Markov Chain Quantal-Response

Nishka Arora*, Kimia Hassibi*, Binyamin (Ben) Wincelberg*, Omer Tamuz

January 12, 2024

Contents

Abstract

Real-world players in games are susceptible to making errors. In the quantal-response model, the players perceive utilities that include random errors. Here, the players' responses are probabilistic, with players having a higher likelihood of choosing better responses than worse responses. We study the dynamics of the quantal-response model by analyzing the stationary distribution of a Markov chain. The probability of transitioning from one strategy profile to another is determined by quantal-responses. The stationary distribution is a joint probability distribution on the players' actions. We consider a game in which the players are divided into two groups, and the utility of a player in one group depends only on the actions of the players in the other group. We find that if we allow all the players in such a game to simultaneously revise their responses,

* listed in alphabetical order

the stationary distribution is the product of the marginal distributions of each group.

1 Definitions and Notation

1.1 Quantal-Response Normal-Form Game

Let $G = (N, S, u)$ be a normal-form game. $N = \{1, 2, ..., n\}$ is the set of players. For each player $i \in N$, there is a strategy set $S_i = \{s_{i1}, s_{i2}, ..., s_{iJ_i}\}$ of J_i pure strategies. Each player $i \in N$ also has a payoff function $u_i : S \to \mathbb{R}$, where $S = \prod_{i \in N} S_i$.

We denote $\Delta = \prod_{i \in N} \Delta_i$ as the set of mixed strategy profiles, where Δ_i is the set of probability measures on S_i. For all $p_i \in \Delta_i$, $p_i : S_i \to \mathbb{R}$, where $\sum_{s_{ij} \in S_i} p_i(s_{ij}) = 1$, and $p_i(s_{ij}) \geq 0$, $\forall s_{ij} \in S_i$. To simplify notation, we denote $p_{ij} = p_i(s_{ij})$.

The payoff function of each player $i \in N$ is $u_i(p) = \sum_{s \in S} p(s) u_i(s)$, where $p(s) = \prod_{i \in N} p_i(s_i)$. We denote $\zeta_i = \mathbb{R}^{J_i}$ as the space of possible payoffs for strategies that player i could adopt. Let $\zeta = \prod_{i=1}^{n} \zeta_i$. We define $\bar{u} : \Delta \to \zeta$, by

$$\bar{u}(p) = (\bar{u}_1(p), ..., \bar{u}_n(p)), \text{ where } \bar{u}_i(p) \in \mathbb{R}^{J_i}$$

and where

$$[\bar{u}_i(p)]_j = \bar{u}_{ij}(p) = u_i(s_{ij}, p_{-i}).$$

We consider the scenario in which each player's utility for each action is subject to random error. For each player $i \in N$, $\forall j \in \{1, ..., J_i\}$, $\forall p \in \Delta$,

$$\hat{u}_{ij}(p) = \bar{u}_{ij}(p) + \varepsilon_{ij}.$$

Player i's error vector is defined as $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{iJ_i})$. The errors ε_i are distributed according to joint density $f_i(\varepsilon_i)$. We say that $f = (f_1, ..., f_n)$ is admissible if the marginal distribution of each f_i exists for each ε_{ij} and $\mathbb{E}[\varepsilon_i] = \mathbf{0}$.

Following the notation in [2], given $p \in \Delta$, for all $\bar{u}(p) = (\bar{u}_1(p), ..., \bar{u}_n(p))$, where $\bar{u}_i(p) \in \mathbb{R}^{J_i}$, $\forall i$, the ij-response set $R_{ij}(\bar{u}_i(p)) \subseteq \mathbb{R}^{J_i}$ is defined as

$$R_{ij}(\bar{u}_i(p)) = \{\varepsilon_i \in \mathbb{R}^{J_i} | \bar{u}_{ij}(p) + \varepsilon_{ij} \geq \bar{u}_{ik}(p) + \varepsilon_{ik}, \forall k = 1, ..., J_i\}.$$

