日本オペレーションズ・リサーチ学会
2021年 秋季研究発表会

透過型平均場ゲームにおける大規模な群衆のランダムな動き

01307080　早稲田大学　豊泉洋　TOYOIZUMI Hiroshi

1.　はじめに

他のプレイヤーを観測できる透過型平均場ゲームを用いて、大規模な群衆の動きが、少数の種になるプレイヤーの動きに依存する確率的な平均場として現れることを示す。

動物や人間の社会行動において、周囲の個体と相互作用を起こしながら全体として動く、群衆の動きはさまざまな場面でその性質が研究されている [1, 2, 3, 4]。一方、多数の独立で理性的なプレイヤーが参加するゲームの解析に平均場ゲームが提案され、その解析手法は金融工学などの分野に多数応用されている [5, 6, 7]。

ここでは、従来の平均場ゲームでの大数の法則に基づく決定的な平均場とは異なり、プレイヤーがゲーム内での相手の行動を観測し自分の行動を変える透過型ゲーム [8] では、Polya の壺のように少数の種になるプレイヤーの動きに応じてランダムな平均場となる均衡状態が生じることを、後退型の Hamilton-Jacobi-Bellman 方程式と前進型の輸送方程式がカップリングした連立方程式を具体的に解くことで示す。

2.　透過型協調ゲームと Polya の壺

2 状態 $Q = \{0,1\}$ 上での多数のプレイヤー $X_{-n},\dots,X_{-1},X_0$ の協調ゲームを考える。以下ではプレイヤーとその状態は同一視し、例えば X_0 は Q に値をとる確率変数と考える。プレイヤー X_0 は、ランダムに選ばれたプレイヤー Y とゲームを行い、$F = F(X_0, Y) = 1_{\{X_0=Y\}}$ に従う報酬を得る。

透過型ゲームでは、相手のプレイヤー Y の状態が見えるため、X_0 の最適条件戦略はミラー戦略 $\pi_0(Y) = \delta_Y$ となる。ここで δ_x はディラック測度とする。時刻 $-n$ に種となる $b+r$ 人のプレイヤーが $(r/(b+r))\delta_0 + (b/(b+r))\delta_1$ のように Q 上に分布しゲームが始まり、ミラー戦略に従うプレイヤーが次々と加わり累積すると考え、$n \to \infty$ とすると、確率変数列 $\dots,X_{-n},\dots,X_{-1},X_0$ は、Polya の壺 [9, 10] からランダムに抜き出されたボールと同じ性質を持つ。その時刻 0 での平均場

$$M = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \delta_{X_{-i}(t)}$$

は、遠い過去を示すランダムな末尾事象 (tail event) であり、ベータ分布（$M(1) = 1 - M(2) \sim \beta(b,r)$）を持つ。また、$(X_{-n},\dots,X_{-1},X_0)$ は、入れかえ可能な事象 (exchangeable event) であり、その意味でこの群衆内の各プレイヤーは匿名性を持つ。

3.　時間発展のある透過型平均場ゲーム

前節で見たような形で初期プレイヤーが既に時刻 0 で無数に累積していたと仮定し、その後、プレイヤーが引き続き透過型協調ゲームを行うことを考える。プレイヤー $X_0(t)$ は、レート 1 の Poisson 時計 $N_0(t)$ に従う行動時刻に透過型協調ゲームを行う。$X_0(t)$ はスイッチングコスト α を支払い、状態を変化できる。その瞬間ペイオフは $F = 1_{\{X_0(t)=Y(t)\}} - \alpha|dX_0(t)|$ となる。ここで、プレイヤー $Y(t)$ は、無限のプレイヤーからなる平均場

$$M(t) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \delta_{X_{-i}(t)}$$

からランダムに選ばれ、$X_0(t)$ は $Y(t)$ を観測した上で、$\pi_0 = \pi_0(t,x,y) = P(X_0(t) \in \cdot | X_0(t-) = x, Y(t) = y)$ という条件戦略に従って状態を変化させる。この条件戦略は、割引率 β を考慮した条件付き期待累積利得

$$U(t,x) = E\left[\int_t^\infty e^{-\beta(s-t)} F dN_0(s) \right.$$
$$\left. | X_0(t) = x, M(s), s \ge t \right],$$

を最大化するように選択する。

他のプレイヤーも同様に最適化された戦略をとると考えると、この問題は、最適戦略 $\pi_0 = \pi_0(t,x,y)$、ランダムな価値関数 $U = U(t,x)$、ランダムな平均場 $M = M(t)$ で以下の Hamilton-Jacobi-Bellman 方程式と輸送方程式を同時に満たすものを見つけ

