When do Best Response Dynamics Converge?

Aaron Roth

University of Pennsylvania

February 2 2021

Overview

- We know best response dynamics (BRD) converges in congestion games.

Overview

- We know best response dynamics (BRD) converges in congestion games.
- Is that it? How much further can we push it?

Overview

- We know best response dynamics (BRD) converges in congestion games.
- Is that it? How much further can we push it?
- Today: study a couple more games in which BRD converges, and try to abstract what is needed.

Overview

- We know best response dynamics (BRD) converges in congestion games.
- Is that it? How much further can we push it?
- Today: study a couple more games in which BRD converges, and try to abstract what is needed.
- Characterize *exactly* when BRD is guaranteed to converge.

Load Balancing Games on Identical Machines

Definition

A *load balancing game on identical machines* models n players $i \in P$ scheduling jobs of size $w_i > 0$ on m identical machines F.
The game has:

1. Action space $A_i = F$ for each player
2. For each machine $j \in F$, a load $\ell_j(a) = \sum_{i:a_i=j} w_i$

The cost of player i is the load of the machine he plays on: $c_i(a) = \ell_{a_i}(a)$.

Load Balancing Games on Identical Machines

Definition

A *load balancing game on identical machines* models n players $i \in P$ scheduling jobs of size $w_i > 0$ on m identical machines F. The game has:

1. Action space $A_i = F$ for each player
2. For each machine $j \in F$, a load $\ell_j(a) = \sum_{i:a_i=j} w_i$

The cost of player i is the load of the machine he plays on: $c_i(a) = \ell_{a_i}(a)$.

Almost a congestion game — but facility costs depend on *which* players are using them.

Load Balancing Games on Identical Machines

Load Balancing Games on Identical Machines

Theorem

Best response dynamics converge in load balancing games on identical machines.

Load Balancing Games on Identical Machines

Theorem

Best response dynamics converge in load balancing games on identical machines.

Corollary

Load balancing games on identical machines have pure strategy Nash equilibria

Proof

Define $\phi(a) = \frac{1}{2}\sum_{j=1}^{m} \ell_j(a)^2$. Suppose player i switches from machine j to machine j'. Then we have:

$$\Delta c_i(a) \equiv c_i(j', a_{-i}) - c_i(j, a_{-i})$$

Proof

Define $\phi(a) = \frac{1}{2}\sum_{j=1}^{m} \ell_j(a)^2$. Suppose player i switches from machine j to machine j'. Then we have:

$$\begin{aligned} \Delta c_i(a) &\equiv c_i(j', a_{-i}) - c_i(j, a_{-i}) \\ &= \ell_{j'}(a) + w_i - \ell_j(a) \end{aligned}$$

Proof

Define $\phi(a) = \frac{1}{2}\sum_{j=1}^{m} \ell_j(a)^2$. Suppose player i switches from machine j to machine j'. Then we have:

$$\begin{aligned}
\Delta c_i(a) &\equiv c_i(j', a_{-i}) - c_i(j, a_{-i}) \\
&= \ell_{j'}(a) + w_i - \ell_j(a) \\
&< 0
\end{aligned}$$

Proof

Define $\phi(a) = \frac{1}{2}\sum_{j=1}^{m} \ell_j(a)^2$. Suppose player i switches from machine j to machine j'. Then we have:

$$\begin{aligned}
\Delta c_i(a) &\equiv c_i(j', a_{-i}) - c_i(j, a_{-i}) \\
&= \ell_{j'}(a) + w_i - \ell_j(a) \\
&< 0
\end{aligned}$$

Similarly, we have:

$$\Delta\phi(a) \equiv \phi(j', a_{-i}) - \phi(j, a_{-i})$$

Proof

Define $\phi(a) = \frac{1}{2}\sum_{j=1}^{m} \ell_j(a)^2$. Suppose player i switches from machine j to machine j'. Then we have:

$$\begin{aligned}
\Delta c_i(a) &\equiv c_i(j', a_{-i}) - c_i(j, a_{-i}) \\
&= \ell_{j'}(a) + w_i - \ell_j(a) \\
&< 0
\end{aligned}$$

