Dynamic Game

Overview

	Single Agent	Multi Agent
Static	Static optimization	Static Game
Dynamic	**Dynamic Optimization**	Dynamic Game

Action space

Time space

Model based	Finite	Infinite
Discrete	Discrete time MDP $P(s_{t+1} \mid s_t, a_t)$	Discrete-time dynamic system $x_{t+1} = f(x_t, u_t)$
Continuous	Continuous time MDP $P(s_{t+h} \mid s_t, a_t)$	Continuous-time dynamic system $\dot{x}_t = f(x_t, u_t)$

Overview

	Single Agent	Multi Agent
Static	Static optimization	Static Game
Dynamic	**Dynamic Optimization**	Dynamic Game

Action space

Model free	Finite	Infinite
Discrete	Value-based Reinforcement Learning	Policy-based Reinforcement Learning
Continuous		

Time space

Overview

	Single Agent	Multi Agent
Static	Static optimization	Static Game
Dynamic	Dynamic Optimization	**Dynamic Game**

Action space

Time space

Model based	Finite	Infinite
Discrete	Markov Game (Stochastic Game)	DT Infinite dynamic game (Stochastic Game)
Continuous	Continuous time Markov Game	CT-time Infinite dynamic game (differential game)

Overview

	Single Agent	Multi Agent
Static	Static optimization	Static Game
Dynamic	Dynamic Optimization	**Dynamic Game**

Action space

Model free	Finite	Infinite
Discrete	Multi-Agent Value-based RL	Multi-Agent Policy-based RL
Continuous		

Time space

Basic Principle to Analyze Dynamic Games

Equilibrium concept:

-Nash; Zero-sum; Stackelberg; Correlated

	Single Agent	Multi Agent
Static	Static optimization	Static Game
Dynamic	Dynamic Optimization	Dynamic Game

Dynamic optimization as a static optimization concept:

- -Minimum principle (necessary condition)
- -Dynamic programming principle (sufficient condition)
- -Need to specify information structure

$$
\begin{aligned}
L^{1*} &\triangleq L^1(u^{1*}; u^{2*}; \dots; u^{N*}) \leq L^1(u^{1}; u^{2*}; \dots; u^{N*}),\\
L^{2*} &\triangleq L^2(u^{1*}; u^{2*}; \dots; u^{N*}) \leq L^2(u^{1*}; u^{2}; \dots; u^{N*}),\\
&\cdots\\
L^{N*} &\triangleq L^N(u^{1*}; u^{2*}; \dots; u^{N*}) \leq L^N(u^{1*}; u^{2*}; \dots; u^{N*})
\end{aligned}
$$

(Think in normal form game setting)

Basic Principle to Analyze Dynamic Games

Equilibrium concept:

(Dynamic) Information structure		Nash	Zero-sum	Stackelberg
	Open-loop (perfect state)	Open-loop Nash-Strategy	Open-loop Zero-sum Strategy	
	Feedback (perfect state)	Feedback Nash-Strategy	Feedback Zero-sum Strategy	
	⋮			

- We need to specify ***information structure***
  - ✓ Open-loop vs. close-loop (feedback)
  - ✓ Perfect vs. imperfect
- We need to ***equilibrium concept***
  - ✓ Nash, Zero-sum, Stackelberg, Correlated,…

Equilibrium concept + information structure → solution method

Non-Cooperative Game:

Contents

Discrete-time Infinite Dynamic Game

DT: Definition

Definition (N-person discrete-time deterministic infinite dynamic game)

N-person discrete-time deterministic infinite dynamic game involves:

- players' index set $\mathbf{N} = \{1, \dots, N\}$
- The stage of game index set $\mathbf{K} = \{1, \dots, K\}$
- state of game at stage $k \in \mathbf{K}$, x_k
- Action of $\mathbf{P}_i$ at stage $k \in \mathbf{K}$, $u_k^i \in U_i^k$ where U_i^k denotes permissible action set
- A state equation of dynamic game, $f_k: X \times U_i^1 \times \cdots \times U_i^N \to X$ defined for $\forall k \in \mathbf{K}$,

$$x_{t+1} = f_k\left(x_k, u_k^1, \dots, u_k^N\right),$$

defined for each $k \in \mathbf{K}$ describes the evolution of the game

- Observation of $\mathbf{P}_i$ at stage k, $y_i^k \in Y_k^i$
- state measurement $h_k^i: X \to Y_k^i$

$$y_k^i = h_k^i(x_k), \qquad k \in \mathbf{K}, i \in \mathbf{N}$$

DT: Definition

Definition (N-person discrete-time deterministic infinite dynamic game)

- Information structure(pattern) $\eta_k^i \in N_k^i$

$$N_k^i \subset \{Y_1^1, \dots, Y_k^1; \dots; Y_1^N, \dots, Y_k^N; U_1^1, \dots U_{k-1}^1; \dots; U_1^N, \dots, U_{k-1}^N\}$$

determines the information gained and recalled by P_i at stage k

- A strategy of $\mathbf{P}_i$ at stage k, $\gamma_k^i : N_k^i \to U_k^i$ maps information to action set, and its aggregation $\gamma^i = \{\gamma_1^i, \dots, \gamma_K^i\}$ defines strategy of $\mathbf{P}_i$ in game

- A cost function of P_i,

$$L^i : (X \times U_1^1 \times \cdots \times U_1^N) \times \cdots \times (X \times U_K^1 \times \cdots \times U_K^N) \to R$$

defined for each $i \in \mathbf{N}$

✓ L^i is the accumulated cost for player i

✓ Goal : $\mathbf{P}_i$ wants to find strategy $\gamma^i = \{\gamma_1^i, \dots, \gamma_K^i\}$ which minimize L_i given available information η_k^i

