The S_j functions are continuously differentiable in $\mathbf{x}$.

Control set: $\mathbf{u}_j(t) \in U_j$, with U_j set of admissible controls (or control set).

- Control set could be:
  - Time-invariant and independent of the state;
  - Depend on the *position of the game* $(t, \mathbf{x}(t))$, i.e., $\mathbf{u}_j(t) \in U_j(t, \mathbf{x}(t))$.
  - Depend also on controls of other players (coupled constraints).

Elements of a differential game: Information

Information structure:

Open loop: players base their decision only on time and an initial condition;

Feedback or Markovian: players use the *position of the game* $(t, \mathbf{x}(t))$ as information basis;

Non-Markovian: players use history when choosing their strategies.

Elements of a differential game

Strategies:

Open-loop strategy: selects the control action according to a decision rule μ_j, which is a function of the initial state $\mathbf{x}^0$:
$$\mathbf{u}_j(t) = \mu_j(\mathbf{x}^0, t).$$

As $\mathbf{x}^0$ is fixed, no need to distinguish between $\mathbf{u}_j(t)$ and $\mu_j(\mathbf{x}^0, t)$. Player commits to a fixed time path for his control.

Markovian strategy: selects the control action according to a feedback rule $\mathbf{u}_j(t) = \sigma_j\left(t, \mathbf{x}(t)\right)$. Player j's reaction to any position of the system is predetermined.

- The decision rule σ_j can be, e.g., linear or quadratic function of $\mathbf{x}$ with coefficients depending on t.
- It also can be a nonsmooth function of x and t (e.g., bang-bang controls). Complicated problem....

Elements of a differential game

State equations (*system dynamics*, *evolution equations* or *equations of motion*):

$$\dot{\mathbf{x}}(t) = \frac{d\mathbf{x}}{dt}(t) = f\left(\mathbf{x}(t), \underline{\mathbf{u}}(t), t\right), \quad \mathbf{x}(t^0) = \mathbf{x}^0,$$

- State vector's rate of change depends on $t, \mathbf{x}(t)$ and $\underline{\mathbf{u}}(t)$.
- OL strategies are piecewise continuous in time. A unique trajectory will be generated from $\mathbf{x}^0$.
- For feedback strategies, we make the following simplifying assumption:

Assumption For every admissible strategy vector $\underline{\sigma} = (\sigma_j : j \in M)$, the DE $\dot{\mathbf{x}}(t)$ admit a unique solution, i.e., a unique state trajectory, which is an absolutely continuous function of t.

Assumption met when: (i) $f(\mathbf{x}(t), \underline{\mathbf{u}}(t))$ is continuous in t for each x and $\mathbf{u}_j, j \in M$; (ii) $f\left(\mathbf{x}(t), \mathbf{u}(t), t\right)$ is uniformly Lipschitz in x, $\mathbf{u}_1, \ldots, \mathbf{u}_m$; and (iii) $\sigma_j(t, \mathbf{x})$ is continuous in t for each $\mathbf{x}$ and uniformly Lipschitz in $\mathbf{x}$.

Elements of a differential game

Time horizon:

- T can be finite or infinite;
- T can be prespecified or endogenous (as in, e.g., pursuit-evasion games and patent-race games).

Nash Equilibrium: The definition

- Normal form representation: Set of players' admissible strategies; payoffs expressed as functions of strategies rather than actions.
- Assume that Player $j, j \in M$, maximizes a stream of discounted gains, that is,

$$J_j \triangleq \int_{t^0}^{T} e^{-\rho_j t} g_j(\mathbf{x}(t), \underline{\mathbf{u}}(t), t)\, dt + e^{-\rho_j T} S_j(\mathbf{x}(T)), \tag{3}$$

where ρ_j is the discount rate satisfying $\rho_j \geq 0$.

Nash Equilibrium: The definition

Open-loop Nash equilibrium.
The payoff functions with the state equations and initial data $(t^0, \mathbf{x}^0)$ define the normal form of an OL differential game:

$$\underline{\mathbf{u}}(\cdot) = (\mathbf{u}_1(\cdot), \dots, \mathbf{u}_j(\cdot), \dots, \mathbf{u}_m(\cdot)) \mapsto J_j(t^0, \mathbf{x}^0; \underline{\mathbf{u}}(\cdot)), \quad j \in M. \tag{4}$$

Definition

The control m-tuple $\underline{\mathbf{u}}^*(\cdot) = (\mathbf{u}_1^*(\cdot), \dots, \mathbf{u}_m^*(\cdot))$ is an **open-loop Nash equilibrium** (OLNE) at $(t^0, \mathbf{x}^0)$ if the following holds:

$$J_j(t^0, \mathbf{x}^0; \underline{\mathbf{u}}^*(\cdot)) \geq J_j(t^0, \mathbf{x}^0; [\mathbf{u}_j(\cdot), \underline{\mathbf{u}}^*_{-j}(\cdot)]), \quad \forall \mathbf{u}_j(\cdot), j \in M,$$

where $\mathbf{u}_j(\cdot)$ is any admissible control of Player j and $[\mathbf{u}_j(\cdot), \underline{\mathbf{u}}^*_{-j}(\cdot)]$ is the m-vector of controls obtained by replacing the j-th component in $\underline{\mathbf{u}}^*(\cdot)$ by $\mathbf{u}_j(\cdot)$.

