Two State Mean Field Games

JTGSS lectures 5 and 6

B. Hajek

Department of Electrical and Computer Engineering
and the Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

February 13, 2018

Abstract:

These slides cover the basics of mean field games, based in large part on the paper of [2], which describes the continuous-time, discrete-state context. Much of that paper can be viewed as the discrete-state version of the paper of Lasry and Lions [4] on mean field games. More recent references are [3] and [1]. (Original version: May 31, 2012. The author thanks Michael Livesay for corrections/discussion.)

Outline

I. One player

II. Two players

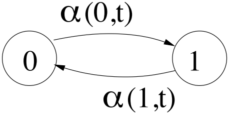
III. $N+1$ players coupled by empirical mean field

IV. The mean field game

Part I. One player

Two-state, continuous-time Markov processes are described and a control problem for them is formulated.

Two-state Markov process with time varying jump rates



Let

$$Q(t) = \begin{pmatrix} -\alpha(0,t) & \alpha(0,t) \\ \alpha(1,t) & -\alpha(1,t) \end{pmatrix}$$

where $\alpha(i,\cdot)$ is continuous, nonnegative over $[0, T]$ for $i \in \{0,1\}$.
Let $\pi(0) = (\pi_0(0), \pi_1(0))$ be a given initial probability distribution.
A two-state, continuous-time Markov process $(X(t) : 0 \le t \le T)$ with initial distribution $\pi(0)$ is said to follow policy α if

$$P(X(t+h) = 1 - i | X(t) = i) = \alpha(i,t)h + o(h)$$

Dynkin formula

For any function $\phi : \{0,1\} \times [0, T]$ that is smooth in its second argument,

$$E\left[\phi(X_t, t) - \phi(X_s, s) - \int_s^t \frac{\partial \phi}{\partial t}(X(r), r) + A\phi(X(r), r)dr \middle| X(s) = i\right] = 0$$

where $A\phi(i, r) = \alpha(i, r)(\phi(1 - i, r) - \phi(i, r))$.

Forward Kolmogorov equations

Let $\pi(i,t) = P\{X(t) = i\}$. The forward Kolmorov equations determine π for given $\pi(0)$ and α:

$$\dot{\pi} = \pi Q \quad \text{or} \quad \dot{\pi}(i,t) = -\alpha(i,t)\pi(i,t) + \alpha(1-i,t)\pi(1-i,t).$$

.

Next, think of α as a control over finite interval $[0, T]$.

$$\min_{\alpha} E\left[\int_0^T c(X(t), \alpha_t, t)dt + \psi(X(T))\right]$$

- Above, α_t is short for $\alpha(X(t), t)$, the rate in force at time t
- Running cost: $c(i, \alpha, t) = f(i, t) + \frac{\alpha^2}{2}$
  - Residence costs per unit time: $f(0, t)$ and $f(1, t)$ are nonnegative functions
  - $\frac{\alpha^2}{2}$ represents cost per unit time of jumping at rate α
  - (A more general form of c is used in [2], but in slides we use this special case.)
- Terminal cost: $\psi(i)$.
- Initial distribution $\pi(0)$ of $X(0)$ specified.

Solution can be found by dynamic programming

Use backwards induction on the cost-to-go functions:

$$u(i,t) = \min_{\alpha} E\left[\int_t^T c(X(s),\alpha_s,s)ds + \psi(X(T)) \middle| X(t)=i\right]$$

$$u(i,t-h) = \min_{a\geq 0} u(i,t)(1-ah)+u(1-i,t)ah+\left(f(i,t)+\frac{a^2}{2}\right)h+o(h)$$

or

$$\frac{u(i,t-h)-u(i,t)}{h} = f(i,t)+\min_{a\geq 0}\left[a(u(1-i,t)-u(i,t))+\frac{a^2}{2}\right]+\frac{o(h)}{\epsilon}$$

or (HJB equations):

$$-\dot{u}(i,t) = f(i,t) - \frac{(u(i,t)-u(1-i,t))_+^2}{2}; \quad u(i,T)=\psi(i)$$