DT: Definition

Normal form description of a dynamic game

- For each fixed initial state x_1 and for each fixed N −tuple permissible strategies $\{\gamma^i \in \Gamma^i; i \in \mathbf{N}\}$ the extensive form description leads to a unique set of vectors $\{u_k^i \triangleq \gamma_k^i(\eta_k^i), x_{k+1}; i \in \mathbf{N}, k \in \mathbf{K}\}$
  - ✓ because of the causal nature of the information structure
  - ✓ the state evolves according to a difference equation.
- Substitution of these quantities into $L^i (i \in \mathbf{N})$ clearly leads to a unique N −tuple of numbers reflecting the corresponding costs to the players.
- This further implies existence of a composite mapping

$$J^i : \Gamma^1 \times \cdots \times \Gamma^N \to \mathbf{R}, \text{for each } i \in \mathbf{N}$$

which is known as the cost functional of $\mathbf{P}i \; (i \in \mathbf{N})$

- Hence, the permissible strategy spaces of the players $(\Gamma^1, \ldots, \Gamma^N)$ together with these cost functions $(J^i, \ldots, J^N)$ constitute the normal form description of the dynamic game for each fixed initial state vector x_1

There is no difference between ***infinite discrete-time dynamic gams*** and ***finite games***

➢ ***allows us to use static game equilibrium concept to analyze the dynamic game***

DT: Definition

Definition (stage-additive cost function)

In a N-person discrete-time deterministic dynamic game of pre-specified fixed duration (i.e., K stages), $\mathbf{P}i$'s cost functional is said to be stage-additive if there exist $g_k^i: X \times X \times U_k^1 \times \cdots \times U_k^N \to \mathbf{R}$, $(k \in K)$, so that

$$L^i(u^1, \dots, u^N) = \sum_{k=1}^{K} g_k^i\left(x_{k+1}, u_k^1, \dots, u_k^N, x_k\right)$$

where

$$u^j = \left(u_1^{j\,\prime}, \dots, u_K^{j\,\prime}\right)'$$

Furthermore, if $L^i(u^1, \dots, u^N)$ depends on only on x_{K+1}, (the termination state), then we call it a terminal cost functional.

- State-additive cost function is widely used for optimal control or dynamic game

DT: Information Structure

We call that $P_i's$ information structure η_k^i is

i. (OL) **open-loop information pattern** if $\quad \eta_k^i = \{x_1\}$

ii. (CLPS) closed-loop perfect state information pattern if $\quad \eta_k^i = \{x_1, \cdots, x_k\}, \; k \in \mathbf{K}$

iii. (CLIS) closed-loop imperfect state information pattern if $\quad \eta_k^i = \{y_1^i, \cdots, y_k^i\}, \; k \in \mathbf{K}$

iv. (MPS) memoryless perfect state information pattern if $\quad \eta_k^i = \{x_1, x_k\}, k \in \mathbf{K}$

v. (FB) **feedback perfect state information** pattern if $\quad \eta_k^i = \{x_k\}, \; k \in \mathbf{K}$

vi. (FIS) feedback imperfect state information pattern if $\quad \eta_k^i = \{y_k^i\}, \; k \in \mathbf{K}$

- With each information structure η_k^i, action $u_k^i \triangleq \gamma_k^i(\eta_k^i)$ can be realized
- Under the information structure, the Nash solution is referred "*open-loop Nash equilibrium solution*" or "*feedback Nash equilibrium solution*"

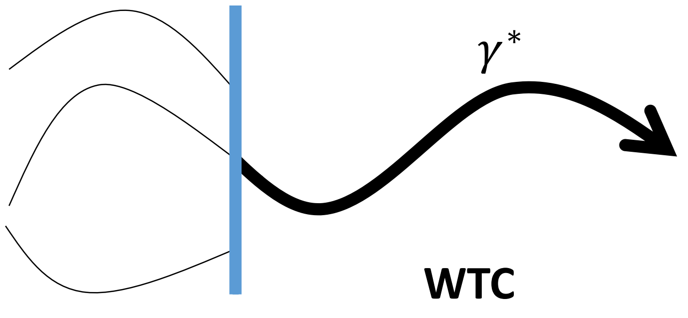
Time Consistency

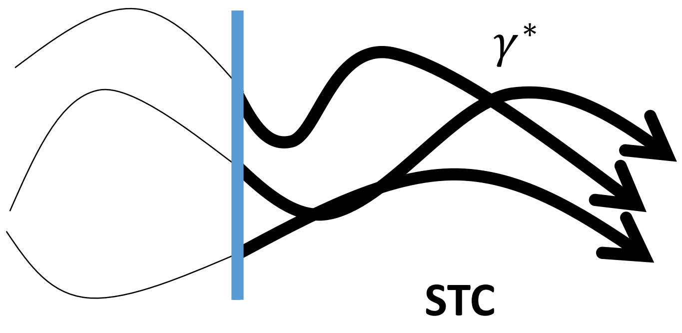
Definition (Weakly time consistent)

An N-tuple of policies γ^* is ***weakly time consistent*** if its truncation to the interval $[s, T], \gamma^*_{[s,T]}$ solves the truncated game $D^{\gamma^*}_{[s,T]}$, this being so for all $s \in (0, T]$

Definition (Strongly time consistent on subgame perfect)

An N-tuple of policies γ^* is ***strongly time consistent*** if its truncation to the interval $[s, T], \gamma^*_{[s,T]}$ solves the truncated game $D^{\gamma}_{[s,T]}$, for every $\gamma_{[0,s)}$, this being so for all $s \in (0, T]$





- In both case, players have no reason to deviate from strategy
- Difference lies in the consistency of past actions with the adopted strategies

DT: Nash Equilibrium Strategy (Formulation)

Definition (Nash equilibrium in discrete time dynamic game : action space)

$N-$ tuple of strategies $\{\gamma^{i*}(\cdot) \in \Gamma^i; i \in N\}$ constitutes a Nash equilibrium(for any information set) if it satisfies following inequalities for all $u^{i*} = \gamma^{i*}(\cdot), i \in N$

$$
\begin{gathered}
L^{1*} \triangleq L^1(u^{1*}; u^{2*}; \dots; u^{N*}) \leq L^1(u^1; u^{2*}; \dots; u^{N*}),\\
L^{2*} \triangleq L^2(u^{1*}; u^{2*}; \dots; u^{N*}) \leq L^2(u^{1*}; u^2; \dots; u^{N*}),\\
\vdots\\
L^{N*} \triangleq L^N(u^{1*}; u^{2*}; \dots; u^{N*}) \leq L^N(u^{1*}; u^{2*}; \dots; u^{N*}),
\end{gathered}
$$

Here, $u^i \triangleq \{u_i^i \dots, u_K^i\}$ is the aggregate action of $\mathbf{P}_i$

Definition (Nash equilibrium in discrete time dynamic game : strategy space)

