Velocidad de convergencia de juego ficticio simultáneo y alternante

Ariel Arbiser[1], Federico Badaloni[2]

[1] Departamento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, arbiser@dc.uba.ar
[2] Departamento de Ciencias de la Computación, Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario, federicobadaloni@hotmail.com

Resumen Extendemos algunos resultados sobre la convergencia de juego ficticio, permitiendo el uso de reglas de desempate, y probamos que existen juegos de 2 jugadores con dimensiones mínimas para los cuales la cantidad de pasos que se requiere antes de alcanzar un equilibrio de Nash puro es exponencial en el tamaño de representación de las matrices de utilidad, tanto para el juego ficticio simultáneo como para el alternante.

Keywords: juego · equilibrio · convergencia.

1. Introducción

El proceso de aprendizaje de juego ficticio (FP) fue propuesto por primera vez por Brown en 1951 [1] en el contexto del juego iterado, como un método para encontrar el valor de un juego finito de suma cero [2]. Consiste en que cada uno de ambos jugadores decida su jugada en cada turno como su mejor respuesta teniendo en cuenta la frecuencia de jugadas del otro y las ganancias esperadas en cada caso. Hacia finales del mismo año, Robinson [3] demostró que este proceso converge para todos los juegos finitos de suma cero. Desde entonces, se han publicado trabajos sobre la convergencia de FP en juegos que no son de suma cero. Miyazawa [4] demostró que esta propiedad vale para todos los juegos con matriz de 2×2 pero, su demostración depende de la incorporación de una regla de desempate particular, sin la cual Monderer y Sela [5] demostraron que no se cumple. Por su parte, Shapley [6] dio el primer ejemplo de un juego (de 3×3) para el cual cierta secuencia no converge. Además, la convergencia de FP fue demostrada para juegos de intereses idénticos [7], juegos potenciales con pesos [8], juegos no degenerados con estrategias complementarias y ganancias disminuyentes [9] y ciertas clases de juegos compuestos [10].

Por otro lado, se han estudiado muchas variantes de FP [1, 11, 12, 13, 14]. Una variante de particular interés es FP con actualización alternante de creencias. Berger [15] planteó que esta forma es en realidad la original y que, si bien la actualización simultánea puede resultar mas intuitiva, es menos potente, dando como ejemplo la clase de los juegos no degenerados con potencial ordinal para la cual la version alternante converge y la simultánea no.

1.1. Velocidad de convergencia de FP

Varios trabajos estudian la eventual convergencia global o pura a un equilibrio de Nash de las distintos familias de juegos. Otro enfoque es la velocidad de convergencia cuando esta ocurre. El interés se debe en gran medida a la equivalencia entre los juegos de suma cero y los problemas de programación lineal [16, 17]. En [18] planteaban que hasta la fecha, lo más eficiente para resolver un juego de suma cero era formularlo como un problema de programación lineal y aplicar el método simplex. En el mismo artículo ofrecen un método mixto y concluyen que se puede acelerar la convergencia en ciertos problemas de programación lineal. En [19] se plantea una variante (con muestreo) y se discute su eficiencia en problemas de optimización a gran escala. En [20] se conjeturó que la velocidad de convergencia de FP es $O(t^{-\frac{1}{2}})$ para todos los juegos (conjetura de Karlin). La idea proviene de que esta cota superior se corresponde con la de la velocidad de convergencia de otro método de aprendizaje relacionado con FP, las dinámicas de no arrepentimiento [21, 22]. En [23] se muestra que una versión fuerte de esta conjetura es falsa (con una regla de desempate arbitraria) pero se deja abierta la pregunta general.

La utilidad práctica de FP como método para calcular equilibrios de Nash fue entonces puesta en duda cuando Brandt, Fischer y Harrenstein [24] demostraron que, para los juegos de suma cero, los no degenerados de $2 \times m$ y los potenciales (tres de las clases mas estudiadas), existen casos en los que el proceso de FP puede requerir una cantidad de rondas exponencial en el tamaño de representación en bits de las utilidades del juego antes de que se juegue algún equilibrio, y mencionan sin demostrar que su resultado puede ser extendido a FP alternante.

2. Juegos bimatriciales y equilibrios de Nash

Sea (A, B) un juego en forma bimatricial de $n \times m$, es decir un juego de dos jugadores finito en el que el jugador 1 (jugador fila) tiene acciones $i \in N = \{i_1, i_2, \ldots, i_n\}$ y el jugador 2 (jugador columna) tiene acciones $j \in M = \{j_1, j_2, \ldots, j_m\}$. A, $B \in \mathbb{R}^{n \times m}$ son las matrices de pago de los jugadores 1 y 2. Cada uno elige su acción sin conocer la elegida por el otro. Si el jugador 1 elige la acción i y el jugador 2 elige la acción j, la ganancia del jugador 1 será $a_{i,j}$ y la ganancia del jugador 2 será $b_{i,j}$. Describimos estos juegos mediante una matriz de pares. Llamamos estrategias mixtas del jugador 1 a los vectores fila $x \in \mathbb{R}^{1 \times n}_{\geq 0}$ tales que $||x||_1 = 1$ y estrategias mixtas del jugador 2 a los vectores columna $y \in \mathbb{R}^{m \times 1}_{\geq 0}$ tales que $||y||_1 = 1$. A las estrategias que asignan probabilidad 1 a una acción y 0 a todas las otras las llamamos *estrategias puras*. Notamos con $\widetilde{i} \in \mathbb{R}^{1 \times n}$ a la estrategia pura del jugador fila correspondiente a la acción i y con $\widetilde{j} \in \mathbb{R}^{m \times 1}$ a la estrategia pura del jugador columna correspondiente a la acción j. La ganancia esperada del jugador 1 al jugar la acción i contra la estrategia mixta y del jugador 2 será $\widetilde{i}Ay$. Análogamente, la ganancia esperada del jugador 2 al jugar la acción j contra la estrategia mixta x del jugador 1 será $xB\widetilde{j}$. Si ambos jugadores juegan las estrategias mixtas x e y respectivamente, sus ganancias esperadas pueden calcularse como xAy para el jugador 1 y xBy para el jugador 2. Si y es una estrategia mixta del jugador 2, definimos el *conjunto de mejores respuestas* a y como $BR_1(y) = argmax_{i \in N}\{\widetilde{i}Ay\}$ y, análogamente, si x es una estrategia mixta del jugador 1, el conjunto de mejores respuestas a x será $BR_2(x) = argmax_{j \in M}\{xB\widetilde{j}\}$. Es decir, son los conjuntos de índices que maximizan las ganancias esperadas contra una estrategia dada. Este concepto puede expandirse de índices a estrategias mixtas. Si x es una estrategia mixta del jugador 1 e y es una estrategia mixta del jugador 2, decimos que $x \in BR_1(y)$ si $\forall i \in N, x_i > 0 \implies i \in BR_1(y)$. Llamamos *equilibrio de Nash* (NE) a todo perfil de estrategias mixtas $(x^*, y^*) \in \mathbb{R}^{1\times n} \times \mathbb{R}^{m\times 1}$ tal que $x^* \in BR_1(y^*)$ e $y^* \in BR_2(x^*)$. En el caso particular de que las estrategias sean puras, al NE lo llamamos también *puro*. Algunas de las familias de juegos más estudiadas en la literatura sobre FP son las siguientes, para juegos de tamaño $n \times m$. (A, B) es *de suma constante* si $\forall i \in N, j \in M : a_{ij} + b_{ij} = c$ con c una constante; si $c = 0$, el juego es de *suma cero*. Los juegos de suma cero representan aquellos en los que un jugador siempre gana tanto como el otro pierde. (A, B) es *degenerado* si $\exists i, i' \in N$ y $j \in M$, $i \neq i'$ / $a_{ij} = a_{i'j}$ o $\exists j, j' \in M$ e $i \in N$, $j \neq j'$ / $b_{ij} = b_{ij'}$. Los juegos no degenerados son de particular interés porque para cada acción del rival, no existen dos acciones con el mismo pago, por tanto, el conjunto de mejor respuesta es siempre unitario. (A, B) es *de intereses idénticos* si $\forall i \in N, j \in M : a_{ij} = b_{ij}$. (A, B) de tamaño $n \times n$ es *simétrico* si $\forall i, j \in N : a_{ij} = -b_{ji}$. Para juegos suma cero, simétricos o de intereses idénticos basta exhibir sólo la matriz A.