Nash Equilibrium: The definition

Open-loop Nash equilibrium.

Player j solves the optimal-control problem

$$\max_{\mathbf{u}_j(\cdot)} \left\{ \int_{t^0}^{T} e^{-\rho_j t} g_j \left(\mathbf{x}(t), [\mathbf{u}_j(t), \underline{\mathbf{u}}^*_{-j}(t)], t \right) dt + e^{-\rho_j T} S_j(\mathbf{x}(T)) \right\},$$

subject to the state equations

$$\dot{\mathbf{x}}(t) = \frac{d\mathbf{x}}{dt}(t) = f\left(\mathbf{x}(t), [\mathbf{u}_j(t), \underline{\mathbf{u}}^*_{-j}(t)], t \right), \quad \mathbf{x}(t^0) = \mathbf{x}^0. \tag{5}$$

Similarly one can define:

Open-loop Stackelberg Equilibria (OLSE)

Nash Equilibrium: The definition

Markovian (feedback)-Nash equilibrium.
Players use feedback strategies $\underline{\sigma}(t, \mathbf{x}) = (\sigma_j(t, \mathbf{x}) : j \in M)$. The normal form of the game, at $(t^0, \mathbf{x}^0)$ is defined by

$$\begin{aligned} J_j(t^0, \mathbf{x}^0; \underline{\sigma}) &= \int_{t^0}^{T} e^{-\rho_j t} g_j\left(\underline{\sigma}(t, \mathbf{x}), t\right) dt + e^{-\rho_j T} S_j(\mathbf{x}(T)), \\ \dot{\mathbf{x}}(t) &= \frac{d\mathbf{x}}{dt}(t) = f\left(\mathbf{x}(t), \underline{\sigma}(t, \mathbf{x}(t)), t\right), \quad \mathbf{x}(t^0) = \mathbf{x}^0. \end{aligned}$$

Define the $(m-1)-$vector

$$\sigma_{-j}(t, \mathbf{x}(t)) \triangleq (\sigma_1(t, \mathbf{x}(t)), \ldots, \sigma_{j-1}(t, \mathbf{x}(t)), \sigma_{j+1}(t, \mathbf{x}(t)), \ldots, \sigma_m(t, \mathbf{x}(t)))$$

Nash Equilibrium: The definition

Markovian (feedback)-Nash equilibrium.

Definition

The feedback m-tuple $\underline{\sigma}^*(\cdot) = (\sigma_1^*(\cdot), \ldots, \sigma_m^*(\cdot))$ is a **feedback or Markovian-Nash equilibrium** (MNE) on $[0, T] \times X$ if for each $(t^0, \mathbf{x}^0)$ in $[0, T] \times X$, the following holds:

$$J_j(t^0, \mathbf{x}^0; \underline{\sigma}^*(\cdot)) \geq J_j(t^0, \mathbf{x}^0; [\sigma_j(\cdot), \boldsymbol{\sigma}_{-j}^*(\cdot)];), \quad \forall \sigma_j(\cdot), j \in M,$$

where $\sigma_j(\cdot)$ is any admissible feedback law for Player j and $[\sigma_j(\cdot), \boldsymbol{\sigma}_{-j}^*(\cdot)]$ is the m-vector of controls obtained by replacing the j-th component in $\boldsymbol{\sigma}^*(\cdot)$ by $\sigma_j(\cdot)$.

Nash Equilibrium: The definition

Markovian (feedback)-Nash equilibrium.
In other words, $\mathbf{u}_j^*(t) \equiv \sigma_j^*(t, \mathbf{x}^*(t))$, where $\mathbf{x}^*(\cdot)$ is generated by $\underline{\sigma}^*$ from $(t^0, \mathbf{x}^0)$, solves the optimal-control problem

$$\max_{\mathbf{u}_j(\cdot)} \left\{ \int_{t^0}^{T} e^{-\rho_j t} g_j\left(\mathbf{x}(t), \left[\mathbf{u}_j(t), \sigma_{-j}^*(t, \mathbf{x}(t))\right], t\right) dt \right.$$
$$\left. + e^{-\rho_j T} S_j(\mathbf{x}(T)), \right\}$$
$$\text{s.t.} \quad \dot{\mathbf{x}}(t) = f\left(\mathbf{x}(t), \left[\mathbf{u}_j(t), \sigma_{-j}^*(t, \mathbf{x}(t))\right], t\right), \; \mathbf{x}(t^0) = \mathbf{x}^0.$$

Similarly one can define:

Feedback Stackelberg Equilibria (FSE)

Outline

- Definition
- Methods

How to solve the game?