Optimal control: $\alpha^*(i,t) = (u(i,t)-u(1-i,t))_+$

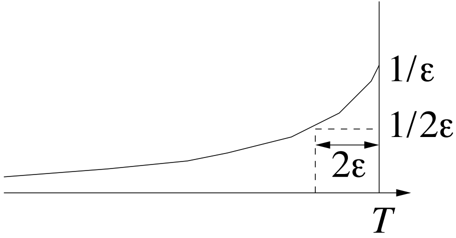
Specialize to terminal cost problem

Suppose $f \equiv 0$ and $\psi(1) > 0 = \psi(0)$. Then $u(0,t) \equiv 0$.
Set $x_t = u(1,t)$. Then the HJB equations become

$$-\dot{x}(t) = -\frac{x(t)^2}{2} \quad \text{with } x(T) = \psi(1)$$

Solution:

$$x(t) = \frac{1}{\epsilon + \frac{T-t}{2}} \quad \text{where} \quad \frac{1}{\epsilon} = \psi(1)$$



Specialize to terminal cost problem (contd)

Let's do accounting of costs starting in state 1 at time t_o
Let τ be the random time of jump from state 1 to state 0.
τ has failure rate function $x(t) = \frac{1}{\epsilon + \frac{T-t}{2}}$ for $t_o \leq t \leq T$
So

$$P(X(T) = 1 | X(t_0) = 1) = \exp\left(-\int_{t_0}^{T} x(s)ds\right) = \left(\frac{x_{t_o}}{x_T}\right)^2$$

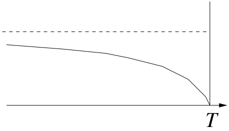
For example, if $t_0 = T - 2\epsilon$ then $P(X(T) = 1|X(t_0) = 1) = \frac{1}{4}$.
yielding

$$x(T-2\epsilon) = \frac{1}{2\epsilon} = \underbrace{\frac{1}{4\epsilon}}_{\text{mean spent on rate}} + \underbrace{\frac{1}{4}\frac{1}{\epsilon}}_{\text{mean spent on terminal cost}}$$

Specialize to fixed running cost problem

Suppose $f(0,t) \equiv 0$, $f(1,t) \equiv \delta > 0$ and $\psi(1) = \psi(0) = 0$.
Then $u(0,t) \equiv 0$. Set $x_t = u(1,t)$. The HJB equations become

$$-\dot{x}(t) = -\frac{x(t)^2}{2} + \delta \quad \text{with } x(T) = 0$$



Set $-\dot{x}(t) = 0$ to find

$$\lim_{T\to\infty} x(0) = \sqrt{2\delta}$$

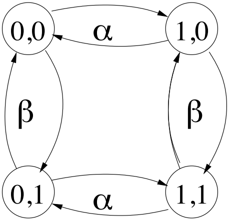
Mean time in state 1 is $\frac{1}{\sqrt{2\delta}}$.
So in the infinite horizon limit, an accounting of mean total cost is

$$\frac{1}{\sqrt{2\delta}} \times \left(\delta + \frac{(\sqrt{2\delta}\ )^2}{2} \right) = \sqrt{2\delta}.$$

Part II. Two players

Suppose two players each control a two-state Markov process, with coupling only through the cost functions. For simplicity, consider a symmetric game. The unique subgame perfect equilibrium (SPE) is determined by backwards differential equations (the continuous time version of dynamic programming.)

Model assumptions



Player one uses strategy $\alpha(i, j, t)$ to control $X_1(t)$.
Player two uses strategy $\beta(j, i, t)$ to control $X_2(t)$.
Running costs $f(i, j) + \frac{\alpha_t^2}{2}$ and $f(j, i) + \frac{\beta_t^2}{2}$
(Assume symmetric running costs and no terminal cost for convenience.)
$u_1(i, j, t)$ is cost to go for player 1 given $(X_1(t), X_2(t)) = (i, j)$.
$u_2(j, i, t)$ is cost to go for player 2 given $(X_2(t), X_1(t)) = (j, i)$.