3. Juego ficticio

Existen dos formas de definir FP. En [15, 11, 5, 7, 23] utilizan una definición semejante a la siguiente, que resulta cómoda para estudiar convergencia. Nosotros le agregamos el manejo de *reglas de desempate*. Para formalizar este concepto, utilizamos una forma de *independencia de alternativas irrelevantes* o condición de Hauthakker [2], donde $\mathbb{P}^+(S)$ denota el conjunto de las partes no vacías del conjunto S. Llamamos *reglas de desempate* de un FP a un par de funciones $d_1 : \mathbb{P}^+(N) \to N$ y $d_2 : \mathbb{P}^+(M) \to M$ (que a cada subconjunto de acciones de un jugador en un eventual empate le asignan la acción que elegirá el proceso), cumpliendo d_1 las condiciones: (1) $\forall S \in \mathbb{P}^+(N), d_1(S) \in S$, y (2) $\forall N_a, N_b \in \mathbb{P}^+(N)$, si $N_b \subseteq N_a, e \in N_b$ y $d_1(N_a) = e \Rightarrow d_1(N_b) = e$; y análogamente para d_2 sobre M. En el resto de este trabajo usaremos entonces la siguiente definición de FP, tanto simultáneo como alternante:

Definición 1. *Sean (A, B) un juego en forma bimatricial de $n \times m$ y una secuencia (i^τ, j^τ) con $i^\tau \in N$, $j^\tau \in M$ para todo $\tau \in \mathbb{N}$. Si d_1 y d_2 son reglas de desempate y tenemos unas secuencias de creencias x^τ e y^τ tales que para todo $\tau \in \mathbb{N}$ $x^\tau = \frac{\sum_{s=1}^{\tau} \widetilde{i^s}}{\tau}$, $y^\tau = \frac{\sum_{s=1}^{\tau} \widetilde{j^s}}{\tau}$. Entonces:*

- *(i^τ, j^τ) es una secuencia de FP simultáneo si (i^1, j^1) es un elemento arbitrario de $N \times M$ y para todo $\tau \in \mathbb{N}$ se cumplen $i^{\tau+1} = d_1(BR_1(y^\tau))$ y $j^{\tau+1} = d_2(BR_2(x^\tau))$.*

- (i^τ, j^τ) *es una secuencia de FP alternante si* i^1 *es un elemento arbitrario de* N *y para todo* $\tau \in \mathbb{N}$ *se cumplen* $i^{\tau+1} = d_1(BR_1(y^\tau))$ *y* $j^\tau = d_2(BR_2(x^\tau))$.

Como los elementos iniciales son arbitrarios, un mismo juego puede tener varios procesos de FP. Puede haber ambigüedad en las acciones si alguno de los conjuntos de mejor respuesta no es unitario[3], por ello algunos autores mencionan informalmente el concepto de *reglas de desempate*. En el presente artículo lo formalizamos. Sobre el juego alternante, vale aclarar que el jugador columna toma su decisión incorporando en sus creencias la información sobre qué jugó el jugador fila en la ronda actual, lo que puede resultar poco intuitivo ya que las decisiones se asumen simultáneas. Diremos que un proceso de FP (simultáneo o alternante) converge de forma pura en la iteración k si (i^k, j^k) es un NE puro.

En [24] se propone una definición similar a la de [3] pero simplificada, que damos a continuación a su vez incorporando reglas de desempate.

Definición 2. *Sea* (A, B) *un juego en forma bimatricial de* $n \times m$*:*

- *Una secuencia de FP simultáneo en* (A, B) *es una secuencia* $(p^\tau, q^\tau)_{\tau\in\mathbb{N}}$ *de pares de vectores no negativos* $(p^\tau, q^\tau) \in \mathbb{N}^n \times \mathbb{N}^m$ *tal que:*

$$p^0 = 0, q^0 = 0$$
$$p^{\tau+1} = p^\tau + \widetilde{i} \text{ donde } i = d_1(argmax_{i'\in N}\{\widetilde{i'}Aq^\tau\})$$
$$q^{\tau+1} = q^\tau + \widetilde{j} \text{ donde } j = d_2(argmax_{j'\in M}\{p^\tau B\widetilde{j'}\})$$

- *Una secuencia de FP alternante en* (A, B) *es una secuencia* $(p^\tau, q^\tau)_{\tau\in\mathbb{N}}$ *de pares de vectores no negativos* $(p^\tau, q^\tau) \in \mathbb{N}^n \times \mathbb{N}^m$ *tal que:*

$$p^0 = 0, q^0 = 0$$
$$p^{\tau+1} = p^\tau + \widetilde{i} \text{ donde } i = d_1(argmax_{i'\in N}\{\widetilde{i'}Aq^\tau\})$$
$$q^{\tau+1} = q^\tau + \widetilde{j} \text{ donde } j = d_2(argmax_{j'\in M}\{p^{\tau+1} B\widetilde{j'}\})$$

La principal diferencia con la definición anterior es que en aquella se define una secuencia de jugadas que cumple una condición contra un historial de creencias sobre la estrategia mixta del otro jugador, y en esta la secuencia es de pares de contadores de jugadas (sin normalizar, por lo que no son estrategias mixtas) que cumplen en cada iteración una condición en referencia a las matrices de pago. Probamos debajo que ambas definiciones son equivalentes.

4. Equivalencia de las definiciones y preservación

La equivalencia de estas dos definiciones de FP no es inmediatamente evidente. Presentamos a continuación dos lemas sobre esta equivalencia, para el caso simultáneo y alternante respectivamente, probando que los historiales de la definición 2 suman en cada iteración la estrategia pura correspondiente a la acción elegida en la definición 1.

[3] Por esto es importante en el estudio de FP la familia de los juegos no degenerados.