Similarly, we have:

$$\begin{aligned}
\Delta \phi(a) &\equiv \phi(j', a_{-i}) - \phi(j, a_{-i}) \\
&= \frac{1}{2}\left((\ell_{j'}(a) + w_i)^2 + (\ell_j(a) - w_i)^2 - \ell_{j'}(a)^2 - \ell_j(a)^2\right)
\end{aligned}$$

Proof

Define $\phi(a) = \frac{1}{2}\sum_{j=1}^{m} \ell_j(a)^2$. Suppose player i switches from machine j to machine j'. Then we have:

$$\begin{aligned}
\Delta c_i(a) &\equiv c_i(j', a_{-i}) - c_i(j, a_{-i}) \\
&= \ell_{j'}(a) + w_i - \ell_j(a) \\
&< 0
\end{aligned}$$

Similarly, we have:

$$\begin{aligned}
\Delta\phi(a) &\equiv \phi(j', a_{-i}) - \phi(j, a_{-i}) \\
&= \frac{1}{2}\left((\ell_{j'}(a) + w_i)^2 + (\ell_j(a) - w_i)^2 - \ell_{j'}(a)^2 - \ell_j(a)^2\right) \\
&= \frac{1}{2}\left(2w_i\ell_{j'}(a) + w_i^2 - 2w_i\ell_j(a) + w_i^2\right)
\end{aligned}$$

Proof

Define $\phi(a) = \frac{1}{2}\sum_{j=1}^{m} \ell_j(a)^2$. Suppose player i switches from machine j to machine j'. Then we have:

$$\begin{aligned}
\Delta c_i(a) &\equiv c_i(j', a_{-i}) - c_i(j, a_{-i}) \\
&= \ell_{j'}(a) + w_i - \ell_j(a) \\
&< 0
\end{aligned}$$

Similarly, we have:

$$\begin{aligned}
\Delta\phi(a) &\equiv \phi(j', a_{-i}) - \phi(j, a_{-i}) \\
&= \frac{1}{2}\left((\ell_{j'}(a) + w_i)^2 + (\ell_j(a) - w_i)^2 - \ell_{j'}(a)^2 - \ell_j(a)^2\right) \\
&= \frac{1}{2}\left(2w_i\ell_{j'}(a) + w_i^2 - 2w_i\ell_j(a) + w_i^2\right) \\
&= w_i\left(\ell_{j'}(a) + w_i - \ell_j(a)\right)
\end{aligned}$$

Proof

Define $\phi(a) = \frac{1}{2}\sum_{j=1}^{m} \ell_j(a)^2$. Suppose player i switches from machine j to machine j'. Then we have:

$$\begin{aligned}
\Delta c_i(a) &\equiv c_i(j', a_{-i}) - c_i(j, a_{-i}) \\
&= \ell_{j'}(a) + w_i - \ell_j(a) \\
&< 0
\end{aligned}$$

Similarly, we have:

$$\begin{aligned}
\Delta\phi(a) &\equiv \phi(j', a_{-i}) - \phi(j, a_{-i}) \\
&= \frac{1}{2}\left((\ell_{j'}(a) + w_i)^2 + (\ell_j(a) - w_i)^2 - \ell_{j'}(a)^2 - \ell_j(a)^2\right) \\
&= \frac{1}{2}\left(2w_i\ell_{j'}(a) + w_i^2 - 2w_i\ell_j(a) + w_i^2\right) \\
&= w_i\left(\ell_{j'}(a) + w_i - \ell_j(a)\right) \\
&= w_i \cdot \Delta c_i(a)
\end{aligned}$$

Proof

Define $\phi(a) = \frac{1}{2}\sum_{j=1}^{m} \ell_j(a)^2$. Suppose player i switches from machine j to machine j'. Then we have:

$$\begin{aligned}
\Delta c_i(a) &\equiv c_i(j', a_{-i}) - c_i(j, a_{-i}) \\
&= \ell_{j'}(a) + w_i - \ell_j(a) \\
&< 0
\end{aligned}$$