ることに帰着される。

$$\pi_0 \in \operatorname*{argmax}_{\pi\in\mathcal{P}(Q)} \int_{x'\in Q} \{F+U(t,x')-U(t,x)\}\pi(dx'),$$
$$-\frac{dU(t,x)}{dt}+\beta U(t,x)$$
$$=\int_{Q^2}\{F+U(t,x')-U(t,x)\}\,\pi_0(dx')dM(y),$$
$$\frac{dM(t)}{dt}=\int_{Q^2}\pi_0 M(t,dx)M(t,dy)-M(t),$$
$$\lim_{t\to\infty}U(t,x)=U_\infty(x), M(0)=M_0, \qquad (1)$$

ここで、終期値 $U_\infty(x)$ は次の定常状態方程式を満たす。

$$\beta U_\infty(x)=\int_{Q^2}\{F+U_\infty(x')-U_\infty(x)\}\,\pi_0(dx')M(\infty,dy).$$

また、初期値 M_0 は、前節で得られたベータ分布に従うランダムな平均場である。

4. ランダムな平均場解による群衆の行動

時間発展のある透過型平均場ゲームでは、他の群衆の動きによって将来得られる利得を考えて、今ミラー戦略を取るのが得策ではないケースも考えられ、プレイヤーはより戦略的な思考が望まれる。

以下では、$\alpha \leq (1+\beta)/(2+\beta)$ であり、スイッチングコストが十分小さい場合を考える。この場合には、最適戦略はミラー戦略

$$\pi_0=\pi_0(t,x,y)=\delta_y,$$

であり、その平均場は $M(t)\equiv M_0$ となり、群衆のランダムな平均場が出現する。またこの場合の価値関数は $U(t,x)\equiv U_\infty(x)=\frac{1-\alpha(1-M_0(x))}{1+\beta}+\frac{1}{1+\beta}\frac{1-2\alpha M_0(1)M_0(0)}{\beta}$ となる。実際にこの解を (1) に代入すると、$\alpha \leq (1+\beta)/(2+\beta)$ ならば、ミラー戦略を行う平均場に対する最適反応がミラー戦略であり、均衡することがわかる。

また、通常の平均場ゲームでは、$\pi_0=\delta_1$（相手プレイヤーの戦略によらず、全員必ず状態 1 を選択する）はナッシュ均衡だが、時間発展のある透過型平均場ゲームでは、プレイヤーの行動時刻に遅れが生じ、状態 0 に留まるプレイヤーがいるため、$\pi_0=\delta_1$ は自身の最適反応にはならない。

参考文献

[1] A. Okubo, “Dynamical aspects of animal grouping: Swarms, schools, flocks, and herds,” *Advances in Biophysics*, vol. 22, pp. 1–94, 1986.

[2] L. Zhao, G. Yang, W. Wang, Y. Chen, J. P. Huang, H. Ohashi, and H. E. Stanley, “Herd behavior in a complex adaptive system,” *Proceedings of the National Academy of Sciences*, vol. 108, p. 15058, 09 2011.

[3] A. V. Banerjee, “A simple model of herd behavior,” *The quarterly journal of economics*, vol. 107, no. 3, pp. 797–817, 1992.

[4] D. S. Scharfstein and J. C. Stein, “Herd behavior and investment,” *The American Economic Review*, vol. 80, no. 3, pp. 465–479, 1990.

[5] J.-M. Lasry and P.-L. Lions, “Mean field games,” *Japanese journal of mathematics*, vol. 2, no. 1, pp. 229–260, 2007.

[6] P. Cardaliaguet, “Notes on mean field games.” from P.-L. Lions’ lectures at Coll‘ege de France, 2010.

[7] P. E. Caines, M. Huang, and R. P. Malhamé, “Mean field games.,” in *Handbook of Dynamic Game Theory*, pp. 706–712, Springer International Publishing, 2018.

[8] A. M. Unakafov, T. Schultze, A. Gail, S. Moeller, I. Kagan, S. Eule, and F. Wolf, “Emergence and suppression of cooperation by action visibility in transparent games,” *PLoS computational biology*, vol. 16, no. 1, p. e1007588, 2020.

[9] D. J. Aldous, “Exchangeability and related topics,” in *École d’Été de Probabilités de Saint-Flour XIII—1983*, pp. 1–198, Springer, 1985.

[10] R. Durrett, *Probability: theory and examples.* Cambridge university press, 2019.