Lema 1 *Sea* (A, B) *un juego en forma bimatricial de* $n \times m$, $(\widetilde{i^\tau}, \widetilde{j^\tau})_{\tau \in \mathbb{N}}$ *una secuencia de FP simultáneo según la definición 1 con secuencias de creencias* y^τ, x^τ *y sea* $(p^\tau, q^\tau)_{\tau \in \mathbb{N}}$ *una secuencia de FP simultáneo según la definición 2, tales que* $p^1 = \widetilde{i^1}$, $q^1 = \widetilde{j^1}$ *y ambas usan las mismas reglas de desempate. Entonces,* $(i^\tau, j^\tau)_{\tau \in \mathbb{N}}$ *y* $(p^\tau, q^\tau)_{\tau \in \mathbb{N}}$ *representan el mismo proceso de aprendizaje. Es decir,* $\forall \tau \in \mathbb{N}$, *se cumplen* $p^\tau = \sum_{s=1}^{\tau} \widetilde{i^s}$ *y* $q^\tau = \sum_{s=1}^{\tau} \widetilde{j^s}$.

Demostración. Por inducción en τ. Para $\tau = 1$, tenemos $p^1 = \widetilde{i^1} = \sum_{s=1}^{1} \widetilde{i^s}$ y $q^1 = \widetilde{j^1} = \sum_{s=1}^{1} \widetilde{j^s}$. Si $\tau > 1$, suponiendo que $p^{\tau-1} = \sum_{s=1}^{\tau-1} \widetilde{i^s}$ y $q^{\tau-1} = \sum_{s=1}^{\tau-1} \widetilde{j^s}$. Por la definición 2, $p^\tau = p^{\tau-1} + \widetilde{i}$ donde $i = d_1(argmax_{i \in N}\{\widetilde{i}Ap^{\tau-1}\})$. Pero, por la definición 1, $i^\tau = d_1(BR_1(y^{\tau-1})) = d_1(BR_1(\frac{\sum_{s=1}^{\tau-1} \widetilde{j^s}}{\tau-1})) =$ $d_1(argmax_{i \in N}\{A\frac{\sum_{s=1}^{\tau-1} \widetilde{j^s}}{\tau-1}\}) = d_1(argmax_{i \in N}\{A\frac{q^{\tau-1}}{\tau-1}\})$. Como escalar un vector no afecta la relación de orden entre sus componentes, entonces $i^\tau = d_1(argmax_{i \in N}\{Aq^{\tau-1}\})$. Luego, aplicando esto a la definición 2, $p^\tau = p^{\tau-1} + \widetilde{i^\tau} = \sum_{s=1}^{\tau-1} \widetilde{i^s} + \widetilde{i^\tau} = \sum_{s=1}^{\tau} \widetilde{i^s}$. Análogamente, $q^\tau = \sum_{s=1}^{\tau} \widetilde{j^s}$.

En el caso alternante, la jugada del jugador columna ya no es análoga a la del jugador fila, y el análisis sigue.

Lema 2 *Sea* (A, B) *un juego en forma bimatricial de* $n \times m$, $(i^\tau, j^\tau)_{\tau \in \mathbb{N}}$ *una secuencia de FP alternante según la definición 1 con secuencias de creencias* y^τ, x^τ *y sea* $(p^\tau, q^\tau)_{\tau \in \mathbb{N}}$ *una secuencia de FP alternante según la definición 2, tales que* $p^1 = \widetilde{i^1}$ *y ambas usan las mismas reglas de desempate. Entonces,* $(i^\tau, j^\tau)_{\tau \in \mathbb{N}}$ *y* $(p^\tau, q^\tau)_{\tau \in \mathbb{N}}$ *representan el mismo proceso de aprendizaje. Es decir, para todo* $\tau \in \mathbb{N}$, *se cumplen* $p^\tau = \sum_{s=1}^{\tau} \widetilde{i^s}$ *y* $q^\tau = \sum_{s=1}^{\tau} \widetilde{j^s}$.

Demostración. Por inducción en τ. Para $\tau = 1$, $p^1 = \widetilde{i^1} = \sum_{s=1}^{1} \widetilde{i^s}$. Por la definición 2, $q^1 = q^0 + \widetilde{j}$ donde q^0 es el vector nulo y $j = d_2(argmax_{j \in M}\{p^1 B\widetilde{j}\}) =$ $d_2(argmax_{j \in M}\{\widetilde{i^1} B\widetilde{j}\})$. Además, por la definición 1, $j^1 = d_2(BR_2(x^1)) =$ $d_2(BR_2(\frac{\sum_{s=1}^{1} \widetilde{i^s}}{\tau})) = d_2(BR_2(i^1)) = d_2(argmax_{j \in M}\{\widetilde{i^1} B\widetilde{j}\})$. Luego, $q^1 = \widetilde{j^1} = \sum_{s=1}^{1} \widetilde{j^s}$. Si $\tau > 1$, suponiendo que $p^{\tau-1} = \sum_{s=1}^{\tau-1} \widetilde{i^s}$ y $q^{\tau-1} = \sum_{s=1}^{\tau-1} \widetilde{j^s}$. Por el mismo argumento que en el caso inductivo del lema 1, $p^\tau = \sum_{s=1}^{\tau} \widetilde{i^s}$. Por otro lado, $q^\tau = y^\tau + \widetilde{j}$ donde $j = d_2(argmax_{j' \in M}\{p^{\tau+1} B\widetilde{j'}\}) =$ $d_2(argmax_{j' \in M}\{\sum_{s=1}^{\tau} \widetilde{i^s} B\widetilde{j'}\})$. Por la definición 1, $j^\tau = d_2(BR_2(x^\tau)) =$ $d_2(BR_2(\frac{\sum_{s=1}^{\tau} \widetilde{i^s}}{\tau})) = d_2(argmax_{j' \in N}\{\frac{\sum_{s=1}^{\tau} \widetilde{i^s}}{\tau} B\widetilde{j}\}) = d_2(argmax_{j' \in N}\{\frac{p^\tau}{\tau} B\widetilde{j}\})$. Como escalar un vector no afecta la relación de orden entre sus componentes, $j^\tau = d_2(argmax_{j' \in N}\{p^\tau B\widetilde{j}\})$. Luego, aplicando esto a la definición 2, $q^\tau = q^{\tau-1} + j^\tau = \sum_{s=1}^{\tau-1} \widetilde{j^s} + j^\tau = \sum_{s=1}^{\tau} \widetilde{j^s}$.

Claramente, en juegos del mismo tamaño, cualquier transformación que sea sólo un cambio de escala en los pagos preservará las secuencias de FP. Interesa en particular cuando el tamaño cambia, para clases de juegos estudiadas. La intención es que, tras el cambio, se preserve la mayoría de los valores. Llamamos *rango de imagen* de una matriz A al cardinal del conjunto de valores tomados por todos sus elementos. Buscamos entonces ejemplos de juegos que tengan rango de imagen mínimo manteniendo la clase, y nos interesará esencialmente el orden del rango de imagen tras el agregado de filas o columnas.

