Mean Field Stackelberg Games: State Feedback Equilibrium

Minyi Huang and Xuwei Yang

School of Mathematics and Statistics
Carleton University
Ottawa, Canada

IPAM Workshop III: Mean Field Games and Applications

UCLA, May 4–8, 2020

Outline of talk

- Major player model and literature
- The limiting model of the mean field Stackelberg game
  - A major player (leader)
  - A representative minor player (follower from an infinite population)
- Dynamic programming
  - t-selves
  - Time consistency
- Explicit solutions in LQ models

System model

1) Dynamics: $N+1$ players $\mathcal{A}_k$, $0 \le k \le N$,

$$
\begin{aligned}
\mathcal{A}_0: \quad & dX_t^0 = f_0(X_t^0, \mu_t^{(N)}, u_t^0)dt + \sigma_0 dW_t^0, \\
\mathcal{A}_i: \quad & dX_t^i = f(X_t^0, X_t^i, \mu_t^{(N)}, u_t^0, u_t^i)dt + \sigma dW_t^i, \quad 1 \le i \le N,
\end{aligned}
$$

- $\mathcal{A}_0$: major player $\qquad \mathcal{A}_i$: minor player
- $X_t^k \in \mathbb{R}^n$: state $\qquad$ initial state: X_0^k; $\qquad u_t^k \in \mathbb{R}^{n_1}$: control
- $W_t^k \in \mathbb{R}^{n_2}$: independent standard Brownian motions
- $\mu_t^{(N)} = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$: empirical distribution of all minor players' states
- All initial states are independent with $E|X_0^k|^2 \le C$ for some fixed C

2) Costs:

J_k^{N+1}: cost functional of player $\mathcal{A}_k$

$$J_0^{N+1}(u^0, u^1, \cdots, u^N) = E\int_0^T e^{-\rho t} L_0(X_t^0, \mu_t^{(N)}, u_t^0)dt,$$

$$J_i^{N+1}(u^0, u^i, u^{-i}) = E\int_0^T e^{-\rho t} L(X_t^0, X_t^i, \mu_t^{(N)}, u_t^0, u_t^i)dt$$

- $\rho > 0$, $1 \leq i \leq N$
- $u^{-i} = (u^1, ..., u^{i-1}, u^{i+1}, ..., u^N)$
- For simplicity, the terminal costs are taken as zero

Related literature

Major-minor player mean field models

- Nash equilibrium solutions
  - Huang (2010), Nguyen and Huang (2012), Nourian and Caines (2013), Firoozi and Caines (2015), X. Huang, Jaimungal and Nourian (2019), Kordonis and Papavassilopoulos (2015), Carmona and Zhu (2016), Carmona and Wang (2017), Ma and Huang (2020)
  - Lasry and Lions (2018), Cardaliaguet, Cirant, and Porretta (2018) – Master equations
- Leadership and Stackelberg equilibrium solutions
  - Wang and Zhang (2014) – Use dynamic programming; simple model
  - Bensoussan et al (2015a, 2016, 2017), Moon and Basar (2016, 2018), Elie et al (2019), Fu and Horst (2018) – Mostly depends on the maximum principle on a fixed horizon $[0, T]$, no time consistency.

Two-player Stackelberg games

- von Stackelberg (1934), Basar and Olsder (1999), Yong (2002), Bensoussan et al (2015)

$\mathcal{P}_2(R^n)$: Borel probability measures with finite second moment; endowed the Wasserstein metric $W_2(\cdot,\cdot)$.
C^2_{b2d}: twice continuously differentiable functions with bounded second derivatives.

Assumptions (for this moment):

- Given any Lipschitz full state (x) feedback u^0 and u^i, the drift functions f_0 and f are Lipschitz in (x,μ).
- $0 \le L_0(x^0,\mu,u^0), L(x^0,x^i,\mu,u^0,u^i) \le C[|x^0|^2 + |x^i|^2 + |u^0|^2 + |u^i|^2 + \int |y|^2\mu(dy)]$.

These assumptions ensure well defined state processes and costs under such strategies (if the control space is restricted to be compact, the growth condition can be relaxed.)

The basic problem

- Develop a Stackelberg equilibrium solution of low complexity

Our plan

- Formalize a “right” mean field limit model
- Combine a notion of t-selves with dynamic programming

The notion of t-selves is useful for obtaining equilibrium solutions with time-consistency (adopted by Ekeland and Lazrak (2006) for optimal control with time inconsistent cost)

Error bounds of mean field approximations will be of interest, but will not be discussed here.

