Game Theoretic Learning: Robustness of Stochastic Stability

Yasuaki WASA

FL12-18-1

December 17, 2012

[1] Y. Lim and J. S. Shamma, "Robust of Stochastic Stability in Game Theoretic Learning," *submitted for conference publication*, 2012.

Introduction

CDC 2012: Game Theory

42 related presentations. (6 related sessions)

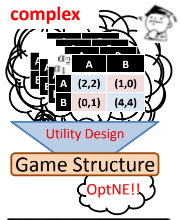
Standard Analysis

under Stationary Environment ➡ Fixed Game

Natural Analysis

under Dynamic/Uncertain Environment

- Environmental Change (Natural Phenomena)
  - ➡ The change of optimal points
- Perturbation (depending on Learning algorithms)
  - ➡ Exploration Parameter ε



Motivation: Following Changing Optimal Equilibria[2]

Problem: Quantitative Analysis using (Potential) Game Theoretic Learning

Analysis[1]: 1. Performance in the neighborhood of perturbed global optima
2. Robustness in game theoretic learning rules

[2] Y. Wasa, T. Goto, T. Hatanaka and M. Fujita, "Seeking Optimal Equilibria for Coverage Games: Payoff-based Learning Approach," *Trans. ISCIE*, Vol. 25, No. 9, pp. 247-255, 2012.

Preliminaries[1]: Stationary Distribution*

Notations

For $\boldsymbol{x} \in \mathbb{R}^n$, $|\boldsymbol{x}|$: the usual Euclidean norm $|\boldsymbol{x}| := \sqrt{\boldsymbol{x}^T\boldsymbol{x}}$

For $M \in \mathbb{R}^{m\times n}$, $\|M\|$: the associated induced matrix norm $\|M\| := \max_{|\boldsymbol{x}|=1} |M\boldsymbol{x}|$

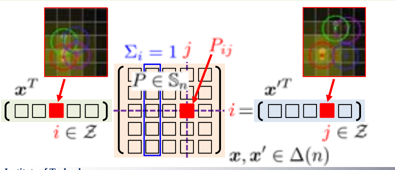
$\Delta(n)$: the probability simplex $\Delta(n) := \{\boldsymbol{x} \in \mathbb{R}^n \mid \mathbf{1}^T\boldsymbol{x} = 1,\ x_i \geq 0\}$

$\mathbb{S}_n$: the set of $n \times n$ stochastic matrices

$P \in \mathbb{S}_n$: a Markov chain over some state space $\mathcal{Z}$

In case P has a single aperiodic recurrent class, there exists a unique stationary distribution $\mu \in \Delta(n)$ such that $\mu^T P = \mu^T$ and

$$\lim_{t\to\infty} \boldsymbol{x}^T P^t = \mu^T,\ \forall \boldsymbol{x} \in \Delta(n)$$



The state space $\mathcal{Z}$ is the same that the action set $\mathcal{A}$ in game theory. P depends on a learning algorithm and an environment.

Preliminaries[1]: Stochastic Stability*

Definition (Regular Perturbation)

*Let $P^0 \in \mathbb{S}_n$ define a Markov chain on $\mathcal{Z}$. Let $\{P^\varepsilon \mid \varepsilon \in (0, \bar{\varepsilon}]\} \subset \mathbb{S}_n$ define a family Markov chains on $\mathcal{Z}$. The family $\{P^\varepsilon\}$ is a **regular perturbation** of P^0 if*

i) each P^ε has a single aperiodic recurrent class

ii) for each $i, j \in \mathcal{Z}$, $\lim_{\varepsilon\to 0} P^\varepsilon_{ij} = P^0_{ij}$

iii) There exists a $\mu^* \in \Delta(n)$ such that $\lim_{\varepsilon\to 0} \mu^\varepsilon = \mu^*$ where μ^ε is the unique stationary distribution of P^ε

Remark: *iii)* is different from the original[3]

If $P^\varepsilon_{ij} > 0$, then $\chi_{ij} \in \mathbb{R}_+$ s.t. $\lim_{\varepsilon\to 0} P^\varepsilon_{ij}/\varepsilon^{\chi_{ij}} \in (0, \infty)$

Definition (Stochastically Stable)

*Let $\{P^\varepsilon\}$ be a regular perturbation of P^0. A state $i \in \mathcal{Z}$ is **stochastically stable** if $\lim_{\varepsilon\to 0} \mu^\varepsilon_i = \mu^*_i > 0$*

Remark: The stochastically stable states are contained in the recurrent communication classes with minimum stochastic potential

[3] H. P. Young, *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princeton University Press, 2001.