Subgame perfect equilibrium via dynamic programming

HJB equations determine unique subgame perfect equilibrium

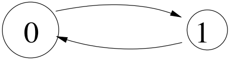
$$
\begin{aligned}
-\dot{u}_1(i,j,t) &= f(i,j) - \frac{(u_1(i,j,t) - u_1(1-i,j,t))_+^2}{2} + \\
&\quad (u_2(j,i,t) - u_2(1-j,i,t))_+ (u_1(i,1-j,t) - u_1(i,j,t))
\end{aligned}
$$

$$
\begin{aligned}
-\dot{u}_2(j,i,t) &= f(j,i) - \frac{(u_2(j,i,t) - u_2(1-j,i,t))_+^2}{2} + \\
&\quad (u_1(i,j,t) - u_1(1-i,j,t))_+ (u_2(j,1-i,t) - u_2(j,i,t))
\end{aligned}
$$

Terminal boundary conditions $u_1(i,j,T) = u_2(j,i,T) = 0$
Represents an ode in eight dimensions, can reduce to four dimensions by symmetry ($u_1(i,j,t) \equiv u_2(i,j,t)$.)

Two frog game

Two frogs wish to sit on separate lily pads; one pad is a bit better



Occupancy cost $f(i,j)$ is given by

$i \backslash j$	0	1
0	1	0
1	ϵ	1

where $0 < \epsilon < 1$.

For $0 < \epsilon << 1$, the frogs mainly want to be apart.
Given they are apart, they'd each prefer state 0.
Think of case $1 << \epsilon T << T$.

What is the subgame perfect equilibrium (SPE)?

Two frog game (continued)

To begin, we expect $u_1(0,1,t) \equiv 0$, $u_1(1,0,t) = \epsilon(T-t)$.

- Once apart, frogs stay apart. Why?
- If frogs start out on the same lily pad, how is it determined which frog gets the better lily pad?

Two frog game (continued)

The SPE equations yield

$$u_1(0,0,t) = \epsilon(T-t) + y_t$$

where y is determined by:

$$-\dot{y}_t = 1 - \epsilon - \frac{3y^2}{2} - \epsilon(T-t)y; \qquad y_T = 0$$

For t near zero, $y_t \approx \frac{1}{\epsilon T}$ Note that y_t is the rate each frog leaves state 0 up until the time one jumps. Frogs are slow to jump. The mean time until the first one jumps is about $\frac{\epsilon T}{2}$. Thus, the lucky frog ends up paying $\frac{\epsilon T}{2}$ on average, and the unlucky frog pays about $\frac{3\epsilon T}{2}$ on average.

Two frog game (continued)

Finally,

$$u_1(1,1,t) = z(t)$$

where z is determined by:

$$-\dot{z} = 1 - \frac{3z^2}{2} + \epsilon(T-t)z; \qquad z(T) = 0$$

For t near zero, $z(t) \approx \frac{2\epsilon T}{3}$. An accounting of the cost-to-go is that both frogs use a large jump intensity $\frac{2\epsilon T}{3}$, so that in average time $\frac{3}{4\epsilon T}$, the lucky one of the frogs jumps to state zero and pays no more. So on average, the frogs each spend about $\frac{3}{4\epsilon T}\frac{1}{2}(\frac{2\epsilon T}{3})^2 = \frac{\epsilon T}{6}$ in an effort to be the lucky frog. Since a frog is lucky with probability one half, he spends on average $\frac{\epsilon T}{2}$ after the jump occurs. Thus the total cost to go for t near zero is $z(t) \approx \frac{2\epsilon T}{3} = \frac{\epsilon T}{6} + \frac{\epsilon T}{2}$.

Two frog game (continued)

- We analyzed the unique SPE
- One strategy would be for frog one, starting from state (1,1), to insist that he will not move, no matter how long frog two stays. In response, frog two leaves in $O(1)$ time units. Why isn't this an SPE? (Besides the fact it would violate uniqueness of SPE, because the roles could be reversed.)
- Are there any other NE besides the SPE?

Part III. $N + 1$ players coupled by empirical mean field

Uniqueness of SPE as for two players holds for N players. Symmetry is not needed. Follows from uniqueness of the HJB ode's. Here we consider a special case of cost function per unit time of a reference player that

- is the same for all players
- depends only on the number of other players in state zero.