Recall

$$
\begin{aligned}
\mathcal{A}_0: &\quad dX_t^0 = f_0(X_t^0, \mu_t^{(N)}, u_t^0)dt + \sigma_0 dW_t^0, \\
\mathcal{A}_i: &\quad dX_t^i = f(X_t^0, X_t^i, \mu_t^{(N)}, u_t^0, u_t^i)dt + \sigma dW_t^i, \quad 1 \le i \le N.
\end{aligned}
$$

The crucial issue here is how to describe a representative minor player in a very compact manner.

So we try mean field approximations.

Denote $\langle \mu, g\rangle = \int g(y)\mu(dy)$. For test function $g \in C^2_{b2d}(\mathbb{R}^n;\mathbb{R})$ and constant vectors $u^0, u \in \mathbb{R}^{n_1}$, by Ito's formula

$$
\begin{aligned}
d\langle \mu_t^{(N)}, g\rangle &= \frac{1}{N}\sum_{i=1}^N dg(X_t^i) \\
&= \frac{1}{N}\sum_{i=1}^N g'(X_t^i) f(X_t^0, X_t^i, \mu_t^{(N)}, u^0, u)dt \\
&\quad + \frac{1}{N}\sum_{i=1}^N \{\frac{1}{2}\mathrm{Tr}[g''(X_t^i)\sigma\sigma^T]dt + g'(X_t^i)\sigma dW_t^i\} \\
&= \langle \mu^{(N)}, g'(\cdot) f(X_t^0, \cdot, \mu^{(N)}, u^0, u) \\
&\quad + \frac{1}{2}\mathrm{Tr}[g''(\cdot)\sigma\sigma^T]\rangle dt + \frac{1}{N}\sum_{i=1}^N g'(X_t^i)\sigma dW_t^i.
\end{aligned}
$$

$\Longrightarrow$ Limiting dynamics

The limiting mean field dynamics

$$d\langle \mu_t, g\rangle = \Big\langle \mu_t, g'(\cdot) f(X_t^0, \cdot, \mu_t, u^0, u) + \frac{1}{2}\mathrm{Tr}[g''(\cdot)\sigma\sigma^T]\Big\rangle dt,$$

where X_t^0 is driven by μ_t instead of $\mu_t^{(N)}$.

Remark on extension: If the constant vector $u \in \mathbb{R}^{n_1}$ is replaced by **individual feedback** $\phi(t, X_t^0, X_t^i, \mu_t)$, the integration should also act on $\phi(t, X_t^0, \cdot, \mu_t)$.

Example (Trivial case): The major player has no effect on μ_t:

$$f(x_0, y, \mu, u^0, u) = \tilde{f}(y, \mu).$$

Then $d\langle \mu_t, g\rangle$ becomes the FPK equation of an underlying McKean Vlasov SDE.

The limiting two-player model: Major player $\mathcal{A}_0$ and representative minor player $\mathcal{A}_1$.

$$
\begin{aligned}
dX_s^0 &= f_0(X_s^0, \mu_s, u_s^0)dt + \sigma_0 dW_s^0, \\
dX_s^1 &= f(X_s^0, X_s^1, \mu_s, u_s^0, u_s^1)ds + \sigma dW_s^1, \\
\Big(\frac{d}{ds}\int_{\mathbb{R}^n} g(y)\mu_s(dy) &= \int_{\mathbb{R}^n} [f^T(X_s^0, y, \mu_s, u_s^0, u_s^1)g'(y) \\
&\qquad + \tfrac{1}{2}\operatorname{Tr}(g''(y)\sigma\sigma^T)]\mu_s(dy)\Big),
\end{aligned}
$$

where $s \geq t$, $X_t^0 = x_0$, $X_t^1 = x_1$, $\mu_t = \mu \in \mathcal{P}_2(\mathbb{R}^n)$, and $g \in C_{b2d}^2(\mathbb{R}^n)$.

- (X_s^0, X_s^1, μ_s): **system state**
- μ_s: also regenerated from empirical distribution of infinite minor players; in general we have neither $\mu_s = \mathcal{L}(X_s^1)$ nor $\mu_s = \mathcal{L}(X_s^1 | X_{[0,s]}^0)$ due to arbitrary choice of x_1 (a further note later).
- The third equation essentially results from the second one. It is informative to list it separately
- Look for **feedback strategies** (u_s^0, u_s^1) where $\mathcal{A}_0$ is the **leader**

