2014 BinN 理論談話会　#13

Congestion pricing and learning in traffic network games

Melo,Emerson.
journal of Public Economic Theory,
Vol. 13(3), pp. 351-367, 2011.

2014.06.28(Fri)　　M1　芝原 貴史

もくじ

1. Introduction
2. Congestion Pricing and Payoff-Based Adaptive Dynamics
3. Global Convergence of the Learning Process
    1. Rest Points and Nash Equilibrium
    2. The Symmetric Case
4. Comments and Final Remarks

1.Introduction

- 交通混雑は典型的な外部不経済の問題
  - 経済学的な解法はPigou（1920）
  - Merginal Cost Pricing　(Beckmann, McGuire, and Winstein (1956))
  - すべての経路に外部不経済に応じた課金＝走行すること自体が他者に影響を与えることが
  - ユーザーは経路の遅延によるコストと，料金によるコスト の2つを考慮しなければならない.

1.Introduction

- この施策は，プランナーがユーザーの情報を得ていることを前提にしている.
- 均衡状態にどのように到達するかがわからない.

- 均衡状態への到達を動的に示した論文は，Sandholm (2005,2007)がある.
  - Sandholm(2005)では進化ゲーム理論を集計人口レベルで用いる.
  - Sandholm(2007)では個人のプレイヤーを想定した.

  - これらの論文は，計画者がmerginal costを導入することでゲームがナッシュ均衡へ収束することを示す．古典的にはRosenthal(1973)で示された.
  - Monderer and Shapley(1996) and Sandholm(2001)で一般化されて示された.

1.Introduction

- 本論文では，学習過程の適用を提案.
- Pay-off based な過程（Cominetti, Melo, and Sorin(2010)）で交通ゲームの文脈で効率的な均衡状態に至る収束過程を証明する.

- ルールは，
  - プレイヤーは前もって各選択肢の利得を推定可.
  - 効用最大化理論に基づき，ロジットモデルを用いて選択確率を表現.
  - プレイヤーは以前通った経路の利得のみ知覚できるが，これは他のプレイヤー全員の選択から決定される混雑度合いに依存.
  - プランナーはプレイヤーの利得を平準化．このときの経済原理で効率的な均衡状態に到達.
  - その利得がその選択肢の次回以降の利得として更新＝学習過程
  - これを繰り返し行う

1.Introduction

- 論文の特徴は以下2つ

- Reinforcemet model に似た学習過程を使用.
  - 確率的選択と緩やかな適用？の性質を持つ
- Merginal cost pricing という単純な施策がプランナーの有効な分散手法であることを理論的に示した.
  - 多くのプレイヤーのいる交通戦略において最小の情報を集めれば良いということ
  - リアプノフ関数を用いてCominetti, Melo, and Sorin(2010)とは異なる結果を得た.

2.Congestion Pricing and Payoff-Based Adaptive Dynamics

N ：プレイヤーの総数

$i \in P$：選択したプレイヤーi

$r \in R$：選択したルート

c_u^r ：u 人のユーザーが通るときの経路 r の平均旅行時間

$S_i = R$：ユーザー i の純粋戦略の集合

$$\Delta = \prod_{i \in P} \Delta^i$$

：純粋戦略S_iの元での混合戦略の集合

$$g_n^i = G^i(r_n) = -c_u^r$$

：stage n におけるユーザー i が経路 r を選択したとき

利得 コスト

$$x_n^i \in \left(x_n^{ir}\right)_{r \in R}$$

：経路rについて個人iが認知している旅行時間

$$\sigma^{ir}(x^i) = \frac{\exp\left(\beta_i x^{ir}\right)}{\sum_{a \in R} \exp\left(\beta_i x^{ia}\right)} \qquad (1)$$

$$\pi_n^i = \sigma^i(x_n^i) \in \Delta^i$$

：期待利得

2.Congestion Pricing and Payoff-Based Adaptive Dynamics

$$\sigma^{ir}(x^i) = \frac{\exp\left(\beta_i x^{ir}\right)}{\sum_{a \in R} \exp\left(\beta_i x^{ia}\right)} \tag{1}$$

- プレイヤーの利己的な戦略は他者の利得を見越したものではない
- ナッシュ均衡は非効率的な均衡状態といえる.
- この選択はプランナーの目的に反する→個々の選択を考慮した，効用の総和の最大化

＝ネットワークの総コスト最小化問題と等価.

2.Congestion Pricing and Payoff-Based Adaptive Dynamics

- 定式化

c_u^r ：u 人のユーザーが経路 r を選択したときのコスト

$$\overline{H}(\pi) = -\mathbf{E}\left[\sum_{r\in R} U^r c_{Ur}^r\right] \qquad \overline{H}:[0,1]^{P\times R} \to \mathbf{R} \qquad (2)$$

$U^r = \sum_{i\in P} X^{ir}$ X^{ir} ：ベルヌーイ分布の確率変数

$\mathbf{P}\left(X^{ir}=1\right) = \pi^{ir}$ $\pi_n^i = \sigma^i(x_n^i) \in \Delta^i$ ：期待利得

$x_n^i \in \left(x_n^{ir}\right)_{r\in R}$ ：経路rについて個人iが認知している旅行時間

- (2)を最大化するように，プランナーはmerginal cost c_{Ur}^r を考える.

