Multi-agent learning

Gradient ascent

Gerard Vreeswijk, Intelligent Software Systems, Computer Science Department, Faculty of Sciences, Utrecht University, The Netherlands.

Saturday 16th May, 2020

Gradient ascent: idea

Idea

Gradient ascent: idea

Idea

- Every opponent is identified with a (possibly mixed) strategy.

Gradient ascent: idea

Idea

- Every opponent is identified with a (possibly mixed) strategy.
- Players can observe the (possibly mixed) strategy of their opponent.

Gradient ascent: idea

Idea

- Every opponent is identified with a (possibly mixed) strategy.
- Players can observe the (possibly mixed) strategy of their opponent.
- After observation, every player changes its strategy a tiny bit in the right direction.

Gradient ascent: idea

Idea

- Every opponent is identified with a (possibly mixed) strategy.
- Players can observe the (possibly mixed) strategy of their opponent.
- After observation, every player changes its strategy a tiny bit in the right direction.
- Comparison with fictitious play.

Gradient ascent: idea

Idea

- Every opponent is identified with a (possibly mixed) strategy.
- Players can observe the (possibly mixed) strategy of their opponent.
- After observation, every player changes its strategy a tiny bit in the right direction.
- Comparison with fictitious play.
    - Like in fictitious play, opponents are modelled through a mixed strategy.

Gradient ascent: idea

Idea

- Every opponent is identified with a (possibly mixed) strategy.
- Players can observe the (possibly mixed) strategy of their opponent.
- After observation, every player changes its strategy a tiny bit in the right direction.
- Comparison with fictitious play.
  - Like in fictitious play, opponents are modelled through a mixed strategy.
  - In fictitious play, players learn projected opponent strategies, and play a best response to it.

Gradient ascent: idea

Idea

- Every opponent is identified with a (possibly mixed) strategy.
- Players can observe the (possibly mixed) strategy of their opponent.
- After observation, every player changes its strategy a tiny bit in the right direction.
- Comparison with fictitious play.
  - Like in fictitious play, opponents are modelled through a mixed strategy.
  - In fictitious play, players learn projected opponent strategies, and play a best response to it.
  - In gradient ascent, players do not project a mixed strategy, and do not play a best response.

Plan for today

Plan for today

1. Two-player, two-action, general sum games with real payoffs.

Plan for today

1. Two-player, two-action, general sum games with real payoffs.
2. Dynamics of (mixed) strategies in such games.

Plan for today

1. Two-player, two-action, general sum games with real payoffs.
2. Dynamics of (mixed) strategies in such games.

   Examples:

Plan for today

1. Two-player, two-action, general sum games with real payoffs.
2. Dynamics of (mixed) strategies in such games.

   Examples:

   (a) Coordination game

Plan for today

1. Two-player, two-action, general sum games with real payoffs.
2. Dynamics of (mixed) strategies in such games.

   Examples:

   (a) Coordination game

   (b) Prisoners' dilemma

Plan for today

1. Two-player, two-action, general sum games with real payoffs.
2. Dynamics of (mixed) strategies in such games.

   Examples:

   (a) Coordination game

   (b) Prisoners' dilemma

   (c) Stag hunt

Plan for today

1. Two-player, two-action, general sum games with real payoffs.
2. Dynamics of (mixed) strategies in such games.

   Examples:

   (a) Coordination game
   (b) Prisoners' dilemma
   (c) Stag hunt
   (d) Other examples

Plan for today

1. Two-player, two-action, general sum games with real payoffs.
2. Dynamics of (mixed) strategies in such games.

   Examples:

   (a) Coordination game
   (b) Prisoners' dilemma
   (c) Stag hunt
   (d) Other examples

3. **IGA**: Infinitesimal Gradient Ascent. Singh, Kearns and Mansour (2000).

Plan for today

1. Two-player, two-action, general sum games with real payoffs.
2. Dynamics of (mixed) strategies in such games.

   Examples:

   (a) Coordination game
   (b) Prisoners' dilemma
   (c) Stag hunt
   (d) Other examples

3. **IGA**: Infinitesimal **G**radient **A**scent. Singh, Kearns and Mansour (2000).

   - Convergence of IGA.

Plan for today

1. Two-player, two-action, general sum games with real payoffs.
2. Dynamics of (mixed) strategies in such games.

   Examples:

   (a) Coordination game
   (b) Prisoners' dilemma
   (c) Stag hunt
   (d) Other examples
3. **IGA**: Infinitesimal Gradient Ascent. Singh, Kearns and Mansour (2000).
   - Convergence of IGA.
4. **IGA-WoLF**: Win or Learn Fast. Bowling and Veloso (2001, 2002).

Plan for today

1. Two-player, two-action, general sum games with real payoffs.
2. Dynamics of (mixed) strategies in such games.

   Examples:

   (a) Coordination game
   (b) Prisoners' dilemma
   (c) Stag hunt
   (d) Other examples

3. **IGA**: Infinitesimal Gradient Ascent. Singh, Kearns and Mansour (2000).
   - Convergence of IGA.
4. **IGA-WoLF**: Win or Learn Fast. Bowling and Veloso (2001, 2002).
   - Convergence of IGA-WoLF

Plan for today

1. Two-player, two-action, general sum games with real payoffs.
2. Dynamics of (mixed) strategies in such games.

   Examples:

   (a) Coordination game
   (b) Prisoners' dilemma
   (c) Stag hunt
   (d) Other examples

3. **IGA**: Infinitesimal Gradient Ascent. Singh, Kearns and Mansour (2000).
   - Convergence of IGA.
4. **IGA-WoLF**: **Win or Learn Fast**. Bowling and Veloso (2001, 2002).
   - Convergence of IGA-WoLF + analysis of the proof of convergence.

Part 1:
Payoffs of general 2x2 games
in normal form

Two-player, two-action, general sum games

In its most general form, a two-player, two-action game in normal form with real-valued payoffs can be represented by

$$M = \begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} r_{11}, c_{11} & r_{12}, c_{12} \\ r_{21}, c_{21} & r_{22}, c_{22} \end{pmatrix} \end{array}$$

Two-player, two-action, general sum games

In its most general form, a two-player, two-action game in normal form with real-valued payoffs can be represented by

$$M = \begin{array}{cc} & \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{array}{c} \text{T} \\ \text{B} \end{array} & \begin{pmatrix} r_{11}, c_{11} & r_{12}, c_{12} \\ r_{21}, c_{21} & r_{22}, c_{22} \end{pmatrix} \end{array}$$

Row plays mixed $(\alpha, 1-\alpha)$. Column plays mixed $(\beta, 1-\beta)$.

Two-player, two-action, general sum games

In its most general form, a two-player, two-action game in normal form with real-valued payoffs can be represented by

$$M = \begin{array}{c c} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} r_{11}, c_{11} & r_{12}, c_{12} \\ r_{21}, c_{21} & r_{22}, c_{22} \end{pmatrix} \end{array}$$

Row plays mixed $(\alpha, 1-\alpha)$. Column plays mixed $(\beta, 1-\beta)$.

Expected payoff:

Two-player, two-action, general sum games

In its most general form, a two-player, two-action game in normal form with real-valued payoffs can be represented by

$$M = \begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} r_{11}, c_{11} & r_{12}, c_{12} \\ r_{21}, c_{21} & r_{22}, c_{22} \end{pmatrix} \end{array}$$

Row plays mixed $(\alpha, 1-\alpha)$. Column plays mixed $(\beta, 1-\beta)$.

Expected payoff:

$$u_1(\alpha, \beta)$$

Two-player, two-action, general sum games

In its most general form, a two-player, two-action game in normal form with real-valued payoffs can be represented by

$$M = \begin{array}{c c} & \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{array}{c} \text{T} \\ \text{B} \end{array} & \begin{pmatrix} r_{11}, c_{11} & r_{12}, c_{12} \\ r_{21}, c_{21} & r_{22}, c_{22} \end{pmatrix} \end{array}$$

Row plays mixed $(\alpha, 1-\alpha)$. Column plays mixed $(\beta, 1-\beta)$.

Expected payoff:

$$u_1(\alpha, \beta) = \alpha[\beta r_{11} + (1-\beta) r_{12}] + (1-\alpha)[\beta r_{21} + (1-\beta) r_{22}]$$

Two-player, two-action, general sum games

In its most general form, a two-player, two-action game in normal form with real-valued payoffs can be represented by

$$M = \begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} r_{11}, c_{11} & r_{12}, c_{12} \\ r_{21}, c_{21} & r_{22}, c_{22} \end{pmatrix} \end{array}$$

Row plays mixed $(\alpha, 1-\alpha)$. Column plays mixed $(\beta, 1-\beta)$.

Expected payoff:

$$\begin{aligned} u_1(\alpha, \beta) &= \alpha[\beta r_{11} + (1-\beta) r_{12}] + (1-\alpha)[\beta r_{21} + (1-\beta) r_{22}] \\ &= u\,\alpha\beta + \alpha(r_{12} - r_{22}) + \beta(r_{21} - r_{22}) + r_{22}. \end{aligned}$$

Two-player, two-action, general sum games

In its most general form, a two-player, two-action game in normal form with real-valued payoffs can be represented by

$$M = \begin{array}{c c} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} r_{11}, c_{11} & r_{12}, c_{12} \\ r_{21}, c_{21} & r_{22}, c_{22} \end{pmatrix} \end{array}$$

Row plays mixed $(\alpha, 1-\alpha)$. Column plays mixed $(\beta, 1-\beta)$.

Expected payoff:

$$\begin{aligned} u_1(\alpha, \beta) &= \alpha[\beta r_{11} + (1-\beta) r_{12}] + (1-\alpha)[\beta r_{21} + (1-\beta) r_{22}] \\ &= u\ \alpha\beta + \alpha(r_{12} - r_{22}) + \beta(r_{21} - r_{22}) + r_{22}. \\ u_2(\alpha, \beta) & \end{aligned}$$

Two-player, two-action, general sum games

In its most general form, a two-player, two-action game in normal form with real-valued payoffs can be represented by

$$M = \begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} r_{11}, c_{11} & r_{12}, c_{12} \\ r_{21}, c_{21} & r_{22}, c_{22} \end{pmatrix} \end{array}$$

Row plays mixed $(\alpha, 1-\alpha)$. Column plays mixed $(\beta, 1-\beta)$.

Expected payoff:

$$\begin{aligned} u_1(\alpha, \beta) &= \alpha[\beta r_{11} + (1-\beta) r_{12}] + (1-\alpha)[\beta r_{21} + (1-\beta) r_{22}] \\ &= u\, \alpha\beta + \alpha(r_{12} - r_{22}) + \beta(r_{21} - r_{22}) + r_{22}. \\ u_2(\alpha, \beta) &= \beta[\alpha c_{11} + (1-\alpha) c_{21}] + (1-\beta)[\alpha c_{12} + (1-\alpha) c_{22}] \\ &= u' \alpha\beta + \alpha(c_{21} - c_{22}) + \beta(c_{12} - c_{22}) + c_{22}. \end{aligned}$$

Two-player, two-action, general sum games

In its most general form, a two-player, two-action game in normal form with real-valued payoffs can be represented by

$$M = \begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} r_{11}, c_{11} & r_{12}, c_{12} \\ r_{21}, c_{21} & r_{22}, c_{22} \end{pmatrix} \end{array}$$

Row plays mixed $(\alpha, 1-\alpha)$. Column plays mixed $(\beta, 1-\beta)$.

Expected payoff:

$$\begin{aligned} u_1(\alpha,\beta) &= \alpha[\beta r_{11} + (1-\beta) r_{12}] + (1-\alpha)[\beta r_{21} + (1-\beta) r_{22}] \\ &= u\,\alpha\beta + \alpha(r_{12} - r_{22}) + \beta(r_{21} - r_{22}) + r_{22}. \\ u_2(\alpha,\beta) &= \beta[\alpha c_{11} + (1-\alpha) c_{21}] + (1-\beta)[\alpha c_{12} + (1-\alpha) c_{22}] \\ &= u'\alpha\beta + \alpha(c_{21} - c_{22}) + \beta(c_{12} - c_{22}) + c_{22}. \end{aligned}$$

where $u = (r_{11} - r_{12}) - (r_{21} - r_{22})$ and $u' = (c_{11} - c_{21}) - (c_{12} - c_{22})$.