$N-$ tuple permissible strategies $\left\{\gamma^{i*} \in \Gamma^i; i \in \mathbf{N}\right\}$ constitutes a Nash equilibrium solution if, and only if, the following inequalities are satisfied for all $\left\{\gamma^i \in \Gamma^i; i \in \mathbf{N}\right\}$

$$
\begin{gathered}
J^{1*} \triangleq J^1(\gamma^{1*}; \gamma^{2*}; \dots; \gamma^{N*}) \leq J^1(\gamma^1; \gamma^{2*}; \dots; \gamma^{N*}),\\
J^{2*} \triangleq J^2(\gamma^{1*}; \gamma^{2*}; \dots; \gamma^{N*}) \leq J^2(\gamma^{1*}; \gamma^2; \dots; \gamma^{N*}),\\
\vdots\\
J^{N*} \triangleq J^N(\gamma^{1*}; \gamma^{2*}; \dots; \gamma^{N*}) \leq J^N(\gamma^{1*}; \gamma^{2*}; \dots; \gamma^N),
\end{gathered}
$$

Here, $\gamma^i \triangleq \{\gamma_i^i \dots, \gamma_K^i\}$ is the aggregate strategy of $\mathbf{P}_i$

DT: Open-loop Nash Equilibrium Strategy (Solution Method)

Open-Loop Nash equilibria : Information set $\eta_k^i = \{x_1\}$

- Identical to optimal control problem for each $\mathbf{P}_i$, since open-loop control does not depend on other's control

- The minimum principle provides optimal control $u^{i*} = (u_1^{i*}, \ldots, u_K^{i*})\ \forall i \in N$ and corresponding state trajectory, $(x_1^{i*}, \ldots, x_K^{i*})$

- Optimal open-loop NE strategy $\gamma^{i*}(x_1)$ is weekly time consistent, as it cannot provide optimal strategy out of optimal trajectory

The Minimum Principle

Consider the following optimal control problem defined by

$$L(u) = \int_0^T g\big(t, x(t), u(t)\big)\mathrm{d}t + q(T, x(T))$$

where the state variable $x(t)$ satisfies the differential equation:

$$\dot{x}(t) = f\big(t, x(t), u(t)\big), \qquad x(0) = x_0, \qquad t \geq 0$$

Theorem (The minimum principle)

In the continuous time dynamic system defined by equation (1) and (2), optimal control $u^*(t)$ and corresponding trajectory $x^*(t)$ satisfy following equations:

$$H(t, \lambda, x, u) \coloneqq g(t, x, u) + \lambda(t) f(t, x, u)$$

$$\dot{x}^*(t) = f(t, x^*, u^*)\left(= \frac{\partial H(t, x^*, u^*, \lambda)}{\partial \lambda}\right), x^*(0) = x_0;$$

$$\dot{\lambda}^*(t) = -\frac{\partial H(t, \lambda, x^*, u^*)}{\partial x}; \lambda^*(T) = \frac{\partial h(x^*(T))}{\partial x}$$

$$u^*(t) = \operatorname*{argmin}_{u \in U} H(t, \lambda^*, x^*, u)$$

DT: Open-loop Nash Equilibrium Strategy (Solution Method)

Definition (open-loop Nash equilibria in discrete time dynamic game)

If $\gamma^{i*}(x_1) = u^{i*}$ provides an open-loop Nash equilibrium, and $\{x_k^*, k \in \mathbf{K}\}$ is corresponding state trajectory, there exists costate vectors $\{p_1^i, \dots, p_K^i\}$ for each $i \in \mathbf{N}$ such that:

$$H_k^i(x_k, u_k^1, \dots, u_k^N, p_{k+1}) := g_k^i(x_k, u_k^1, \dots, u_k^N) + p_{k+1}^i f_k(x_k, u_k^1, \dots, u_k^N)$$

$$x_{k+1}^* = f_k(x_k^*, u_k^{1*}, \dots, u_k^{N*}), \qquad x_1^* = x_1$$

$$\gamma_k^{i*}(x_1) = u_k^{i*} = \arg\min_{u_k^i \in U_k^i} H_k^i(x_k^*, u_k^{1^*}, \dots, u_k^{i-1^*}, u_k^i, u_k^{i+1^*}, \dots, u_k^{N*}, p_{k+1})$$

$$p_k^i = \frac{\partial}{\partial x_k} f_k(x_k^*, u_k^{1*}, \dots, u_k^{N*}) p_{k+1}^i + \frac{\partial}{\partial x_k} g_k^i(x_k^*, u_k^{1*}, \dots, u_k^{N*}), \qquad p_K^i = 0$$

$\forall k \in \mathbf{K}, i \in \mathbf{N}$

상대방의 action이 optimal하게 고정되었다는 가정하에 minimum principle이 모든 agent와 모든 시간 instance에 대해 정의된다.

DT: Feedback Nash Equilibrium Strategy (Solution Method)

- **Feedback Nash equilibria**
  - initial state information is known a priori
  - *depend only on the time variable and current value of the state*
  - $x_k \in \eta_k^i$
  - Feedback NE solution provides NE for any subgame defined in $\{s, s+1, \ldots, K\}$ for all $s \in \mathbf{K}$
- N person feedback game in extensive form
  - Recursive procedure to obtain NE of finite game
- Feedback strategy $\gamma^{i*}(\cdot)$ is strongly time consistent

DT: Feedback Nash Equilibrium Strategy (Solution Method)

Definition (feedback Nash equilibria in discrete time dynamic game)