2.Congestion Pricing and Payoff-Based Adaptive Dynamics

- (2)を最大化するためにプランナーは各経路に税 p_u^r を課す.

$$p_u^r = (u-1)(c_u^r - c_{u-1}^r) \qquad \forall r \in R \tag{3}$$

p_u^r ：経路rをu人のユーザーが通っているときの税

- この定式化は古典的な merginal cost pricing を動的に発展させたもの

2.Congestion Pricing and Payoff-Based Adaptive Dynamics

- Stage n において割引された利得とコストは,

$$\tilde{g}_n^i = \tilde{G}^i\left(r_n\right) = G^i\left(r_n\right) - p_u^r = -\tilde{c}_u^r \qquad \forall r \in R$$

$$\tilde{c}_u^r = c_u^r + p_u^r$$

- $\tilde{c}_u^r$ が凸で増加関数なら, $\tilde{c}_1^r \leq \tilde{c}_2^r \cdots \leq \tilde{c}_N^r$

2.Congestion Pricing and Payoff-Based Adaptive Dynamics

- 学習過程のフローと定式化

x_n^{ir} ：n期，個人iの経路rについて認知している利得

↓

π_n^{ir} ：n期に個人iが経路rを選択する確率

↓

r_n^i ：n期に個人iに選択された経路

↓

u_n^r ：n期，経路rの混雑

↓

$\tilde{g}_n^i$ ：混雑度による課税を考慮した利得

↓

x_{n+1}^{ir} ：n+1期，個人iの経路rについて認知している情報（更新）

$$x_{n+1}^{ir} = \begin{cases} (1-\gamma_n)x_n^{ir} + \gamma_n \tilde{g}_n^i & \text{if} \quad r_n^i = r \\ x_n^{ir} & \text{otherwise} \end{cases}$$

$\gamma_n \in (0,1)$ Averaging factor

$\sum_n \gamma_n = \infty \quad \sum_n \gamma_n^2 < \infty$

$\gamma_n = \dfrac{1}{n}$ 普通は←と書く

2.Congestion Pricing and Payoff-Based Adaptive Dynamics

- 学習過程の定式化

$$x_{n+1}^{ir} = \begin{cases} (1-\gamma)x_n^{ir} + \gamma_n \tilde{g}_n^i & \text{if} \quad r_n^i = r \\ x_n^{ir} & \text{otherwise} \end{cases}$$

$$x_{n+1}^{ir} - x_n^{ir} = \gamma_n \left[w_n - x_n \right] \tag{4}$$

$$w_n^{ir} = \begin{cases} \tilde{g}_n^i & \text{if} \quad r_n^i = r \\ x_n^{ir} & \text{otherwise} \end{cases} \tag{5}$$

- (4)の式は確率的アルゴリズムを含み, w_n は個人のロジット型選択確率（以前の認知 x_n に基づき決定）から決まる

2.Congestion Pricing and Payoff-Based Adaptive Dynamics

- (4)式の１階微分は以下のように書ける．(appendix参照)

$$\frac{dx}{dt} = \mathbf{E}\left(w \mid x\right) - x \tag{6}$$

- この式を， $\pi^{ir} = \mathbf{P}\left(X^{ir} = 1\right)$ の条件の元で書き下すと，？

$$\overline{F}^{ir}\left(\pi\right) = \mathbf{E}\left[-\tilde{c}^{r}_{U^r} \mid X^{ir} = 1\right] = \mathbf{E}\left[-\left(c^{r}_{U_i^r+1} + p^{r}_{U_i^r+1}\right)\right] \tag{7}$$

$$U_i^r = \sum_{k \neq i} X^{kr}$$

Fは他者がπで選択することを前提としたときの，ユーザーiによって観測される平均コストを表す.

2.Congestion Pricing and Payoff-Based Adaptive Dynamics

- 認知旅行時間の空間を $\Omega = \prod_{i \in P} \mathbf{R}^R$ と定義
- 写像 $\overline{C} : \Omega \to \Omega$ によって利得ベクトルは以下のように書ける.

$$\overline{C}^{ir}(x) = \overline{F}^{ir}(\Sigma(x)) \tag{8}$$

- ロジット型選択確率は写像 $\Sigma : \Omega \to \Delta$ で表現することで組み入れることができる.

$$\Sigma(x) = \left(\sigma^i(x^i)\right)_{i \in P} \tag{9}$$

2.Congestion Pricing and Payoff-Based Adaptive Dynamics

- 補助命題１
  - $U_{ij}^r = \sum_{k \neq i,j} X^{kr}$　とするとき，

$$\overline{F}^{ir}(\pi) = \mathbf{E}\left(-c^r_{U_i^r+1}\right) + \sum_{j \neq i} \sigma^{jr}\left(x^j\right)\mathbf{E}\left(-\Delta c^r_{U_{ij}^r+2}\right)$$

$$\overline{F}^{ir}(\pi) = \mathbf{E}\left[-\tilde{c}^r_{U^r} \mid X^{ir} = 1\right] = \mathbf{E}\left[-\left(c^r_{U_i^r+1} + p^r_{U_i^r+1}\right)\right]$$

$$\mathbf{E}\left(-\left(c^r_{U_i^r+1} + p^r_{U_i^r+1}\right)\right) = \mathbf{E}\left(-c^r_{U_i^r+1}\right) + \mathbf{E}\left(-p^r_{U_i^r+1}\right)$$