Lema 3 *Sea $A = (a_{i,j})$ un juego de intereses idénticos de tamaño $n \times m$ con un único NE puro, y sean $n' \geq n$, $m' \geq m$. Entonces existe un juego de intereses idénticos A' de tamaño $n' \times m'$ y rango de imagen $O(n+m)$ con un único NE puro tal que: (1) toda secuencia de FP (simultáneo o alternante) en A lo es también en A'; y (2) toda secuencia de FP alternante sobre A' es, a partir del paso 2, una secuencia de FP alternante sobre A. Además, si A es no degenerado, puede pedirse que A' también lo sea.*

Demostración. Sea $v = min(A)$, valor a utilizar en todo el proceso que sigue (no cambia a lo largo de las iteraciones). Sin pérdida de generalidad y para simplificar la construcción, supondremos[4] que el NE está en la posición $(1,1)$. Para obtener un juego de intereses idénticos de $(n+1) \times m$ con la propiedad indicada, definimos $a'_{i,j} = a_{i,j}$ para $1 \leq i \leq n$, $1 \leq j \leq m$, $a'_{n+1,j} = v - n - j$ para $1 \leq j \leq m$. Se itera esta construcción $n' - n$ veces. Análogamente, para obtener un juego de intereses idénticos de $n \times (m+1)$ con la propiedad indicada, definimos[5] $a'_{i,j} = a_{i,j}$ para $1 \leq i \leq n$, $1 \leq j \leq m$, $a'_{i,m+1} = v - m - i$ para $1 \leq i \leq n$. Se itera esta construcción $m' - m$ veces. A' resulta un juego de intereses idénticos de $n' \times m'$ con la propiedad indicada: si $i^\tau \in BR_1(y^{\tau-1})$ en A, también $i^\tau \in BR_1(y^{\tau-1})$ en A', y similarmente para j^τ considerando $BR_2(x^{\tau-1})$ o $BR_2(x^\tau)$ en los casos simultáneo o alternante, respectivamente.

A priori no vale un lema similar para la reducción en tamaño de las matrices: puede haber secuencias que pasen por todas las acciones de cada jugador, de modo que podría no ser factible eliminar acciones de alguno de ellos.

5. Estabilidad de FP

En [11] se enuncia un resultado que da una intuición sobre el comportamiento de FP. Al ser este un mecanismo usado para encontrar un NE, es esperable un comportamiento estable alrededor de ellos. Allí se prueba este *principio de estabilidad* usando otros principios. Damos a continuación una prueba alternativa más directa considerando posibles empates, y agregamos otra para el caso de FP alternante.

Lema 4 *Sea $(i^\tau, j^\tau)_{\tau \in \mathbb{N}}$ una secuencia de aprendizaje de FP alternante (resp., simultáneo) en el juego en forma bimatricial (A, B) de tamaño $n \times m$, con secuencias de creencias $(x^\tau, y^\tau)_{\tau \in \mathbb{N}}$ y reglas de desempate (d_1, d_2). Si en la iteración k se jugó el NE puro (i^*, j^*), entonces $i^* \in BR_1(y^k)$ y $j^* \in BR_2(x^{k+1})$ (resp., $j^* \in BR_2(x^k)$).*

[4] Consecuentemente con la construcción que sigue.

[5] Los términos $-j$ y $-i$ aseguran la condición de no degenerado, requerida por el teorema 4, y los términos $-n$ y $-m$ aseguran rango de imagen $O(n+m)$.

Demostración. Para el jugador fila, veamos que i^* es una mejor respuesta a las creencias del jugador fila sobre la estrategia del jugador columna. Veamos primero entonces qué forma tienen estas creencias según como se actualizan.

$$y^k = \frac{\sum_{s=1}^{k-1} \widetilde{j^s} + \widetilde{j^*}}{k} = \frac{\sum_{s=1}^{k-1} \widetilde{j^s}}{k} + \frac{\widetilde{j^*}}{k} = \frac{\sum_{s=1}^{k-1} \widetilde{j^s}(k-1)}{k(k-1)} + \frac{\widetilde{j^*}}{k} = \frac{k-1}{k} y^{k-1} + \frac{\widetilde{j^*}}{k}$$

Luego, $BR_1(y^k) = BR_1(\frac{k-1}{k} y^{k-1} + \frac{\widetilde{j^*}}{k}) = \text{argmax}_{i \in N} \{\widetilde{i} A (\frac{(k-1)y^{k-1}}{k} + \frac{\widetilde{j^*}}{k})\} = \text{argmax}_{i \in N} \{\frac{(k-1)}{k} \widetilde{i} A y^{k-1} + \frac{1}{k} \widetilde{i} A \widetilde{j^*}\}$. Como en la iteración k se jugó el perfil (i^*, j^*), por la definición 1, $i^* \in BR_1(y^{k-1}) = \text{argmax}_{i \in N} \{\widetilde{i} A y^{k-1}\}$. Es decir que para cualquier $i \in N$, $\widetilde{i^*} A y^{k-1} \geq \widetilde{i} A y^{k-1}$ y también (multiplicando en ambos lados por una constante positiva) $\frac{(k-1)}{k} \widetilde{i^*} A y^{k-1} \geq \frac{(k-1)}{k} \widetilde{i} A y^{k-1}$. Al ser un NE, $i^* \in BR_1(\widetilde{j^*}) = \text{argmax}_{i \in N} \{\widetilde{i} A \widetilde{j^*}\}$. Es decir que para cualquier $i \in N$, $\widetilde{i^*} A \widetilde{j^*} \geq \widetilde{i} A \widetilde{j^*}$ y consecuentemente $\frac{1}{k} \widetilde{i^*} A \widetilde{j^*} \geq \frac{1}{k} \widetilde{i} A \widetilde{j^*}$. Sumando estas dos desigualdades, $\frac{(k-1)}{k} \widetilde{i^*} A y^{k-1} + \frac{1}{k} \widetilde{i^*} A \widetilde{j^*} \geq \frac{(k-1)}{k} \widetilde{i} A y^{k-1} + \frac{1}{k} \widetilde{i} A \widetilde{j^*}$ para cualquier $i \in N$ y por lo tanto $i^* \in BR_1(y^k)$. Similarmente para el jugador columna, $j^* \in BR_2(x^{k+1})$. Para FP simultáneo, el caso para el jugador fila es igual al detallado, y el del jugador columna, análogo al del jugador fila.