Level K

$$\begin{cases} L^1(\gamma_1^1,\dots,\gamma_{K-1}^1,\gamma_K^{1*};\gamma_1^2,\dots,\gamma_{K-1}^2,\gamma_K^{2*};\dots;\gamma_1^N,\dots,\gamma_{K-1}^N,\gamma_K^{N*}) \le L^1(\gamma_1^1,\dots,\gamma_{K-1}^1,\gamma_K^{1};\gamma_1^2,\dots,\gamma_{K-1}^2,\gamma_K^{2*};\dots;\gamma_1^N,\dots,\gamma_{K-1}^N,\gamma_K^{N*}) \\ L^2(\gamma_1^1,\dots,\gamma_{K-1}^1,\gamma_K^{1*};\gamma_1^2,\dots,\gamma_{K-1}^2,\gamma_K^{2*};\dots;\gamma_1^N,\dots,\gamma_{K-1}^N,\gamma_K^{N*}) \le L^2(\gamma_1^1,\dots,\gamma_{K-1}^1,\gamma_K^{1*};\gamma_1^2,\dots,\gamma_{K-1}^2,\gamma_K^{2};\dots;\gamma_1^N,\dots,\gamma_{K-1}^N,\gamma_K^{N*}) \\ \vdots \\ L^K(\gamma_1^1,\dots,\gamma_{K-1}^1,\gamma_K^{1*};\gamma_1^2,\dots,\gamma_{K-1}^2,\gamma_K^{2*};\dots;\gamma_1^N,\dots,\gamma_{K-1}^N,\gamma_K^{N*}) \le L^K(\gamma_1^1,\dots,\gamma_{K-1}^1,\gamma_K^{1*};\gamma_1^2,\dots,\gamma_{K-1}^2,\gamma_K^{2*};\dots;\gamma_1^N,\dots,\gamma_{K-1}^N,\gamma_K^{N}) \end{cases}$$

Level $K-1$

$$\begin{cases} L^1(\gamma_1^1,\dots,\gamma_{K-1}^{1*},\gamma_K^{1*};\gamma_1^2,\dots,\gamma_{K-1}^{2*},\gamma_K^{2*};\dots;\gamma_1^N,\dots,\gamma_{K-1}^{N*},\gamma_K^{N*}) \le L^1(\gamma_1^1,\dots,\gamma_{K-1}^{1},\gamma_K^{1*};\gamma_1^2,\dots,\gamma_{K-1}^{2*},\gamma_K^{2*};\dots;\gamma_1^N,\dots,\gamma_{K-1}^{N*},\gamma_K^{N*}) \\ L^2(\gamma_1^1,\dots,\gamma_{K-1}^{1*},\gamma_K^{1*};\gamma_1^2,\dots,\gamma_{K-1}^{2*},\gamma_K^{2*};\dots;\gamma_1^N,\dots,\gamma_{K-1}^{N*},\gamma_K^{N*}) \le L^2(\gamma_1^1,\dots,\gamma_{K-1}^{1*},\gamma_K^{1*};\gamma_1^2,\dots,\gamma_{K-1}^{2},\gamma_K^{2*};\dots;\gamma_1^N,\dots,\gamma_{K-1}^{N*},\gamma_K^{N*}) \\ \vdots \\ L^K(\gamma_1^1,\dots,\gamma_{K-1}^{1*},\gamma_K^{1*};\gamma_1^2,\dots,\gamma_{K-1}^{2*},\gamma_K^{2*};\dots;\gamma_1^N,\dots,\gamma_{K-1}^{N*},\gamma_K^{N*}) \le L^K(\gamma_1^1,\dots,\gamma_{K-1}^{1*},\gamma_K^{1*};\gamma_1^2,\dots,\gamma_{K-1}^{2*},\gamma_K^{2*};\dots;\gamma_1^N,\dots,\gamma_{K-1}^{N},\gamma_K^{N*}) \end{cases}$$

Level 1

$$\begin{cases} L^1(\gamma_1^{1*},\gamma_2^{1*}\dots,\gamma_K^{1*};\gamma_1^{2*},\gamma_2^{2*}\dots,\gamma_K^{2*};\cdots;\gamma_1^{N*},\gamma_2^{N*}\dots,\gamma_K^{N*}) \le L^1(\gamma_1^{1},\gamma_2^{1*}\dots,\gamma_K^{1*};\gamma_1^{2*},\gamma_2^{2*}\dots,\gamma_K^{2*};\cdots;\gamma_1^{N*},\gamma_2^{N*}\dots,\gamma_K^{N*}) \\ L^2(\gamma_1^{1*},\gamma_2^{1*}\dots,\gamma_K^{1*};\gamma_1^{2*},\gamma_2^{2*}\dots,\gamma_K^{2*};\cdots;\gamma_1^{N*},\gamma_2^{N*}\dots,\gamma_K^{N*}) \le L^2(\gamma_1^{1*},\gamma_2^{1*}\dots,\gamma_K^{1*};\gamma_1^{2},\gamma_2^{2*}\dots,\gamma_K^{2*};\cdots;\gamma_1^{N*},\gamma_2^{N*}\dots,\gamma_K^{N*}) \\ \vdots \\ L^N(\gamma_1^{1*},\gamma_2^{1*}\dots,\gamma_K^{1*};\gamma_1^{2*},\gamma_2^{2*}\dots,\gamma_K^{2*};\cdots;\gamma_1^{N*},\gamma_2^{N*}\dots,\gamma_K^{N*}) \le L^N(\gamma_1^{1*},\gamma_2^{1*}\dots,\gamma_K^{1*};\gamma_1^{2*},\gamma_2^{2*}\dots,\gamma_K^{2*};\cdots;\gamma_1^{N},\gamma_2^{N*}\dots,\gamma_K^{N*}) \end{cases}$$

- **Backward Induction:**

DT: Feedback Nash Equilibrium Strategy (Solution Method)

Definition (feedback Nash equilibria in discrete time dynamic game)

For N-person discrete time infinite dynamic game, the set of strategies $\{\gamma_k^{i*}(x_k); k \in K, i \in N\}$ provides **feedback** **Nash equilibrium** solution if and only if there exists functions $V^i(k,\cdot): R^n \to R$ such that following recursive relations are satisfied:

$$V^i(k,x) = \min_{u_k^i \in U_k^i} \left[g_k^i\left(x, \gamma_k^{1*}(x), \dots, u_k^i, \dots, \gamma_k^{N*}(x)\right) + V^i\left(k+1, \tilde{f}_k^{i*}(x, u_k^i)\right)\right]$$
$$= g_k^i\left(x, \gamma_k^{1*}(x), \dots, \gamma_k^{i*}(x), \dots, \gamma_k^{N*}(x)\right) + V^i\left(k+1, \tilde{f}_k^{i*}\left(x, \gamma_k^{i*}(x)\right)\right);$$
$$V^i(K+1, x) = 0, \qquad \forall i \in N$$

where

$$\tilde{f}_k^{i*}(x, u_k^i) \triangleq f_k\left(x, \gamma_k^{1*}(x), \dots, \gamma_k^{i-1*}(x), u_k^i, \gamma_k^{i+1*}(x), \dots, \gamma_k^{N*}(x)\right)$$

Every such equilibrium solution is strongly time consistent, and corresponding NE cost for $\mathbf{P}_i$ is $V^i(1, x_1)$

- **Employ HJB equation (Dynamic Programming Principle)**
- Apply Best response principle for every time step (**rationality**), and the best responses of all the players are consistent (**consistency**)

DT: Feedback Nash Equilibrium Strategy (Solution Method)

- **Proof Sketch:**

In definition, the first set of N inequalities have to hold for all $\gamma_k^i \in \Gamma_k^i$ implies that they have to hold for all state x_k which are reachable by combination of strategies.