Gradient of expected payoff

Gradient:

$$\frac{\partial u_1(\alpha,\beta)}{\partial\alpha} = \beta u \ + (r_{12} - r_{22})$$
$$\frac{\partial u_2(\alpha,\beta)}{\partial\beta} = \alpha u' + (c_{21} - c_{22})$$

Gradient of expected payoff

Gradient:

$$\frac{\partial u_1(\alpha,\beta)}{\partial \alpha} = \beta u \; + (r_{12} - r_{22})$$
$$\frac{\partial u_2(\alpha,\beta)}{\partial \beta} = \alpha u' + (c_{21} - c_{22})$$

As an affine map:

$$\begin{bmatrix} \partial u_1 / \partial \alpha \\ \partial u_2 / \partial \beta \end{bmatrix} = \begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} = U \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + C$$

Gradient of expected payoff

Gradient:

$$\frac{\partial u_1(\alpha,\beta)}{\partial \alpha} = \beta u + (r_{12} - r_{22})$$
$$\frac{\partial u_2(\alpha,\beta)}{\partial \beta} = \alpha u' + (c_{21} - c_{22})$$

As an affine map:

$$\begin{bmatrix} \partial u_1/\partial\alpha \\ \partial u_2/\partial\beta \end{bmatrix} = \begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} = U \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + C$$

Stationary point:

$$(\alpha^*, \beta^*) = \left(\frac{c_{22} - c_{21}}{u'}, \frac{r_{22} - r_{12}}{u}\right)$$

Gradient of expected payoff

Gradient:

$$\frac{\partial u_1(\alpha,\beta)}{\partial \alpha} = \beta u \ + (r_{12} - r_{22})$$

$$\frac{\partial u_2(\alpha,\beta)}{\partial \beta} = \alpha u' + (c_{21} - c_{22})$$

As an affine map:

$$\begin{bmatrix} \partial u_1/\partial \alpha \\ \partial u_2/\partial \beta \end{bmatrix} = \begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} = U \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + C$$

Stationary point:

$$(\alpha^*, \beta^*) = \left(\frac{c_{22} - c_{21}}{u'}, \frac{r_{22} - r_{12}}{u}\right)$$

Remarks:

Gradient of expected payoff

Gradient:

$$\frac{\partial u_1(\alpha,\beta)}{\partial \alpha} = \beta u \; + (r_{12} - r_{22})$$
$$\frac{\partial u_2(\alpha,\beta)}{\partial \beta} = \alpha u' + (c_{21} - c_{22})$$

As an affine map:

$$\begin{bmatrix} \partial u_1/\partial\alpha \\ \partial u_2/\partial\beta \end{bmatrix} = \begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} = U \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + C$$

Stationary point:

$$(\alpha^*, \beta^*) = \left(\frac{c_{22} - c_{21}}{u'}, \frac{r_{22} - r_{12}}{u}\right)$$

Remarks:

- There is at most one stationary point.

Gradient of expected payoff

Gradient:

$$\frac{\partial u_1(\alpha,\beta)}{\partial \alpha} = \beta u \ + (r_{12} - r_{22})$$
$$\frac{\partial u_2(\alpha,\beta)}{\partial \beta} = \alpha u' + (c_{21} - c_{22})$$

As an affine map:

$$\begin{bmatrix} \partial u_1/\partial\alpha \\ \partial u_2/\partial\beta \end{bmatrix} = \begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} = U \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + C$$

Stationary point:

$$(\alpha^*, \beta^*) = \left(\frac{c_{22} - c_{21}}{u'}, \frac{r_{22} - r_{12}}{u}\right)$$

Remarks:

- There is at most one stationary point.
- If a stationary point exists, it may lie outside $[0,1]^2$.

Gradient of expected payoff

Gradient:

$$\frac{\partial u_1(\alpha,\beta)}{\partial \alpha} = \beta u \; + (r_{12} - r_{22})$$
$$\frac{\partial u_2(\alpha,\beta)}{\partial \beta} = \alpha u' + (c_{21} - c_{22})$$

As an affine map:

$$\begin{bmatrix} \partial u_1/\partial\alpha \\ \partial u_2/\partial\beta \end{bmatrix} = \begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} = U \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + C$$

Stationary point:

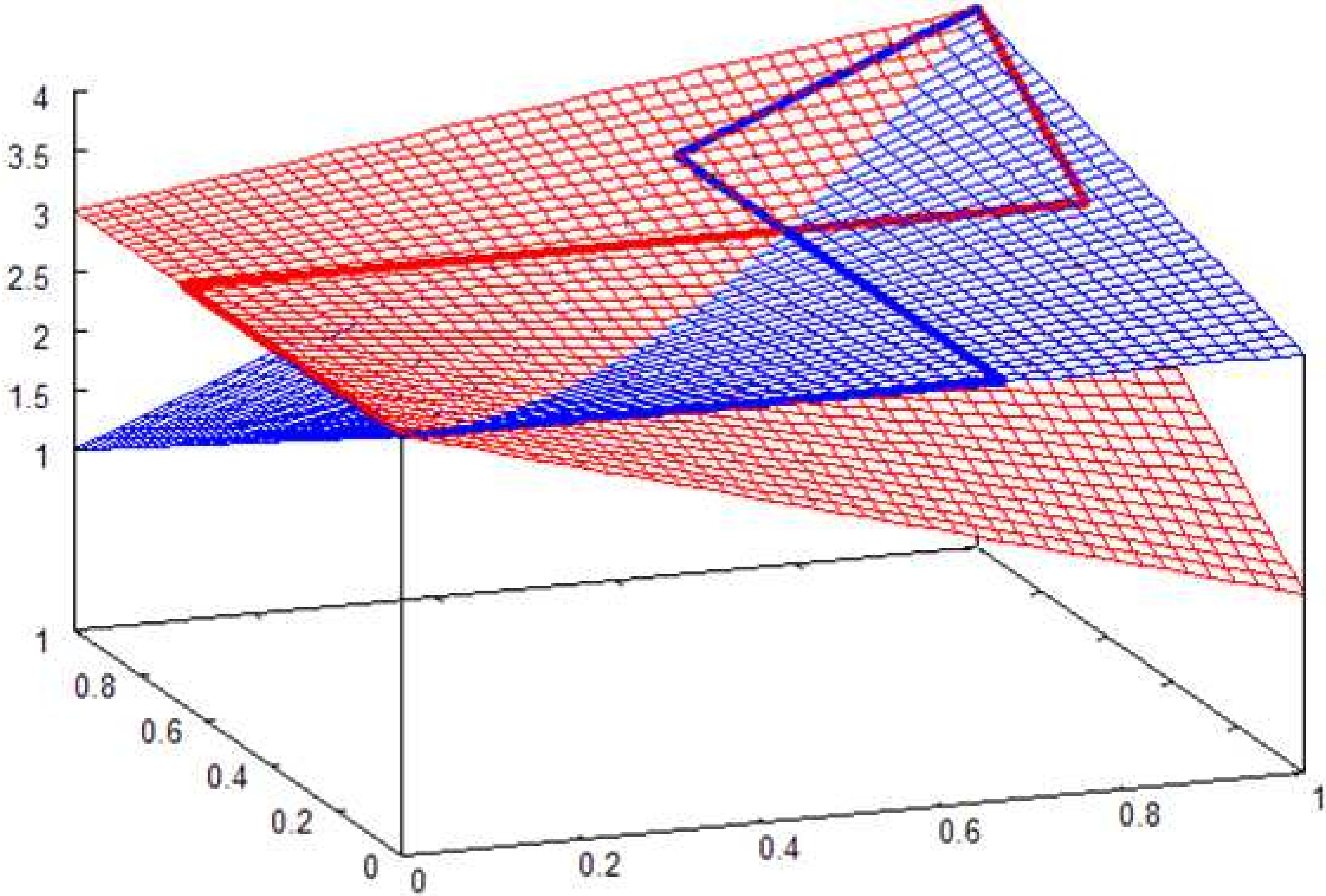
$$(\alpha^*, \beta^*) = (\frac{c_{22} - c_{21}}{u'}, \frac{r_{22} - r_{12}}{u})$$

Remarks:

- There is at most one stationary point.
- If a stationary point exists, it may lie outside $[0,1]^2$.
- If there is a stationary point inside $[0,1]^2$, it is a weak (i.e., non-strict) Nash equilibrium.

Example: payoffs in Stag Hunt (r=4, t=3, s=1, p=3)

Example: payoffs in Stag Hunt (r=4, t=3, s=1, p=3)



Player 1 may only move "back – front"; Player 2 may only move "left – right".

Part 2: IGA

Gradient ascent

Affine differential map:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \begin{bmatrix} \partial u_1 / \partial \alpha \\ \partial u_2 / \partial \beta \end{bmatrix}_t$$

Gradient ascent

Affine differential map:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \begin{bmatrix} \partial u_1 / \partial \alpha \\ \partial u_2 / \partial \beta \end{bmatrix}_t$$

- Because $\alpha, \beta \in [0, 1]$, the dynamics must be confined to $[0, 1]^2$.

Gradient ascent

Affine differential map:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \begin{bmatrix} \partial u_1 / \partial \alpha \\ \partial u_2 / \partial \beta \end{bmatrix}_t$$

- Because $\alpha, \beta \in [0, 1]$, the dynamics must be confined to $[0, 1]^2$.
- Suppose the state (α, β) is on the boundary of the probability space $[0, 1]^2$, and the gradient vector points outwards.

  Intuition: one of the players has an incentive to improve, but cannot improve further.

Gradient ascent

Affine differential map:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \begin{bmatrix} \partial u_1 / \partial \alpha \\ \partial u_2 / \partial \beta \end{bmatrix}_t$$

- Because $\alpha, \beta \in [0, 1]$, the dynamics must be confined to $[0, 1]^2$.
- Suppose the state (α, β) is on the boundary of the probability space $[0, 1]^2$, and the gradient vector points outwards.

  Intuition: one of the players has an incentive to improve, but cannot improve further.
- To maintain dynamics within $[0, 1]^2$, the gradient is projected back on to $[0, 1]^2$.

  Intuition: if one of the players has an incentive to improve, but *cannot* improve, then he *will not* improve.

Gradient ascent

Affine differential map:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \begin{bmatrix} \partial u_1 / \partial \alpha \\ \partial u_2 / \partial \beta \end{bmatrix}_t$$

- Because $\alpha, \beta \in [0, 1]$, the dynamics must be confined to $[0, 1]^2$.
- Suppose the state (α, β) is on the boundary of the probability space $[0, 1]^2$, and the gradient vector points outwards.

  Intuition: one of the players has an incentive to improve, but cannot improve further.
- To maintain dynamics within $[0, 1]^2$, the gradient is projected back on to $[0, 1]^2$.

  Intuition: if one of the players has an incentive to improve, but *cannot* improve, then he *will not* improve.
- If nonzero, the projected gradient is parallel to the (closest) boundary of $[0, 1]^2$.

Infinitesimal Gradient Ascent : IGA (Singh *et al.*, 2000)

Affine differential map:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t} + \eta \left(\begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} \right)$$

Infinitesimal Gradient Ascent : IGA (Singh *et al.*, 2000)

Affine differential map:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \left(\begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} \right)$$

Theorem (Singh, Kearns and Mansour, 2000) *If players follow IGA, where* $\eta \to 0$*, their average payoffs will converge to the (expected) payoffs of a NE. If their strategies converge, they will converge to that same NE.*

Infinitesimal Gradient Ascent : IGA (Singh *et al.*, 2000)

Affine differential map:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t} + \eta \left(\begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} \right)$$

Theorem (Singh, Kearns and Mansour, 2000) *If players follow IGA, where* $\eta \to 0$*, their average payoffs will converge to the (expected) payoffs of a NE. If their strategies converge, they will converge to that same NE.*
The proof is based on a qualitative result in the the theory of differential equations, which says that the behaviour of an affine differential map is determined by the multiplicative matrix U:

Infinitesimal Gradient Ascent : IGA (Singh *et al.*, 2000)

Affine differential map:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \left(\begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} \right)$$

Theorem (Singh, Kearns and Mansour, 2000) *If players follow IGA, where* $\eta \to 0$*, their average payoffs will converge to the (expected) payoffs of a NE. If their strategies converge, they will converge to that same NE.*
The proof is based on a qualitative result in the the theory of differential equations, which says that the behaviour of an affine differential map is determined by the multiplicative matrix U:

1. If U is invertible, and its eigenvalue λ (solution of $Ux = \lambda x \Leftrightarrow$ solution of $\mathrm{Det}[U - \lambda I] = 0$) is real, $\exists$ stationary point

Infinitesimal Gradient Ascent : IGA (Singh *et al.*, 2000)

Affine differential map:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \left(\begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} \right)$$

Theorem (Singh, Kearns and Mansour, 2000) *If players follow IGA, where* $\eta \to 0$*, their average payoffs will converge to the (expected) payoffs of a NE. If their strategies converge, they will converge to that same NE.*
The proof is based on a qualitative result in the the theory of differential equations, which says that the behaviour of an affine differential map is determined by the multiplicative matrix U:

1. If U is invertible, and its eigenvalue λ (solution of $Ux = \lambda x \Leftrightarrow$ solution of $\mathrm{Det}[U - \lambda I] = 0$) is real, $\exists$ stationary point, w.i. is a saddle point.