Similarly, we have:

$$\begin{aligned}
\Delta\phi(a) &\equiv \phi(j', a_{-i}) - \phi(j, a_{-i}) \\
&= \frac{1}{2}\left((\ell_{j'}(a) + w_i)^2 + (\ell_j(a) - w_i)^2 - \ell_{j'}(a)^2 - \ell_j(a)^2\right) \\
&= \frac{1}{2}\left(2w_i\ell_{j'}(a) + w_i^2 - 2w_i\ell_j(a) + w_i^2\right) \\
&= w_i\left(\ell_{j'}(a) + w_i - \ell_j(a)\right) \\
&= w_i \cdot \Delta c_i(a) \\
&< 0
\end{aligned}$$

Proof

Define $\phi(a) = \frac{1}{2}\sum_{j=1}^{m} \ell_j(a)^2$. Suppose player i switches from machine j to machine j'. Then we have:

$$\begin{aligned}
\Delta c_i(a) &\equiv c_i(j', a_{-i}) - c_i(j, a_{-i}) \\
&= \ell_{j'}(a) + w_i - \ell_j(a) \\
&< 0
\end{aligned}$$

Similarly, we have:

$$\begin{aligned}
\Delta\phi(a) &\equiv \phi(j', a_{-i}) - \phi(j, a_{-i}) \\
&= \frac{1}{2}\left((\ell_{j'}(a) + w_i)^2 + (\ell_j(a) - w_i)^2 - \ell_{j'}(a)^2 - \ell_j(a)^2\right) \\
&= \frac{1}{2}\left(2w_i\ell_{j'}(a) + w_i^2 - 2w_i\ell_j(a) + w_i^2\right) \\
&= w_i\left(\ell_{j'}(a) + w_i - \ell_j(a)\right) \\
&= w_i \cdot \Delta c_i(a) \\
&< 0
\end{aligned}$$

Note: $\Delta c_i \neq \Delta\phi$.

Red State/Blue State Game

And now — a game that doesn't look like a congestion game.

Red State/Blue State Game

And now — a game that doesn't look like a congestion game.

Definition

The Red State/Blue State game is played on a graph $G = (V, E)$.

1. The players are vertices $P = V$.
2. Each edge $e = (i, j) \in E$ has weight w_e
3. Actions $A_i = \{-1, 1\}$ (read {red, blue})
4. $u_i(a) = \sum_{e=(i,j)\in E} w_e \cdot a_i \cdot a_j = \sum_{j:a_i=a_j} w_{i,j} - \sum_{j:a_i\neq a_j} w_{i,j}$

Red State/Blue State Game

And now — a game that doesn't look like a congestion game.

Definition

The Red State/Blue State game is played on a graph $G = (V, E)$.

1. The players are vertices $P = V$.
2. Each edge $e = (i, j) \in E$ has weight w_e
3. Actions $A_i = \{-1, 1\}$ (read {red, blue})
4. $u_i(a) = \sum_{e=(i,j)\in E} w_e \cdot a_i \cdot a_j = \sum_{j:a_i=a_j} w_{i,j} - \sum_{j:a_i\neq a_j} w_{i,j}$

"Everyone picks an affiliation, and obtains utility equal to the weight of friends who pick the same affiliation, and disutility equal to the weight of friends who don't."

Red State/Blue State Game

Red State/Blue State Game

Theorem

The Red-State/Blue-State game always has a pure strategy Nash equilibrium.

Proof

Define:

$$\phi(a) = \sum_{j<i} w_{i,j} a_i a_j$$

Proof

Define:

$$\phi(a) = \sum_{j<i} w_{i,j} a_i a_j$$

Now consider a best response move made by player i. We have:

Proof

Define:

$$\phi(a) = \sum_{j<i} w_{i,j} a_i a_j$$

Now consider a best response move made by player i. We have:

$$\Delta u_i \quad = \quad \sum_{j \neq i} w_e \cdot a_i \cdot a_j - \sum_{j \neq i} w_e \cdot (-a_i) \cdot a_j$$

Proof

Define:

$$\phi(a) = \sum_{j<i} w_{i,j} a_i a_j$$

Now consider a best response move made by player i. We have:

$$\begin{aligned} \Delta u_i &= \sum_{j \neq i} w_e \cdot a_i \cdot a_j - \sum_{j \neq i} w_e \cdot (-a_i) \cdot a_j \\ &= 2 \sum_{j \neq i} w_e \cdot a_i \cdot a_j \end{aligned}$$