In other words, the ij-response set $R_{ij}(\bar{u}_i(p))$ is the set of all error vectors ε_i that make action j the dominant strategy for player i when all other players perform strategy profile p. The probability that player $i \in N$ will select strategy $j \in S_i$ when given $\bar{u}_i(p) \in \mathbb{R}^{J_i}$ is the quantal-response function $\sigma_{ij}(\bar{u}_i(p))$, which is defined as

$$\sigma_{ij}(\bar{u}_i(p)) = \int_{R_{ij}(\bar{u}_i(p))} f(\varepsilon) d\varepsilon.$$

1.2 Logit-Response (note: not used in Theorem)

A particularly interesting case of the quantal-response function is the logistic quantal-response function. Recall $f = (f_1, ..., f_n)$, where the errors $\varepsilon_i \in \mathbb{R}^{J_i}$ are distributed according to joint density $f_i(\varepsilon_i)$. For some $\beta \geq 0$, the logistic quantal-response function gives the optimal choice behavior when f_i, $\forall i \in N$, is an extreme value distribution with the cumulative distribution function $F_i(\varepsilon_{ij}) = e^{-e^{-\beta\varepsilon_{ij}-\gamma}}$ and the errors ε_{ij} are independent. Specifically, for some $\bar{u}_i(x) \in \mathbb{R}^{J_i}$, the logistic quantal-response function is

$$\sigma_{ij}(\bar{u}_i(x)) = \frac{\mathrm{e}^{\beta u_i(s_{ij},x_{-i})}}{\sum\limits_{k=1}^{J_i} \mathrm{e}^{\beta u_i(s_{ik},x_{-i})}}.$$

1.3 Markov Chain

A Markov chain is a discrete-time stochastic process defined by a sequence of random variables $X_1, X_2, ..., X_n$ that satisfies the Markov property

$$\mathbb{P}(X_{i+1} = x | X_1 = x_1, X_2 = x_2, ...X_n = x_n) = \mathbb{P}(X_{i+1} = x | X_i = x_i),$$

where $\mathbb{P}(X_1 = x_1, X_2 = x_2, ..., X_n = x_n) > 0$. The random variables $X_1, X_2, ..., X_n$ take values on the countable state space, which we define as the set of action profiles S in Section 1.1.

The initial state X_1 is chosen at random according to a distribution λ on S, which assigns probability λ_i to state i via

$$\mathbb{P}(X_1 = i) = \lambda_i.$$

The Markov chain has an associated stochastic transition matrix $\mathcal{P} = [p_{ij}]$, where $i, j \in S$. Here, p_{ij} is the probability of state X_k being action profile i and transitioning to action profile j. That is to say $p_{ij} = \mathbb{P}(X_{k+1} = j | X_k = i)$. For each $i \in S$, $\sum\limits_{j \in S} p_{ij} = 1$. Also, $p_{ij} \geq 0$, $\forall p_{ij}$.

1.4 Transition Matrix of the Markov Chain

We consider a few cases of the transition matrix $\mathcal{P}$ that are differentiated by which players receive revision opportunities in a period. Just as in [1], we will define learning dynamics as a behavioral rule (such as logit responding) and a revision process.

1.4.1 Asynchronous Learning (note: not used in Theorem)

We first consider asynchronous learning, in which exactly one player receives a revision opportunity in a period. For any player $i \in N$, we define $m_i(x, y)$ as the likelihood of player i receiving a revision opportunity and changing the current action profile x into the action profile y. In other words, we have

$$m_i(x,y) = \begin{cases} \sigma_{iy_i}(\bar{u}_i(x)) & \text{if } x_{-i} = y_{-i} \\ 0 & \text{if } x_{-i} \neq y_{-i} \end{cases}, \tag{1}$$

where x_{-i} and y_{-i} are the action profiles without the action of player i and y_i is the action of player i in strategy profile y. Then, transition matrix $\mathcal{P}$ is defined as $\mathcal{P} = [p_{xy}]$, where $p_{xy} = \frac{1}{n} \sum\limits_{i=1}^{n} m_i(x,y)$.

1.4.2 Simultaneous Learning

We also consider the case in which all players receive revision opportunities in every period. Here, the transition matrix $\mathcal{P}$ is defined as $\mathcal{P} = [p_{xy}]$, where $p_{xy} = \prod\limits_{i \in N} \sigma_{iy_i}(\bar{u}_i(x))$.