At time k, set of inequalities becomes equivalent to the problem of seeking Nash equilibria of N-person static game with cost functional

$$g_{k-1}^i\left(x_{k-1}, u_{k-1}^1, \dots, u_{k-1}^N\right) + V^i(k, x_k), \qquad i \in N,$$

where

$$x_k = f_{k-1}(x_{k-1}, u_{k-1}^1, \dots, u_{k-1}^N)$$

Here, we observe that the Nash equilibrium controls can only be functions of x_{k-1}, and previous theorem provides a set of necessary and sufficient conditions for $\{\gamma_{k-1}^{i*}(x_{k-1}); i \in N\}$ to solve this static Nash game.

Continuous-time Infinite Dynamic Game

CT: Definition

Definition (N-person continuous-time deterministic infinite dynamic game, differential game)

N-person differential game involves:

- players' index set $\mathbf{N} = \{1, \dots, N\}$
- A time interval $[0, T]$ which is specified *a priori,* duration of the evolution of game
- Permissible state trajectories of the game, $\{x(t), 0 \le t \le T\}$
- control function(or simply control) of P_i, $\{u_i(t), 0 \le t \le T\}$
- A differential equation

$$\frac{dx(t)}{dt} = f\left(t, x(t), u^1(t), \dots, u^N(t)\right), \qquad x(0) = x_0$$

  whose solution describes the state trajectory of the game

- state information gained and recalled by $\mathbf{P}_i$ at time t, $\eta^i(t)$
- strategy of $\mathbf{P}_i$ γ^i, with property $u^i(t) = \gamma^i(t, \eta^i(t))$ cost function of $\mathbf{P}_i$ in the differential game L^i,

$$L^i(u^1, \cdots, u^N) = \int_0^T g^i\left(t, x(t), u^1(t), \dots, u^N(t)\right)dt + q^i(x(T))$$

Goal :player $\mathbf{P}_i$ wants to find strategy γ^i which minimize L_i given available information $\eta^i(t)$

CT: Information Structure

- We call that $\mathbf{P}_i's$ information structure is

i. (OL) open-loop information pattern if	$\eta^i(t) = \{x_0\},\ t \in [0, T]$
ii. (CLPS) closed-loop perfect state information pattern if	$\eta^i(t) = \{x(s), 0 \leq s \leq t\},\ t \in [0, T]$
iii. (MPS) memoryless perfect state information pattern if	$\eta^i(t) = \{x_0, x(t)\},\ \ t \in [0, T]$
iv. (FB) feedback pattern if	$\eta^i(t) = \{x(t)\},\ t \in [0, T]$

- With each information structure $\eta^i(t),\ u^i(t) \triangleq \gamma^i\left(t, \eta^i(t)\right)$ can be realized
- Under the information structure, the Nash solution is referred "*open-loop Nash equilibrium solution*" or "*feedback Nash equilibrium solution*"

CT: Open-loop Nash Equilibrium Strategy (Formulation)

- We consider $N-$ person dynamic game defined in continuous time

$$\frac{dx(t)}{dt} = f\big(t, x(t), u^1(t), \dots, u^N(t)\big), \qquad x(0) = x_0$$

and cost functional

$$L^i(u^1, \cdots, u^N) = \int_0^T g^i\big(t, x(t), u^1(t), \dots, u^N(t)\big)dt + q^i(x(T))$$

CT: Open-loop Nash Equilibrium Strategy (Formulation)

Definition (Nash equilibrium in discrete time dynamic game : action space)

$N-$tuple of strategies $\{\gamma^{i*}(\cdot) \in \Gamma^i; i \in N\}$ constitutes a Nash equilibrium(for any information set) if it satisfies following inequalities for all $u^{i*} = \gamma^{i*}(\cdot), i \in N$

$$
\begin{gathered}
L^{1*} \triangleq L^1(u^{1*}; u^{2*}; \dots; u^{N*}) \leq L^1(u^1; u^{2*}; \dots; u^{N*}), \\
L^{2*} \triangleq L^2(u^{1*}; u^{2*}; \dots; u^{N*}) \leq L^2(u^{1*}; u^2; \dots; u^{N*}), \\
\vdots \\
L^{N*} \triangleq L^N(u^{1*}; u^{2*}; \dots; u^{N*}) \leq L^N(u^{1*}; u^{2*}; \dots; u^{N*}),
\end{gathered}
$$

Here, $u^i(t) \in S^i$ is the action of $\mathbf{P}_i$ chosen at time $t \in [0, T]$

Definition (Nash equilibrium in discrete time dynamic game : strategy space)

$N-$ tuple permissible strategies $\{\gamma^{i*} \in \Gamma^i; i \in \mathbf{N}\}$ constitutes a Nash equilibrium solution if, and only if, the following inequalities are satisfied for all $\{\gamma^i \in \Gamma^i; i \in \mathbf{N}\}$

$$
\begin{gathered}
J^{1*} \triangleq J^1(\gamma^{1*}; \gamma^{2*}; \dots; \gamma^{N*}) \leq J^1(\gamma^1; \gamma^{2*}; \dots; \gamma^{N*}), \\
J^{2*} \triangleq J^2(\gamma^{1*}; \gamma^{2*}; \dots; \gamma^{N*}) \leq J^2(\gamma^{1*}; \gamma^2; \dots; \gamma^{N*}), \\
\vdots \\
J^{N*} \triangleq J^N(\gamma^{1*}; \gamma^{2*}; \dots; \gamma^{N*}) \leq J^N(\gamma^{1*}; \gamma^{2*}; \dots; \gamma^N),
\end{gathered}
$$