$p^r_u = (u-1)(c^r_u - c^r_{u-1})$　より

$$\mathbf{E}\left(-p^r_{U_i^r+1}\right) = \mathbf{E}\left(-U_i^r\left(c^r_{U_i^r+1} - c^r_{U_i^r}\right)\right)$$

$$= \mathbf{E}\left(-\sum_{j \neq i} X^{jr}\Delta c^r_{U_i^r+1}\right) \qquad = \sum_{j \neq i} \sigma^{jr}\left(x^j\right)\mathbf{E}\left(-\Delta c^r_{U_{ij}^r+2}\right)$$

2.Congestion Pricing and Payoff-Based Adaptive Dynamics

- 提案１：(6)は下のように書けるのではないか.

$$\frac{dx}{dt} = \mathbf{E}(w \mid x) - x \tag{6}$$

$$\frac{dx^{ir}}{dt} = \sigma^{ir}\left(x^{i}\right)\left[\overline{C}^{ir}(x) - x^{ir}\right] \tag{10}$$

$$\overline{C}^{ir}(x) = \overline{F}^{ir}\left(\Sigma(x)\right) \quad \Sigma(x) = \left(\sigma^{i}\left(x^{i}\right)\right)_{i \in P} \quad \pi_{n}^{i} = \sigma^{i}(x_{n}^{i}) \in \Delta^{i}$$ より

$$\overline{C}^{ir}(x) = \mathbf{E}\left(-c_{U_i^r+1}^{r}\right) + \sum_{j \neq i} \sigma^{jr}\left(x^{j}\right)\mathbf{E}\left(-\Delta c_{U_{ij}^r+2}^{r}\right)$$

3. Global Convergence of the Learning Process

- (7)の形で表される利得の写像は，ポテンシャル関数の勾配としても表せる．(Monderer and Shapley (1996))

- 今回提案する関数形は，Sandholm(2001)に近い．ポテンシャル関数をHで表せることが興味深い．

(7)

$$\overline{F}^{ir}(\pi) = \mathbf{E}\left[-\tilde{c}^{r}_{U^r} \mid X^{ir} = 1\right] = \mathbf{E}\left[-\left(c^{r}_{U_i^r+1} + p^{r}_{U_i^r+1}\right)\right]$$

$$\overline{H} : [0,1]^{P\times R} \qquad \overline{H}(\pi) = -\mathbf{E}\left[\sum_{r\in R} U^r c^r_{U^r}\right] \quad (2)$$

3. Global Convergence of the Learning Process

- 提案2

$$\overline{F}(\pi) = \nabla \overline{H}(\pi) \quad \text{for all} \quad \pi \in \Delta$$

- 証明

$$\overline{F}^{ir}(\pi) = \mathbf{E}\left[-\tilde{c}^{r}_{U^r} \mid X^{ir} = 1\right] = \mathbf{E}\left[-\left(c^{r}_{U^r_i+1} + p^{r}_{U^r_i+1}\right)\right] \tag{7}$$

$$\overline{H}(\pi) = -\sum_{r \in R} \mathbf{E}\left[U^r c^r_{Ur}\right]$$

$$\overline{H}(\pi) = -\sum_{r \in R}\left[\pi^{ir}\mathbf{E}\left(\left(U^r_i + 1\right)c^r_{U^r_i+1}\right) + \left(1 - \pi^{ir}\right)\mathbf{E}\left(U^r c^r_{Ur}\right)\right]$$

$$\frac{\partial \overline{H}}{\partial \pi^{ir}}(\pi) = -\mathbf{E}\left(\left(U^r_i + 1\right)c^r_{U^r_i+1}\right) + \mathbf{E}\left(U^r c^r_{Ur}\right)$$

$$= \mathbf{E}\left[-\left(c^r_{U^r_i+1} + p^r_{U^r_i+1}\right)\right] = \overline{F}^{ir}(\pi)$$

3. Global Convergence of the Learning Process

- 補助命題2

$$\frac{\partial^2 \overline{H}}{\partial \pi^{jr} \partial \pi^{ir}}(\pi) = 2\mathbf{E}\left[c^r_{U^r_{ij}+1} - c^r_{U^r_{ij}+2}\right] \in [-2\delta, 0],$$

$$with\ j \neq i, where\ U^r_{ij} = \sum_{k \neq i,j} X_{kr}$$

- 補助命題3

$$\left\| \nabla \pi^{ir}\left(x^i\right) \right\|_1 = \beta_i \pi^{ir} \sum_{a \in R} \left| \delta_{ar} - \pi^{ia} \right| = 2\beta_i \pi^{ir}\left(1 - \pi^{ir}\right) \leq \frac{1}{2}\beta_i$$