1.5 Update Function Representation of Markov Chain (note: not used in Theorem)

An equivalent representation of the Markov chain replaces the probabilistic transitions of a Markov chain with a deterministic update function that takes as input realizations of a uniform random variable. Let U_n, $n \geq 1$ be a sequence of independent and identically distributed random variables, in which $U_n \sim$ Uniform$[0,1]$, $\forall n \geq 1$. For clarity, let us enumerate the elements in the state space S from 1 to $|S|$. Consider the piecewise function $\phi : S \times [0,1] \to S$, such that

$$\phi(i, U_n) = \begin{cases} 1 & \text{if } U_n \in [0, p_{i1}] \\ 2 & \text{if } U_n \in (p_{i1}, p_{i1} + p_{i2}] \\ ... & \\ j & \text{if } U_n \in \left(\sum\limits_{t=1}^{j-1} p_{it}, \sum\limits_{t=1}^{j} p_{it} \right] \\ ... & \\ |S| & \text{if } U_n \in \left(\sum\limits_{t=1}^{|S|-1} p_{it}, 1 \right]. \end{cases}$$

Starting in state X_n, $\forall n \geq 1$, the next state X_{n+1} is determined by the following

$$X_{n+1} = \phi(X_n, U_n).$$

1.6 Graphical Games

Consider the game $G = (V, S, u)$ with $V = \{1, ..., n\}$, where $n = |V|$ players. Also, consider the simple, undirected graph $\mathcal{G} = (V, E)$ with V as the set of vertices and E as the set of edges. We define the game G as a graphical game on $\mathcal{G}$ if each player's utility is a function of its neighbors on $\mathcal{G}$. In other words, G is a graphical game on $\mathcal{G}$ if for all $i, j \in V$, if $i \neq j$ and $(i,j) \notin E$, then

$\forall a \in S$, $u_i(a) = u_i(a^j)$ where action profile a^j is the same as a in all but the j^{th} position.

2 Theorem

Let $\mathcal{G} = (V, E)$ be a bipartite graph with partitions $\mathcal{G}_0$ and $\mathcal{G}_1$. Let $G = (V, S, u)$ be a graphical game on $\mathcal{G}$. If the quantal-response dynamics exhibit simultaneous learning, then the stationary distribution ($\pi = \lim_n \lambda \mathcal{P}^n$, for arbitrary initial distribution λ) of the quantal-response dynamics is a product measure. Specifically,

$$\pi \mathcal{P} = \pi,$$

such that

$$\pi = \pi_0 \times \pi_1,$$

where π_0 is a measure on $\mathcal{S}_0 = \Pi_{k \in \mathcal{G}_0} S_k$ and π_1 is a measure on $\mathcal{S}_1 = \Pi_{k \in \mathcal{G}_1} S_k$.

2.1 Proof

In the bipartite graph G, we know that the utilities of players in one partition depend only on the actions of the players in the other partition. More specifically, for player $i \in \mathcal{G}_j$, $j \in \{0, 1\}$, and for $p \in \Delta$,

$$\bar{u}_i(p) = g_i(p^{-j}),$$

for some function $g_i : \left(\prod_{k \in \mathcal{G}_{(j+1) mod 2}} S_k \right) \to \mathbb{R}^{J_i}$, where p^{-j} is the mixed strategy profile of all players that are not in partition $\mathcal{G}_j$. Therefore, for $x, y \in S$,

$$\sigma_{iy_i}(\bar{u}_i(x)) = \sigma_{iy_i}(g_i(x^{-j})).$$

Therefore,

$$p_{xy} = \prod_{i \in \mathcal{G}_0} \sigma_{iy_i}(g_i(x^1)) \times \prod_{i \in \mathcal{G}_1} \sigma_{iy_i}(g_i(x^0)),$$

where x^0 and x^1 are the mixed strategy profiles of all the players in partition $\mathcal{G}_0$ and $\mathcal{G}_1$, respectively.

