A Mean Field Competition

Yuchong Zhang
(With Marcel Nutz)

Department of Statistics
Columbia University

IPAM WFM2017

Introduction: mean field games

- Introduced by Lasry–Lions and Huang–Malhamé–Caines, 2006
- Nash equilibria for $N \to \infty$ players
- Interaction through empirical distribution ν of the private states
- Typical setting: each player controls a diffusion with some reward and cost-of-effort which depend on ν

- Coupled system: HJB and Kolmogorov PDEs, or FBSDEs
- Only Linear-Quadratic control can be solved explicitly

- Cardaliaguet, Carmona, Delarue, Fouque, Lacker, Yam, ...
- Mean field games of timing: Carmona–Delarue–Lacker, Nutz
- Principal-agent problem: Élie, Mastrolia & Possamai

Mean field competition

- A continuum of agents $i \in (I, \mathcal{I}, \mu)$
- Each agent controls intensity λ of her independent Poisson process
- Quadratic instantaneous cost $c\lambda_t^2\, dt$
- Goal is reached if process jumps (once)

Agents...

- are ranked according to their completion times
- are paid a rank-based reward $R(r)$, $r \in [0, 1]$
- maximize expected reward minus cost

Interaction:

- $\rho(t) = \mu\{i : \tau^i \leq t\}$ = proportion of agents arrived at time t
- Agents observe $\rho(t)$ and choose feedback control $\lambda(\rho(t))$
- Reward $R(\rho(\tau_i))$ and cost $c(\rho(t))$ depend on rank/proportion

State equation

Suppose that all agents use control $\lambda \in \Lambda =$ set of feedback controls (piecewise Lipschitz). Then:

- There exists a unique continuous function $\rho : \mathbb{R}_+ \to [0,1)$ satisfying

$$\rho(t) = \int_0^t \lambda(\rho(s))(1-\rho(s))\,ds, \quad t \geq 0$$

- $\rho(t)$ is the proportion of agents that have finished by time t:

$$\rho(t) = \mu\{i : \tau_\lambda^i(\omega) \in [0,t]\} \quad P\text{-a.s.}$$

- $\rho(t)$ is also the probability that a given agent i has finished by time t:

$$\rho(t) = P\{\tau_\lambda^i \in [0,t]\} \quad \mu\text{-a.s.}$$

This uses the “Exact Law of Large Numbers”

The single-agent problem

- Fix one agent and suppose that all other agents use $\bar{\lambda} \in \Lambda$
- Let $\rho(t)$ be the state as determined by $\bar{\lambda}$
- The reward scheme $R : [0, 1] \to \mathbb{R}_+$ is decreasing, piecewise Lipschitz and left-continuous at $r = 1$
- Before completion, value function of our agent is

$$v(r) = \sup_{\lambda \in \Lambda} E\left[R(\rho(\tau_\lambda)) - \int_0^{\tau_\lambda} c(\rho(t))\lambda(\rho(t))^2 \, dt \middle| \rho(0) = r \right]$$

  Agents never complete are paid $R(1)$
- Nash equilibrium corresponds to a fixed point of

$$\bar{\lambda} \mapsto \rho \mapsto \text{optimal } \lambda$$

Existence and uniqueness

Theorem

There exists a unique (a.e.) equilibrium optimal control $\lambda^* \in \Lambda$,

$$\lambda^*(r) = \frac{R(r) - \frac{1}{2\sqrt{1-r}} \int_r^1 \frac{R(y)}{\sqrt{1-y}}\, dy}{2c(r)}, \quad r \in [0,1)$$

In equilibrium, the value function of any agent before completion is

$$v(r) = \frac{1}{2\sqrt{1-r}} \int_r^1 \frac{R(y)}{\sqrt{1-y}}\, dy, \quad r \in [0,1)$$

Closed-form examples

- Reward budget $B = \int_0^1 R(r)\,dr \geq 0$
- Shape parameter $q \geq 0$
- Cut-off parameter $\alpha \in (0, 1]$