Accordingly, define costs

$$J_0(t, x_0, \mu, u^0(\cdot), u^1(\cdot)) = E\int_t^T e^{-\rho(s-t)} L_0(X_s^0, \mu_s, u_s^0) ds,$$

$$J_1(t, x_0, x_1, \mu, u^0(\cdot), u^1(\cdot)) = E\int_t^T e^{-\rho(s-t)} L(X_s^0, X_s^1, \mu_s, u_s^0, u_s^1) ds.$$

Example (linear model)

$$f(x_0,\mu,u^0) = A_0^0(s)x_0 + F^0(s)\int z\mu(dz),$$
$$f(x_0,y,\mu,u^0,u) = A_0(s)x_0 + A(s)y + F(s)\int z\mu(dz)$$

(which may have absorbed feedback like $u^0(s,x_0,\mu)$ and $u(s,x_0,y,\mu)$).

Denote
$$\langle z\rangle_\mu = \int z\mu(dz) \in \mathbb{R}^n.$$

Then we have the linear system for the major player

$$dX_s^0 = [A_0^0(s)X_s^0 + F^0(s)\langle z\rangle_{\mu_s}]ds + \sigma_0 dW_s^0,$$
$$d\langle z\rangle_{\mu_s} = [A_0(s)X_s^0 + A(s)\langle z\rangle_{\mu_s} + F(s)\langle z\rangle_{\mu_s}]ds$$

with initial condition $(x_0,\langle z\rangle_\mu)$, for which we **uniquely** solve $(X_s^0,\langle z\rangle_{\mu_s})$.

Recall

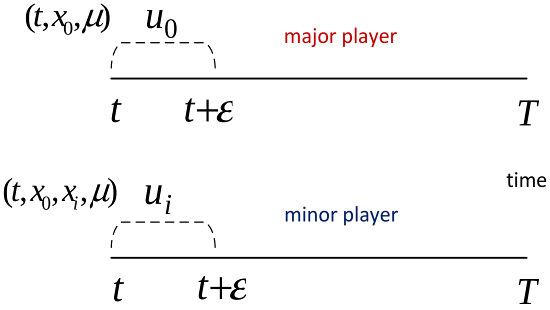
$$
\begin{aligned}
dX_s^0 &= f_0(X_s^0, \mu_s, u_s^0)dt + \sigma_0 dW_s^0, \\
dX_s^1 &= f(X_s^0, X_s^1, \mu_s, u_s^0, u_s^1)ds + \sigma dW_s^1, \\
\frac{d}{ds}\int_{\mathbb{R}^n} g(y)\mu_s(dy) &= \int_{\mathbb{R}^n} [f^T(X_s^0, y, \mu_s, u_s^0, u_s^1)g'(y) \\
&\qquad + \tfrac{1}{2}\operatorname{Tr}(g''(y)\sigma\sigma^T)]\mu_s(dy),
\end{aligned}
$$

where $s \geq t$, $X_t^0 = x_0$, $X_t^1 = x_1$, $\mu_t = \mu \in \mathcal{P}_2(\mathbb{R}^n)$, and $g \in C_{b2d}^2(\mathbb{R}^n)$.

Further remark:

- Here no constraint on the choice of X_t^1 at the initial time t.
- However, for performance analysis in the finite player model, the initial empirical distribution of all minor players should match μ_t in order for the latter to be relevant.
- Existence question (need more information on the two control laws); then use a special McKean-Vlasov interpretation – fix random measure flow $(\mu_s, 0 \leq s \leq T)$; specify a fixed point.

Dynamic programming:



- State feedback
- t-selves
- Two agents (one leader and one follower) optimize for one short period $[t, t+\epsilon]$; after that, the game is taken over by other agents due to the t-self notion.

Value functions for t-$\mathcal{A}_0$ and t-$\mathcal{A}_1$ agents:

$$V_0(t, x_0, \mu) = J_0(t, x_0, \mu, u^{0*}, u^{1*}),$$
$$V_1(t, x_0, x_1, \mu) = J_1(t, x_0, x_1, \mu, u^{0*}, u^{1*}).$$

Task below: determine the Stackelberg strategies (u^{0*}, u^{1*})

Ref. Basar and Olsder (1999) – closed-loop feedback information and dynamic programming in a two-player Stackelberg game

