Stochastic Differential Games in Finite and Infinite Population Regimes

Part III: *Going from finite to infinite population and back: Mean field games under Stackelberg-Nash equilibrium*

Tamer Başar

Coordinated Science Laboratory
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
Urbana, IL - 61801

June 19-20, 2018

Outline

- Introduction
- Mean Field Games with Major and Minor Agents
- Mean Field Stackelberg Differential Games
- Conclusions

Introduction

- Lecture II addressed MFGs where players do not have disproportionate levels of power, and each player alone has only infinitesimal effect on the game
  - Players chose their strategies *simultaneously* and *non-cooperatively*
  - Approach led to decentralized ϵ-Nash equilibria

- In this Lecture III, at least one player has disproportionate power in the game, whose impact is not infinitesimal. There are two approaches:
  - Major/minor player games: game exhibits strong influence of the major player, but the equilibrium solution is symmetric–Nash equilibrium
  - Stackelberg games: hierarchical decision making with leader(s) and followers, where leader has a global advantage over the followers

Introduction

- Lecture II addressed MFGs where players do not have disproportionate levels of power, and each player alone has only infinitesimal effect on the game
  - Players chose their strategies *simultaneously* and *non-cooperatively*
  - Approach led to decentralized ϵ-Nash equilibria

- In this Lecture III, at least one player has disproportionate power in the game, whose impact is not infinitesimal. There are two approaches:
  - Major/minor player games: game exhibits strong influence of the major player, but the equilibrium solution is symmetric–Nash equilibrium
  - Stackelberg games: hierarchical decision making with leader(s) and followers, where leader has a global advantage over the followers

Introduction

- Stackelberg games: Games that have a hierarchical decision structure (leader(s) and follower(s))
  - The leader first announces his optimal strategy by taking into account the rational reactions of the followers.
  - The follower chooses his optimal strategy based on the leader's announced strategy.
  - The leader implements his optimal strategy (with optimal response of the followers).
  - Stackelberg (stochastic) differential and dynamic games have been studied since 1970's. No general theory exists for the case when leader has dynamic information, except for some special structures.

Outline

MFGs with Major and Minor Agents

- Huang (2010), Nguyen and Huang (2012), Nourian and Caines (2013)
- Major agent's SDE (linear dynamics)

$$dx_0(t) = (A_0 x_0(t) + B_0 u_0(t) + F x^{(N)}(t))dt + D_0 dW_0(t)$$

- Minor agent's linear SDE ($i = 1, \ldots, N$)

$$dx_i(t) = (A_i x_i(t) + B_i u_i(t) + F_i x^{(N)}(t) + G x_0(t))dt + D_i dW_i(t)$$

- Mean field of N minor agents: $x^{(N)}(t) = \frac{1}{N} \sum_{i=1}^{N} x_i(t)$

MFGs with Major and Minor Agents

- Cost function of the major agent (quadratic)

$$J_0^N(u_0, u_{-0}) = \mathbb{E} \int_0^T \|x_0(t) - x^{(N)}(t)\|_{Q_0}^2 + \|u_0(t)\|_{R_0}^2 dt$$

- Cost function of the minor agent (quadratic)

$$J_i^N(u_i, u_{-i}) = \mathbb{E} \int_0^T \|x_i(t) - x^{(N)}(t)\|_{Q_i}^2 + \|u_i(t)\|_{R_i}^2 dt$$

- Major agent is coupled with N minor agents through the mean field term of the N minor agents
- Minor agents are coupled with each other through the mean field term
- Minor agents are coupled with the major agent through the major agent's SDE (major agent's state could also enter into minor agent's cost function)

MFGs with Major and Minor Agents

- The major and minor agents choose their optimal decisions simultaneously
- (decentralized) ϵ-Nash equilibrium

$$J_i^N(u_i^*, u_{-i}^*) \leq J_i^N(u_i, u_{-i}^*) + \epsilon$$

- u_i^*: decentralized ϵ-Nash strategy
  - u_i^* is a function of local state information
- Due to the major agent's state included in the minor agent's SDE and the cost function, the approximated mean field process is a stochastic process that is dependent on major agent's SDE.
- Under certain technical conditions [Huang, SICON 2010], there exists a mean field process in the infinite population game, and the SDG becomes one of a 2-player NE game between major agent and a generic minor agent.
- Extension to SDG where minor agents have partial information on major agent's state is possible [Sen-Caines, SICON 2016].