The running cost of a reference player

The running cost of a player i is taken to be $c(i, \theta, \alpha) = f(i, \theta) + \frac{\alpha^2}{2}$ where θ is the fraction of other players in state 0.
The HJB equations reduce to a $2(n+1)$-dimensional system of ode's.

Transition rates of controlled Markov process

Reference player uses $\alpha(i, n, t)$ and others use $\beta(j, n, t)$
$(i(t), n(t))_{0 \le t \le T}$ forms a controlled Markov process on
$\{0, 1\} \times \{0, 1, \dots, N\}$ with transition rates:

transition	rate
$(i, n) \to (1 - i, n)$	$\alpha(i, n, t)$
$(i, n) \to (i, n + 1)$	$\gamma^+(i, n, t) = (N - n)\beta(1, n + 1 - i, t).$
$(i, n) \to (i, n - 1)$	$\gamma^-(i, n, t) = n\beta(0, n - i, t).$

(A player $j \neq i$ counts i among its set of other players.)

HJB equation for $N+1$ player system

Denote cost-to-go for reference player by $u(i,n,t)$
The HJB equations become (here $(i,n) \in \{0,1\} \times \{0,1,\ldots,N\}$):

$$\begin{aligned} -\dot{u}(i,n,t) &= f(i,n,t) - \frac{((\alpha^*(i,n,t))^2}{2} \\ &\quad + \gamma^+(i,n,t)(u(i,n+1,t) - u(i,n,t)) \\ &\quad + \gamma^-(i,n,t)(u(i,n-1,t) - u(i,n,t)), \end{aligned}$$

where the corresponding control policy is

$$\alpha^*(i,n,t) = (u(i,n,t) - u(1-i,n,t))_+,$$

with terminal boundary conditions $u(i,n,t) = \psi(i,n)$.

HJB equation for $N + 1$ player system

The HJB equations can be viewed in two different ways.

- For policy β of the other N players fixed, the equations determine the optimal policy for the reference player. i.e. $\alpha^* = BR(\beta)$.
- To find the SPE, replace $\beta(\cdot, \cdot, t)$ by $\alpha(\cdot, \cdot, t)$ for each t. Equations are still ode's with Lipschitz continuous RHS. Solution yields the unique SPE.

The $N+1$ player Markov chain for Markov perfect equilibrium

Suppose all $N+1$ players use the Markov perfect policy α^*. The number of players on state 0, $(Z_t : 0 \leq t \leq T)$ is a continuous time Markov chain on $\{0, \ldots, N+1\}$. If $Z_t = n$, the n players on state 0 each see $n-1$ other players on state 0 and use rate $\alpha^*(0, n-1, t)$. And the $N+1-n$ players on state 1 each see n players on state 0 and use rate $\alpha^*(1, n, t)$. So the transition rates of Z are given by

transition	rate
$n \to n-1$	$q_{n,n-1} = n\alpha^*(0, n-1, t)$
$n \to n+1$	$q_{n,n+1} = (N+1-n)\alpha^*(1, n, t)$

As usual, let $q_{n,n} = -(q_{n,n-1} + q_{n,n+1})$ and all other $q_{n,n'} = 0$. The Kolmogorov forward equations for $\pi_n(t) = P\{Z_t = n\}$ are $\dot{\pi}(t) = \pi(t)Q$ for inital distribution $\pi(0)$ specified. Fluid equations for $\theta_t = Z_t/(N+1)$ are
$\dot{\theta} = -\theta\alpha^*(0, \lfloor (N+1)\theta \rfloor, t) + (1-\theta)\alpha^*(1, \lfloor (N+1)\theta \rfloor, t)$.

Part IV.The mean field game.

Symmetric subgame perfect equilibria of the limiting game as $N \to \infty$ are examined. Mean field limit equations are odes with a mixture of initial and terminal constraints. We examine uniqueness for two examples: one *follow the crowd* and one *avoid the crowd*.