Perturbed System and Theorems in [1]

A perturbed stochastic system[1] At each stage $k \in \mathbb{Z}_+$,

the state $z(k) \in \mathcal{Z}$, the transition probability matrix $P(k) \in \mathbb{S}_n$

Update Rule: Markov Chain (discrete time system)

$$\Pr[z(k+1)] \mid z(0), z(1), \cdots, z(k)] = \Pr[z(k+1) \mid z(k)] = P_{z(k)z(k+1)}(k) \quad (1)$$

Theorem 1

Let $P^* \in \mathbb{S}_n$ have a single aperiodic recurrent class, and let $\mu^* \in \Delta(n)$ be the associated stationary distribution over finite state space $\mathcal{Z}$. For any $\delta_1 > 0$, there exists a $\delta_2 > 0$ such that for the dynamic process defined by (1)

$$\limsup_{k\in\mathbb{Z}_+} \|P(k) - P^*\| < \delta_2 \Rightarrow \limsup_{k\in\mathbb{Z}_+} |\Pr[z(k) = i] - \mu^*_i| < \delta_1,\ \forall i \in \mathcal{Z}$$

Theorem 2

Let $\Theta \subset \mathbb{R}^m$ be a compact set. For each $\theta \in \Theta$, let $\{P^\varepsilon_\theta\}$ be a regular perturbation of P^0_θ, and let μ^*_θ be the associated distribution characterizing stochastic stability. Furthermore, for each ε, let $\theta \mapsto P^\varepsilon_\theta$ be continuous. Let the dynamic process (1) satisfy $P(k) \in \{P^\varepsilon_\theta \mid \theta \in \Theta\}$ and $\|P(k+1) - P(k)\| < \delta_2$. For any $\delta_1 > 0$, there exist $\varepsilon > 0$ and $\delta_2 > 0$ such that $\limsup_{k\in\mathbb{Z}_+} |\Pr[z(k) = i] - \mu_\theta(k)^*| \leq \delta_1,\ \forall i \in \mathcal{Z}$

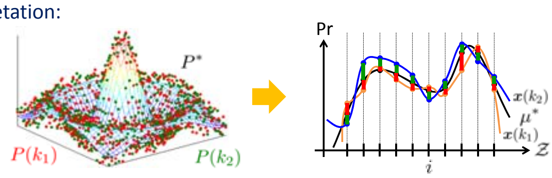
Theorem 1 in [1]: Analysis

Theorem 1

Let $P^* \in \mathbb{S}_n$ have a single aperiodic recurrent class, and let $\mu^* \in \Delta(n)$ be the associated stationary distribution over finite state space $\mathcal{Z}$. For any $\delta_1 > 0$, there exists a $\delta_2 > 0$ such that for the dynamic process defined by (1)

$$\limsup_{k\in\mathbb{Z}_+} \|P(k) - P^*\| < \delta_2 \Rightarrow \limsup_{k\in\mathbb{Z}_+} |\Pr[z(k) = i] - \mu^*_i| < \delta_1,\ \forall i \in \mathcal{Z}$$

Interpretation:



limsup($\|P(k_1) - P^*\|, \|P(k_2) - P^*\|, \ldots) < \delta_2$ limsup(, , …) $< \delta_1$

Remark: Even in case $P^* \to P^\varepsilon$, the above theorem is satisfied. (Corollary 1 in [1])

($\because \exists\delta(\varepsilon) > 0$ s.t. $|\mu^\varepsilon - \mu^*| < \delta(\varepsilon)$)

Theorem1 in [1]: Sketch of Proof*

prob. dis. $x(k) \in \Delta(n)$, state $z(k) \in \mathcal{Z}$, $\Pr[z(k)=i] = x_i(k)$, $\forall i \in \mathcal{Z}$ $\forall k \in \mathbb{Z}_+$

Notations $E(k) := P(k) - P^*$, $x(k) = \mu^* + W w(k)$ (error)

$\left[(P^*)^T \mu^* = \mu^*,\ W: n \times (n-1)$ matrix with orthonormal columns, $W^T W = I$, $w(k) \in \mathbb{R}^{n-1} \right]$

Switching Linear System

$$x(k+1) = P^T(k) x(k) = ((P^*)^T + E^T(k))(\mu^* + W w(k)), \quad \forall k \in \mathbb{Z}_+$$

Error equation

$$w(k+1) = (\underbrace{W^T (P^*)^T W}_{=:A} + W^T E^T(k) W) w(k) + W^T E^T(k) \mu^*$$

$\left[x(k+1) - (P^*)^T \mu^* = x(k+1) - \mu^* = W w(k+1) \right]$

Point: P^* : a single aperiodic recurrent class ⇒ A : a stability matrix

(∵ [4] and the property of matrix transformation)