Infinitesimal Gradient Ascent : IGA (Singh *et al.*, 2000)

Affine differential map:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \left(\begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} \right)$$

Theorem (Singh, Kearns and Mansour, 2000) *If players follow IGA, where* $\eta \to 0$, *their average payoffs will converge to the (expected) payoffs of a NE. If their strategies converge, they will converge to that same NE.*
The proof is based on a qualitative result in the the theory of differential equations, which says that the behaviour of an affine differential map is determined by the multiplicative matrix U:

1. If U is invertible, and its eigenvalue λ (solution of $Ux = \lambda x \Leftrightarrow$ solution of $\mathrm{Det}[U - \lambda I] = 0$) is real, $\exists$ stationary point, w.i. is a saddle point.
2. If U is invertible, and its eigenvalue λ is imaginary, there is a stationary point

Infinitesimal Gradient Ascent : IGA (Singh *et al.*, 2000)

Affine differential map:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t} + \eta \left(\begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} \right)$$

Theorem (Singh, Kearns and Mansour, 2000) *If players follow IGA, where* $\eta \to 0$, *their average payoffs will converge to the (expected) payoffs of a NE. If their strategies converge, they will converge to that same NE.*
The proof is based on a qualitative result in the the theory of differential equations, which says that the behaviour of an affine differential map is determined by the multiplicative matrix U:

1. If U is invertible, and its eigenvalue λ (solution of $Ux = \lambda x \Leftrightarrow$ solution of $\mathrm{Det}[U - \lambda I] = 0$) is real, $\exists$ stationary point, w.i. is a saddle point.

2. If U is invertible, and its eigenvalue λ is imaginary, there is a stationary point, which, in particular, is a centric point.

Infinitesimal Gradient Ascent : IGA (Singh *et al.*, 2000)

Affine differential map:

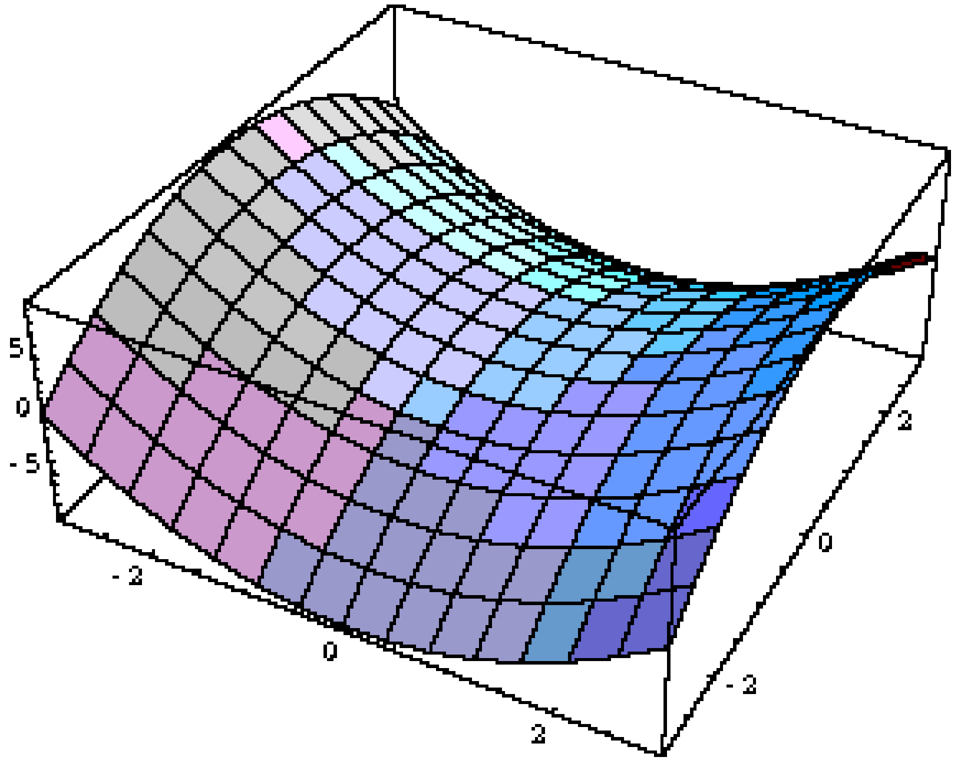
$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \left(\begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} \right)$$

Theorem (Singh, Kearns and Mansour, 2000) *If players follow IGA, where* $\eta \to 0$*, their average payoffs will converge to the (expected) payoffs of a NE. If their strategies converge, they will converge to that same NE.*

The proof is based on a qualitative result in the the theory of differential equations, which says that the behaviour of an affine differential map is determined by the multiplicative matrix U:

1. If U is invertible, and its eigenvalue λ (solution of $Ux = \lambda x \Leftrightarrow$ solution of $\mathrm{Det}[U - \lambda I] = 0$) is real, $\exists$ stationary point, w.i. is a saddle point.
2. If U is invertible, and its eigenvalue λ is imaginary, there is a stationary point, which, in particular, is a centric point.
3. If U is not invertible (iff $u = 0$ or $u' = 0$), there is no stationary point.

Saddle point



Gradient ascent: Coordination game

Gradient ascent: Coordination game

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{array}{c} \text{T} \\ \text{B} \end{array} & \begin{pmatrix} 1,1 & 0,0 \\ 0,0 & 1,1 \end{pmatrix} \end{array}$$

Gradient ascent: Coordination game

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 1,1 & 0,0 \\ 0,0 & 1,1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} 2\cdot \beta - 1 \\ 2\cdot \alpha - 1 \end{bmatrix}$$

Gradient ascent: Coordination game

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 1,1 & 0,0 \\ 0,0 & 1,1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} 2\cdot \beta - 1 \\ 2\cdot \alpha - 1 \end{bmatrix}$$

- Stationary at $(1/2, 1/2)$.

Gradient ascent: Coordination game

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{array}{c} \text{T} \\ \text{B} \end{array} & \begin{pmatrix} 1,1 & 0,0 \\ 0,0 & 1,1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} 2 \cdot \beta - 1 \\ 2 \cdot \alpha - 1 \end{bmatrix}$$

- Stationary at $(1/2, 1/2)$.
- Matrix

$$U = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$$

has real eigenvalues: $\lambda^2 - 4 = 0$.

Gradient ascent: Coordination game

- Symmetric, but not zero sum:

$$\begin{array}{c} \\ \mathrm{T} \\ \mathrm{B} \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{pmatrix} 1,1 & 0,0 \\ 0,0 & 1,1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} 2\cdot \beta - 1 \\ 2\cdot \alpha - 1 \end{bmatrix}$$

- Stationary at $(1/2, 1/2)$.
- Matrix

$$U = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$$

has real eigenvalues: $\lambda^2 - 4 = 0$.
Saddle point inside $[0,1]^2$.

Gradient ascent: Coordination game

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 1,1 & 0,0 \\ 0,0 & 1,1 \end{pmatrix} \end{array}$$

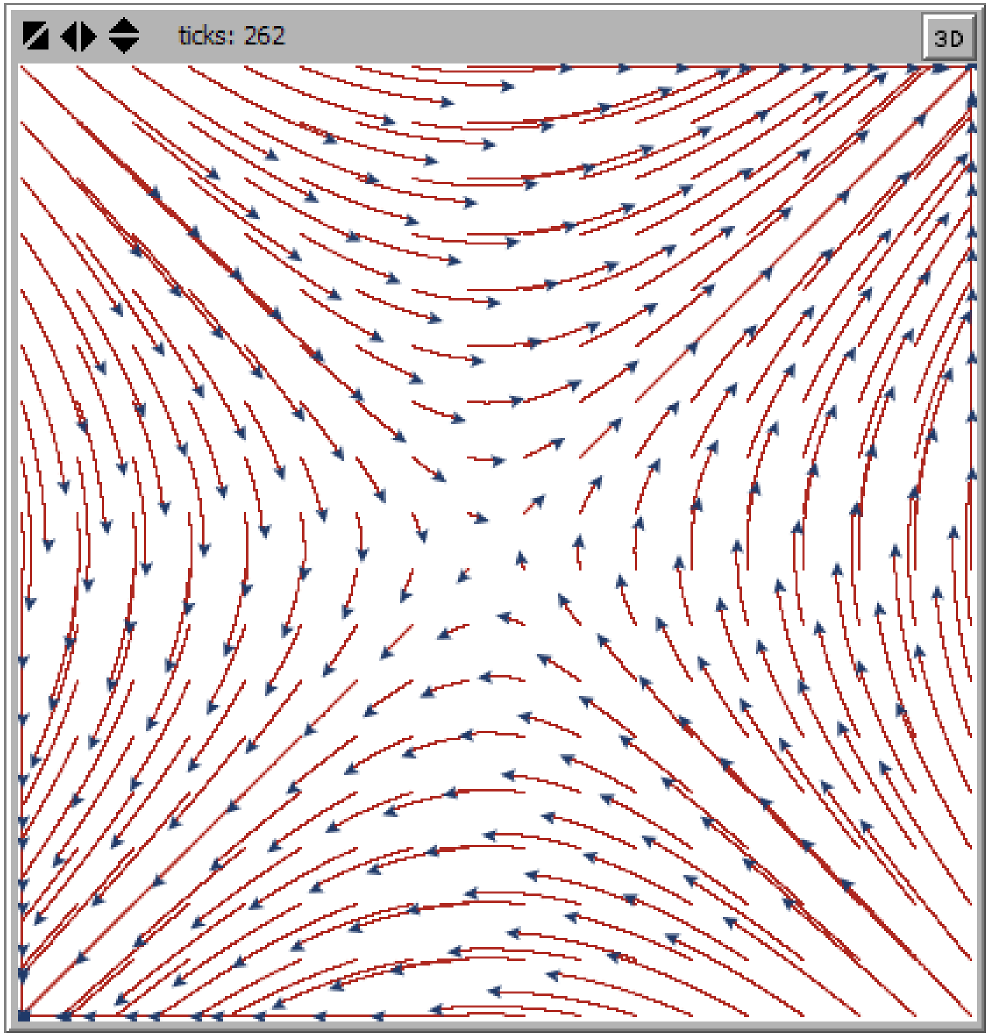
- Gradient:

$$\begin{bmatrix} 2\cdot \beta - 1 \\ 2\cdot \alpha - 1 \end{bmatrix}$$

- Stationary at $(1/2, 1/2)$.
- Matrix

$$U = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$$

has real eigenvalues: $\lambda^2 - 4 = 0$.
Saddle point inside $[0,1]^2$.



Gradient ascent: Prisoners' Dilemma

Gradient ascent: Prisoners' Dilemma

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 3,3 & 0,5 \\ 5,0 & 1,1 \end{pmatrix} \end{array}$$

Gradient ascent: Prisoners' Dilemma

- Symmetric, but not zero sum:

$$\begin{array}{c} \\ T \\ B \end{array} \begin{array}{c} \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{pmatrix} 3,3 & 0,5 \\ 5,0 & 1,1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} -1\cdot \beta - 1 \\ -1\cdot \alpha - 1 \end{bmatrix}$$

Gradient ascent: Prisoners' Dilemma

- Symmetric, but not zero sum:

$$\begin{array}{c} \\ T \\ B \end{array} \begin{array}{c} \begin{array}{cc} \quad L & \quad R \end{array} \\ \begin{pmatrix} 3,3 & 0,5 \\ 5,0 & 1,1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} -1 \cdot \beta - 1 \\ -1 \cdot \alpha - 1 \end{bmatrix}$$

- Stationary at $(-1, -1)$.

Gradient ascent: Prisoners' Dilemma

- Symmetric, but not zero sum:

$$\begin{array}{c c} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} T \\ B \end{array} & \begin{pmatrix} 3,3 & 0,5 \\ 5,0 & 1,1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} -1 \cdot \beta - 1 \\ -1 \cdot \alpha - 1 \end{bmatrix}$$

- Stationary at $(-1, -1)$.
- Matrix

$$U = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

has real eigenvalues: $\lambda^2 - 1 = 0$.

Gradient ascent: Prisoners' Dilemma

- Symmetric, but not zero sum:

$$
\begin{array}{c} \\ T \\ B \end{array}
\begin{array}{c}
\begin{array}{cc} \;\; L & \;\;\; R \end{array} \\
\begin{pmatrix} 3,3 & 0,5 \\ 5,0 & 1,1 \end{pmatrix}
\end{array}
$$

- Gradient:

$$
\begin{bmatrix} -1 \cdot \beta - 1 \\ -1 \cdot \alpha - 1 \end{bmatrix}
$$

- Stationary at $(-1, -1)$.
- Matrix

$$
U = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}
$$

has real eigenvalues: $\lambda^2 - 1 = 0$.
Saddle point outside $[0, 1]^2$.