Lema 5 *Sea $(i^\tau, j^\tau)_{\tau \in \mathbb{N}}$ una secuencia de aprendizaje de FP alternane (resp., simultáneo) en el juego en forma bimatricial (A, B) de tamaño $n \times m$, con secuencias de creencias $(x^\tau, y^\tau)_{\tau \in \mathbb{N}}$ y reglas de desempate (d_1, d_2). Si en la iteración k se jugó el NE puro (i^*, j^*), entonces $BR_1(y^k) \subseteq BR_1(y^{k-1})$ y $BR_2(x^{k+1}) \subseteq BR_2(x^k)$ (resp., $BR_2(x^k) \subseteq BR_2(x^{k-1})$).*

Demostración. Veamos el caso del jugador fila. Debemos probar que para todo $i \in N$, $i \in BR_1(y^k) \implies i \in BR_1(y^{k-1})$. Supongamos por el absurdo un $i' \in N$ tal que $i' \notin BR_1(y^{k-1})$. Como $BR_1(y^{k-1}) = \text{argmax}_{i \in N} \{\widetilde{i} A y^{k-1}\}$, si $i^* \in BR_1(y^{k-1})$ pero $i' \notin BR_1(y^{k-1})$, entonces $\widetilde{i^*} A y^{k-1} > \widetilde{i'} A y^{k-1}$ y luego (multiplicando ambos lados por una constante positiva) $\frac{(k-1)}{k} \widetilde{i^*} A y^{k-1} > \frac{(k-1)}{k} \widetilde{i'} A y^{k-1}$. Además, como (i^*, j^*) es un NE puro, $i^* \in BR_1(\widetilde{j^*}) = \text{argmax}_{i \in N} \{\widetilde{i} A \widetilde{j^*}\}$. Es decir que $\widetilde{i^*} A \widetilde{j^*} \geq \widetilde{i'} A \widetilde{j^*}$ y también $\frac{k-1}{k} \widetilde{i^*} A \widetilde{j^*} \geq \frac{k-1}{k} \widetilde{i'} A \widetilde{j^*}$. Entonces:

$$\begin{aligned} \widetilde{i^*} A y^k &= \widetilde{i^*} A (\frac{k-1}{k} y^{k-1} + \frac{\widetilde{j^*}}{k}) = \frac{(k-1)}{k} \widetilde{i^*} A y^{k-1} + \frac{1}{k} \widetilde{i^*} A \widetilde{j^*} \\ &> \frac{(k-1)}{k} \widetilde{i'} A y^{k-1} + \frac{1}{k} \widetilde{i'} A \widetilde{j^*} = \widetilde{i'} A (\frac{k-1}{k} y^{k-1} + \frac{\widetilde{j^*}}{k}) = \widetilde{i'} A y^k \end{aligned}$$

Puesto que por el lema anterior $i^* \in BR_1(y^k) = \text{argmax}_{i \in N} \{\widetilde{i} A y^k\}$, entonces $i' \notin BR_1(y^k)$. Similarmente para el jugador columna, $BR_2(x^{k+1}) \subseteq BR_2(x^k)$. Para FP simultáneo, el caso para el jugador fila es igual al detallado, y el del jugador columna, análogo al del jugador fila.

Teorema 1 *Sea $(i^\tau, j^\tau)_{\tau \in \mathbb{N}}$ una secuencia de aprendizaje de FP alternante (resp., simultáneo) en el juego en forma bimatricial (A, B) de tamaño $n \times m$, con secuencias de creencias $(x^\tau, y^\tau)_{\tau \in \mathbb{N}}$ y reglas de desempate (d_1, d_2). Si en la iteración k*

se jugó el NE puro (i^*, j^*), entonces este perfil se repetirá en todas las iteraciones siguientes.

Demostración. Nuevamente, veamos el caso del jugador fila. En la iteración k se jugó (i^*, j^*), luego $d_1(BR_1(y^{k-1})) = i^*$. Pero, por los lemas anteriores, $i^* \in BR_1(y^k) \subseteq BR_1(y^{k-1})$. Luego, por la definición de reglas de desempate, $d_1(BR_1(y^k)) = i^*$. Similarmente para el jugador columna, $d_2(BR_2(x^k)) = j^*$. Por los lemas anteriores, $j^* \in BR_2(x^{k+1}) \subseteq BR_2(x^k)$. Luego, por la definición de reglas de desempate, $d_2(BR_2(x^{k+1})) = j^*$. Para FP simultáneo, ambos casos siguen esta línea y son análogos entre sí.

6. Velocidad de convergencia de FP alternante

En esta sección extendemos los resultados de [24] sobre cotas superiores de la velocidad de convergencia de FP simultáneo en clases de juegos de interés. Veamos primero que para el caso de los juegos de suma constante simétricos, el resultado efectivamente es trasladable en forma bastante directa a la forma alternante. En el enunciado alcanzará con pedir que el último perfil sea un NE puro ya que, en ese caso, por el teorema 1, todos los siguientes lo serán.

Teorema 2 *Para todo $k \geq 2$ existe un juego simétrico A de 3×3 representable en $O(k)$ bits con al menos un NE puro y una secuencia de FP alternante $(i^\tau, j^\tau)_{\tau \in \mathbb{N}}$ sobre A tal que (i^{2^k-1}, j^{2^k-1}) no es un NE puro.*

Demostración. Sea A el juego simétrico siguiente:

$$\begin{array}{c|ccc|}
\cdot & j_1 & j_2 & j_3 \\
\hline
i_1 & 0 & \text{-1} & -\epsilon \\
i_2 & 1 & 0 & -\epsilon \\
i_3 & \epsilon & \epsilon & 0 \\
\hline
\end{array}$$

Si $0 < \epsilon < 1$, (i_3, j_3) es el único NE puro [6]. Consideremos un $k > 1$ arbitrario y sea $\epsilon = 2^{-k}$. Para estos valores, ϵ puede codificarse en $O(k)$ bits, mientras que las otras utilidades del juego son constantes, por lo que la representación del juego será tambien del orden de $O(k)$ bits. Por lo tanto, basta con probar que un proceso de FP alternante puede requerir $2^k - 1$ rondas antes de que se juegue (i_3, j_3). Si el proceso comienza con el jugador fila jugando i_1, entonces las utilidades esperadas del jugador columna serán $-x^1 A = (0, 1, \epsilon)$ y elegirá j_2. En la siguiente ronda, el jugador 1 reaccionará con i_3, dando que el jugador 2 tendrá una creencia sobre su estrategia de $x^2 = (\frac{1}{2}, 0, \frac{1}{2})$, por lo que las utilidades esperadas del jugador 2 serán $-x^2 A = (\frac{-\epsilon}{2}, \frac{1-\epsilon}{2}, \frac{\epsilon}{2})$ y volverá a elegir j_2. Este perfil (i_3, j_2) se repetirá $2^k - 2$ rondas, ya que tendremos que mientras $2 \leq \tau < 2^k$, se cumplirán $x^\tau = \frac{\widetilde{i_1} + (\tau-1)\widetilde{i_3}}{\tau} = (\frac{1}{\tau}, 0, \frac{\tau-1}{\tau})$, $-x^\tau A = (-\frac{(\tau-1)\epsilon}{\tau}, \frac{1-(\tau-1)\epsilon}{\tau}, \frac{\epsilon}{\tau})$, $y^\tau = \widetilde{j_2} = (0, 1, 0)$, $A y^\tau = (-1, 0, \epsilon)$.

[6] Puede verse en forma directa o bien al ser el único perfil resultante después de la eliminación iterada de estrategias estrictamente dominadas [2]; y lo mismo para los dos teoremas siguientes.