$\therefore\ \exists X > 0$ s.t. $X = X^T$ and $A^T X A - X < 0$

(∵ discrete Lyapunov equation)

⇒ Lyapunov function $V(w) := w^T(k) X w(k)$

(Lyapunov argument)

then any $\exists \delta > 0$ s.t. $\limsup_{k \in \mathbb{Z}_+} |w(k)| < \delta$

$\left[∵ \limsup_{k \in \mathbb{Z}_+} \|P(k) - P^*\| < \delta_2 \Leftrightarrow \limsup_{k \in \mathbb{Z}_+} \|E\| < \delta_2 \right]$

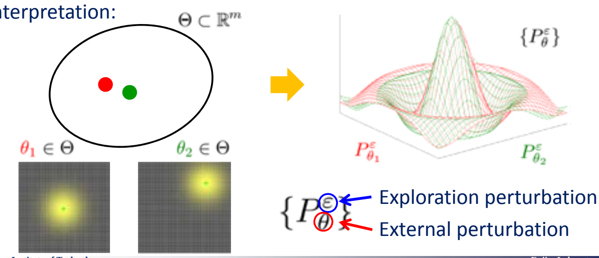
[4] M. Vidyasagar, *(Hidden) Markov Processes: Theory and Applications to Biology*, Princeton University Press, Theorem 4.21 (4.24), Corollary 5.12, under preparation.

Theorem 2 in [1]: Analysis

Theorem 2

Let $\Theta \subset \mathbb{R}^m$ be a compact set. For each $\theta \in \Theta$, let $\{P_\theta^\varepsilon\}$ be a regular perturbation of P_θ^0, and let μ_θ^* be the associated distribution characterizing stochastic stability. Furthermore, for each ε, let $\theta \mapsto P_\theta^\varepsilon$ be continuous. Let the dynamic process (1) satisfy $P(k) \in \{P_\theta^\varepsilon \mid \theta \in \Theta\}$ and $\|P(k+1) - P(k)\| < \delta_2$. For any $\delta_1 > 0$, there exist $\varepsilon > 0$ and $\delta_2 > 0$ such that $\limsup_{k \in \mathbb{Z}_+} |\Pr[z(k) = i] - \mu_{\theta(k)}^*| \le \delta_1$, $\forall i \in \mathcal{Z}$

Interpretation:



Theorem 2 in [1]: Sketch of Proof*

Notations Same as Theorem 1 in [1] $\forall k \in \mathbb{Z}_+$

Switching Linear System $x(k+1) = P^T(k) x(k)$

$x(k) = \mu_{\theta(k)}^* + W w(k)$ (error), $\mu_{\theta(k)}^\varepsilon = (P_{\theta(k)}^\varepsilon)^T \mu_{\theta(k)}^\varepsilon$

⇒ $\mu_{\theta(k+1)}^* + W w(k+1) = (P_{\theta(k)}^\varepsilon)^T (\mu_{\theta(k)}^* + W w(k))$

⇒
$$\begin{aligned} w(k+1) &= W^T (P_{\theta(k)}^\varepsilon)^T W w(k) - W^T \mu_{\theta(k+1)}^* + W^T (P_{\theta(k)}^\varepsilon)^T \mu_{\theta(k)}^* \\ &= W^T (P_{\theta(k)}^\varepsilon)^T W w(k) + W^T (\mu_{\theta(k+1)}^\varepsilon - \mu_{\theta(k+1)}^*) \\ &\quad + W^T (\mu_{\theta(k)}^\varepsilon - \mu_{\theta(k+1)}^\varepsilon) + W^T (P_{\theta(k)}^\varepsilon)^T (\mu_{\theta(k)}^* - \mu_{\theta(k)}^\varepsilon) \end{aligned}$$

(add $0 = W^T \mu_{\theta(k+1)}^\varepsilon - W^T \mu_{\theta(k+1)}^\varepsilon + W^T \mu_{\theta(k)}^\varepsilon - W^T (P_{\theta(k)}^\varepsilon)^T \mu_{\theta(k)}^\varepsilon$)

$\exists \delta_\mu(\delta_2) > 0$ s.t. $|\mu_{\theta(k)}^\varepsilon - \mu_{\theta(k+1)}^\varepsilon| < \delta_\mu(\delta_2)$ $\left[∵ \theta \mapsto P_\theta^\varepsilon \text{ : continuous}, \ \|P(k+1) - P(k)\| < \delta_2 \right]$

$\exists \delta(\varepsilon) > 0$ s.t. $|\mu_{\theta(k)}^* - \mu_{\theta(k)}^\varepsilon| < \delta(\varepsilon)$, $\forall k \in \mathbb{Z}_+$ [∵ Stochastically stable]

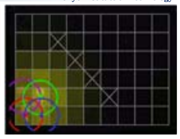
$W^T (P_{\theta(k)}^\varepsilon)^T W$ Similarly analysis in Theorem 1 in [1]

(cf. slowly varying Linear Parameter Varying (LPV) system[5])

[5] J. S. Shamma, "An overview of LPV Systems," in *Control of Linear Parameter Varying Systems with Applications*, J. Mohammadpour and C. Scherer, Eds. Springer, pp. 3-26, 2012.