There exist two main approaches to optimal control and dynamic games:

- 1. via the Calculus of Variations (making use of the Maximum Principle)
- 2. via Dynamic Programming (making use of the Principle of Optimality)

[Note] Based on lecture notes by Claire J. Tomlin (UC Berkeley).

Calculus of Variations

t_0 is the initial time (fixed), t_f the final time (free):

minimizing $J = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} L(x(t), u(t), t)dt$

$$\dot{x} = f(x, u, t) \qquad x(t) \in \mathbb{R}^n \qquad u \in \mathbb{R}^{n_i}$$

$\psi(x(t_f), t_f) = 0 \qquad \psi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^p$ is a smooth map

using the Lagrange multipliers $\lambda \in \mathbb{R}^p, p(t) \in \mathbb{R}^n$,

$$\tilde{J} = \phi(x(t_f), t_f) + \lambda^T \psi(x(t_f), t_f) + \int_{t_0}^{t_f} \left[\underline{L(x, u, t) + p^T (f(x, u, t)} - \dot{x})\right] dt$$

Hamiltonian H(x, u, p, t)

The so-called Legendre transformation

Calculus of Variations

assuming independent variations in $\delta u(), \delta x(), \delta p(), \delta\lambda$, and δt_f:

$$\begin{aligned}
\delta \tilde{J} = & \ (D_1\phi + D_1\psi^T\lambda)\delta x|_{t_f} + (D_2\phi + D_2\psi^T\lambda)\delta t|_{t_f} + \psi^T\delta\lambda \\
+ & \ (H - p^T\dot{x})\delta t|_{t_f} \\
+ & \ \int_{t_0}^{t_f} \left[D_1 H\delta x + D_3 H\delta u \underline{- p^T\delta\dot{x}} + (D_2 H^T - \dot{x})^T\delta p\right] dt
\end{aligned}$$

$D_i H$ stands for the derivative of H with respect to the i th argument.

Integrating by parts for $\int p^T\delta\dot{x}dt$ yields

$$\begin{aligned}
\delta \tilde{J} = & \ (D_1\phi + D_1\psi^T\lambda - p^T)\delta x(t_f) + (D_2\phi + D_2\psi^T\lambda + H)\delta t_f + \psi^T\delta\lambda \\
+ & \ \int_{t_0}^{t_f} \left[(D_1 H + \dot{p}^T)\delta x + D_3 H\delta u + (D_2^T H - \dot{x})^T\delta p\right] dt
\end{aligned}$$

An extremum of $\tilde{J}$ is achieved when $\delta\tilde{J} = 0$ for all independent variations

Necessary conditions for optimality

Table 1

Description	Equation	Variation
Final State constraint	$\psi(x(t_f), t_f) = 0$	$\delta\lambda$
State Equation	$\dot{x} = \frac{\partial H}{\partial p}^T$	δp
Costate equation	$\dot{p} = -\frac{\partial H}{\partial x}^T$	δx
Input stationarity	$\frac{\partial H}{\partial u} = 0$	δu
Boundary conditions	$D_1\phi - p^T = -D_1\psi^T\lambda\vert_{t_f}$	$\delta x(t_f)$
	$H + D_2\phi = -D_2\psi^T\lambda\vert_{t_f}$	δt_f

Final state
$\psi(x(t_f), t_f) = 0$

transversality condition
(横截条件)

此外：boundary conditions $x(t_0) = x_0$

… Written explicitly as

$$\dot{x} = \frac{\partial H}{\partial p}^T(x, u^*, p)$$
$$\dot{p} = -\frac{\partial H}{\partial x}^T(x, u^*, p)$$
$$\frac{\partial H}{\partial u}(x, u^*, p) = 0$$

The key point:
functional minimization problem
➔
Static optimization problem on the function H(x, u, p, t)

$$u^*(t) = \underset{\{\text{ min over u }\}}{\operatorname{argmin}} \quad H(x^*(t), u, p^*(t), t)$$

When they are also sufficient conditions?

e.g. the convexity of the Hamiltonian H(x; u; p; t) in u.