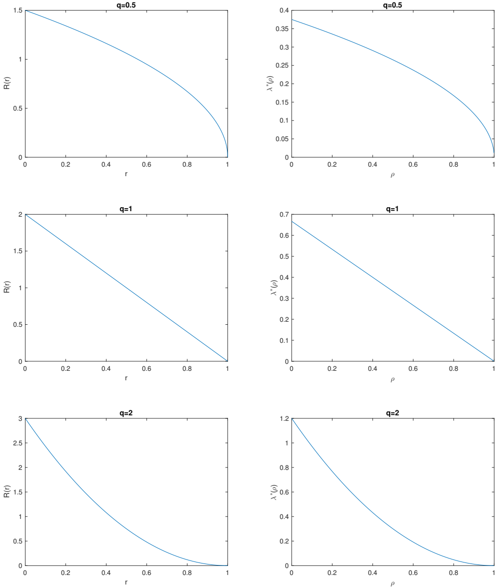
$$R(r) = \kappa(1-r)^q \mathbf{1}_{[0,\alpha]}(r), \quad \kappa = \frac{B(1+q)}{1-(1-\alpha)^{1+q}}$$

Closed-form value function and equilibrium effort:

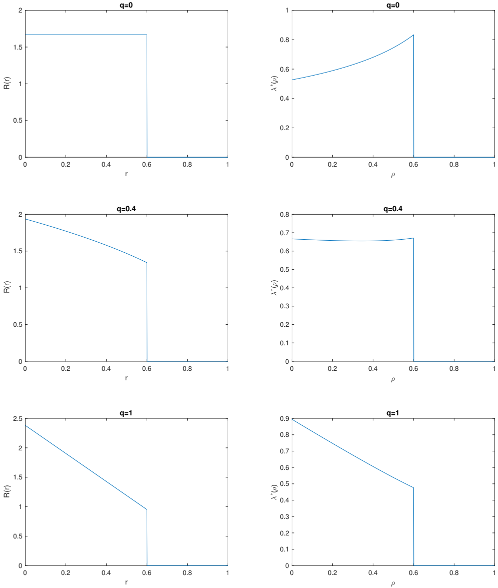
$$v(r) = \frac{\kappa}{(1+2q)}\left((1-r)^q - (1-\alpha)^q\sqrt{\frac{1-\alpha}{1-r}}\right)^+,$$

$$\lambda^*(r) = \mathbf{1}_{\{r\leq\alpha\}}\frac{\kappa}{2c(r)(1+2q)}\left(2q(1-r)^q + (1-\alpha)^q\sqrt{\frac{1-\alpha}{1-r}}\right)$$

Example: no cut-off



Example: cut-off at $\alpha = 60\%$



Planner's problem

- Given R, there exists a unique, deterministic equilibrium state ρ
- Time until α-fraction of the population has completed:

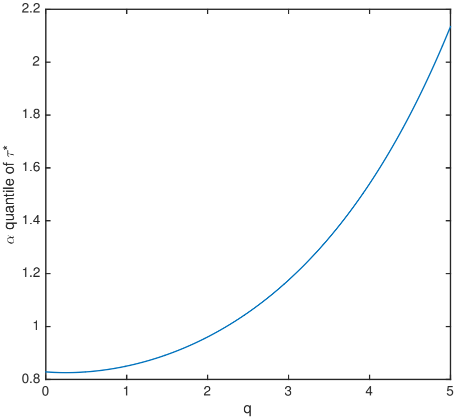
$$T_\alpha(R) = \inf\{t \geq 0 : \rho(t) \geq \alpha\} \in (0, \infty]$$