Ronda	(i,j)	x	xB	y	Ay
1	$(i_1,\ j_2)$	$(1,\ 0,\ 0)$	$(0,1,2^{-k})$	$(0,\ 1,\ 0)$	$(-1,\ 0,\ 2^{-k})$
2	$(i_3,\ j_2)$	$(\frac{1}{2},\ 0,\ \frac{1}{2})$	$(-2^{-(k+1)},\ \frac{1-2^{-k}}{2},\ 2^{-(k+1)})$	$(0,\ 1,\ 0)$	$(-1,\ 0,\ 2^{-k})$
3	$(i_3,\ j_2)$	$(\frac{1}{3},\ 0,\ \frac{2}{3})$	$(-\frac{2^{-k+1}}{3},\ \frac{1-2^{-k+1}}{3},\ \frac{2^{-k}}{3})$	$(0,\ 1,\ 0)$	$(-1,\ 0,\ 2^{-k})$
	$\vdots$				
τ	$(i_3,\ j_2)$	$(\frac{1}{\tau},\ 0,\ \frac{\tau-1}{\tau})$	$(-\frac{(\tau-1)2^{-k}}{\tau},\ \frac{1-(\tau-1)2^{-k}}{\tau},\ \frac{2^{-k}}{\tau})$	$(0,\ 1,\ 0)$	$(-1,\ 0,\ 2^{-k})$
	$\vdots$				
2^k	$(i_3,\ j_2)$	$(\epsilon,\ 0,\ \frac{\epsilon-1}{\epsilon})$	$(\epsilon(\epsilon-1),\ \epsilon^2,\ \epsilon^2)$	$(0,\ 1,\ 0)$	$(-1,\ 0,\ 2^{-k})$

Tabla 1: Proceso de FP alternante en el juego del teorema 2

La tabla 1 muestra el proceso. Para el jugador fila, justificar su decisión es trivial ya que la estrategia percibida de su oponente es pura e i_3 es la única acción con utilidad esperada positiva. Para entender por qué el jugador columna no cambia su estrategia, notemos que[7] $\tau < 2^k = \frac{1}{\epsilon} \Rightarrow \epsilon(\tau - 1 + 1) < 1 \Rightarrow 1 - (\tau-1)\epsilon > \epsilon \Rightarrow \frac{1-(\tau-1)\epsilon}{\tau} > \frac{\epsilon}{\tau}$.

Concluimos entonces que $(i_1, j_2), \underbrace{(i_3, j_2), ...(i_3, j_2)}_{2^k-2 \text{ veces}}$ es una secuencia de aprendizaje de FP alternante de este juego, exponencialmente larga en k y en la cual no se juega ningún NE puro[8].

La demostración para el caso de los juegos no degenerados de 2×3 es un poco menos directa y requiere plantear una variante del juego propuesto en [24], que podemos extender a otros tamaños.

Teorema 3 *Para todo $k \geq 2$, $n \geq 2$, $m \geq 3$, existe un juego no degenerado de intereses idénticos A de $n \times m$ representable en $O(k + nm.log(max\{n, m\}))$ bits con al menos un NE puro, con rango de imagen $O(n + m)$, y una secuencia de FP alternante $(i^\tau, j^\tau)_{\tau \in \mathbb{N}}$ sobre A tal que para todo $\tau \leq 2^k$, (i^τ, j^τ) no es un NE puro.*

Demostración. Sea A el juego de intereses idénticos siguiente:

$$\begin{array}{c|ccc|} \cdot & j_1 & j_2 & j_3 \\ \hline i_1 & 1 & 2 & 0 \\ i_2 & 0 & 2+\epsilon & 2+2\epsilon \\ \hline \end{array}$$

[7] Como interpretación informal, el incentivo resultante de la única vez que se jugó i_1 es muy fuerte por la gran diferencia de utilidades para el jugador 2 entre (i_1, j_2) e (i_3, j_3), por lo que deberán transcurrir 2^{k-1} iteraciones de i_3 luego de i_1 para que las utilidades esperadas se compensen.

[8] Que en la ronda 2^k se juegue un NE o aún no, dependerá de la regla de desempate usada.

Si $0 < \epsilon < 1$, el juego es no degenerado e (i_2, j_3) es el único NE puro.

Consideremos un $k > 1$ arbitrario. Mostraremos que para $\epsilon = 2^{-k}$, un proceso de FP alternante puede tomar 2^k rondas antes de que se juegue (i_2, j_3). Al igual que en la demostración anterior, el juego puede codificarse en $O(k)$ bits. Si el proceso comienza con el jugador fila jugando i_1, entonces $x^1 A = (1, 2, 0)$ y por lo tanto el jugador columna elegirá j_2. En la siguiente ronda, el jugador 1 reaccionará con i_2, dando que el jugador 2 tendrá una creencia sobre su estrategia de $x^2 = (\frac{1}{2}, \frac{1}{2})$, por lo que las utilidades esperadas del jugador 2 serán $x^2 A = (\frac{1}{2}, 2 + \frac{\epsilon}{2}, 1 + \epsilon)$ y volverá a elegir j_2. El perfil (i_2, j_2) se repetirá $2^k - 1$ rondas, ya que tendremos que mientras $2 \leq \tau \leq 2^k$, se cumplirán $x^\tau = \frac{\widetilde{i_1} + (\tau-1)\widetilde{i_2}}{\tau} = (\frac{1}{\tau}, \frac{\tau-1}{\tau})$, $x^\tau A = (\frac{1}{\tau}, 2 + \frac{\tau-1}{\tau}\epsilon, 2\frac{\tau-1}{\tau}(1+\epsilon))$, $y^\tau = \widetilde{j_2} = (0, 1, 0)$, $Ay^\tau = (2, 2+\epsilon)$. La tabla 2 muestra el proceso. Para el jugador fila, justificar su decisión es trivial ya que la estrategia percibida de su oponente es pura y la utilidad esperada de i_2 es siempre marginalmente mayor que la de i_1. Para entender por qué el jugador columna no cambia su estrategia, notemos que[9] $\tau \leq 2^k = \frac{1}{\epsilon} \Rightarrow 2 + \frac{\tau-1}{\tau}\epsilon > \frac{1}{\tau}$ y $2 + \frac{\tau-1}{\tau}\epsilon > 2\frac{\tau-1}{\tau}(1+\epsilon)$.

	(i,j)	x	xB	y	Ay
Ronda					
1	$(i_1,\ j_2)$	$(1,0)$	$(1,2,0)$	$(0,1,0)$	$(2,2+\epsilon)$
2	$(i_2,\ j_2)$	$(\frac{1}{2},\frac{1}{2})$	$(\frac{1}{2},2+\frac{\epsilon}{2},1+\epsilon)$	$(0,1,0)$	$(2,2+\epsilon)$
3	$(i_2,\ j_2)$	$(\frac{1}{3},\frac{2}{3})$	$(\frac{1}{3},2+\frac{2}{3}\epsilon,\frac{4+4\epsilon}{3})$	$(0,1,0)$	$(2,2+\epsilon)$
	$\vdots$				
τ	$(i_2,\ j_2)$	$(\frac{1}{\tau},\frac{\tau-1}{\tau})$	$(\frac{1}{\tau},2+\frac{\tau-1}{\tau}\epsilon,2\frac{\tau-1}{\tau}(1+\epsilon))$	$(0,1,0)$	$(2,2+\epsilon)$
	$\vdots$				
2^k	$(i_2,\ j_2)$	$(\epsilon,1-\epsilon)$	$(\epsilon,2+\epsilon-\epsilon^2,2(1-\epsilon)(1+\epsilon))$	$(0,1,0)$	$(2,2+\epsilon)$

Tabla 2: Proceso de FP alternante en el juego del teorema 3

Concluimos entonces que $(i_1, j_2), \underbrace{(i_2, j_2), ...(i_2, j_2)}_{2^k - 1 \text{ veces}}$ es una secuencia de aprendizaje de FP alternante de este juego, exponencialmente larga en k y en la que no aparece ningún NE puro, y el teorema sigue del lema 3 con a lo sumo nm elementos que requieren $O(log(max\{n, m\}))$ bits cada uno.