The generators:

$$
\begin{aligned}
\mathcal{L}_0^{u^0}\cdot =& f_0^T(x_0,\mu,u^0)\frac{\partial}{\partial x_0}\cdot+\frac{1}{2}\operatorname{Tr}[(\frac{\partial^2}{\partial x_0^2}\cdot)\sigma_0\sigma_0^T],\\
\mathcal{L}_1^{u^0,u^1}\cdot =& f^T(x_0,x_1,\mu,u^0,u^1)\frac{\partial}{\partial x_1}\cdot+\frac{1}{2}\operatorname{Tr}[(\frac{\partial^2}{\partial x_1^2}\cdot)\sigma\sigma^T],\\
\mathcal{L}_{\mathrm{mf}}^{u^0,u^1}\cdot =& f^T(x_0,y,\mu,u^0,u^1)\frac{\partial}{\partial y}\cdot+\frac{1}{2}\operatorname{Tr}[(\frac{\partial^2}{\partial y^2}\cdot)\sigma\sigma^T].
\end{aligned}
$$

The Stackelberg equilibrium (u^{0*}, u^{1*}), if it exists, is characterized by the Hamilton-Jacobi-Bellman (HJB) equation system for $(t, x_0, x_1, \mu) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)$,

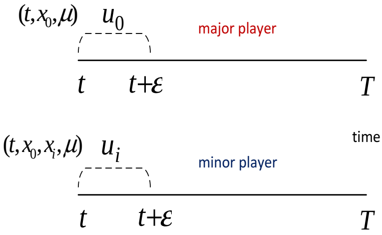
$$\begin{cases} \rho V_1 = \frac{\partial V_1}{\partial t} + L(x_0, x_1, \mu, u^{0*}, u^{1*}) + (\mathcal{L}_0^{u^{0*}} + \mathcal{L}_1^{u^{0*}, u^{1*}}) V_1 \\ \qquad\qquad + \int_{\mathbb{R}^n} \mathcal{L}_{\mathrm{mf}}^{u^{0*}, u^{1*}} \partial_\mu V_1(t, x_0, x_1, \mu; {\color{red}y}) \mu(dy), \\ \rho V_0 = \frac{\partial V_0}{\partial t} + L_0(x_0, \mu, u^{0*}) + \mathcal{L}_0^{u^{0*}} V_0 \\ \qquad\qquad + \int_{\mathbb{R}^n} \mathcal{L}_{\mathrm{mf}}^{u^{0*}, u^{1*}} \partial_\mu V_0(t, x_0, \mu; {\color{red}y}) \mu(dy), \end{cases}$$

where $V_1 = 0$ and $V_0 = 0$ at time T.

Here $\mathcal{L}_{\mathrm{mf}}^{u^0, u^1}$ acts on $\partial_\mu V_i$, $i = 0, 1$ via the y variable, with (t, x_0, x_1, μ) fixed. Note that $\partial_\mu V_i$ has the extra independent variable y.

Assumptions: smoothness and growth conditions on V_1, V_0, and $\partial_\mu V_1(t, x_0, x_1, \mu; \cdot)$, $\partial_\mu V_0(t, x_0, \mu; \cdot)$

For ref. on differentiation w.r.t. μ and the master equation of MFGs, see e.g. Cardaliaguet, Delarue, Lasry, and Lions (2015)



How to *formally* derive the dynamic programming equation

- Assume some smoothness and growth conditions for the value functions; do local expansion on $[t, t+\epsilon]$
- Choose two optimizers on $[t, t+\epsilon]$

$$V_0(t, x_0, \mu) \approx \{L_0\epsilon + e^{-\rho\epsilon} EV_0(t+\epsilon, X^0(t+\epsilon), \mu_{t+\epsilon})\}_{\text{optimize}_{u^0,u^1}},$$

$$V_1(t, x_0, x_1, \mu) \approx \{L\epsilon + e^{-\rho\epsilon} EV_1(t+\epsilon, X^0(t+\epsilon), X^1(t+\epsilon), \mu_{t+\epsilon})\}_{\text{optimize }u^0,u^1}.$$

- Below we explain the choice of the optimizers directly based on the HJB equations

The minor player's optimizer:
Given u^0, denote the minor players' best response by $\hat{u}^1$ as a feedback parametrized by u^0.
For the minor player, denote

$$
\begin{aligned}
H_1 = & L(x_0, x_1, \mu, u^0, u^1) + (\mathcal{L}_0^{u^0} + \mathcal{L}_1^{u^0, u^1}) V_1(t, x_0, x_1, \mu) \\
& + \int_{\mathbb{R}^n} \mathcal{L}_{\mathrm{mf}}^{u^0, \hat{u}^1} \partial_\mu V_1(t, x_0, x_1, \mu; y) \mu(dy).
\end{aligned}
$$

We determine

$$
\hat{u}^1 = \operatorname{argmin}_{u^1} H_1 = \varphi_1(x_0, x_1, \mu, u^0, \tfrac{\partial V_1}{\partial x_1}).
$$

Question: Why u^1 in $\mathcal{L}_{mf}^{u^0, u^1}$ does not contribute to selecting the optimizer $\hat{u}^1$?
Denote $h(y) = \partial_\mu V_1(t, x_0, \mu; y)$. The integral term in H_1 arises from $\langle \mu_{t+\epsilon} - \mu, h \rangle \approx \frac{1}{N} \sum_{k=1}^{N} [h(X_{t+\epsilon}^k) - h(X_t^k)]$, which is nearly unaffected by the control of player $\mathcal{A}_1$ on $[t, t+\epsilon]$.