Gradient ascent: Prisoners' Dilemma

- Symmetric, but not zero sum:

$$\begin{array}{c} \\ \mathrm{T} \\ \mathrm{B} \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \left(\begin{array}{cc} 3,3 & 0,5 \\ 5,0 & 1,1 \end{array} \right) \end{array}$$

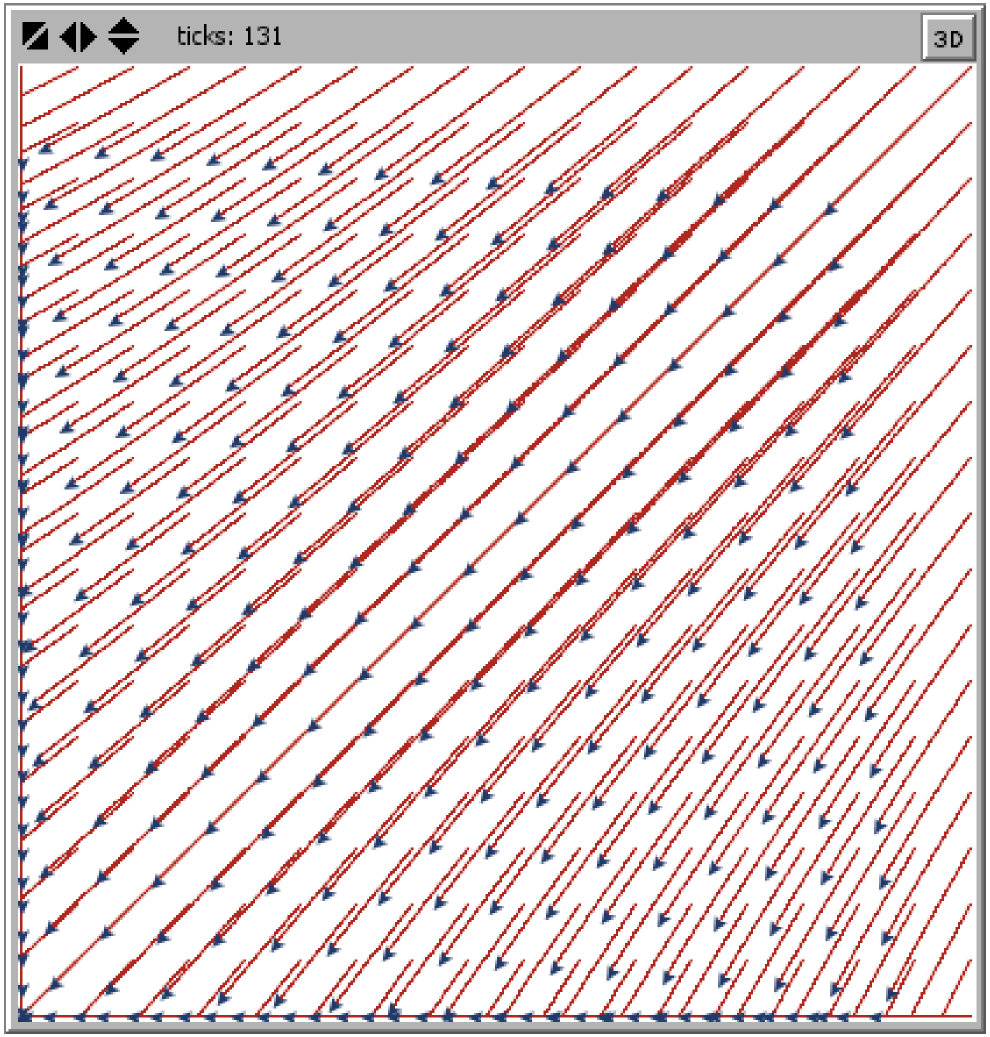
- Gradient:

$$\left[\begin{array}{c} -1\cdot\beta - 1 \\ -1\cdot\alpha - 1 \end{array} \right]$$

- Stationary at $(-1, -1)$.
- Matrix

$$U = \left[\begin{array}{rr} 0 & -1 \\ -1 & 0 \end{array} \right]$$

has real eigenvalues: $\lambda^2 - 1 = 0$.
Saddle point outside $[0, 1]^2$.



Gradient ascent: Stag hunt

Gradient ascent: Stag hunt

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 5,5 & 0,3 \\ 3,0 & 2,2 \end{pmatrix} \end{array}$$

Gradient ascent: Stag hunt

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 5,5 & 0,3 \\ 3,0 & 2,2 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} 4 \cdot \beta - 2 \\ 4 \cdot \alpha - 2 \end{bmatrix}$$

Gradient ascent: Stag hunt

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 5,5 & 0,3 \\ 3,0 & 2,2 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} 4\cdot \beta - 2 \\ 4\cdot \alpha - 2 \end{bmatrix}$$

- Stationary at $(1/2, 1/2)$.

Gradient ascent: Stag hunt

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \left(\begin{array}{cc} 5,5 & 0,3 \\ 3,0 & 2,2 \end{array} \right) \end{array}$$

- Gradient:

$$\left[\begin{array}{c} 4 \cdot \beta - 2 \\ 4 \cdot \alpha - 2 \end{array} \right]$$

- Stationary at $(1/2, 1/2)$.
- Matrix

$$U = \left[\begin{array}{cc} 0 & 4 \\ 4 & 0 \end{array} \right]$$

has real eigenvalues:
$\lambda^2 - 16 = 0$.

Gradient ascent: Stag hunt

- Symmetric, but not zero sum:

$$\begin{array}{c} \\ \mathrm{T} \\ \mathrm{B} \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{pmatrix} 5,5 & 0,3 \\ 3,0 & 2,2 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} 4 \cdot \beta - 2 \\ 4 \cdot \alpha - 2 \end{bmatrix}$$

- Stationary at $(1/2, 1/2)$.

- Matrix

$$U = \begin{bmatrix} 0 & 4 \\ 4 & 0 \end{bmatrix}$$

has real eigenvalues: $\lambda^2 - 16 = 0$. Saddle point inside $[0, 1]^2$.

Gradient ascent: Stag hunt

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} T \\ B \end{array} & \begin{pmatrix} 5,5 & 0,3 \\ 3,0 & 2,2 \end{pmatrix} \end{array}$$

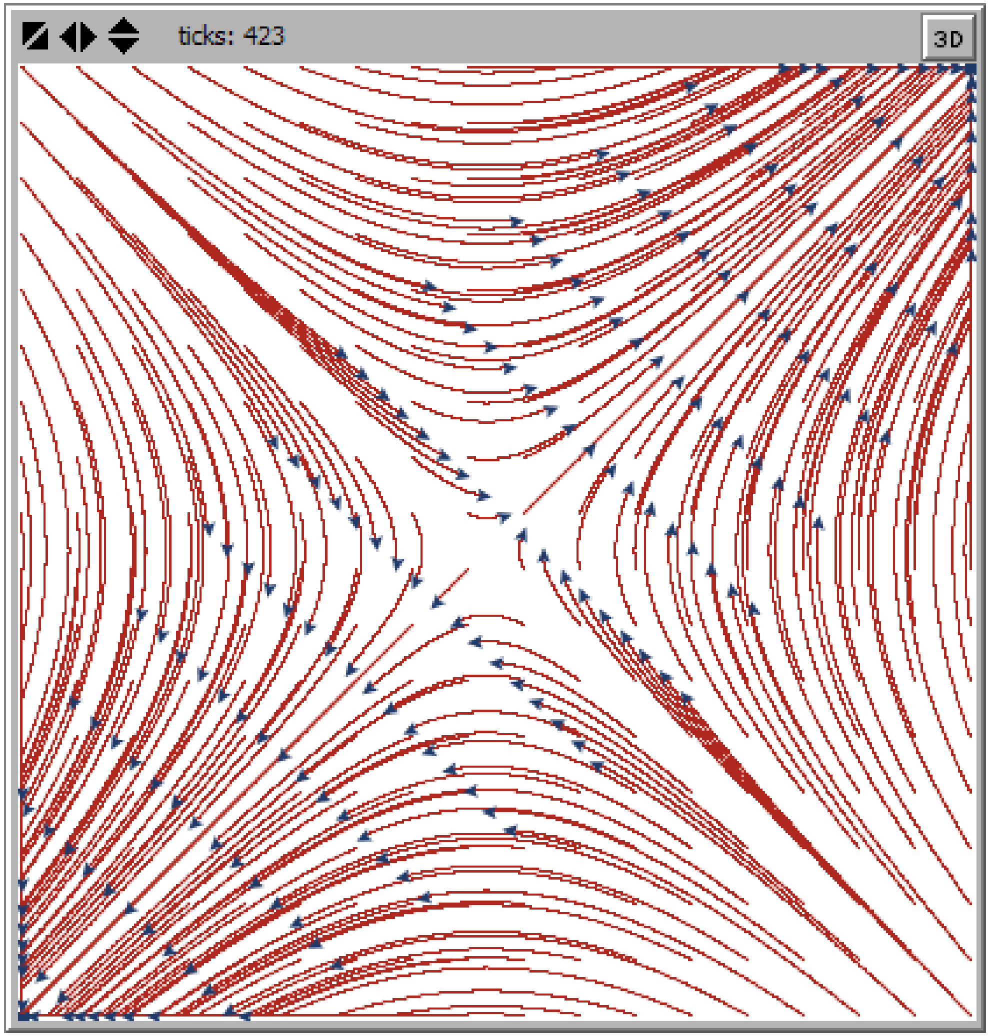
- Gradient:

$$\begin{bmatrix} 4\cdot \beta - 2 \\ 4\cdot \alpha - 2 \end{bmatrix}$$

- Stationary at $(1/2, 1/2)$.
- Matrix

$$U = \begin{bmatrix} 0 & 4 \\ 4 & 0 \end{bmatrix}$$

has real eigenvalues: $\lambda^2 - 16 = 0$. Saddle point inside $[0,1]^2$.



Gradient ascent: Game of Chicken

Gradient ascent: Game of Chicken

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 0,0 & -1,1 \\ 1,-1 & -3,-3 \end{pmatrix} \end{array}$$

Gradient ascent: Game of Chicken

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{array}{c} \text{T} \\ \text{B} \end{array} & \begin{pmatrix} 0,0 & -1,1 \\ 1,-1 & -3,-3 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} -3\cdot \beta + 2 \\ -3\cdot \alpha + 2 \end{bmatrix}$$

Gradient ascent: Game of Chicken

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 0,0 & -1,1 \\ 1,-1 & -3,-3 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} -3\cdot\beta+2 \\ -3\cdot\alpha+2 \end{bmatrix}$$

- Stationary at $(2/3, 2/3)$.

Gradient ascent: Game of Chicken

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 0,0 & -1,1 \\ 1,-1 & -3,-3 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} -3\cdot\beta + 2 \\ -3\cdot\alpha + 2 \end{bmatrix}$$

- Stationary at $(2/3, 2/3)$.
- Matrix

$$U = \begin{bmatrix} 0 & -3 \\ -3 & 0 \end{bmatrix}$$

has real eigenvalues: $\lambda^2 - 9 = 0$.

Gradient ascent: Game of Chicken

- Symmetric, but not zero sum:

$$\begin{array}{c c} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 0,0 & -1,1 \\ 1,-1 & -3,-3 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} -3\cdot\beta + 2 \\ -3\cdot\alpha + 2 \end{bmatrix}$$

- Stationary at $(2/3, 2/3)$.
- Matrix

$$U = \begin{bmatrix} 0 & -3 \\ -3 & 0 \end{bmatrix}$$

has real eigenvalues: $\lambda^2 - 9 = 0$.
Saddle point inside $[0,1]^2$.

Gradient ascent: Game of Chicken

- Symmetric, but not zero sum:

$$\begin{array}{c c} & \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{array}{c} \text{T} \\ \text{B} \end{array} & \begin{pmatrix} 0,0 & -1,1 \\ 1,-1 & -3,-3 \end{pmatrix} \end{array}$$

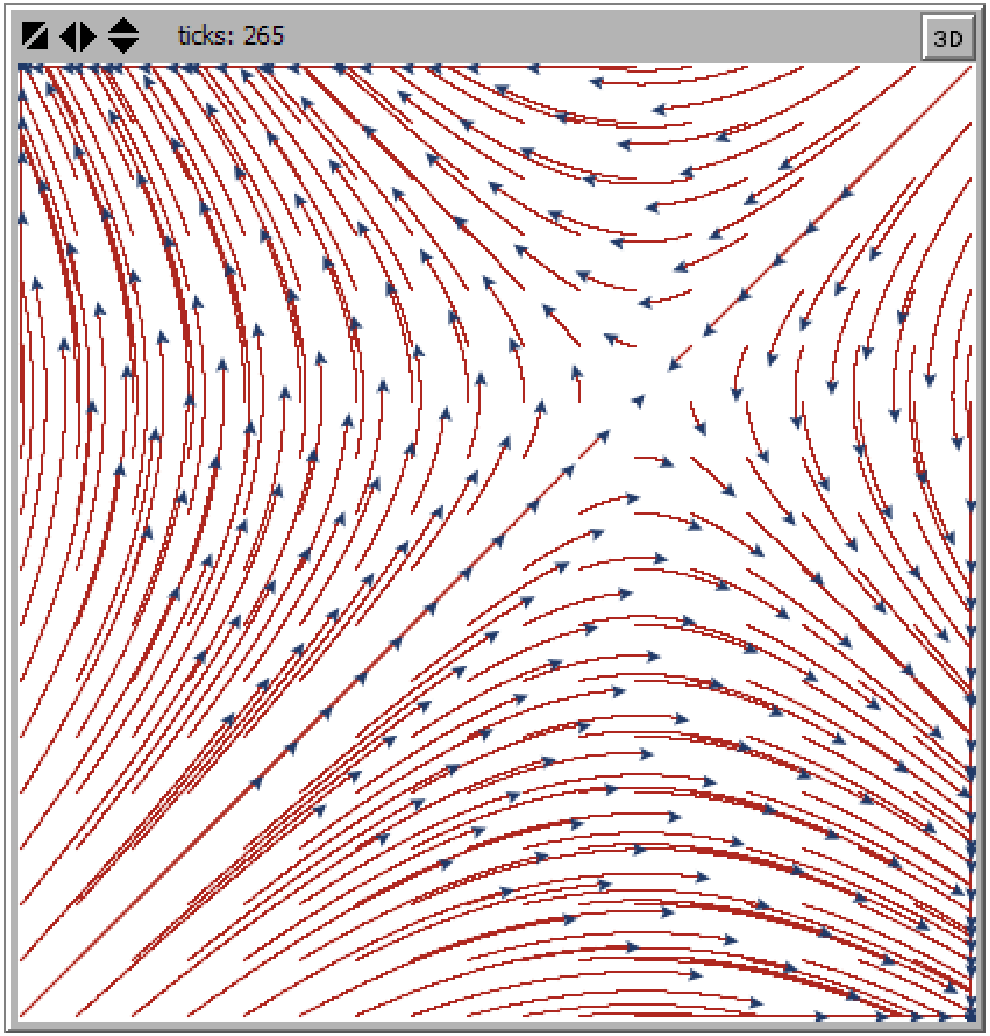
- Gradient:

$$\begin{bmatrix} -3 \cdot \beta + 2 \\ -3 \cdot \alpha + 2 \end{bmatrix}$$

- Stationary at $(2/3, 2/3)$.
- Matrix

$$U = \begin{bmatrix} 0 & -3 \\ -3 & 0 \end{bmatrix}$$

has real eigenvalues: $\lambda^2 - 9 = 0$.
Saddle point inside $[0,1]^2$.