Here, $\gamma^i\left(t, \eta^i(t)\right)$ is the strategy of $\mathbf{P}_i$ at time $t \in [0, T]$

CT: Open-loop Nash Equilibrium Strategy (Solution Method)

- Open-Loop Nash equilibria
  - Information set $\eta_k^i = \{x_0\}$

- Identical to optimal control problem for each $\mathbf{P}_i$, since open-loop control does not depend on other's control

$$\text{minimize } J^i\left(u^{1*}(t), \ldots, u^{i-1*}(t), \boldsymbol{u^i}(t), u^{i+1*}(t), \ldots u^{N*}(t)\right)$$
$$\text{s.t } \dot{x}^* = f\left(t, x^*(t), u^{1*}(t), \ldots, u^{i-1*}(t), \boldsymbol{u^i}(t), u^{i+1*}(t), \ldots u^{N*}(t)\right)$$

- The minimum principle provides optimal control $u^{i*}(t)$ $\forall i \in N$ and state trajectory $x^*(t)$

- Optimal open-loop NE strategy $\gamma^{i*}(t, x_0) = u^{i*}(t)$ is weekly time consistent, as it cannot provide optimal strategy out of optimal trajectory

The Minimum Principle

Consider the following optimal control problem defined by

$$L(u) = \int_0^T g\big(t, x(t), u(t)\big)\mathrm{d}t + q(T, x(T))$$

where the state variable $x(t)$ satisfies the differential equation:

$$\dot{x}(t) = f\big(t, x(t), u(t)\big), \qquad x(0) = x_0, \qquad t \geq 0$$

Theorem (The minimum principle)

In the continuous time dynamic system defined by equation (1) and (2), optimal control $u^*(t)$ and corresponding trajectory $x^*(t)$ satisfy following equations:

$$H(t, \lambda, x, u) \coloneqq g(t, x, u) + \lambda(t) f(t, x, u)$$

$$\dot{x}^*(t) = f(t, x^*, u^*)\left(= \frac{\partial H(t, x^*, u^*, \lambda)}{\partial \lambda}\right), x^*(0) = x_0;$$

$$\dot{\lambda}^*(t) = -\frac{\partial H(t, \lambda, x^*, u^*)}{\partial x}; \lambda^*(T) = \frac{\partial h(x^*(T))}{\partial x}$$

$$u^*(t) = \underset{u \in U}{\operatorname{argmin}}\, H(t, \lambda^*, x^*, u)$$

CT: Open-loop Nash Equilibrium Strategy (Solution Method)

Definition (Open-loop Nash equilibria in continuous time dynamic game)

If $\gamma^{i*}(t, x_0) = u^{i*}(t)$ provides an open-loop Nash equilibrium, there exists N co-state functions $p^i(\cdot)\colon [0, T] \to R^n$ for each $i \in N$ such that:

$$\dot{x}^i = f\big(t, x^*(t), u^{1*}(t), \dots, u^{N*}(t)\big), x^*(0) = x_0$$

$$\gamma^{i*}(t, x_0) = u^{i*}(t) = \underset{u^i(t)}{\operatorname{argmin}} H^i\big(t, p^i(t), x^*(t), u^{1*}(t), \dots, u^i(t), \dots, u^{N*}(t)\ \big)$$

$$\dot{p}^i = -\frac{\partial}{\partial x} H^i\left(t, p^i(t), x^*, u^{1*}(t), \dots, u^{N*}(t)\right)$$

$$\dot{p}^i(T) = -\frac{\partial}{\partial x} q^i(x^*(T))$$

where

$$H^i\big(t, p^i, x, u^1, \dots, u^N\big) \coloneqq g^i(t, x, u^1, \dots, u^N) + p^i f(t, x, u^1, \dots, u^N)$$

CT: Feedback Nash Equilibrium Formulation

- **Feedback Nash equilibria**
  - ✓ initial state information is known a priori
  - ✓ *depend only on the time variable and current value of the state*
  - ✓ $x(t) \in \eta_t$
  - ✓ Feedback NE solution provides NE for any subgame defined in $[t, T]$ for all t ∈ $[0, T)$

- Definition of the feedback NE leads to a recursive derivation
  - Value function $V^i(t, x)$, minimum cost-to-go for player i at time t on state x

- Optimal feedback NE strategy $\gamma^{i*}(t, \eta_t)$ is strongly time consistent

CT: Feedback Nash Equilibrium Formulation

Definition (Feedback NE Solution)

An N-tuple of strategies $\{\gamma^{i*} \in \Gamma^i; i \in N\}$ constitutes a feedback Nash equilibrium solution if there exists $V^i(\cdot,\cdot)$ on $[0,T] \times R^n$ s.t.

$$V^i(t,x) = \int_t^T g^i\left(s, x^*(s), \gamma^{1*}(s,\eta_s), \cdots, \gamma^{i*}(s,\eta_s), \cdots, \gamma^{N*}(s,\eta_s)\right) ds + q^i\left(x^*(T)\right)$$

$$\le \int_t^T g^i\left(s, x^*(s), \gamma^{1*}(s,\eta_s), \cdots, \gamma^{i}(s,\eta_s), \cdots, \gamma^{N*}(s,\eta_s)\right) ds + q^i\left(x^i(T)\right), \forall \gamma^i \in \Gamma^i, x \in R^n$$

Where, on $[t,T]$,

$$\dot{x}^i(s) = f\left(s, x^i(s), \gamma^{1*}(s,\eta_s), \cdots, \gamma^{i}(s,\eta_s), \cdots, \gamma^{N*}(s,\eta_s)\right); \quad x^i(t) = x,$$

$$\dot{x}^*(s) = f\left(s, x^*(s), \gamma^{1*}(s,\eta_s), \cdots, \gamma^{i*}(s,\eta_s), \cdots, \gamma^{N*}(s,\eta_s)\right); \quad x^*(t) = x$$

η_s stands for data set $\{x(s), x_0\}$ or $\{x(\sigma), \sigma \le s\}$, depending on information pattern is MPS or CLPS

CT: Feedback Nash Equilibrium Strategy (Solution Method)

- Time interval restriction, $[t, T]$ provides same differential game with initial state $x(t), \forall t$

- Under either MPS or CLPS information pattern, *feedback NE will depend only on the time variable and current value of the state*, but not on memory.