The major player's optimizer: Recall we have specified

$$\hat{u}^1 = \varphi_1(x_0, x_1, \mu, u^0, \tfrac{\partial V_1}{\partial x_1}).$$

Denote

$$\begin{aligned} H_0 = & L_0(x_0, \mu, u^0) + \mathcal{L}_0^{u^0} V_0(t, x_0, \mu) \\ & + \int_{\mathbb{R}^n} \mathcal{L}_{\mathrm{mf}}^{u^0, \hat{u}^1} \partial_\mu V_0(t, x_0, \mu; y) \mu(dy). \end{aligned}$$

Let the minimizer of H_0 be

$$\begin{aligned} u^{0*} &= \mathrm{argmin} H_0 \\ &= \varphi_0(x_0, \mu, \tfrac{\partial V_0}{\partial x_0}, \partial_\mu V_0(t, x_0, \mu; \cdot), \tfrac{\partial V_1}{\partial x_1}(t, x_0, \cdot, \mu)). \end{aligned} \tag{3.1}$$

Substituting (3.1) into $\hat{u}^1$ gives

$$u^{1*} = \varphi_1(x_0, x_1, \mu, u^{0*}, \tfrac{\partial V_1}{\partial x_1}).$$

We summarize the above procedure.

Given (t, x_0, x_1, μ),

$$u^0 \in \mathbb{R}^{n_1} \longrightarrow \hat{u}^1 \longrightarrow u^{0*} \longrightarrow u^{1*}$$

- For nonlinear dynamics and costs where the control enters the dynamics linearly and the cost in a quadratic form, these optimizers can be explicitly computed.

The closed loop dynamics

$$
\begin{aligned}
dX_s^{0*} &= f_0(X_s^{0*}, \mu_s^*, u_s^{0*})dt + \sigma_0 dW_s^0, \\
dX_s^{1*} &= f(X_s^{0*}, X_s^{1*}, \mu_s^*, u_s^{0*}, u_s^{1*})ds + \sigma dW_s^1, \\
\frac{d}{ds}\int_{\mathbb{R}^n} g(y)\mu_s^*(dy) &= \int_{\mathbb{R}^n} [f^T(X_s^{0*}, y, \mu_s^*, u_s^{0*}, u_s^{1*})g'(y) \\
&\qquad + \tfrac{1}{2}\operatorname{Tr}(g''(y)\sigma\sigma^T)]\mu_s^*(dy),
\end{aligned}
$$

where the initial condition is (x_0, x_1, μ).

Continuous time LQ model

The drift terms in dynamics

$$
\begin{aligned}
f_0(X_t^0, \mu_t, u_t^0) &= A_0 X_t^0 + B_0 u_t^0 + F_0 \langle y \rangle_{\mu_t}, \\
f(X_t^0, X_t^1, \mu_t, u_t^0, u_t^1) &= A X_t^1 + B u_t^1 + D u_t^0 + F \langle y \rangle_{\mu_t} + G X_t^0.
\end{aligned}
$$

Quadratic cost integrands

$$
\begin{aligned}
L_0(X_t^0, \mu_t, u_t^0) &= |X_t^0 - \Gamma_0 \langle y \rangle_{\mu_t}|_{Q_0}^2 + |u_t^0|_{R_0}^2, \\
L(X_t^0, X_t^1, \mu_t, u_t^0, u_t^1) &= |X_t^1 - \Gamma_1 X_t^0 - \Gamma_2 \langle y \rangle_{\mu_t}|_Q^2 \\
&\quad + |u_t^1|_R^2 + |u_t^0|_{R_1}^2 + 2 u_t^{0\,T} R_2 u_t^1.
\end{aligned}
$$

The matrices A_0, A, B_0, B, F_0, F, D, G, Γ_0, Γ_1, Γ_2, Q_0, Q, $R > 0$, $R_0 > 0$, R_1 and R_2 have compatible dimensions.

Denote $\widetilde{D} = D - BR^{-1}R_2^T$.