MFGs with Major and Minor Agents

- The major and minor agents choose their optimal decisions simultaneously
- (decentralized) ϵ-Nash equilibrium

$$J_i^N(u_i^*, u_{-i}^*) \leq J_i^N(u_i, u_{-i}^*) + \epsilon$$

- u_i^*: decentralized ϵ-Nash strategy
  - u_i^* is a function of local state information
- Due to the major agent's state included in the minor agent's SDE and the cost function, the approximated mean field process is a stochastic process that is dependent on major agent's SDE.
- Under certain technical conditions [Huang, SICON 2010], there exists a mean field process in the infinite population game, and the SDG becomes one of a 2-player NE game between major agent and a generic minor agent.
- Extension to SDG where minor agents have partial information on major agent's state is possible [Sen-Caines, SICON 2016].

MFGs with Major and Minor Agents

- The major and minor agents choose their optimal decisions simultaneously
- (decentralized) ϵ-Nash equilibrium

$$J_i^N(u_i^*, u_{-i}^*) \leq J_i^N(u_i, u_{-i}^*) + \epsilon$$

- u_i^*: decentralized ϵ-Nash strategy
  - u_i^* is a function of local state information
- Due to the major agent's state included in the minor agent's SDE and the cost function, the approximated mean field process is a stochastic process that is dependent on major agent's SDE.
- Under certain technical conditions [Huang, SICON 2010], there exists a mean field process in the infinite population game, and the SDG becomes one of a 2-player NE game between major agent and a generic minor agent.
- Extension to SDG where minor agents have partial information on major agent's state is possible [Sen-Caines, SICON 2016].

Outline

2. Mean Field Stackelberg Differential Games

Problem Formulation

- SDEs for the leader and N followers

$$dx_0(t) = [A_0 x_0(t) + B_0 u_0(t)]dt + D_0 dW_0(t)$$
$$dx_i(t) = [A x_i(t) + B u_i(t) + F u_0(t)]dt + D dW_i(t), \; 1 \leq i \leq N$$

- Cost functions (leader and N followers)

$$J_0^N(u_0, u^N) = \mathbb{E} \int_0^T \Big[\| x_0(t) - H_0 \frac{1}{N} \sum_{i=1}^N x_i(t) \|_{Q_0}^2 + \| u_0(t) \|_{R_0}^2 \Big] dt$$
$$J_i^N(u_i, u_{-i}, u_0) = \mathbb{E} \int_0^T \Big[\| x_i(t) - \frac{H}{N} \sum_{i=1}^N x_i(t) \|_Q^2 + \| u_i(t) \|_R^2 + 2 u_i^T(t) L u_0(t) \Big] dt$$

- Leader and followers are coupled through the mean field term and the control of the leader
- Stochastic mean field approximation

Moon and Başar, CDC (2015), Automatica (under review, 2018)

Problem Formulation

- Information structure: Adapted open loop
- Hierarchy of the (global) Stackelberg game
  - The leader has a global advantage over the followers
  - At the outset of the game, the leader announces his strategy over the whole planning horizon ($[0, T]$), and commits to it
  - With the knowledge of the leader's strategy, the followers determine their strategies non-cooperatively over the entire horizon by minimizing the individual cost functions (the followers play a Nash game)
  - The leader takes into account the followers' optimal response to minimize the individual cost functions

Difficulty

- Two different optimization problems
  - Nash game of the N followers given an arbitrary u_0 of the leader
  - Leader's optimization problem with additional constraints induced by the N Nash followers

- Characterization of a Nash-Stackelberg equilibrium is difficult
  - Complexity of the Nash game of the followers increases with respect to N
  - Number of additional constraints for the leader's optimization problem induced by the N Nash followers increases with respect to N
  - Informational complexity (centralized information)