Lemma 1. *$\mathcal{P}$ maps product measures to product measures.*

Proof. Let $\mathcal{S}_0 = \Pi_{k \in \mathcal{G}_0} S_k$ and $\mathcal{S}_1 = \Pi_{k \in \mathcal{G}_1} S_k$ denote the sets of strategy profiles restricted to $\mathcal{G}_0$ and $\mathcal{G}_1$. Suppose $\mu \in \Delta$ can be decomposed as $\mu_0 \times \mu_1$, where μ_0 is a measure on $\mathcal{S}_0$ and μ_1 is a measure on $\mathcal{S}_1$. Then

$$\begin{aligned}
\mu \mathcal{P} &= \left(\sum_{x \in S} \mu_1(x^1) \prod_{i \in \mathcal{G}_0} \sigma_{iy_i}(g_i(x^1)) \times \mu_0(x^0) \prod_{i \in \mathcal{G}_1} \sigma_{iy_i}(g_i(x^0)) \right)_{y \in S} \\
&= \left(\left[\sum_{x^1 \in \mathcal{S}_1} \mu_1(x^1) \prod_{i \in \mathcal{G}_0} \sigma_{iy_i}(g_i(x^1)) \right] \times \left[\sum_{x^0 \in \mathcal{S}_0} \mu_0(x^0) \prod_{i \in \mathcal{G}_1} \sigma_{iy_i}(g_i(x^0)) \right] \right)_{y \in S} \\
&= \nu_0 \times \nu_1,
\end{aligned}$$

where

$$\nu_0(y^0) = \sum_{x^1 \in \mathcal{S}_1} \mu_1(x^1) \prod_{i \in \mathcal{G}_0} \sigma_{iy_i^0}(g_i(x^1)),$$

$$\nu_1(y^1) = \sum_{x^0 \in \mathcal{S}_0} \mu_0(x^0) \prod_{i \in \mathcal{G}_1} \sigma_{iy_i^1}(g_i(x^0))$$

as desired. □

Lemma 2. *The set of product probability measures is sequentially closed.*

Proof. Let $X = X_0 \times X_1$, and Π denote the set of corresponding product probability measures on X, and fix a sequence $(\mu_n)_n = (\mu_{0_n})_n \times (\mu_{1_n})_n \in \Pi$ with limit μ. For $A \subset X_0, B \subset X_1$, $\mu(A \times B) = \lim_n \mu_n(A \times B) = \lim_n(\mu_{0_n}(A)\mu_{1_n}(B)) = \lim_n(\mu_n(A \times X_1)\mu_n(X_0 \times B)) = \lim_n(\mu_n(A \times X_1)) \cdot \lim_n(\mu_n(X_0 \times B)) = \lim_n(\mu_{0_n}(A)) \cdot \lim_n(\mu_{1_n}(B)) \in \Pi$.

□

Since $\mathcal{P}$ is the transition matrix of an ergodic Markov chain, there is a unique equilibrium distribution π, such that for all initial distributions λ, $\pi = \lim_n \lambda \mathcal{P}^n$. Let λ_0 denote a measure in Δ that can be decomposed as a product of measures on $\mathcal{S}_0$ and $\mathcal{S}_1$ (such as any pure strategy profile). By the above lemmas, $\pi = \lim_n \lambda_0 \mathcal{P}^n$ is a product measure.

References

[1] Carlos Alós-Ferrer and Nick Netzer. “The logit-response dynamics”. en. In: *Games and Economic Behavior* 68.2 (Mar. 2010), pp. 413–427. ISSN: 08998256. DOI: 10.1016/j.geb.2009.08.004. URL: https://linkinghub.elsevier.com/retrieve/pii/S0899825609001651 (visited on 09/30/2023).

[2] Jakub Černý, Bo An, and Allan N. Zhang. “Quantal Correlated Equilibrium in Normal Form Games”. en. In: *Proceedings of the 23rd ACM Conference on Economics and Computation*. Boulder CO USA: ACM, July 2022, pp. 210–239. ISBN: 978-1-4503-9150-4. DOI: 10.1145/3490486.3538350. URL: https://dl.acm.org/doi/10.1145/3490486.3538350 (visited on 09/30/2023).