The Stackelberg equilibrium strategy is

$$
\begin{aligned}
u^{0*} &= -\tfrac{1}{2}R_0^{-1}[B_0^T \tfrac{\partial V_0}{\partial x_0} + \widetilde{D}^T \int_{\mathbb{R}^n} \tfrac{\partial}{\partial y}(\partial_\mu V_0)\mu(dy)], \\
u^{1*} &= -\tfrac{1}{2}R^{-1}[B^T \tfrac{\partial V_1}{\partial x_1} + 2R_2^T u^{0*}].
\end{aligned}
$$

The HJB equations

$$\begin{aligned}\rho V_0 =& \frac{\partial V_0}{\partial t} + (\frac{\partial V_0}{\partial x_0})^T (A_0 x_0 + B_0 u^{0*} + F_0 \langle y \rangle_\mu) \\ &+ \int_{\mathbb{R}^n} (Gx_0 + Ax_1 + F\langle y \rangle_\mu + Bu^{1*} + Du^{0*})|^T_{x_1=y} \cdot \\ &\frac{\partial}{\partial y} \partial_\mu V_0 \mu(dy) + L_0(x_0, \mu, u^{0*}) \\ &+ \frac{1}{2} \mathrm{Tr}[(\frac{\partial^2 V_0}{\partial x_0^2}) \sigma_0 \sigma_0^T + \int_{\mathbb{R}^n} \frac{\partial^2}{\partial y^2} \partial_\mu V_0 \sigma \sigma^T \mu(dy)],\end{aligned}$$

$$\begin{aligned}\rho V_1 =& \frac{\partial V_1}{\partial t} + (\frac{\partial V_1}{\partial x_0})^T (A_0 x_0 + B_0 u^{0*} + F_0 \langle y \rangle_\mu) \\ &+ (\frac{\partial V_1}{\partial x_1})^T (Gx_0 + Ax_1 + F\langle y \rangle_\mu + Bu^{1*} + Du^{0*}) \\ &+ \int_{\mathbb{R}^n} (Gx_0 + Ax_1 + F\langle y \rangle_\mu + Bu^{1*} + Du^{0*})|^T_{x_1=y} \cdot \\ &\frac{\partial}{\partial y} \partial_\mu V_1 \mu(dy) + L(x_0, x_1, \mu, u^{0*}, u^{1*}) \\ &+ \frac{1}{2} \mathrm{Tr}[\frac{\partial^2 V_1}{\partial x_0^2} \sigma_0 \sigma_0^T + \frac{\partial^2 V_1}{\partial x_1^2} \sigma \sigma^T + \int_{\mathbb{R}^n} \frac{\partial^2}{\partial y^2} \partial_\mu V_1 \sigma \sigma^T \mu(dy)].\end{aligned}$$

Assume V_0 and V_1 take the form

$$\begin{aligned}
V_0(t, x_0, \mu) &= x_0^T P_0^0(t) x_0 + \langle y \rangle_\mu^T P_1^0(t) \langle y \rangle_\mu \\
&\quad + 2 x_0^T P_{01}^0(t) \langle y \rangle_\mu + r_0(t), \\
V_1(t, x_0, x_1, \mu) &= x_0^T P_0(t) x_0 + x_1^T P_1(t) x_1 + \langle y \rangle_\mu^T P_2(t) \langle y \rangle_\mu \\
&\quad + 2(x_0^T P_{01}(t) x_1 + x_0^T P_{02}(t) \langle y \rangle_\mu + x_1^T P_{12}(t) \langle y \rangle_\mu) \\
&\quad + r_1(t).
\end{aligned}$$

The feedback Stackelberg strategies are given by

$$\begin{aligned}
u^{0*} &= K_0^0 x_0 + K_1^0 \langle y \rangle_\mu, \\
u^{1*} &= K_0 x_0 + K_1 x_1 + K_2 \langle y \rangle_\mu,
\end{aligned}$$

where the coefficients can be determined (see next page).

$$
\begin{aligned}
K_0^0 &= -R_0^{-1}(B_0^T P_0^0 + \widetilde{D}^T P_{01}^{0T}),\\
K_1^0 &= -R_0^{-1}(B_0^T P_{01}^0 + \widetilde{D}^T P_1^0),\\
K_0 &= -R^{-1}(R_2^T K_0^0 + B^T P_{01}^T),\\
K_1 &= -R^{-1} B^T P_1,\\
K_2 &= -R^{-1}(R_2^T K_1^0 + B^T P_{12}).
\end{aligned}
$$