The HJB equations of the mean field game

As $N \to \infty$ we may expect in some cases that the process (θ_t) has a deterministic limit. Then for some initial value $\overline{\theta}$ of θ_0, for a given value of t we do not need to consider all values of θ_t.
Yields mean field game (MFG) equations:

$$\begin{aligned}
\dot{\theta}_t &= (1-\theta_t)(u(1,t)-u(0,t))_+ - \theta_t(u(0,t)-u(1,t))_+ \\
-\dot{u}(i,t) &= f(i,\theta_t,t) - \frac{((u(i,t)-u(0-i,t))_+)^2}{2}
\end{aligned}$$

$$\begin{aligned}
\theta_0 &= \overline{\theta} && \text{boundary condition at } 0 \\
u(i,T) &= \psi(i,\theta_T). && \text{boundary condition at } T
\end{aligned}$$

Example: Follow the crowd cost function

Suppose the cost per time spent in state i is

$$f(i,\theta) = |1-\theta-i| = \begin{cases} 1-\theta & i=0 \\ \theta & i=1 \end{cases}$$

where θ is the fraction of other players in state 0.

Example: Follow the crowd cost function (continued)

The MFG equations become:

$$\begin{aligned}
\dot{\theta}_t &= (1-\theta)(u(1,t)-u(0,t))_+ - \theta(u(0,t)-u(1,t))_+ \\
-\dot{u}(0,t) &= 1-\theta-\frac{((u(0,t)-u(1,t))_+)^2}{2} \\
-\dot{u}(1,t) &= \theta-\frac{((u(1,t)-u(0,t))_+)^2}{2}
\end{aligned}$$

$$\theta_0 = \overline{\theta}$$

$$u(0,T) = u(1,T) = 0$$

The case $\overline{\theta} = \frac{1}{2}$ is interesting. One solution is $\theta \equiv \frac{1}{2}$ and $u(0,t) = u(1,t) = \frac{T-t}{2}$. Are there any other solutions?

Example: Follow the crowd cost function (continued)

Let $y_t = u(1,t) - u(0,t)$. Then the MFG equations become:

$$\begin{aligned} \dot{\theta}_t &= (1-\theta)(y_t)_+ - \theta(-y_t)_+ : \quad \theta_0 = \overline{\theta} \\ -\dot{y}_t &= 2\theta_t - 1 - \frac{y_t^2 sign(y_t)}{2}; \quad y_T = 0 \end{aligned}$$

θ_t represents the faction of players in state 0

y_t is the cost-to-go differential for state 1 minus state 0

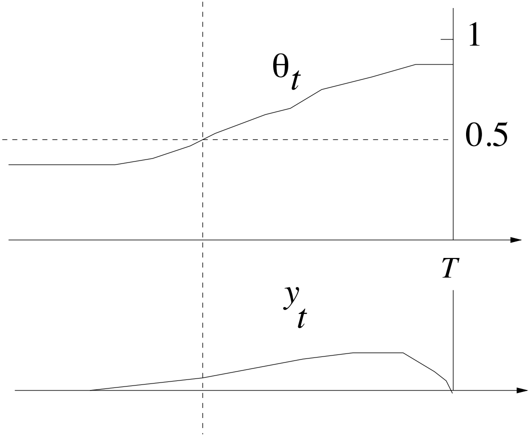
Example: Follow the crowd cost function (continued)

Seek solutions with $y_t \geq 0$ throughout $[0, T]$. Then

$$\begin{aligned} -\dot{y} &= 2\theta_t - 1 - \frac{y_t^2}{2}; \quad y_T = 0 \\ \dot{\theta}_t &= (1-\theta)y_t; \quad \theta_0 = \overline{\theta} \\ (or - \dot{\theta}_t &= -(1-\theta)y_t; \quad \theta_0 = \overline{\theta} \end{aligned}$$

Idea: Think of as backwards equation using a condition such as $\theta_T = \theta_f > 0.5$. Then adjust θ_f to satisfy the initial condition $\theta_0 = \overline{\theta}$.