El hecho de que el juego originalmente propuesto por Brandt, Fischer y Harrenstein no sirva para el teorema anterior no es menor. En el teorema que sigue

[9] Similarmente a lo que ocurre en el teorema anterior, si bien en las iteraciones recientes el jugador 1 jugó i_2, el incentivo resultante de la única vez que jugó i_1 es muy alto por la gran diferencia de utilidades entre (i_1, j_3) e (i_2, j_3), por lo que deberán transcurrir $2^k - 1$ iteraciones de i_2 luego de i_1 para que las utilidades esperadas se compensen.

vemos que FP simultáneo puede requerir una cantidad de rondas exponencial mientras que toda secuencia de FP alternante converge inmediatamente.

Teorema 4 *Para todo $k \geq 2$, $n \geq 2$, $m \geq 3$, existe un juego no degenerado de intereses idénticos A de $n \times m$ representable en $O(k + nm.log(max\{n, m\}))$ bits, con rango de imagen $O(n + m)$, y una secuencia de FP simultáneo $(i^\tau, j^\tau)_{\tau\in\mathbb{N}}$ sobre A tal que, para todo $\tau \leq 2^k$, (i^τ, j^τ) no es un NE puro, y para todo proceso de FP alternante $(\widehat{i^\tau}, \widehat{j^\tau})_{\tau\in\mathbb{N}}$ en A, $(\widehat{i^4}, \widehat{j^4})$ es un NE puro.*

Demostración. Sea A el juego de intereses idénticos siguiente:

$\cdot$	j_1	j_2	j_3
i_1	1	2	0
i_2	0	$2+\epsilon$	3

Si $0 < \epsilon < 1$, el juego es no degenerado e (i_2, j_3) es el único NE puro.

Consideremos un $k > 1$ arbitrario. Mostraremos que para $\epsilon = 2^{-k}$, un proceso de FP simultáneo puede tomar 2^k rondas antes de que se juegue (i_2, j_3), mientras que todo proceso de FP alternante converge de forma pura en un número constante de rondas. Al igual que en los teoremas anteriores, el juego puede codificarse en $O(k)$ bits.

Veamos primero el caso alternante. Existen dos posibles secuencias de FP alternante para este juego, dado que el jugador 1 elegirá primero y el jugador 2 reaccionará. Si el jugador fila juega i_2, el jugador columna responderá con j_3, siendo este el NE puro. Si el jugador fila comienza con i_1, el jugador columna responderá con j_2. Esto llevará al jugador fila a jugar i_2 en la segunda ronda por ser $A\widetilde{j^2} = (2, 2 + \epsilon)$, mientras que el jugador fila continuará jugando j_2. Esta situación se repetirá una ronda más, tras la cual, la jugador columna se verá incentivado a jugar j_3. Estos dos procesos pueden verse en las tablas 3 y 4. Vemos entonces que, independientemente de k, ambos procesos convergen de forma pura en 4 rondas o menos.

Para el caso simultáneo, nos basamos en la prueba de [24]. Si el proceso comienza con (i_1, j_1), las utilidades esperadas serán $Ay^1 = (1, 0)$ y $x^1B = (1, 2, 0)$ respectivamente. Luego, en la segunda iteración el jugador fila elegirá i_1 y el jugador columna j_2. Las utilidades esperadas se actualizarán entonces como $Ay^2 = (\frac{3}{2}, \frac{2+\epsilon}{2})$ y $x^2B = (1, 2, 0)$. En total, por al menos $2^k - 1$ rondas, los jugadores elegirán las mismas jugadas que en la iteración 2, dado que para todo $2 \leq \tau \leq 2^k$ tendremos $Ay^\tau = (2 - \frac{1}{\tau}, (2 + \epsilon)\frac{\tau-1}{\tau})$ y $x^\tau B = (1, 2, 0)$, y $\tau \leq 2^k \Rightarrow 1 > \epsilon(\tau - 1) \Rightarrow 2 - \frac{1}{\tau} > (2 + \epsilon)\frac{\tau-1}{\tau}$. La tabla 5 muestra el proceso.

Concluimos entonces que $(i_1, j_1), \underbrace{(i_1, j_2), ...(i_1, j_2)}_{2^k - 1 \text{ veces}}$ es una secuencia de aprendizaje de FP simultáneo de este juego exponencialmente larga en k y en la cual no se juega ningún NE puro, y el teorema sigue del lema 3 con a lo sumo nm elementos que requieren $O(log(max\{n, m\}))$ bits cada uno.

	(i,j)	x	xB	y	Ay
Iteración					
1	(i_2, j_3)	$(0,1)$	$(0, 2+\epsilon, 3)$	$(0,0,1)$	$(0,3)$

Tabla 3: Proceso de FP alternante sobre el juego del teorema 4 comenzando por i_2

	(i,j)	x	xB	y	Ay
Iteración					
1	$(i_1,\ j_2)$	$(1,0)$	$(1,2,0)$	$(0,1,0)$	$(2, 2+\epsilon)$
2	$(i_2,\ j_2)$	$(\frac{1}{2}, \frac{1}{2})$	$(\frac{1}{2}, 2+\frac{\epsilon}{2}, \frac{3}{2})$	$(0,1,0)$	$(2, 2+\epsilon)$
3	$(i_2,\ j_2)$	$(\frac{1}{3}, \frac{2}{3})$	$(\frac{1}{3}, 2+\frac{2}{3}\epsilon, 2)$	$(0,1,0)$	$(2, 2+\epsilon)$
4	$(i_2,\ j_3)$	$(\frac{1}{4}, \frac{3}{4})$	$(\frac{1}{4}, 2+\frac{3}{4}\epsilon, \frac{9}{4})$	$(0, \frac{3}{4}, \frac{1}{4})$	$(\frac{3}{2}, \frac{3(3+\epsilon)}{4})$

Tabla 4: Proceso de FP alternante sobre el juego del teorema 4 comenzando por i_1