Denote $[\![Y]\!]^2_M = Y^T M Y$. And we have the Riccati equations:

$$\begin{aligned}\dot{P}_0 = {} & \rho P_0 - P_0(A_0 + B_0 K_0^0) - (A_0 + B_0 K_0^0)^T P_0 \\ & - (P_{01} + P_{02})(G + BK_0 + DK_0^0) \\ & - (G + BK_0 + DK_0^0)^T (P_{01} + P_{02})^T \\ & - K_0^{0T} R_2 K_0 - K_0^T R_2^T K_0^0 - [\![K_0]\!]^2_R - [\![K_0^0]\!]^2_{R_1} - [\![\Gamma_1]\!]^2_Q,\end{aligned}$$

$$\dot{P}_1 = \rho P_1 - P_1(A + BK_1) - (A + BK_1)^T P_1 - [\![K_1]\!]^2_R - Q,$$

$$\begin{aligned}\dot{P}_2 = {} & \rho P_2 - P_{02}^T(B_0 K_1^0 + F_0) - (B_0 K_1^0 + F_0)^T P_{02} \\ & - P_{12}^T(BK_2 + DK_1^0 + F) - (BK_2 + DK_1^0 + F)^T P_{12} \\ & - (A + BK_1 + BK_2 + DK_1^0 + F)^T P_2 \\ & - P_2(A + BK_1 + BK_2 + DK_1^0 + F) \\ & - [\![K_2]\!]^2_R - [\![K_1^0]\!]^2_{R_1} - K_1^{0T} R_2 K_2 - K_2^T R_2^T K_1^0 - [\![\Gamma_2]\!]^2_Q,\end{aligned}$$

$$
\begin{aligned}
\dot{P}_{01} = \rho P_{01} &- (A_0 + B_0 K_0^0)^T P_{01} - P_{01}(A + BK_1) \\
&- (G + BK_0 + DK_0^0)^T (P_1 + P_{12}^T) \\
&- K_0^T R K_1 - K_0^{0T} R_2 K_1 + \Gamma_1^T Q,
\end{aligned}
$$

$$
\begin{aligned}
\dot{P}_{02} = \rho P_{02} &- 2P_0(B_0 K_1^0 + F_0) - (A_0 + B_0 K_0^0)^T P_{02} \\
&- P_{01}(BK_2 + DK_1^0 + F) \\
&- (G + BK_0 + DK_0^0)^T (P_{12} + P_2) \\
&- P_{02}(A + BK_1 + BK_2 + DK_1^0 + F) - K_0^T R K_2 \\
&- K_0^{0T} R_1 K_1^0 - K_0^{0T} R_2 K_2 - K_0^T R_2^T K_1^0 - \Gamma_1^T Q \Gamma_2,
\end{aligned}
$$

$$
\begin{aligned}
\dot{P}_{12} = \rho P_{12} &- P_{01}^T (B_0 K_1^0 + F_0) - P_1(BK_2 + DK_1^0 + F) \\
&- P_{12}(A + BK_1 + BK_2 + DK_1^0 + F) \\
&- (A + BK_1)^T P_{12} - K_1^T R K_2 - K_1^T R_2^T K_1^0 + Q\Gamma_2,
\end{aligned}
$$

$$\begin{aligned}\dot{P}_0^0 &= \rho P_0^0 - P_0^0(A_0 + B_0 K_0^0) - (A_0 + B_0 K_0^0)^T P_0^0 \\ &\quad - P_{01}^0(G + BK_0 + DK_0^0) - (G + BK_0 + DK_0^0)^T P_{01}^{0T} \\ &\quad - [\![K_0^0]\!]_{R_0}^2 - Q_0,\end{aligned}$$

$$\begin{aligned}\dot{P}_1^0 &= \rho P_1^0 - P_{01}^{0T}(B_0 K_0^0 + F_0) - (B_0 K_0^0 + F_0)^T P_{01}^0 \\ &\quad - (A + BK_1 + BK_2 + DK_1^0 + F)^T P_1^0 \\ &\quad - P_1^0(A + BK_1 + BK_2 + DK_1^0 + F) - [\![K_1^0]\!]_{R_0}^2 - [\![\Gamma_0]\!]_{Q_0}^2,\end{aligned}$$

$$\begin{aligned}\dot{P}_{01}^0 &= \rho P_{01}^0 - P_0^0(B_0 K_1^0 + F_0) - (A_0 + B_0 K_0^0)^T P_{01}^0 \\ &\quad - P_{01}^0(A + BK_1 + BK_2 + DK_1^0 + F) \\ &\quad - (G + BK_0 + DK_0^0)^T P_1^0 - K_0^{0T} R_0 K_1^0 + Q_0 \Gamma_0.\end{aligned}$$

All 9 equations have 0 terminal conditions at T.