Proof

Define:

$$\phi(a) = \sum_{j<i} w_{i,j} a_i a_j$$

Now consider a best response move made by player i. We have:

$$\begin{aligned}
\Delta u_i &= \sum_{j \neq i} w_e \cdot a_i \cdot a_j - \sum_{j \neq i} w_e \cdot (-a_i) \cdot a_j \\
&= 2 \sum_{j \neq i} w_e \cdot a_i \cdot a_j
\end{aligned}$$

Similarly:

$$\Delta \phi = \sum_{j \neq i} w_e \cdot a_i \cdot a_j - \sum_{j \neq i} w_e \cdot (-a_i) \cdot a_j$$

Proof

Define:

$$\phi(a) = \sum_{j<i} w_{i,j} a_i a_j$$

Now consider a best response move made by player i. We have:

$$\begin{aligned} \Delta u_i &= \sum_{j \neq i} w_e \cdot a_i \cdot a_j - \sum_{j \neq i} w_e \cdot (-a_i) \cdot a_j \\ &= 2 \sum_{j \neq i} w_e \cdot a_i \cdot a_j \end{aligned}$$

Similarly:

$$\begin{aligned} \Delta \phi &= \sum_{j \neq i} w_e \cdot a_i \cdot a_j - \sum_{j \neq i} w_e \cdot (-a_i) \cdot a_j \\ &= 2 \sum_{j \neq i} w_e \cdot a_i \cdot a_j \end{aligned}$$

Proof

Define:

$$\phi(a) = \sum_{j<i} w_{i,j} a_i a_j$$

Now consider a best response move made by player i. We have:

$$\begin{aligned}
\Delta u_i &= \sum_{j\neq i} w_e \cdot a_i \cdot a_j - \sum_{j\neq i} w_e \cdot (-a_i) \cdot a_j \\
&= 2\sum_{j\neq i} w_e \cdot a_i \cdot a_j
\end{aligned}$$

Similarly:

$$\begin{aligned}
\Delta \phi &= \sum_{j\neq i} w_e \cdot a_i \cdot a_j - \sum_{j\neq i} w_e \cdot (-a_i) \cdot a_j \\
&= 2\sum_{j\neq i} w_e \cdot a_i \cdot a_j \\
&= \Delta u_i
\end{aligned}$$

Abstracting Away...

What do we need to make the proof work?

Abstracting Away...

What do we need to make the proof work?

Definition

A function $\phi : A \to \mathbb{R}_{\geq 0}$ is an *exact potential function* for a game G if for all $a \in A$, all i, and all $a_i, b_i \in A_i$:

$$\phi(b_i, a_{-i}) - \phi(a_i, a_{-i}) = c_i(b_i, a_{-i}) - c_i(a_i, a_{-i})$$

Abstracting Away...

What do we need to make the proof work?

Definition

A function $\phi : A \to \mathbb{R}_{\geq 0}$ is an *exact potential function* for a game G if for all $a \in A$, all i, and all $a_i, b_i \in A_i$:

$$\phi(b_i, a_{-i}) - \phi(a_i, a_{-i}) = c_i(b_i, a_{-i}) - c_i(a_i, a_{-i})$$

Definition

$\phi : A \to \mathbb{R}_{\geq 0}$ is an *ordinal potential function* for a game G if for all $a \in A$, all i, and all $a_i, b_i \in A_i$:

$$(c_i(b_i, a_{-i}) - c_i(a_i, a_{-i}) < 0) \Rightarrow (\phi(b_i, a_{-i}) - \phi(a_i, a_{-i}) < 0)$$

i.e. the change in utility is always equal *in sign* to the change in potential.

A Characterization

Theorem

Best response dynamics is guaranteed to converge in a game G if and only if the game has an ordinal potential function.

A Characterization

Theorem

Best response dynamics is guaranteed to converge in a game G if and only if the game has an ordinal potential function.

So our proof technique is without loss of generality!

Proof

1. We've already seen the forward direction (ordinal potential function $\Rightarrow$ BRD converges) several times now, so lets prove the reverse direction.