Objective

- Mean field analysis
  - Complexity of the Nash game of the followers is reduced
  - The number of additional constraints for the leader's optimization problem is independent of N

Main Objectives

- Obtain an *approximated* stochastic mean field process
- Characterize a (decentralized) ϵ-Nash-Stackelberg equilibrium
  - The optimal controls of the leader and the followers are functions of their local information

A Generic Follower's Local Optimal Control Problem

- SDE and cost function (given u_0 and z)

$$dx_i(t) = [Ax_i(t) + Bu_i(t)]dt + DdW_i(t)$$

$$\bar{J}_i(u_i) = \mathbb{E}\int_0^T \Big[\|x_i(t) - Hz(t)\|_Q^2 + \|u_i(t)\|_R^2 + 2u_i^T(t)Lu_0(t)\Big]dt$$

- $z(t)$ is an arbitrary stochastic process
- Local optimal decentralized controller **[Rec]**

$$u_i^*(t) = R^{-1}B^T p_i(t) - R^{-1}Lu_0(t)$$

$$dx_i^*(t) = \Big[Ax_i^*(t)dt + BR^{-1}B^T p_i(t) - BR^{-1}Lu_0(t)\Big]dt + DdW_i(t)$$

$$dp_i(t) = \Big[-A^T p_i(t) + Q(x_i^*(t) - Hz(t))\Big]dt + r_i(t)dW_i(t), \;\; p_i(T) = 0$$

- Given z and u_0, the FBSDE admits a unique solution

Stochastic Mean Field Approximation

- $(1/N)\sum_{i=1}^{N} r_i(t)dW_i(t)$ and $(1/N)\sum_{i=1}^{N} DdW_i(t)$ are negligible when N goes to infinity (SLLN)
- Candidate stochastic mean field process (given u_0) **[Rec]**

$$dz(t) = \Big[Az(t) + BR^{-1}B^T p(t) - BR^{-1}Lu_0(t)\Big]dt,\ z(0) = 0$$

$$dp(t) = \Big[-A^T p(t) + Q(z(t) - Hz(t))\Big]dt + l_0(t)dW_0(t),\ p(T) = 0$$

- The mean field stochastic process z depends on the strategy of the leader (as expected): z is not a deterministic process!

Main Result for N-Nash Followers

Optimality of N-Nash Followers

- Stochastic mean field approximation: For any u_0,

$$\mathbb{E}\int_0^T \Big\|\frac{1}{N}\sum_{i=1}^N x_i^*(t) - z(t)\Big\|^2 dt = O(\frac{1}{N})$$

- ϵ-Nash equilibrium: For any u_0, $u^{N*} = \{u_i^*, 1 \le i \le N\}$ constitutes an ϵ-Nash equilibrium for the N followers. That is, for any i, $1 \le i \le N$,

$$J_i(u_i^*, u_{-i}^*, u_0) \le \inf_{u_i} J_i(u_i, u_{-i}^*, u_0) + \epsilon,\ \ \epsilon = O(1/\sqrt{N}).$$

- z is an approximated (stochastic) mean field behavior
- Optimality of the N followers holds for any u_0

Leader's Local Optimal Control Problem

- Leader's local optimal control problem

$$\bar{J}_0(u_0) = \mathbb{E} \int_0^T \left[\|x_0(t) - H_0 z(t)\|^2_{Q_0} + \|u_0(t)\|^2_{R_0} \right] dt$$

$$\min_{u_0} \bar{J}_0(u_0) \text{ such that}$$

$$dx_0(t) = [A_0 x_0(t) + B_0 u_0(t)]dt + D_0 dW_0(t)$$

$$dz(t) = [Az(t) + BR^{-1}B^T p(t) - BR^{-1} L u_0(t)]dt, \; z(0) = 0$$

$$dp(t) = [-A^T p(t) + Q(z(t) - Hz(t))]dt, \; p(T) = 0$$

- Two additional constraints due to the mean field approximation that are independent of N
- *Non-standard* optimal control problem (hard but still tractable)