Recall the feedback strategies (at time t):

$$u^{0*} = K_0^0 X^0(t) + K_1^0 \langle y \rangle_{\mu_t},$$
$$u^{1*} = K_0 X^0(t) + K_1 X^1(t) + K_2 \langle y \rangle_{\mu_t},$$

where $\langle y \rangle_\mu = \int y\mu(dy)$.

Theorem

If the Riccati ODE system of $(P_0, P_1, \cdots, P_{01}^0)$ has a solution on $[0, T]$, then (u^{0*}, u^{1*}) is a feedback Stackelberg equilibrium on $[0, T]$.

A set of equilibrium strategies determined for $[0, T]$ is **time-consistent** if it has been applied up to time $t \in (0, T)$ and is still an equilibrium solution for the remaining period $[t, T]$ for any t.

Related literature

- Ekeland and Lazrak (2006), Ekeland and Pirvu (2008), Bjork and Murgoci (2008), Yong (2017), ...

Theorem

If the Riccati ODE system of $(P_0, P_1, \cdots, P_{01}^0)$ *has a solution on* $[0, T]$*, then the set of strategies is time consistent.*

The major player $\mathcal{A}_0$, a representative minor player $\mathcal{A}_1$, and the mean field state $\overline{X}$ have the dynamics:

$$
\begin{aligned}
X^0_{t+1} =& A_0 X^0_t + B_0 u^0_t + F_0 \overline{X}_t + W^0_t, \quad t = 0, \cdots, T-1,\\
X^1_{t+1} =& A X^1_t + B u^1_t + D u^0_t + F \overline{X}_t + G X^0_t + W^1_t,\\
\overline{X}_{t+1} =& (A+F)\overline{X}_t + B\overline{u}_t + D u^0_t + G X^0_t,
\end{aligned}
$$

where W^i_t, $0 \le t \le T-1$ are i.i.d. with zero mean and finite variance. The control mean field $\overline{u}_t = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N u^i_t$, where each u^i_t is a minor player's control. Cost functionals are

$$
\begin{aligned}
J_0(u^0, u^1) &= E[\sum_{t=0}^{T-1} \alpha^t (|X^0_t - \Gamma_0 \overline{X}_t|^2_{Q_0} + |u^0_t|^2_{R_0})],\\
J_1(u^0, u^1) &= E[\sum_{t=0}^{T-1} \alpha^t (|X^1_t - \Gamma_1 X^0_t - \Gamma_2 \overline{X}_t|^2_Q\\
&\qquad + |u^1_t|^2_R + |u^0_t|^2_{R_1} + 2 {u^0_t}^T R_2 u^1_t)],
\end{aligned}
$$

where $\alpha \in (0,1)$ is the discount factor.

Main result on the discrete time LQ model

- The Stackelberg strategies are determined from dynamic programming and Riccati equations.

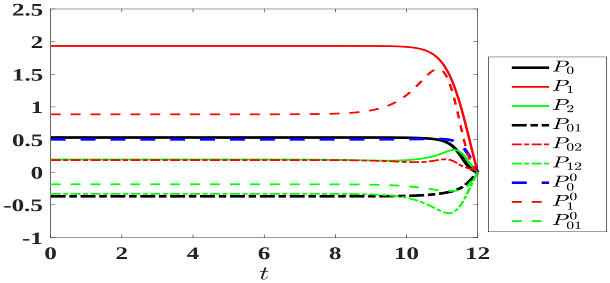
Difference of the continuous and discrete LQ models:

- To maintain leadership, the continuous time case requires the cross term $2{u_t^0}^T R_2 u_t^1$.
- The discrete time case may still have leadership even if the coefficient $R_2 = 0$.

Riccati equations

Continuous time LQ model

The parameter values are given by $A_0 = 1$, $B_0 = 2$, $F_0 = 0.5$, $A = 0.5$, $B = 1$, $D = 1$, $F = 0.2$, $G = 0.4$, $\Gamma_0 = 0.8$, $\Gamma_1 = 0.3$, $\Gamma_2 = 0.5$, $Q = 2$, $Q_0 = 1$, $R = 1$, $R_0 = 0.5$, $R_1 = 1$, $R_2 = 0.5$, $T = 12$, and $\rho = 0.1$.



Summary:

- Mean field Stackelberg games and dynamic programming
- Explicit solutions in LQ models
- To do: performance estimates with finite players

Thank you!