Extensions 2, 3: Convergence Problem

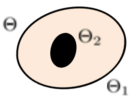
2. In the theorem 1, suppose $\Theta \subset \mathbb{R}^m$: a compact set.
   If there are some obstacles, does its assumption hold?
   And how does "$\theta \mapsto P_\theta^\varepsilon$: continuous" change?



Solution: Is the following component possible?

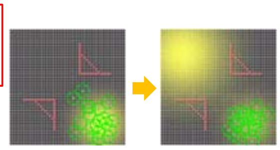
$\Theta = \Theta_1 \backslash \Theta_2$ and $\theta \mapsto P_\theta^\varepsilon$: continuous

$\Theta_2 \subset \Theta_1 \subset \mathbb{R}^m$, Θ_1: a compact set, Θ_2: an open set, $\theta \in \Theta$



If it does not hold, the following assumption does not hold for all area. $\|P(k+1) - P(k)\| < \delta_2$

3. If θ changes dramatically (in other words, "$\theta \mapsto P_\theta^\varepsilon$: discontinuous" exists), how do we evaluate it?



Solution:

We should consider that the changed state is a new initial state. So, its case is excluded.

Extensions 4, 5: Learning Algorithm(LA)

Corollary 3 [1]

Suppose that any $\theta \in \Theta$ results in a potential game with potential function $\phi(a;\theta)$. Then for either LLL or binary LLL, for every $\delta_1 > 0$, there exists a $\delta_2 > 0$ and $\tau > 0$ such that $|\theta(k+1) - \theta(k)| < \delta_2$, $\forall k \in \mathbb{Z}_+$ implies that

$$\liminf_{k \in \mathbb{Z}_+} \Pr\left[\phi(a(k);\theta(k)) = \max_{a \in \mathcal{A}} \phi(a;\theta(k))\right] > 1 - \delta_1$$

4. Corollaries 2-4 hold for LLL or Binary LLL(with a constrained action set). Do they hold for general LAs, especially PIPIP?

Solution : LAs (including PIPIP) which guarantees irreducible and aperiodic process and potential function maximizers are OK.
Do the authors in [1] write Corollaries 2-4 for simulations?

5. In [1], the statement "if a state $z \in \mathcal{Z}$ is *not* stochastically stable, then small perturbations limit the long run probability of visiting this state." is written. At least, how long/steps must we run a LA?

Solution : No idea.

Extensions 6, 7: Similarity of Approach

6. [1] uses Time-variant environmental function θ. So, is it compatible with Estimated-state based potential game[8] or Bayesian (potential) game[9,10]?

Component of Potential Game (normal model[7])

$$U_i(a_i', a_{-i};\theta) - U_i(a_i, a_{-i};\theta) = \phi(a_i', a_{-i};\theta) - \phi(a_i, a_{-i};\theta)$$

Role of parameter θ [7] Fixed (no role)

- [1] Environmental Parameter (nature/uncontrollable) } $a(k,\theta) \to a^*(k',\theta')$
- [8] Estimation Values (depending on players' actions)
- [9,10] Types (depending on players' incentive/belief) } $(a;\theta) \to (a^*;\theta^*)$

7. Can we prove the convergence of Error parameter w based on [11]?

Switching Linear System $x(k+1) = P^T(k) x(k)$ $\quad w(k) = W^T (x(k) - \mu^*)$ (error)

Error equation $w(k+1) = A(k) w(k) + B(k)$

[7] D. Monderer and L. Shapley, "Potential Games," *Games and Economic Behavior*, Vol. 14, No. 1, pp. 124-143, 1996.
[8] J. R. Marden, "State Based Potential Games," *submitted for journal publication*, 2011.
[9] G. Facchini, F. V. Megen, P Borm and S. Tijs, "Congestion models and weighted Bayesian potential games," *Theory and Decision*, Springer, Vol. 42, No. 2, pp. 193-206, 1997.
[10] T. Ui, "Robust Equilibria of Potential Games," *Econometrica*, Vol. 69, No. 5, pp. 1373-1380, 2001.
[11] H. K. Khalil, *Nonlinear Systems*, *Third Edition*, Prentice Hall, 2002.