$n_i \times n_i$ Hessian matrix $D_2^2 H(x, u, p, t)$ be positive definite

Example: producer-consumer game

Let p(t) denote the price of a good at time t.

- the good can be produced by a player, at rate $u_1(t)$,
- and consumed by the other player at rate $u_2(t)$.

In a very simplified model, the variation of the price in time can be described by the differential equation

$$\dot{p} = (u_2 - u_1)p, \qquad (4.19)$$

$$c(s) = \frac{s^2}{2}, \quad \phi(s) = 2\sqrt{s}.$$

Payoffs:

$$J^{prod} = \int_0^T \Big[p(t)u_2(t) - c(u_1(t))\Big]\,dt, \qquad (4.20)$$

$$J^{cons} = \int_0^T \Big[\phi(u_2(t)) - p(t)u_2(t)\Big]\,dt. \qquad (4.21)$$

Example: solving the game

Steps 1-2:

STEP 1: the optimal controls are determined in terms of the adjoint variables:

$$u_1^\sharp(x, q_1, q_2) \;=\; \operatorname*{argmax}_{\omega \geq 0} \left\{ q_1 \cdot (-\omega p) - \frac{\omega^2}{2} \right\} \;=\; -\, q_1 p \,,$$

$$u_2^\sharp(x, q_1, q_2) \;=\; \operatorname*{argmax}_{\omega \geq 0} \left\{ q_2 \cdot (\omega p) + 2\sqrt{\omega} - p\omega \right\} \;=\; \frac{1}{(1 - q_2)^2\, p^2}\,.$$

Notice that here we are assuming $p > 0$, $q_1 \leq 0$, $q_2 < 1$.

Example: solving the game

STEP 2: the state $p(\cdot)$ and the adjoint variables $q_1(\cdot), q_2(\cdot)$ are determined by solving the boundary value problem

$$\left\{\begin{array}{rcccl} \dot{p} &=& (u_2^\sharp - u_1^\sharp)p &=& \dfrac{1}{(q_2-1)^2\, p} + q_1 p^2\,, \\[2ex] \dot{q}_1 &=& -q_1(u_2^\sharp - u_1^\sharp) - u_2^\sharp &=& -\, q_1^2\, p - \dfrac{q_1+1}{(1-q_2)^2\, p^2}\,, \\[2ex] \dot{q}_2 &=& -q_2(u_2^\sharp - u_1^\sharp) + u_2^\sharp &=& -\, q_1 q_2 p + \dfrac{1}{(1-q_2)p}\,, \end{array}\right. \tag{4.23}$$

with initial and terminal conditions

$$\left\{\begin{array}{rcl} x(0) &=& x_0\,, \\ q_1(T) &=& 0\,, \\ q_2(T) &=& 0\,. \end{array}\right. \tag{4.24}$$

Dynamic Programming

t_0 is the initial time (fixed), t_f the final time (free):

minimizing $J = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} L(x(t), u(t), t)dt$

$$\dot{x} = f(x, u, t) \qquad x(t) \in \mathbb{R}^n \qquad u \in \mathbb{R}^{n_i}$$

$$\psi(x(t_f), t_f) = 0 \qquad \psi : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^p \text{ is a smooth map}$$

Define the "cost-to-go":

$$J(x(t), t) = \phi(x(t_f), t_f) + \int_t^{t_f} L(x(\tau), u(\tau), \tau)d\tau$$

Define the "optimal Hamiltonian":

$$H^*(x, p, t) := H(x, u^*, p, t)$$

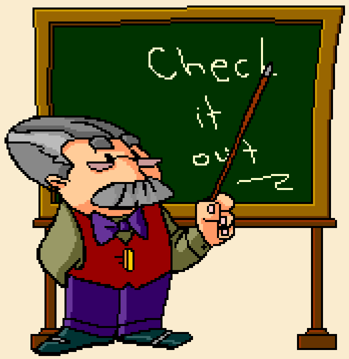
Hamilton Jacobi Bellman equation

Theorem:

Consider, the time varying optimal control problem of (2) with fixed endpoint t_f and time varying dynamics. If the optimal value function, i.e. $J^*(x(t_0), t_0)$ is a smooth function of x, t, then $J^*(x, t)$ satisfies the Hamilton Jacobi Bellman *partial differential equation*

$$\frac{\partial J^*}{\partial t}(x,t) = -H^*\left(x, \frac{\partial J^*}{\partial x}(x,t), t\right) \tag{19}$$

with boundary conditions given by $J^*(x, t_f) = \phi(x, t_f)$ for all $x \in \{x : \psi(x, t_f) = 0\}$.



作业

内容 阅读有关材料（自选）

GAME NEVER OVER