Example: Follow the crowd cost function (continued)



Must have $\theta_T \geq \frac{1}{2}$ to have nonnegative y. Going backwards in time, up until θ reaches $\frac{1}{2}$, y decreases no slower than as const/time. Thus θ crosses $\frac{1}{2}$ in finite reverse time. Demonstrates multiple solutions. Not only from initial condition $\overline{\theta} = \frac{1}{2}$.

Example: Avoid the crowd cost function

Suppose the cost per time spent in state i is

$$f(i, \theta) = |i - \theta| = \begin{cases} \theta & i = 0 \\ 1 - \theta & i = 1 \end{cases}$$

where θ is the fraction of other players in state 0.
Same as follow the crowd except swap θ and $1 - \theta$.

Example: Avoid the crowd cost function (continued)

Let $y_t = u(1,t) - u(0,t)$. Then the MFG equations become:

$$\begin{aligned} \dot{\theta}_t &= (1-\theta)(y_t)_+ - \theta(-y_t)_+; \quad \theta_0 = \overline{\theta} \\ -\dot{y} &= 1 - 2\theta_t - \frac{y_t^2 sign(y_t)}{2}; \quad y_T = 0 \end{aligned}$$

θ_t represents the faction of players in state 0
y_t is the cost-to-go differential for state 1 minus state 0

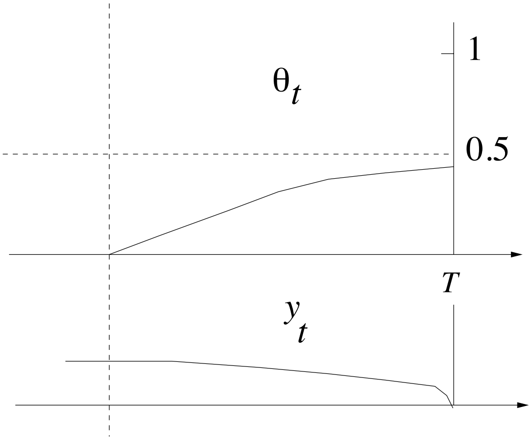
Example: Avoid the crowd cost function (continued)

Seek solutions with $y_t \geq 0$ throughout $[0, T]$. Then

$$
\begin{aligned}
-\dot{y} &= 1 - 2\theta - \frac{y_t^2}{2}; \quad y_T = 0 \\
\dot{\theta}_t &= (1-\theta)y_t; \quad \theta_0 = \overline{\theta} \\
(or - \dot{\theta}_t &= -(1-\theta)y_t; \quad \theta_0 = \overline{\theta}
\end{aligned}
$$

Idea: Think of as backwards equation using a condition such as $\theta_T = \theta_f < 0.5$. Then adjust θ_f to satisfy the initial condition $\theta_0 = \overline{\theta}$.

Example: Avoid the crowd cost function (continued)



Must have $\theta_T \leq \frac{1}{2}$ to have nonnegative y. Backwards trajectories sweep space as the terminal value of θ varies, yielding unique solutions.

Discussion

- See [2] for uniqueness of MFG solutions for more general avoid the crowd condition, and rate of convergence analysis.
- We've seen that for finite N the paper deals with (unique) subgame perfect equilibrium.
- I suspect there are other Nash equilibria even for the finite games. The two frog game could have an ultimatum aspect.
- Nonuniqueness in $N \to \infty$ reflects sensitivity to initial condition in folllow the crowd situations.
- Interestingly, some of the solutions for follow the crowd MFG system can have initial state $\overline{\theta} < \frac{1}{2}$ and yet the crowd congregates to $\theta_f > \frac{1}{2}$. It is as if a coordinator announces that state 0 is the place to be.

P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions.
The master equation and the convergence problem in mean field games.
arXiv 1509.02505, 09 2015.

D. Gomes, J. Mohr, and R.R. Souza.
Mean field limit of a continuous time finite state game.
ArXiv 1001.2918.

D.A. Gomes, J. Mohr, and R.R. Souza.
Continuous time finite state mean field games.
Applied Mathematics and Optimization, 68(1):99–143, August 2013.

J.-M. Lasry and P.-L. Lions.
Mean field games.
Japanese Journal of Mathematics, 2(1):229–260, 2007.