Gradient ascent: Battle of the Sexes

Gradient ascent: Battle of the Sexes

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{array}{c} \text{T} \\ \text{B} \end{array} & \begin{pmatrix} 0,0 & 2,3 \\ 3,2 & 1,1 \end{pmatrix} \end{array}$$

Gradient ascent: Battle of the Sexes

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{array}{c} \text{T} \\ \text{B} \end{array} & \begin{pmatrix} 0,0 & 2,3 \\ 3,2 & 1,1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} -4 \cdot \beta + 1 \\ -4 \cdot \alpha + 1 \end{bmatrix}$$

Gradient ascent: Battle of the Sexes

- Symmetric, but not zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 0,0 & 2,3 \\ 3,2 & 1,1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} -4\cdot \beta + 1 \\ -4\cdot \alpha + 1 \end{bmatrix}$$

- Stationary at $(1/4, 1/4)$.

Gradient ascent: Battle of the Sexes

- Symmetric, but not zero sum:

$$\begin{array}{c} \\ \mathrm{T} \\ \mathrm{B} \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{pmatrix} 0,0 & 2,3 \\ 3,2 & 1,1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} -4\cdot\beta + 1 \\ -4\cdot\alpha + 1 \end{bmatrix}$$

- Stationary at $(1/4, 1/4)$.
- Matrix

$$U = \begin{bmatrix} 0 & -4 \\ -4 & 0 \end{bmatrix}$$

has real eigenvalues:
$\lambda^2 - 16 = 0$.

Gradient ascent: Battle of the Sexes

- Symmetric, but not zero sum:

$$\begin{array}{c c} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 0,0 & 2,3 \\ 3,2 & 1,1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} -4\cdot\beta + 1 \\ -4\cdot\alpha + 1 \end{bmatrix}$$

- Stationary at $(1/4, 1/4)$.
- Matrix

$$U = \begin{bmatrix} 0 & -4 \\ -4 & 0 \end{bmatrix}$$

has real eigenvalues:
$\lambda^2 - 16 = 0$. Saddle point
inside $[0,1]^2$.

Gradient ascent: Battle of the Sexes

- Symmetric, but not zero sum:

$$\begin{array}{c} \\ T \\ B \end{array} \begin{array}{c} \begin{array}{cc} L & R \end{array} \\ \left(\begin{array}{cc} 0,0 & 2,3 \\ 3,2 & 1,1 \end{array} \right) \end{array}$$

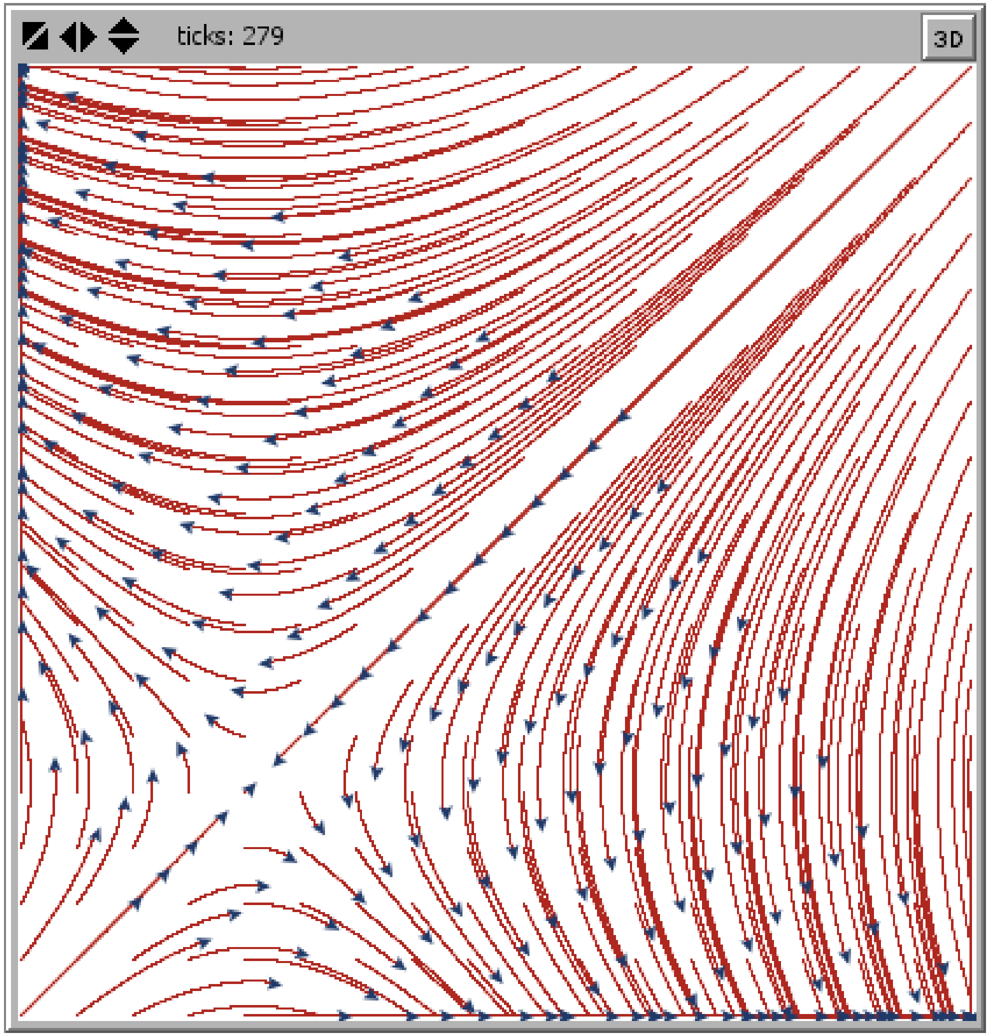
- Gradient:

$$\left[\begin{array}{c} -4\cdot \beta + 1 \\ -4\cdot \alpha + 1 \end{array} \right]$$

- Stationary at $(1/4, 1/4)$.
- Matrix

$$U = \left[\begin{array}{rr} 0 & -4 \\ -4 & 0 \end{array} \right]$$

has real eigenvalues: $\lambda^2 - 16 = 0$. Saddle point inside $[0,1]^2$.



Gradient ascent: Matching pennies

Gradient ascent: Matching pennies

- Symmetric, zero sum:

$$
\begin{array}{cc}
 & \begin{array}{cc} \text{L} & \text{R} \end{array} \\
\begin{array}{c} \text{T} \\ \text{B} \end{array} & \begin{pmatrix} 1,-1 & -1,1 \\ -1,1 & 1,-1 \end{pmatrix}
\end{array}
$$

Gradient ascent: Matching pennies

- Symmetric, zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 1,-1 & -1,1 \\ -1,1 & 1,-1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} 4\cdot\beta - 2 \\ -4\cdot\alpha + 2 \end{bmatrix}$$

Gradient ascent: Matching pennies

- Symmetric, zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} 1,-1 & -1,1 \\ -1,1 & 1,-1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} 4\cdot \beta - 2 \\ -4\cdot \alpha + 2 \end{bmatrix}$$

- Stationary at $(1/2, 1/2)$.

Gradient ascent: Matching pennies

- Symmetric, zero sum:

$$
\begin{array}{cc}
 & \begin{array}{cc} \text{L} & \text{R} \end{array} \\
\begin{array}{c} \text{T} \\ \text{B} \end{array} & \begin{pmatrix} 1,-1 & -1,1 \\ -1,1 & 1,-1 \end{pmatrix}
\end{array}
$$

- Gradient:

$$
\begin{bmatrix} 4\cdot\beta - 2 \\ -4\cdot\alpha + 2 \end{bmatrix}
$$

- Stationary at $(1/2, 1/2)$.
- Matrix

$$
U = \begin{bmatrix} 0 & 4 \\ -4 & 0 \end{bmatrix}
$$

has imaginary eigenvalues:
$\lambda^2 + 16 = 0$.

Gradient ascent: Matching pennies

- Symmetric, zero sum:

$$\begin{array}{c|cc} & \text{L} & \text{R} \\ \hline \text{T} & 1,-1 & -1,1 \\ \text{B} & -1,1 & 1,-1 \end{array}$$

- Gradient:

$$\begin{bmatrix} 4\cdot\beta - 2 \\ -4\cdot\alpha + 2 \end{bmatrix}$$

- Stationary at $(1/2, 1/2)$.
- Matrix

$$U = \begin{bmatrix} 0 & 4 \\ -4 & 0 \end{bmatrix}$$

has imaginary eigenvalues: $\lambda^2 + 16 = 0$. Centric point inside $[0, 1]^2$.

Gradient ascent: Matching pennies

- Symmetric, zero sum:

$$\begin{array}{c} \\ \mathrm{T} \\ \mathrm{B} \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{pmatrix} 1,-1 & -1,1 \\ -1,1 & 1,-1 \end{pmatrix} \end{array}$$

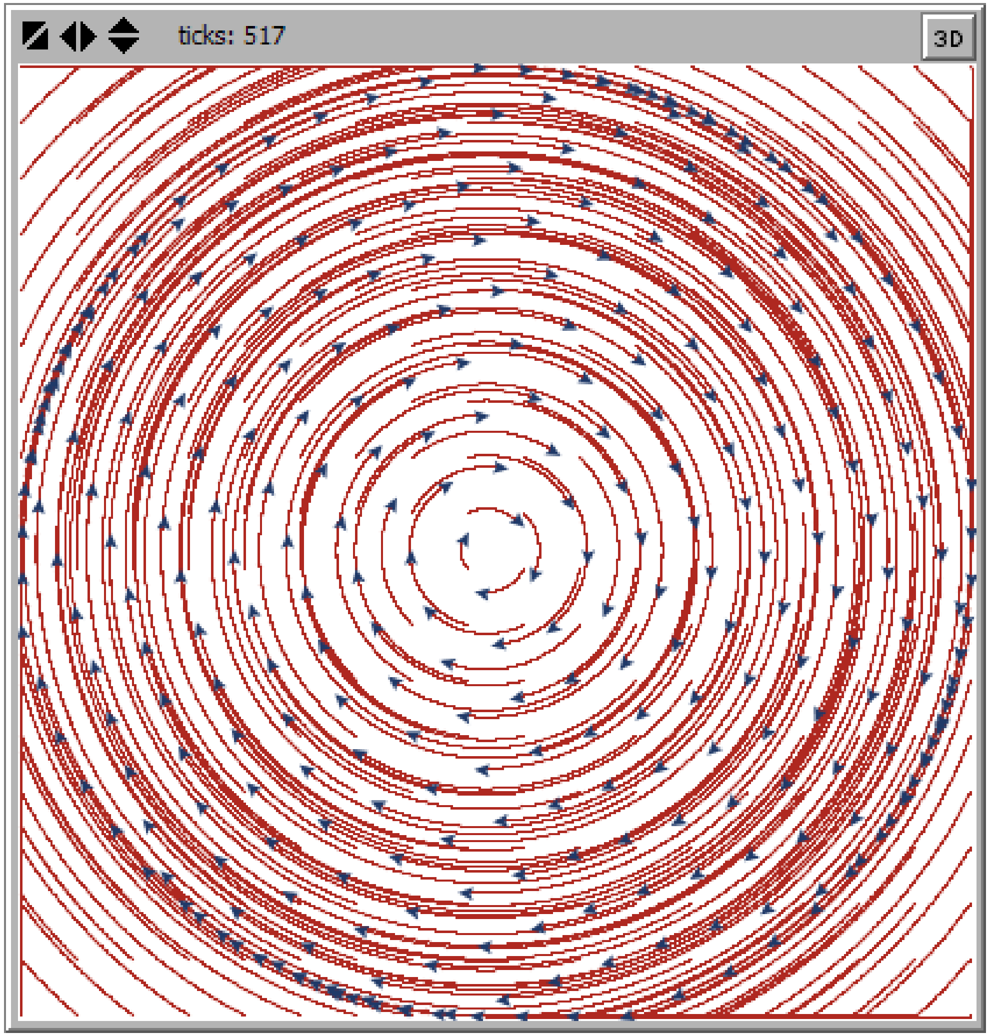
- Gradient:

$$\begin{bmatrix} 4\cdot \beta - 2 \\ -4\cdot \alpha + 2 \end{bmatrix}$$

- Stationary at $(1/2, 1/2)$.
- Matrix

$$U = \begin{bmatrix} 0 & 4 \\ -4 & 0 \end{bmatrix}$$

has imaginary eigenvalues: $\lambda^2 + 16 = 0$. Centric point inside $[0, 1]^2$.