- If value functions V^i are continuously differentiable in x and t, then N partial differential equations replace previous equation (HJB equation)

Obtaining the optimal control strategy from HJB equation

Theorem

If a continuously differentiable function $V(t,x)$ can be found that satisfies the HJB equation (9), then it generate the optimal strategy through the static (pointwise) minimization problem defined by the RHS of (9)

$$-\frac{\partial V(t,x)}{\partial t} = \min_{u\in U}\left[\frac{\partial V(t,x)}{\partial x} f(t,x,u) + g(t,x,u)\right] \quad (9)$$

$$u^*(t) = \operatorname*{argmin}_{u\in U}\left[\frac{\partial V(t,x)}{\partial x} f(t,x,u) + g(t,x,u)\right] \quad (10)$$

CT: Feedback Nash Equilibrium Strategy (Solution Method)

Definition (Feedback NE solution with value function)

For an N person differential game for $[0, T]$ and under either MPS or CLPS, N tuple of strategies $\{\gamma^{i*} \in \Gamma^i, i \in N\}$ provides a feedback Nash equilibrium solution if there exists functions $V^i : \{0, T\} \times R^n \to R, i \in N$ satisfying the partial differential equations

$$-\frac{\partial V^i(t,x)}{\partial t} = \min_{u^i \in S^i} \left[\frac{\partial V^i(t,x)}{\partial t} \tilde{f}^{i*}(t,x,u^i) + \tilde{g}^{i*}(t,x,u^i) \right]$$

$$= \frac{\partial V^i(t,x)}{\partial t} \tilde{f}^{i*}\left(t,x,\gamma^{i*}(t,x)\right) + \tilde{g}^{i*}\left(t,x,\gamma^{i*}(t,x)\right)$$

$$V^i(T,x) = q^i(x), \qquad \forall i \in N$$

where

$$\tilde{f}^{i*}(t,x,u^i) \triangleq f\left(t,x,\{\gamma^*_{-i}(t,x),u^i\}\right),$$

$$\tilde{g}^{i*}(t,x,u^i) \triangleq g^i\left(t,x,\{\gamma^*_{-i}(t,x),u^i\}\right),$$

$$\{\gamma^*_{-i}(t,x),u^i\} \triangleq \gamma^{1*}(t,x), \dots, u^i, \dots, \gamma^{N*}(t,x)$$

Every such equilibrium solution is STC, and Nash equilibrium cost of P_i is $V^i(0,x_0)$

Example

Structural Dynamic Control using Game Theory

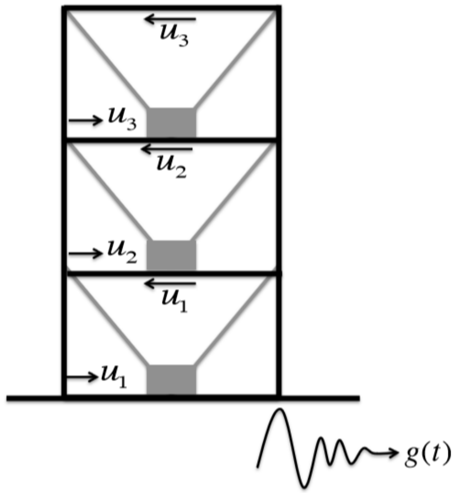
- Linear Quadratic Differential game is defined as follows
  - Cost function for each agent:

$$J_i = \frac{1}{2}\int_0^T \left\{ x^T(t) Q_i x(t) + \sum_{j=1}^{N} u_j(t)^T R_{ij} u_j(t) \right\} dt + \frac{1}{2} x^T(T) Q_T x(T)$$

  - Dynamics of joint state

$$\dot{x}(t) = Ax(t) + \sum_{j=1}^{N} B_j u_j(t) + Eg(t)$$

- Dynamic control of the three story building is expressed as:



$$J_1 = \int_0^\infty \left\{ x^T Q_1 x + u_1^T R_{11} u_1 + u_2^T R_{12} u_2 + u_3^T R_{13} u_3 \right\} dt$$
$$J_2 = \int_0^\infty \left\{ x^T Q_3 x + u_1^T R_{21} u_1 + u_2^T R_{32} u_2 + u_3^T R_{23} u_3 \right\} dt$$
$$J_3 = \int_0^\infty \left\{ x^T Q_3 x + u_1^T R_{31} u_1 + u_2^T R_{32} u_2 + u_3^T R_{33} u_3 \right\} dt$$

$$\dot{x} = Ax + B_1 u_1 + B_2 u_2 + B_3 u_3 + Eg$$

Cooperative Control

- The cooperative control policy is derived assuming each agent tries to minimize the commonly shared objective

$$J_1 + J_2 + J_3 = \int_0^{\infty} \left\{ x^T(Q_1 + Q_2 + Q_3)x + \left[u_1 u_2 u_3\right]^T \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \right\} dt$$

$$J_1 + J_2 + J_3 = \int_0^{\infty} \{z^T Q z + u^T R u\}\, dt$$

  - ✓ where $u = [u_1, u_2, u_3]$ is aggregated control action vector
  - ✓ Q and R are aggregated accordingly

- Due to the cooperation, the cooperative control problem can be formulated as an optimal control problem and can be solved using optimal control theory

$$PA + A^T P - PSP + Q = 0$$

$$S = BR^{-1}B^T$$

$$u(x) = -R^{-1}B^T P x(t) = F^* x(t)$$

Nash Feedback Control

- Assuming the control actions are expressed as linear function of constant gain matrix such that $u_i = F_i x$, we can re-write the cost functions

$$J_1(F_1, F_2, F_3, x_0) = \int_0^\infty \{x^T(Q_1 + F_1^T R_{11} F_1 + F_2^T R_{12} F_2 + F_3^T R_{13} F_3)x\}\, dt$$
$$J_2(F_1, F_2, F_3, x_0) = \int_0^\infty \{x^T(Q_2 + F_1^T R_{21} F_1 + F_2^T R_{22} F_2 + F_3^T R_{23} F_3)x\}\, dt$$
$$J_3(F_1, F_2, F_3, x_0) = \int_0^\infty \{x^T(Q_3 + F_1^T R_{31} F_1 + F_2^T R_{32} F_2 + F_3^T R_{33} F_3)x\}\, dt$$