Leader's Local Optimal Control Problem

- Optimal decentralized controller: **[Rec]**

$$u_0^*(t) = R_0^{-1} B_0^T p_0(t) - R_0^{-1} L^T R^{-1} B^T \lambda_1(t)$$

where

$$\begin{cases} dx_0^*(t) = [A_0 x_0^*(t) + B_0 R_0^{-1} B_0^T p_0(t)]dt \\ \qquad\qquad - B_0 R_0^{-1} L^T R^{-1} B^T \lambda_1(t) dt + D_0 dW_0(t) \\ dp_0(t) = [-A_0^T p_0(t) + Q_0(x_0^*(t) - H_0 z(t))]dt + q_0(t) dW_0(t) \\ d\lambda_1(t) = [-A^T \lambda_1(t) + H_0^T Q_0(H_0 z(t) - x_0^*(t))]dt \\ \qquad\qquad + (H^T - I) Q \lambda_2(t) dt + q_1(t) dW_0(t) \\ d\lambda_2(t) = [A\lambda_2(t) - BR^{-1} B^T \lambda_1(t)]dt \\ dz(t) = [Az(t) + BR^{-1} B^T p(t) - BR^{-1} L R_0^{-1} B_0^T p_0(t)]dt \\ \qquad\qquad + [BR^{-1} L R_0^{-1} L^T R^{-1} B^T \lambda_1(t)]dt \\ dp(t) = [-A^T p(t) + Q(z(t) - Hz(t))]dt \\ x_0^*(0) = x_0(0),\ \lambda_2(0) = 0,\ z(0) = 0,\ \ p_0(T) = 0,\ \lambda_1(T) = 0,\ p(T) = 0 \end{cases}$$

- x_0^*, λ_2, z are forward SDEs, and p_0, p, λ_1 are backward SDEs

Leader's Local Optimal Control Problem

- Recall that given z and u_0, a generic follower's local optimal control problem admits a unique solution, given in terms of the unique solution of the FBSDE
- Existence of solutions of FBSDEs for the leader's problem:

$$\mathcal{X}(t) = \left(x_0^{*T}(t) \quad \lambda_2^T(t) \quad z^T(t)\right)^T, \; \mathcal{Y}(t) = \left(p_0^T(t) \quad \lambda_1^T(t) \quad p^T(t)\right)^T$$
$$d\mathcal{X}(t) = [\mathcal{A}_1\mathcal{X}(t) + \mathcal{B}_1\mathcal{Y}(t)]dt + \mathcal{D}_1 dW_0(t)$$
$$d\mathcal{Y}(t) = [\mathcal{A}_2\mathcal{X}(t) + \mathcal{B}_2\mathcal{Y}(t)]dt + \mathcal{D}_2(t) dW_0(t)$$

Leader's Local Optimal Control Problem

- Linear transformation: $\mathcal{Y}(t) = -\Lambda(t)\mathcal{X}(t) + V(t)$

$$-\frac{d\Lambda(t)}{dt} = \Lambda(t)\mathcal{A}_1 + \mathcal{A}_1^T\Lambda(t) + \mathcal{A}_2 - \Lambda(t)\mathcal{B}_1\Lambda(t), \quad \Lambda(T) = 0$$

$$dV(t) = [-\mathcal{A}_1^T + \Lambda(t)\mathcal{B}_1]V(t)dt + [\mathcal{D}_2(t) + \Lambda(t)\mathcal{D}_1]dW_0(t), \quad V(T) = 0$$

- Due to the linear transformation, the solution of the FBSDEs exists if the solution of the RDE $\Lambda(t)$ exists for $t \in [0, T]$ with $\Lambda(T) = 0$

- $\Lambda(t)$: *Non-symmetric* RDE ($\mathcal{B}_1$ and $\mathcal{A}_2$ are not symmetric)

Leader's Local Optimal Control Problem

Existence of solution to the RDE II

Let

$$\Upsilon(W_1, W_2) = \begin{pmatrix} W_1\mathcal{A}_1 - W_2\mathcal{A}_2 & -W_1\mathcal{B}_1 + \mathcal{A}_1^T W_2 - W_2\mathcal{A}_1^T - \mathcal{A}_2^T W_2 \\ 0 & -\mathcal{B}_1^T W_2 - \mathcal{A}_1 W_2 \end{pmatrix},$$

where $W_1 = W_1^T > 0$ and $W_2 = W_2^T$. If $\Upsilon + \Upsilon^T < 0$, then the RDE has a unique solution $\Lambda(t)$ for all $t \in [0, T]$ with $\Lambda(T) = 0$.