Gradient ascent: other game with centric

Gradient ascent: other game with centric

- Symmetric, zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \text{L} & \text{R} \end{array} \\ \begin{array}{c} \text{T} \\ \text{B} \end{array} & \begin{pmatrix} -2,2 & 1,1 \\ 3,-3 & -2,1 \end{pmatrix} \end{array}$$

Gradient ascent: other game with centric

- Symmetric, zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} -2,2 & 1,1 \\ 3,-3 & -2,1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} -8\cdot \beta + 3 \\ 5\cdot \alpha - 4 \end{bmatrix}$$

Gradient ascent: other game with centric

- Symmetric, zero sum:

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{array}{c} \mathrm{T} \\ \mathrm{B} \end{array} & \begin{pmatrix} -2,2 & 1,1 \\ 3,-3 & -2,1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} -8\cdot \beta + 3 \\ 5\cdot \alpha - 4 \end{bmatrix}$$

- Stationary at $(4/5, 3/8)$.

Gradient ascent: other game with centric

- Symmetric, zero sum:

$$\begin{array}{c} \\ \mathrm{T} \\ \mathrm{B} \end{array} \begin{array}{c} \begin{array}{cc} \quad\mathrm{L} & \quad\mathrm{R} \end{array} \\ \begin{pmatrix} -2,2 & 1,1 \\ 3,-3 & -2,1 \end{pmatrix} \end{array}$$

- Gradient:

$$\begin{bmatrix} -8\cdot\beta + 3 \\ 5\cdot\alpha - 4 \end{bmatrix}$$

- Stationary at $(4/5, 3/8)$.

- Matrix

$$U = \begin{bmatrix} 0 & -8 \\ 5 & 0 \end{bmatrix}$$

has imaginary eigenvalues:
$\lambda^2 + 40 = 0$.

Gradient ascent: other game with centric

- Symmetric, zero sum:

$$\begin{array}{c} \\ \mathrm{T} \\ \mathrm{B} \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \left(\begin{array}{cc} -2,2 & 1,1 \\ 3,-3 & -2,1 \end{array} \right) \end{array}$$

- Gradient:

$$\begin{bmatrix} -8\cdot \beta + 3 \\ 5\cdot \alpha - 4 \end{bmatrix}$$

- Stationary at $(4/5, 3/8)$.
- Matrix

$$U = \begin{bmatrix} 0 & -8 \\ 5 & 0 \end{bmatrix}$$

has imaginary eigenvalues: $\lambda^2 + 40 = 0$. Centric point inside $[0,1]^2$.

Gradient ascent: other game with centric

- Symmetric, zero sum:

$$\begin{array}{c} \\ \mathrm{T} \\ \mathrm{B} \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{L} & \mathrm{R} \end{array} \\ \begin{pmatrix} -2,2 & 1,1 \\ 3,-3 & -2,1 \end{pmatrix} \end{array}$$

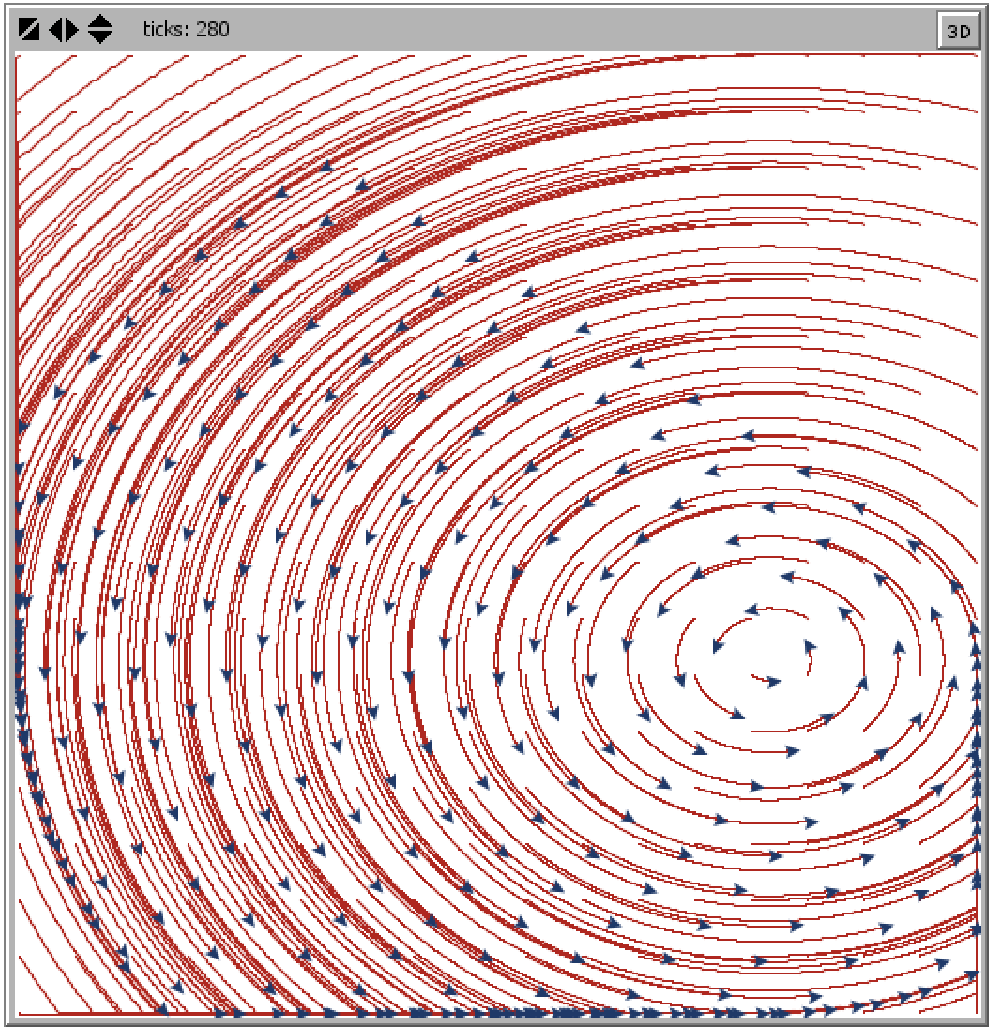
- Gradient:

$$\begin{bmatrix} -8\cdot\beta + 3 \\ 5\cdot\alpha - 4 \end{bmatrix}$$

- Stationary at $(4/5, 3/8)$.
- Matrix

$$U = \begin{bmatrix} 0 & -8 \\ 5 & 0 \end{bmatrix}$$

has imaginary eigenvalues: $\lambda^2 + 40 = 0$. Centric point inside $[0,1]^2$.



Convergence of IGA (Singh *et al.*, 2000)

Proof outline. There are two main cases:

Convergence of IGA (Singh *et al.*, 2000)

Proof outline. There are two main cases:

1. **There is no stationary point, or the stationary point lies outside $[0,1]^2$.** *Then there is movement everywhere in* $[0,1]^2$.

   Since movement is caused by an affine differential map the flow is in one direction, hence gets stuck somewhere at the boundary.

Convergence of IGA (Singh *et al.*, 2000)

Proof outline. There are two main cases:

1. **There is no stationary point, or the stationary point lies outside $[0,1]^2$.** *Then there is movement everywhere in $[0,1]^2$.*

   Since movement is caused by an affine differential map the flow is in one direction, hence gets stuck somewhere at the boundary.

2. **There is a stationary point inside $[0,1]^2$.**

Convergence of IGA (Singh *et al.*, 2000)

Proof outline. There are two main cases:

1. **There is no stationary point, or the stationary point lies outside $[0, 1]^2$.** *Then there is movement everywhere in* $[0, 1]^2$.

   Since movement is caused by an affine differential map the flow is in one direction, hence gets stuck somewhere at the boundary.

2. **There is a stationary point inside $[0, 1]^2$.**

   (a) The stationary point is an attractor. Then it attracts movement which then becomes stationary.

Convergence of IGA (Singh *et al.*, 2000)

Proof outline. There are two main cases:

1. **There is no stationary point, or the stationary point lies outside $[0, 1]^2$.** *Then there is movement everywhere in* $[0, 1]^2$.

   Since movement is caused by an affine differential map the flow is in one direction, hence gets stuck somewhere at the boundary.

2. **There is a stationary point inside $[0, 1]^2$.**

   (a) The stationary point is an attractor. Then it attracts movement which then becomes stationary.

   (b) The stationary point is a repellor. Then it repels movement towards the boundary.

Convergence of IGA (Singh *et al.*, 2000)

Proof outline. There are two main cases:

1. **There is no stationary point, or the stationary point lies outside $[0,1]^2$.** *Then there is movement everywhere in $[0,1]^2$.*

   Since movement is caused by an affine differential map the flow is in one direction, hence gets stuck somewhere at the boundary.

2. **There is a stationary point inside $[0,1]^2$.**

   (a) The stationary point is an attractor. Then it attracts movement which then becomes stationary.

   (b) The stationary point is a repellor. Then it repels movement towards the boundary.

   (c) Both (2a) and (2b): saddle point.

Convergence of IGA (Singh *et al.*, 2000)

Proof outline. There are two main cases:

1. **There is no stationary point, or the stationary point lies outside $[0, 1]^2$.** *Then there is movement everywhere in* $[0, 1]^2$.

   Since movement is caused by an affine differential map the flow is in one direction, hence gets stuck somewhere at the boundary.

2. **There is a stationary point inside $[0, 1]^2$.**

   (a) The stationary point is an attractor. Then it attracts movement which then becomes stationary.

   (b) The stationary point is a repellor. Then it repels movement towards the boundary.

   (c) Both (2a) and (2b): saddle point.

   (d) None of the above. Then strategies do not converge.

Convergence of IGA (Singh *et al.*, 2000)

Proof outline. There are two main cases:

1. **There is no stationary point, or the stationary point lies outside $[0,1]^2$.** *Then there is movement everywhere in* $[0,1]^2$.

   Since movement is caused by an affine differential map the flow is in one direction, hence gets stuck somewhere at the boundary.

2. **There is a stationary point inside $[0,1]^2$.**

   (a) The stationary point is an attractor. Then it attracts movement which then becomes stationary.

   (b) The stationary point is a repellor. Then it repels movement towards the boundary.

   (c) Both (2a) and (2b): saddle point.

   (d) None of the above. Then strategies do not converge.

In three out of four cases, the dynamics ends, hence ends in Nash.

Convergence of IGA (Singh *et al.*, 2000)

Proof outline. There are two main cases:

1. **There is no stationary point, or the stationary point lies outside $[0,1]^2$.** *Then there is movement everywhere in* $[0,1]^2$.

   Since movement is caused by an affine differential map the flow is in one direction, hence gets stuck somewhere at the boundary.

2. **There is a stationary point inside $[0,1]^2$.**

   (a) The stationary point is an attractor. Then it attracts movement which then becomes stationary.

   (b) The stationary point is a repellor. Then it repels movement towards the boundary.

   (c) Both (2a) and (2b): saddle point.

   (d) None of the above. Then strategies do not converge.

In three out of four cases, the dynamics ends, hence ends in Nash.

Part 3: IGA-WoLF

IGA-WoLF (Bowling *et al.*, 2001)

Bowling and Veloso modify IGA so as to ensure convergence in Case 2d.
Idea: Win or Learn Fast

IGA-WoLF (Bowling *et al.*, 2001)

Bowling and Veloso modify IGA so as to ensure convergence in Case 2d.
Idea: Win or Learn Fast (**WoLF**).

IGA-WoLF (Bowling *et al.*, 2001)

Bowling and Veloso modify IGA so as to ensure convergence in Case 2d. Idea: Win or Learn Fast (**WoLF**).