- Assuming the control actions are expressed as linear function of constant gain matrix such that $u_i = F_i x$, we can re-write the transition functions

$$\dot{x} = (A + B_1 F_1 + B_2 F_2 + B_3 F_3)x + Eg$$

Nash Feedback Control

- The Nash equilibrium (F_1^*, F_2^*, F_3^*) satisfies the following conditions:

$$
\begin{aligned}
J_1(F_1^*, F_2^*, F_3^*, x_0) &\leq J_1(F_1, F_2^*, F_3^*, x_0) \\
J_2(F_1^*, F_2^*, F_3^*, x_0) &\leq J_2(F_1^*, F_2, F_3^*, x_0) \\
J_3(F_1^*, F_2^*, F_3^*, x_0) &\leq J_3(F_1^*, F_2^*, F_3, x_0)
\end{aligned}
$$

- Nash equilibrium (F_1^*, F_2^*, F_3^*) strategy can be computed by solving the coupled inequality equations
- When agent 2 and 3 are assumed to follow the optimum strategy, the first agent should best respond to the fixes strategies as:

$$
F_1^* = \arg\min_{F_1} J_1(F_1, F_2^*, F_3^*, x_0)
$$

- When $F_2^* = -R_{22}^{-1} B_2^T P_2$ and $F_3^* = -R_{33}^{-1} B_3^T P_3$ (P_2 and P_3 are unknown matrices), the objection function of player 1 becomes:

$$
\begin{aligned}
J_1(F_1, F_2^*, F_3^*, x_0) = \int_0^\infty & [x^T\{Q_1 + F_1^T R_{11} F_1 + (-R_{22}^{-1} B_2^T P_2)^T R_{12} (-R_{22}^{-1} B_2^T P_2) \\
& (-R_{33}^{-1} B_3^T P_3)^T R_{13} (-R_{33}^{-1} B_3^T P_3)\} x] dt
\end{aligned}
$$

Nash Feedback Control

$$J_1(F_1, F_2^*, F_3^*, x_0) = \int_0^\infty [x^T\{Q_1 + F_1^T R_{11} F_1 + (-R_{22}^{-1} B_2^T P_2)^T R_{12} (-R_{22}^{-1} B_2^T P_2)$$
$$(-R_{33}^{-1} B_3^T P_3)^T R_{13} (-R_{33}^{-1} B_3^T P_3)\} x] dt$$

- Setting $S_{ij} = B_i R_{ii}^{-1} R_{ij} R_{ii}^{-1} B_i^T$, Agent 1 needs to maximizes

$$\bar{J}_1(F_1, x_0) = \int_0^\infty \{x^T (Q_1 + P_2 S_{12} P_2 + P_3 S_{13} P_3) x + x^T F_1^T R_{11} F_1 x\} dt$$

- Assuming the following state dynamics

$$\begin{aligned}
\dot{x} &= (A + B_1 F_1 + B_2 F_2 + B_3 F_3) x + Eg \\
&= \left\{A + B_1 F_1 + B_2(-R_{22}^{-1} B_2^T P_2) + B_3(-R_{33}^{-1} B_3^T P_3)\right\} x + Eg \\
&= (A - S_{22} P_2 - S_{33} P_3) x + B_1 F_1 x + Eg
\end{aligned}$$

- Having formulated the cost function and the system dynamics, while assuming the gain matrices for other controller, we can derive the Riccati equation for player 1:

$$(A')^T P_1 + P_1 A' - P_1 S_{11} P_1 + Q' = 0$$
$$(A - S_{22} P_2 - S_{33} P_3)^T P_1 + P_1 (A - S_{22} P_2 - S_{33} P_3) - P_1 S_{11} P_1 + (Q_1 + P_2 S_{12} P_2 + P_{33} S_{13} P_3) = 0$$

Two Person Min-Max Game

- To explicitly account for a external load (earthquake), a two player zero-sum game framework can be used. In it, the external load is treated as a fictitious agent competing with controllers. The cost function for the controller is redefined as

$$J_c(u, g, x_0) = \int_0^\infty \{x^T Q x + u^T R u - g^T V g\} dt$$

or the external load is defined as the negative of that for the controlle

$$J_g(u, g, x_0) = -J_c(u, g, x_0) = \int_0^\infty \{-x^T Q x - u^T R u + g^T V g\} dt$$

- the Min-Max control problem is to find the gain matrix $F^*(u^*(t) \ = F^*(x(t))$ that minimizes the worst-case cost function incurred by the external load as follows [5]:

$$F^* = \min_{F \in F} \sup_{g \in L_2^q(0,\infty)} J_c(F, g, x_0)$$

$$J_c(F, g, x_0) = \int_0^\infty \{x^T (Q + F^T R F) x - g^T V g\} dt$$

Two Person Min-Max Game

- Policies for the controllers and the earthquake are

$$u^*(t) = -R^{-1}B^T P_1 x(t)$$
$$g^*(t) = -V^{-1}E^T P_2 x(t)$$

- Inserting the policy into the cost function:

$$\bar{J}_c(F, x_0) = \int_0^\infty \left[x^T \left\{Q + F^T RF - (-V^{-1}E^T P_2)^T V(-V^{-1}E^T P_2)\right\} x\right] dt$$

- Setting $M = EV^{-1}ET$

$$\bar{J}_c(F, x_0) = \int_0^\infty \left[x^T \left\{(Q - P_2 M P_2) + TF^T RF\right\} x\right] dt$$

- State dynamics is expressed as:

$$\begin{aligned}\dot{x} &= Ax + BFx + Eg \\ &= \left\{A - E(V^{-1}E^T P_2)\right\} x + BFx \quad = (A - MP_2)x + BFx\end{aligned}$$

Two Person Min-Max Game

- Setting $S = BR^{-1}B^T,\ M = EV^{-1}E^T$,

$$(A - MP_2)^T P_1 + P_1(A - MP_2) - P_1 S P_1 + (Q - P_2 M P_2) = 0$$

$$(A - SP_1)^T P_2 + P_2(A - SP_1) - P_2 M P_2 + (-Q - P_1 S P_1) = 0$$

- Adding Two equations:

$$(P_1 + P_2)(A - SP_1 - MP_2) + (A - SP_1 - MP_2)^T (P_1 + P_2) = 0$$

- Substituting $P_1 = -P_2 = P$, the above equation becomes

$$A^T P + PA - P(S - M)P + Q = 0$$