Proof

1. We've already seen the forward direction (ordinal potential function $\Rightarrow$ BRD converges) several times now, so lets prove the reverse direction.
2. Consider a graph $G = (V, E)$:
   2.1 Let each $a \in A$ be a vertex in the graph: i.e. $V = A$.
   2.2 For each pair of vertices $a, b \in V$, add a directed edge (a, b) if it is possible to get to get from b to a via a best response move – i.e. if there is some index i such that $b = (b_i, a_{-i})$, and $c_i(b_i, a_{-i}) < c_i(a)$.

Proof

1. We've already seen the forward direction (ordinal potential function $\Rightarrow$ BRD converges) several times now, so lets prove the reverse direction.
2. Consider a graph $G = (V, E)$:
   2.1 Let each $a \in A$ be a vertex in the graph: i.e. $V = A$.
   2.2 For each pair of vertices $a, b \in V$, add a directed edge (a, b) if it is possible to get to get from b to a via a best response move – i.e. if there is some index i such that $b = (b_i, a_{-i})$, and $c_i(b_i, a_{-i}) < c_i(a)$.
3. BRD can be viewed as traversing this graph: Start at an arbitrary vertex a, and then traverse arbitrary outgoing edges.

Proof

1. We've already seen the forward direction (ordinal potential function $\Rightarrow$ BRD converges) several times now, so lets prove the reverse direction.
2. Consider a graph $G = (V, E)$:
   2.1 Let each $a \in A$ be a vertex in the graph: i.e. $V = A$.
   2.2 For each pair of vertices $a, b \in V$, add a directed edge (a, b) if it is possible to get to get from b to a via a best response move – i.e. if there is some index i such that $b = (b_i, a_{-i})$, and $c_i(b_i, a_{-i}) < c_i(a)$.
3. BRD can be viewed as traversing this graph: Start at an arbitrary vertex a, and then traverse arbitrary outgoing edges.
4. Nash Equilibria are the sinks in this graph.

Proof

1. We've already seen the forward direction (ordinal potential function $\Rightarrow$ BRD converges) several times now, so lets prove the reverse direction.
2. Consider a graph $G = (V, E)$:
   2.1 Let each $a \in A$ be a vertex in the graph: i.e. $V = A$.
   2.2 For each pair of vertices $a, b \in V$, add a directed edge (a, b) if it is possible to get to get from b to a via a best response move – i.e. if there is some index i such that $b = (b_i, a_{-i})$, and $c_i(b_i, a_{-i}) < c_i(a)$.
3. BRD can be viewed as traversing this graph: Start at an arbitrary vertex a, and then traverse arbitrary outgoing edges.
4. Nash Equilibria are the sinks in this graph.
5. BRD converges = there are no cycles in this graph.

Proof

Proof

1. So suppose BRD converges (i.e. G is acyclic). We construct a potential function ϕ.

Proof

1. So suppose BRD converges (i.e. G is acyclic). We construct a potential function ϕ.
2. The graph is acyclic, so: from every state a there is some sink s that is reachable. (why?)

Proof

1. So suppose BRD converges (i.e. G is acyclic). We construct a potential function ϕ.
2. The graph is acyclic, so: from every state a there is some sink s that is reachable. (why?)
3. For each vertex a, define $\phi(a)$ to be the length of the *longest* finite path from a to any sink s.

Proof

1. So suppose BRD converges (i.e. G is acyclic). We construct a potential function ϕ.
2. The graph is acyclic, so: from every state a there is some sink s that is reachable. (why?)
3. For each vertex a, define $\phi(a)$ to be the length of the *longest* finite path from a to any sink s.
4. We need: for any edge $a \to b$, $\phi(b) < \phi(a)$.

Proof

1. So suppose BRD converges (i.e. G is acyclic). We construct a potential function ϕ.
2. The graph is acyclic, so: from every state a there is some sink s that is reachable. (why?)
3. For each vertex a, define $\phi(a)$ to be the length of the *longest* finite path from a to any sink s.
4. We need: for any edge $a \to b$, $\phi(b) < \phi(a)$.
5. Its true! $\phi(a) \geq \phi(b) + 1$. (why?)

Thanks!

See you next class — stay healthy, and wear a mask!