To this end, IGA-WoLF uses a variable learning rate:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t} + \eta \begin{bmatrix} l_t^1 \cdot \partial u_1 / \partial \alpha \\ l_t^2 \cdot \partial u_2 / \partial \beta \end{bmatrix}_{t}$$

IGA-WoLF (Bowling *et al.*, 2001)

Bowling and Veloso modify IGA so as to ensure convergence in Case 2d. Idea: Win or Learn Fast (**WoLF**).

To this end, IGA-WoLF uses a variable learning rate:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \begin{bmatrix} l_t^1 \cdot \partial u_1 / \partial \alpha \\ l_t^2 \cdot \partial u_2 / \partial \beta \end{bmatrix}_t$$

where $l_t^{\{1,2\}} \in \{l_{\min}, l_{\max}\}$ all positive.

IGA-WoLF (Bowling *et al.*, 2001)

Bowling and Veloso modify IGA so as to ensure convergence in Case 2d. Idea: Win or Learn Fast (**WoLF**).

To this end, IGA-WoLF uses a variable learning rate:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t} + \eta \begin{bmatrix} l_t^1 \cdot \partial u_1 / \partial \alpha \\ l_t^2 \cdot \partial u_2 / \partial \beta \end{bmatrix}_{t}$$

where $l_t^{\{1,2\}} \in \{l_{\min}, l_{\max}\}$ all positive.

$$l_t^1 =_{Def} \begin{cases} l_{\min} & \text{if } u_1(\alpha_t, \beta_t) > u_1(\alpha^e, \beta_t) & \text{Winning} \\ l_{\max} & \text{otherwise} & \text{Losing} \end{cases}$$

IGA-WoLF (Bowling *et al.*, 2001)

Bowling and Veloso modify IGA so as to ensure convergence in Case 2d. Idea: Win or Learn Fast (**WoLF**).

To this end, IGA-WoLF uses a variable learning rate:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \begin{bmatrix} l_t^1 \cdot \partial u_1 / \partial \alpha \\ l_t^2 \cdot \partial u_2 / \partial \beta \end{bmatrix}_t$$

where $l_t^{\{1,2\}} \in \{l_{\min}, l_{\max}\}$ all positive.

$$l_t^1 =_{Def} \begin{cases} l_{\min} & \text{if } u_1(\alpha_t, \beta_t) > u_1(\alpha^e, \beta_t) & \text{Winning} \\ l_{\max} & \text{otherwise} & \text{Losing} \end{cases}$$

$$l_t^2 =_{Def} \begin{cases} l_{\min} & \text{if } u_2(\alpha_t, \beta_t) > u_2(\alpha_t, \beta^e) & \text{Winning} \\ l_{\max} & \text{otherwise} & \text{Losing} \end{cases}$$

IGA-WoLF (Bowling *et al.*, 2001)

Bowling and Veloso modify IGA so as to ensure convergence in Case 2d. Idea: Win or Learn Fast (**WoLF**).

To this end, IGA-WoLF uses a variable learning rate:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \begin{bmatrix} l_t^1 \cdot \partial u_1 / \partial \alpha \\ l_t^2 \cdot \partial u_2 / \partial \beta \end{bmatrix}_t$$

where $l_t^{\{1,2\}} \in \{l_{\min}, l_{\max}\}$ all positive.

$$l_t^1 =_{Def} \begin{cases} l_{\min} & \text{if } u_1(\alpha_t, \beta_t) > u_1(\alpha^e, \beta_t) \quad \text{Winning} \\ l_{\max} & \text{otherwise} \quad \text{Losing} \end{cases}$$

$$l_t^2 =_{Def} \begin{cases} l_{\min} & \text{if } u_2(\alpha_t, \beta_t) > u_2(\alpha_t, \beta^e) \quad \text{Winning} \\ l_{\max} & \text{otherwise} \quad \text{Losing} \end{cases}$$

where α^e is a row strategy belonging to some NE, chosen by the row player.

IGA-WoLF (Bowling *et al.*, 2001)

Bowling and Veloso modify IGA so as to ensure convergence in Case 2d. Idea: Win or Learn Fast (**WoLF**).

To this end, IGA-WoLF uses a variable learning rate:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \begin{bmatrix} l_t^1 \cdot \partial u_1 / \partial \alpha \\ l_t^2 \cdot \partial u_2 / \partial \beta \end{bmatrix}_t$$

where $l_t^{\{1,2\}} \in \{l_{\min}, l_{\max}\}$ all positive.

$$l_t^1 =_{Def} \begin{cases} l_{\min} & \text{if } u_1(\alpha_t, \beta_t) > u_1(\alpha^e, \beta_t) \quad \text{Winning} \\ l_{\max} & \text{otherwise} \quad \text{Losing} \end{cases}$$

$$l_t^2 =_{Def} \begin{cases} l_{\min} & \text{if } u_2(\alpha_t, \beta_t) > u_2(\alpha_t, \beta^e) \quad \text{Winning} \\ l_{\max} & \text{otherwise} \quad \text{Losing} \end{cases}$$

where α^e is a row strategy belonging to some NE, chosen by the row player. Similarly for β^e and column player.

IGA-WoLF (Bowling *et al.*, 2001)

Bowling and Veloso modify IGA so as to ensure convergence in Case 2d. Idea: Win or Learn Fast (**WoLF**).

To this end, IGA-WoLF uses a variable learning rate:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \begin{bmatrix} l_t^1 \cdot \partial u_1 / \partial \alpha \\ l_t^2 \cdot \partial u_2 / \partial \beta \end{bmatrix}_t$$

where $l_t^{\{1,2\}} \in \{l_{\min}, l_{\max}\}$ all positive.

$$l_t^1 =_{Def} \begin{cases} l_{\min} & \text{if } u_1(\alpha_t, \beta_t) > u_1(\alpha^e, \beta_t) & \text{Winning} \\ l_{\max} & \text{otherwise} & \text{Losing} \end{cases}$$

$$l_t^2 =_{Def} \begin{cases} l_{\min} & \text{if } u_2(\alpha_t, \beta_t) > u_2(\alpha_t, \beta^e) & \text{Winning} \\ l_{\max} & \text{otherwise} & \text{Losing} \end{cases}$$

where α^e is a row strategy belonging to some NE, chosen by the row player. Similarly for β^e and column player. So (α^e, β^e) need not be Nash!

Case 2d: revolution around Nash equilibrium

Lemma 1. *With fixed l^1 and l^2, the trajectory of the strategy pair (α, β) is an elliptic orbit around (α^*, β^*) with axes*

$$\begin{bmatrix} 0 \\ \sqrt{l^2|u|/l^1|u'|} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Case 2d: revolution around Nash equilibrium

Lemma 1. *With fixed l^1 and l^2, the trajectory of the strategy pair (α, β) is an elliptic orbit around (α^*, β^*) with axes*

$$\left[\begin{array}{c} 0 \\ \sqrt{l^2|u|/l^1|u'|} \end{array} \right], \left[\begin{array}{c} 1 \\ 0 \end{array} \right]$$

Remarks:

Case 2d: revolution around Nash equilibrium

Lemma 1. *With fixed l^1 and l^2, the trajectory of the strategy pair (α, β) is an elliptic orbit around (α^*, β^*) with axes*

$$\begin{bmatrix} 0 \\ \sqrt{l^2|u|/l^1|u'|} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Remarks:

- For ellipses with center (α^*, β^*) there are four possibilities, depending on u, u', and whether $\sqrt{l^2|u|/l^1|u'|} > 1$ or < 1.

Case 2d: revolution around Nash equilibrium

Lemma 1. *With fixed l^1 and l^2, the trajectory of the strategy pair (α, β) is an elliptic orbit around (α^*, β^*) with axes*

$$\begin{bmatrix} 0 \\ \sqrt{l^2|u|/l^1|u'|} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Remarks:

- For ellipses with center (α^*, β^*) there are four possibilities, depending on u, u', and whether $\sqrt{l^2|u|/l^1|u'|} > 1$ or < 1.

  1. Lies flat and axes < 1.

Case 2d: revolution around Nash equilibrium

Lemma 1. *With fixed l^1 and l^2, the trajectory of the strategy pair (α, β) is an elliptic orbit around (α^*, β^*) with axes*

$$\begin{bmatrix} 0 \\ \sqrt{l^2|u|/l^1|u'|} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Remarks:

- For ellipses with center (α^*, β^*) there are four possibilities, depending on u, u', and whether $\sqrt{l^2|u|/l^1|u'|} > 1$ or < 1.

  1. Lies flat and axes < 1.

  2. Stands and axes < 1.

Case 2d: revolution around Nash equilibrium

Lemma 1. *With fixed l^1 and l^2, the trajectory of the strategy pair (α, β) is an elliptic orbit around (α^*, β^*) with axes*

$$\begin{bmatrix} 0 \\ \sqrt{l^2|u|/l^1|u'|} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Remarks:

- For ellipses with center (α^*, β^*) there are four possibilities, depending on u, u', and whether $\sqrt{l^2|u|/l^1|u'|} > 1$ or < 1.

1. Lies flat and axes < 1.
2. Stands and axes < 1.
3. Lies flat and axes > 1.

Case 2d: revolution around Nash equilibrium

Lemma 1. *With fixed l^1 and l^2, the trajectory of the strategy pair (α, β) is an elliptic orbit around (α^*, β^*) with axes*

$$\begin{bmatrix} 0 \\ \sqrt{l^2|u|/l^1|u'|} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Remarks:

- For ellipses with center (α^*, β^*) there are four possibilities, depending on u, u', and whether $\sqrt{l^2|u|/l^1|u'|} > 1$ or < 1.

1. Lies flat and axes < 1.
2. Stands and axes < 1.
3. Lies flat and axes > 1.
4. Stands and axes > 1.

Case 2d: revolution around Nash equilibrium

Lemma 1. *With fixed l^1 and l^2, the trajectory of the strategy pair (α, β) is an elliptic orbit around (α^*, β^*) with axes*

$$\begin{bmatrix} 0 \\ \sqrt{l^2|u|/l^1|u'|} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Remarks:

- For ellipses with center (α^*, β^*) there are four possibilities, depending on u, u', and whether $\sqrt{l^2|u|/l^1|u'|} > 1$ or < 1.

  1. Lies flat and axes < 1.
  2. Stands and axes < 1.
  3. Lies flat and axes > 1.
  4. Stands and axes > 1.

- Bowling *et al.* do not prove this result but refer to Sing *et al.*, who, on their turn refer to a work on differential equations by Reinhard (1987).

Case 2d: revolution around Nash equilibrium

Lemma 2. *A player is "winning" if and only if that player's strategy is moving away from the center.*

Case 2d: revolution around Nash equilibrium

Lemma 2. *A player is "winning" if and only if that player's strategy is moving away from the center.*

Proof. When play revolves around the center, there can only be one equilibrium.

Case 2d: revolution around Nash equilibrium

Lemma 2. *A player is "winning" if and only if that player's strategy is moving away from the center.*

Proof. When play revolves around the center, there can only be one equilibrium. So (α^e, β^e) **is** in fact an equilibrium.

Case 2d: revolution around Nash equilibrium

Lemma 2. *A player is "winning" if and only if that player's strategy is moving away from the center.*

Proof. When play revolves around the center, there can only be one equilibrium. So (α^e, β^e) **is** in fact an equilibrium.

Consider the row player, who wins if and only if

$$u_1(\alpha_t, \beta_t) - u_1(\alpha^e, \beta_t) > 0.$$

Simplifying and using $\beta u - (r_{12} - r_{22}) = \partial u_1 / \partial \alpha$ yields

$$(\alpha - \alpha^e) \frac{\partial u_1}{\partial \alpha} > 0.$$

Case 2d: revolution around Nash equilibrium

Lemma 2. *A player is "winning" if and only if that player's strategy is moving away from the center.*

Proof. When play revolves around the center, there can only be one equilibrium. So (α^e, β^e) **is** in fact an equilibrium.

Consider the row player, who wins if and only if

$$u_1(\alpha_t, \beta_t) - u_1(\alpha^e, \beta_t) > 0.$$

Simplifying and using $\beta u - (r_{12} - r_{22}) = \partial u_1 / \partial \alpha$ yields

$$(\alpha - \alpha^e)\frac{\partial u_1}{\partial \alpha} > 0.$$

Thus, row player "wins" iff either $\alpha > \alpha^e$ and α increases, or else $\alpha < \alpha^e$ and α decreases. $\square$

Case 2d: revolution around Nash equilibrium

Lemma 2. *A player is "winning" if and only if that player's strategy is moving away from the center.*

Proof. When play revolves around the center, there can only be one equilibrium. So (α^e, β^e) **is** in fact an equilibrium.

Consider the row player, who wins if and only if

$$u_1(\alpha_t, \beta_t) - u_1(\alpha^e, \beta_t) > 0.$$

Simplifying and using $\beta u - (r_{12} - r_{22}) = \partial u_1 / \partial \alpha$ yields

$$(\alpha - \alpha^e) \frac{\partial u_1}{\partial \alpha} > 0.$$

Thus, row player "wins" iff either $\alpha > \alpha^e$ and α increases, or else $\alpha < \alpha^e$ and α decreases. $\square$

Corollary. *The learning rate is constant throughout any one quadrant.*

Case 2d: revolution around Nash equilibrium

Lemma 3. *Let* C *be the center. For every initial strategy pair* (α, β) *that is sufficiently close to* C*, the* l_{min} / l_{max} *dynamics will bring that pair to* C.