- Let $\mathcal{F}(W_1, W_2) = \text{diag}\{\Upsilon + \Upsilon^T, -W_1\}$
- $\mathcal{F}(W_1, W_2) < 0$ is a linear matrix inequality (LMI) and therefore can be checked easily via standard semidefinite programming

Main Result: (ϵ_1, ϵ_2)-Stackelberg Equilibrium

Main Result: (ϵ_1, ϵ_2)-Stackelberg Equilibrium

Assume that the leader's local problem admits an optimal solution u_0^*. Then $\{u_0^*, u_1^*(u_0^*), \ldots, u_N^*(u_0^*)\}$ constitutes an (ϵ_1, ϵ_2)-Stackelberg equilibrium, where $\epsilon_1 = \epsilon = O(1/\sqrt{N})$, i.e.

- $u^{N*}(u_0^*) = \{u_1^*(u_0^*), \ldots, u_N^*(u_0^*)\}$:
  ϵ_1-Nash equilibrium under u_0^* with $\epsilon_1 = O(1/\sqrt{N})$
- Leader's optimality

$$J_0^N(u_0^*, u^{N*}(u_0^*)) \leq \inf_{u_0} J_0^N(u_0, u^{N*}(u_0)) + \epsilon_2, \quad \epsilon_2 = O(1/\sqrt{N})$$

- u_0^*: ϵ-Stackelberg decentralized control of the leader (function of x_0 & z)
- $u_i^*(u_0^*)$: ϵ-Nash decentralized control of the follower i (function of x_0 & u_0^*)

Simulations

$$dx_0(t) = [2x_0(t) + u_0(t)]dt + dW_0(t), \; dx_i(t) = [1.3x_i(t) + 2u_i(t)]dt + dW_i(t)$$
$$J_0 = \mathbb{E}\int_0^{10}[(x_0(t) - 0.8x^N(t))^2 + 2u_0^2(t)]dt$$
$$J_i = \mathbb{E}\int_0^{10}[(x_i(t) - 0.7x^N(t))^2 + 2u_0^2(t) + u_i(t)u_0(t)]dt$$
$$\epsilon(N) := (\mathbb{E}\int_0^T \|\frac{1}{N}\sum_{i=1}^N x_i^*(t) - z(t)\|^2 dt)^{1/2} = O(\frac{1}{\sqrt{N}})$$

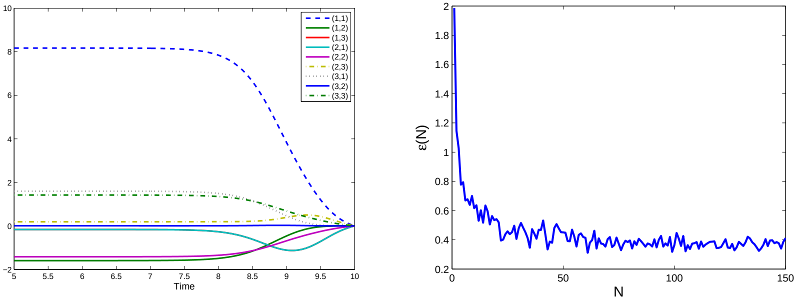

Figure: Left: $\Lambda(t)$, Right: $\epsilon(N)$ with respect to N

Conclusions

- Two approaches to MFGs with disproportionately positioned players
- Major/minor player MFGs with local state information and NE
- Mean field Stackelberg games with adapted open loop
- Characterized the decentralized (ϵ_1, ϵ_2)-Stackelberg equilibrium with $\epsilon_1 = \epsilon_2 = O(1/\sqrt{N})$ for the leader and the followers
- Identified the LMI condition under which the game admits the (ϵ_1, ϵ_2)-Stackelberg equilibrium
- Other types of information structures (many challenges with dynamic strategic information for leader(s))