- Given $\alpha \in (0,1)$ and budget $B > 0$,

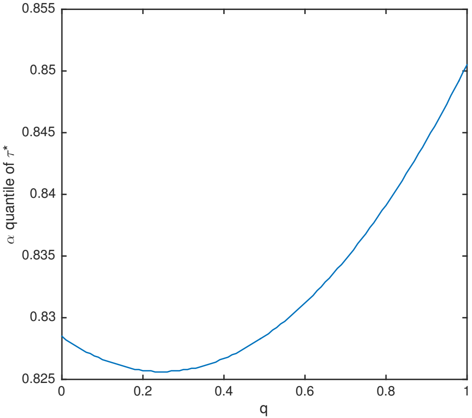
$$\text{minimize } T_\alpha(R) \quad \text{subject to} \quad \int_0^1 R(r)\, dr \leq B$$

- What reward scheme R^* can attain $T^*_\alpha = \inf_R T_\alpha(R)$?
- Assumption: $R \geq 0$, i.e. no punishment is allowed.

$T_\alpha(R)$ for power functions with cut-off at $\alpha = 50\%$



$T_\alpha(R)$ for power functions with cut-off at $\alpha = 50\%$



A calculus of variation problem

- Can show that $T_\alpha(R)$ admits the integral representation:

$$T_\alpha(R) = \int_0^\alpha \frac{2c(r)}{\sqrt{1-r}\left(R(r)\sqrt{1-r} - \int_r^\alpha \frac{R(s)}{2\sqrt{1-s}} ds\right)} dr$$

- A constrained calculus of variation problem.
- Constraints:
  - R is nonnegative
  - R is non-increasing
  - $\int_0^\alpha R(r)dr \leq B$

Main theorem

Theorem

Suppose $\frac{c(r)(1-r)}{2-r}$ is decreasing. Given a fixed reward budget $B > 0$ and $\alpha \in (0,1)$, the optimal non-negative rank-based reward scheme to minimize the α-quantile of the completion time is

$$R^*(r) = \frac{B}{C}\left\{\sqrt{\frac{c(r)}{2-r}} + \frac{1}{2}\int_r^\alpha \frac{1}{1-s}\sqrt{\frac{c(s)}{2-s}}ds\right\}\mathbf{1}_{[0,\alpha]}(r),$$

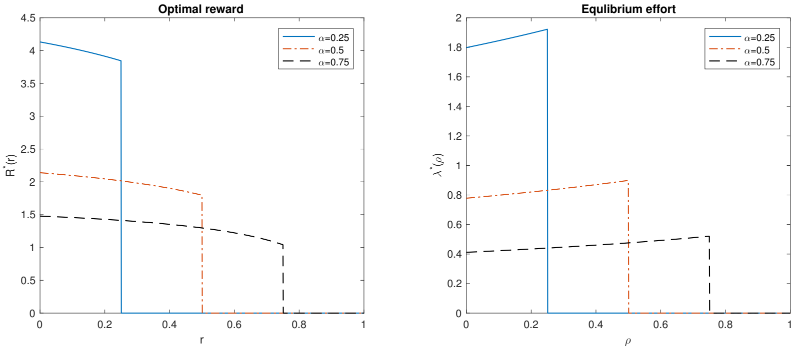
and the minimum α-quantile is

$$T_\alpha^* = \frac{4C^2}{B},$$

where

$$C := \frac{1}{2}\int_0^\alpha \frac{\sqrt{c(r)(2-r)}}{1-r}dr.$$

Plots for $\alpha = 25\%, 50\%, 75\%$, with c constant



General cost function

- What if $\frac{c(r)(1-r)}{2-r}$ is increasing or not monotone?
- By piecewise constant approximation of R and c, we get

$$T_\alpha^\Pi = \sum_{j=1}^n \frac{4c_j\left(\sqrt{1-r_{j-1}}-\sqrt{1-r_j}\right)}{R_j\sqrt{1-r_j}-\sum_{i=j+1}^n R_i\left(\sqrt{1-r_{i-1}}-\sqrt{1-r_i}\right)}$$