	(i,j)	x	xB	y	Ay
Iteración					
1	(i_1, j_1)	$(1,0)$	$(1,2,0)$	$(1,0,0)$	$(1,0)$
2	(i_1, j_2)	$(1,0)$	$(1,2,0)$	$(\frac{1}{2}, \frac{1}{2}, 0)$	$(\frac{3}{2}, \frac{2+\epsilon}{2})$
3	(i_1, j_2)	$(1,0)$	$(1,2,0)$	$(\frac{1}{3}, \frac{2}{3}, 0)$	$(\frac{5}{3}, \frac{4+2\epsilon}{3})$
	$\vdots$				
τ	$(i_1,\ j_2)$	$(1,0)$	$(1,2,0)$	$(\frac{1}{\tau}, \frac{\tau-1}{\tau}, 0)$	$(2-\frac{1}{\tau}, (2+\epsilon)\frac{\tau-1}{\tau})$
	$\vdots$				
2^k	$(i_1,\ j_2)$	$(1,0)$	$(1,2,0)$	$(\epsilon, 1-\epsilon, 0)$	$(2-\epsilon, (2+\epsilon)(1-\epsilon))$

Tabla 5: Proceso de FP simultáneo sobre el juego del teorema 4 comenzando por (i_1, j_1)

7. Conclusiones y trabajo futuro

Hemos probado que FP alternante podría requerir una cantidad exponencial de rondas antes de alcanzar un NE puro, para lo cual hemos definido juegos de tamaño minimal que pueden ser extendidos manteniendo buenas propiedades. Hay varias posibilidades de trabajo futuro. Dado un juego que tenga algún NE puro, ¿cuál es el costo computacional de determinar si FP simultáneo (o alternante) va a requerir un número exponencial de rondas antes de alcanzar un NE? En este trabajo mostramos patrones para ello, pero interesaría si en el caso de matrices cualesquiera se puede mejorar la cota exponencial que ya tenemos mediante la simulación del proceso. Una variante involucra generalizar FP considerando $BR(x^{\tau-k})$ con k una constante, es decir, que los jugadores sólo puedan observar las decisiones tomadas varias rondas atrás. Son también relevantes otras nociones de convergencia más débiles y el tratamiento de equilibrios mixtos. Eventualmente se podría combinar FP con aprendizaje por refuerzo. Dejamos finalmente abierto el planteo de los problemas tratados en este artículo para juegos de más de dos jugadores.

Referencias

[1] G. Brown. “Iterative solution of games by fictitious play”. En: *Activity Analysis of Production and Allocation* 13 (ene. de 1951).

[2] M. Osborne y A. Rubinstein. *A course in Game Theory.* MIT Press, 1994.

[3] J. Robinson. “An Iterative Method of Solving a Game”. En: *Annals of Mathematics. Second Series* 54 (sep. de 1951). DOI: 10.2307/1969530.

[4] K. Miyasawa. “On the Convergence of Learning Processes in a 2x2 Non-Zero-Person Game”. En: *Research Memo* 33 (oct. de 1961).

[5] D. Monderer y A. Sela. “A 2x2 Game without the Fictitious Play Property”. En: *Games and Economic Behavior* 14 (feb. de 1996), págs. 144-148. DOI: 10.1006/game.1996.0045.

[6] L. Shapley. “Some Topics in Two-Person Games”. En: *Annals of Mathematics Studies.* 52 (ene. de 1964).

[7] D. Monderer y L. Shapley. “Fictitious Play Property for Games with Identical Interests”. En: *Journal of Economic Theory* 68 (feb. de 1996), págs. 258-265. DOI: 10.1006/jeth.1996.0014.

[8] D. Monderer y L. S. Shapley. “Potential Games”. En: *Games and Economic Behavior* 14.1 (1996), págs. 124-143. ISSN: 0899-8256. DOI: https://doi.org/10.1006/game.1996.0044. URL: http://www.sciencedirect.com/science/article/pii/S0899825696900445.

[9] U. Berger. “Learning in games with strategic complementarities revisited”. En: *Journal of Economic Theory* 143 (nov. de 2008), págs. 292-301. DOI: 10.1016/j.jet.2008.01.007.

[10] A. Sela. "Fictitious play in 'one-against-all' multi-player games". En: *Economic Theory* 14 (nov. de 1999), págs. 635-651. DOI: 10.1007/s001990050345.

[11] D. Monderer y A. Sela. *Fictitious play and- no-cycling conditions.* Sonderforschungsbereich 504 Publications 97-12. University of Mannheim, jun. de 1997. URL: https://ideas.repec.org/p/xrs/sfbmaa/97-12.html.

[12] U. Berger. "Fictitious play in 2xn games". En: *Journal of Economic Theory* 120 (abr. de 2003).

[13] R. Chu y G. Vreeswijk. "Extending fictitious play with pattern recognition". En: *CEUR Workshop Proceedings* 1113 (ene. de 2013), págs. 40-53.

[14] A. Washburn. "A new kind of fictitious play". En: *Naval Research Logistics (NRL)* 48 (jun. de 2001), págs. 270-280. DOI: 10.1002/nav.7.

[15] U. Berger. "Brown's original fictitious play". En: *Journal of Economic Theory* 135 (feb. de 2007), págs. 572-578. DOI: 10.1016/j.jet.2005.12.010.

[16] G. B. Dantzig. *Linear Programming and Extensions.* Santa Monica, CA: RAND Corporation, 1963. DOI: 10.7249/R366.

[17] G. Dantzig. "A proof of the equivalence of the programming problem and the game problem". En: *Activity Analysis of Production and Allocation* (1951), págs. 330-335.

[18] S. I. Gass y P. M. Zafra. "Modified fictitious play for solving matrix games and linear-programming problems". En: *Computers & Operations Research* 22.9 (1995), págs. 893-903. ISSN: 0305-0548. DOI: https://doi.org/10.1016/0305-0548(94)00075-J. URL: http://www.sciencedirect.com/science/article/pii/030505489400075J.

[19] T. III, M. Epelman y R. Smith. "A Fictitious Play Approach to Large-Scale Optimization". En: *Operations Research* 53 (jun. de 2005), págs. 477-489. DOI: 10.1287/opre.1040.0178.

[20] S. Karlin. *Mathematical Methods and Theory in Games, Programming, and Economics.* Vol. 1-2. Addison-Wesley, 1959.

[21] Y. Viossat y A. Zapechelnyuk. "No-Regret Dynamics and Fictitious Play". En: *Journal of Economic Theory* 148 (jul. de 2012). DOI: 10.1016/j.jet.2012.07.003.

[22] A. Jafari, A. Greenwald, D. Gondek y G. Ercal. "On No-Regret Learning, Fictitious Play, and Nash Equilibrium". En: *ICML* (jul. de 2001).

[23] C. Daskalakis y Q. Pan. "A Counterexample to Karlin's Strong Conjecture for Fictitious Play". En: *Proceedings - Annual IEEE Symposium on Foundations of Computer Science, FOCS* (dic. de 2014). DOI: 10.1109/FOCS.2014.10.

[24] F. Brandt, F. Fischer y P. Harrenstein. "On the Rate of Convergence of Fictitious Play". En: *Theory of Computing Systems* 53.1 (jul. de 2013), págs. 41-52. ISSN: 1433-0490. DOI: 10.1007/s00224-013-9460-5. URL: https://doi.org/10.1007/s00224-013-9460-5.