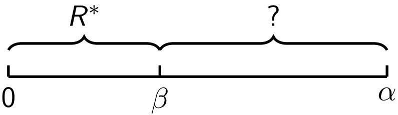
- Letting $\Delta_n = R_n$ and $\Delta_i = R_i - R_{i+1}$, $i < n$, the problem can be equivalently written as

$$\begin{aligned}\underset{\Delta_1,\ldots,\Delta_n}{\text{minimize}} \quad & \sum_{j=1}^n \frac{4c_j(\sqrt{1-r_{j-1}}-\sqrt{1-r_j})}{\sum_{i=j}^n \Delta_i\sqrt{1-r_i}} \\ \text{subject to} \quad & \Delta_1,\ldots,\Delta_n \geq 0,\ \sum_{i=1}^n \Delta_i r_i \leq B.\end{aligned}$$

- A convex programming problem!

Time-inconsistency

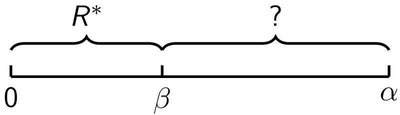
- Is the optimal scheme optimal as the race progresses?



- The optimal remaining scheme R_β can be explicitly computed, and is different from R^*.
- Agents ranked higher than β would have behaved differently under $\mathbf{1}_{[0,\beta]}R^* + \mathbf{1}_{[\beta,1]}R_\beta$

Time-inconsistency

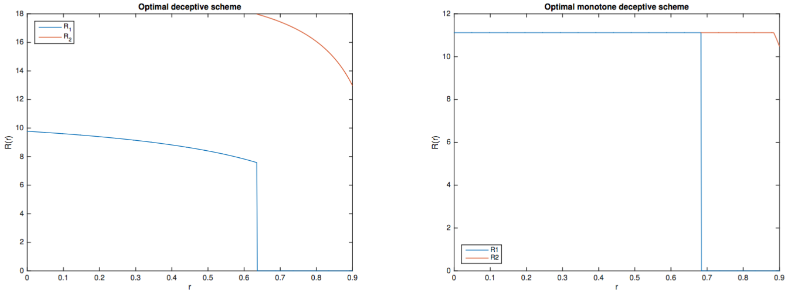
- Is the optimal scheme optimal as the race progresses?



- The optimal remaining scheme R_β can be explicitly computed, and is different from R^*.
- Agents ranked higher than β would have behaved differently under $\mathbf{1}_{[0,\beta]}R^* + \mathbf{1}_{[\beta,1]}R_\beta$

Optimal ways of cheating

- What if the planner cheats and changes the scheme during the competition (once)?
- The optimal way to cheat can be computed in closed form
- With or without monotonicity/regret



$T^{d}_{0.9} = 5.11$ $\quad$ $T^{md}_{0.9} = 5.60$

(without cheating: $T^{*}_{0.9} = 7.32$)

Summary and extensions

- A tractable mean field model of competition
- Unique equilibrium in semi-closed form
- Principal-agent problem explicitly solvable via calculus of variation
- Optimal reward scheme is time-inconsistent
- Future work
    - Relation with N-player game
    - Common noise
    - Multi-stage race
    - Heterogeneous agents

Thank you for your attention!

Exact Law of Large Numbers

Following Yeneng Sun (1998-2014): If $(I, \mathcal{I}, \mu)$ and $(\Omega, \mathcal{F}, P)$ are suitable (in particular atomless), one can construct an extension $(I \times \Omega, \Sigma, \nu)$ of the product which has a Fubini property and supports Σ-measurable i.i.d. families f with any given distribution.

Exact Law of Large Numbers

Let $f : I \times \Omega \to \mathbb{R}$ be Σ-measurable and ν-integrable. If $f(i, \cdot)$, $i \in I$ are essentially pairwise i.i.d. with a distribution having mean m, then $\int f(\cdot, \omega)\, d\mu = m$ for P-almost all $\omega \in \Omega$.