Case 2d: revolution around Nash equilibrium

Lemma 3. *Let* C *be the center. For every initial strategy pair* (α, β) *that is sufficiently close to* C*, the* l_{min} / l_{max} *dynamics will bring that pair to* C*.*

Proof.

Let (α, β) be a strategy pair. According to Lemma 1, its trajectory forms an ellipse with center C.

Case 2d: revolution around Nash equilibrium

Lemma 3. *Let C be the center. For every initial strategy pair (α, β) that is sufficiently close to C, the l_{min} / l_{max} dynamics will bring that pair to C.*

Proof.

Let (α, β) be a strategy pair. According to Lemma 1, its trajectory forms an ellipse with center C.

If (α, β) is sufficiently close to C, this ellipse will entirely be in $[0, 1]^2$ and its trajectory will not be disrupted.

Case 2d: revolution around Nash equilibrium

Lemma 3. *Let C be the center. For every initial strategy pair (α, β) that is sufficiently close to C, the l_{min} / l_{max} dynamics will bring that pair to C.*

Proof.

Let (α, β) be a strategy pair. According to Lemma 1, its trajectory forms an ellipse with center C.

If (α, β) is sufficiently close to C, this ellipse will entirely be in $[0, 1]^2$ and its trajectory will not be disrupted.

There are two cases.

Case 2d: revolution around Nash equilibrium

Lemma 3. *Let C be the center. For every initial strategy pair (α, β) that is sufficiently close to C, the l_{min} / l_{max} dynamics will bring that pair to C.*

Proof.

Let (α, β) be a strategy pair. According to Lemma 1, its trajectory forms an ellipse with center C.

If (α, β) is sufficiently close to C, this ellipse will entirely be in $[0, 1]^2$ and its trajectory will not be disrupted.

There are two cases.

1. *Strategy pair moves clockwise.*

Case 2d: revolution around Nash equilibrium

Lemma 3. *Let C be the center. For every initial strategy pair (α, β) that is sufficiently close to C, the $l_{\min}$ / $l_{\max}$ dynamics will bring that pair to C.*

Proof.

Let (α, β) be a strategy pair. According to Lemma 1, its trajectory forms an ellipse with center C.

If (α, β) is sufficiently close to C, this ellipse will entirely be in $[0, 1]^2$ and its trajectory will not be disrupted.

There are two cases.

1. *Strategy pair moves clockwise.*

(a) We will thus have to ensure that learning parameters are set such that ellipse that forms the trajectory "stands" when (α, β) is in $Q1$ and $Q3$.

Case 2d: revolution around Nash equilibrium

Lemma 3. *Let C be the center. For every initial strategy pair (α, β) that is sufficiently close to C, the $l_{\min}$ / $l_{\max}$ dynamics will bring that pair to C.*

Proof.

Let (α, β) be a strategy pair. According to Lemma 1, its trajectory forms an ellipse with center C.

If (α, β) is sufficiently close to C, this ellipse will entirely be in $[0, 1]^2$ and its trajectory will not be disrupted.

There are two cases.

1. *Strategy pair moves clockwise.*

(a) We will thus have to ensure that learning parameters are set such that ellipse that forms the trajectory "stands" when (α, β) is in $Q1$ and $Q3$.

(b) Similarly, we will have to ensure that learning parameters are such that the ellipse "lies flat" when (α, β) is in $Q2$ and $Q4$.

Case 2d: revolution around Nash equilibrium

Lemma 3. *Let C be the center. For every initial strategy pair (α, β) that is sufficiently close to C, the $l_{\min}$ / $l_{\max}$ dynamics will bring that pair to C.*

Proof.

Let (α, β) be a strategy pair. According to Lemma 1, its trajectory forms an ellipse with center C.

If (α, β) is sufficiently close to C, this ellipse will entirely be in $[0, 1]^2$ and its trajectory will not be disrupted.

There are two cases.

1. *Strategy pair moves clockwise.*
   (a) We will thus have to ensure that learning parameters are set such that ellipse that forms the trajectory "stands" when (α, β) is in $Q1$ and $Q3$.
   (b) Similarly, we will have to ensure that learning parameters are such that the ellipse "lies flat" when (α, β) is in $Q2$ and $Q4$.
2. *Strategy pair moves counter-clockwise*. Similar reasoning.

Case 2d: revolution around Nash equilibrium

Lemma 3. *Let C be the center. For every initial strategy pair (α, β) that is sufficiently close to C, the $l_{\min}$ / $l_{\max}$ dynamics will bring that pair to C.*

Proof.

Let (α, β) be a strategy pair. According to Lemma 1, its trajectory forms an ellipse with center C.

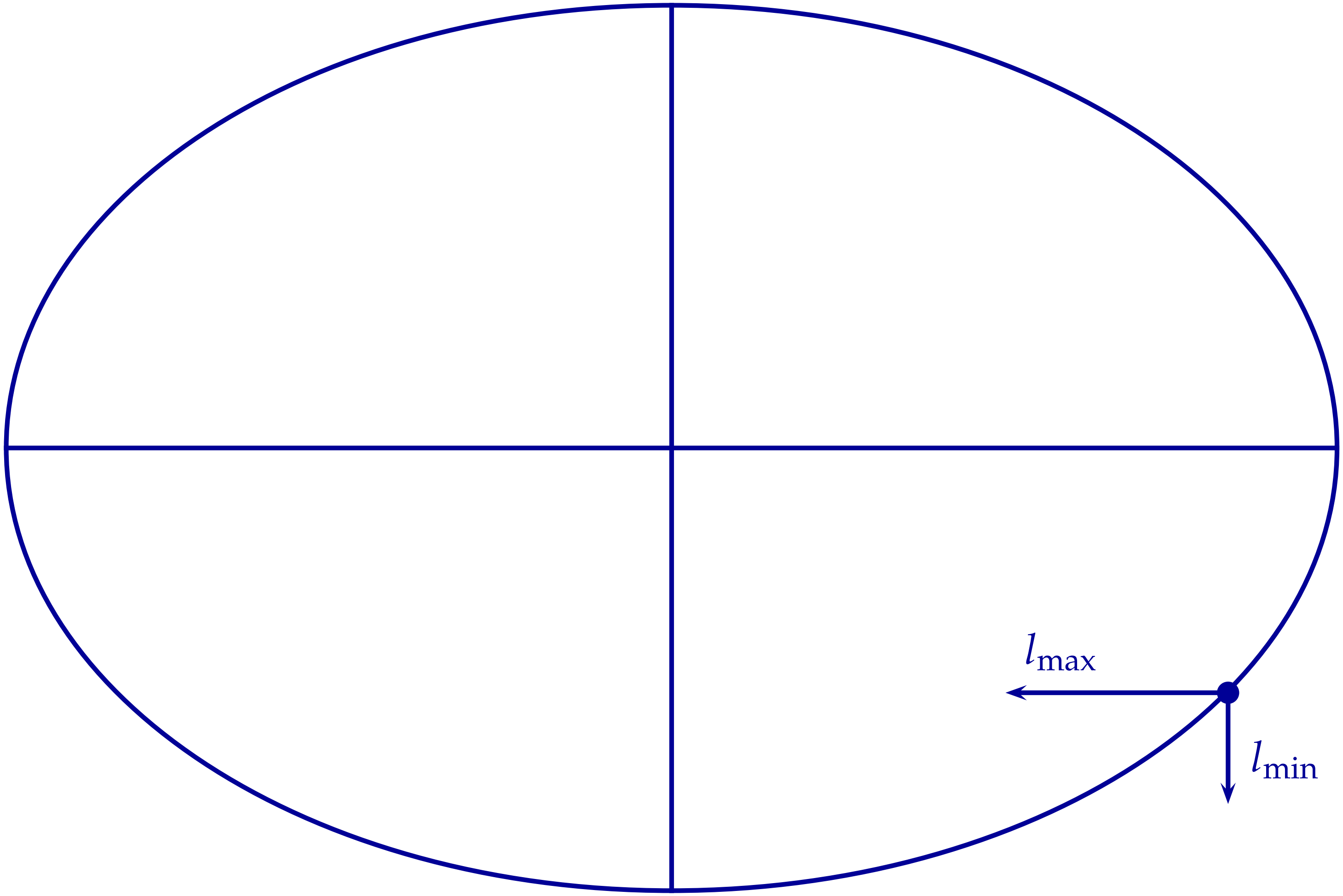
If (α, β) is sufficiently close to C, this ellipse will entirely be in $[0, 1]^2$ and its trajectory will not be disrupted.

There are two cases.

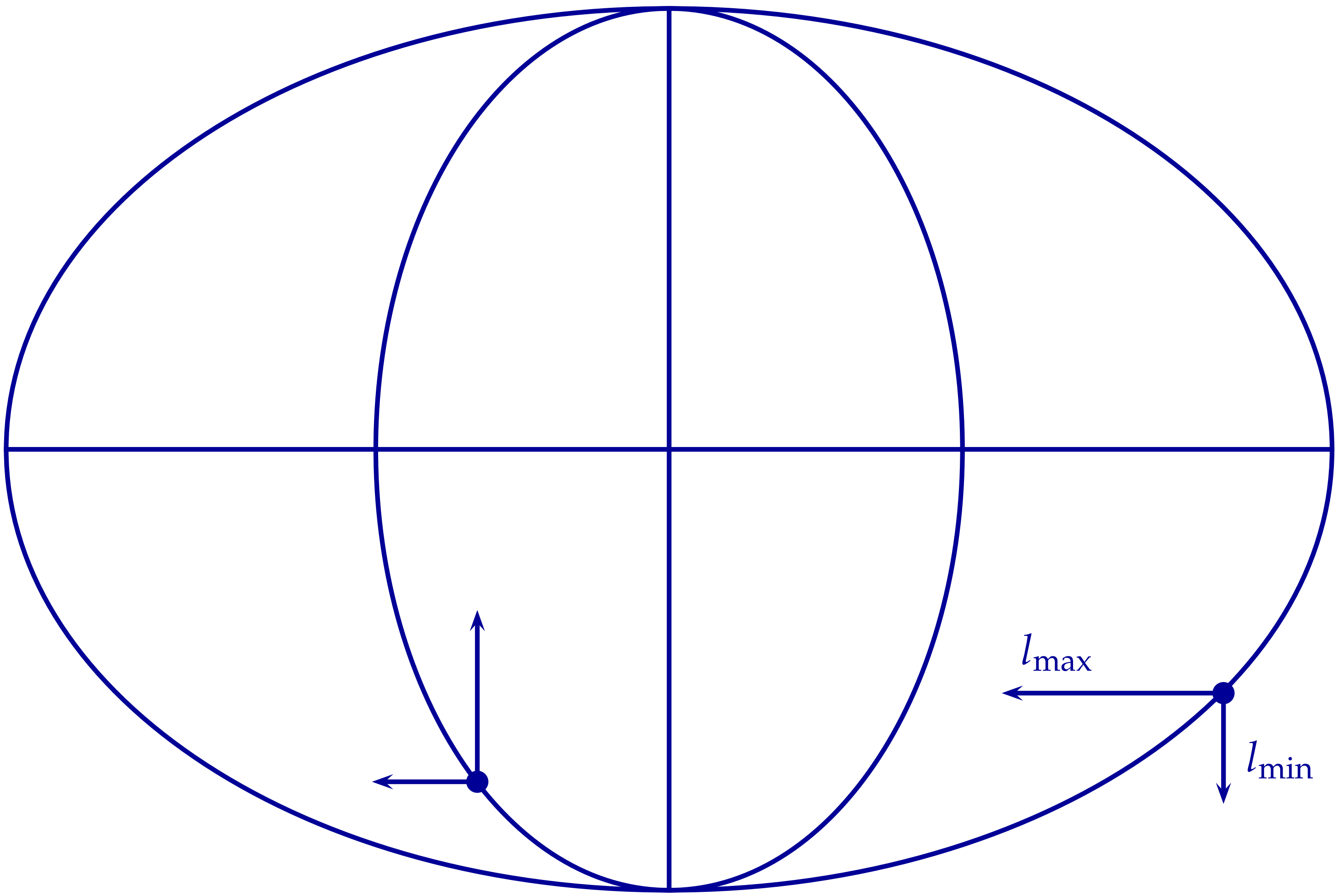
1. *Strategy pair moves clockwise.*
   (a) We will thus have to ensure that learning parameters are set such that ellipse that forms the trajectory “stands” when (α, β) is in $Q1$ and $Q3$.
   (b) Similarly, we will have to ensure that learning parameters are such that the ellipse “lies flat” when (α, β) is in $Q2$ and $Q4$.
2. *Strategy pair moves counter-clockwise*. Similar reasoning.

(First a suggestive picture, then the rest of the proof.)

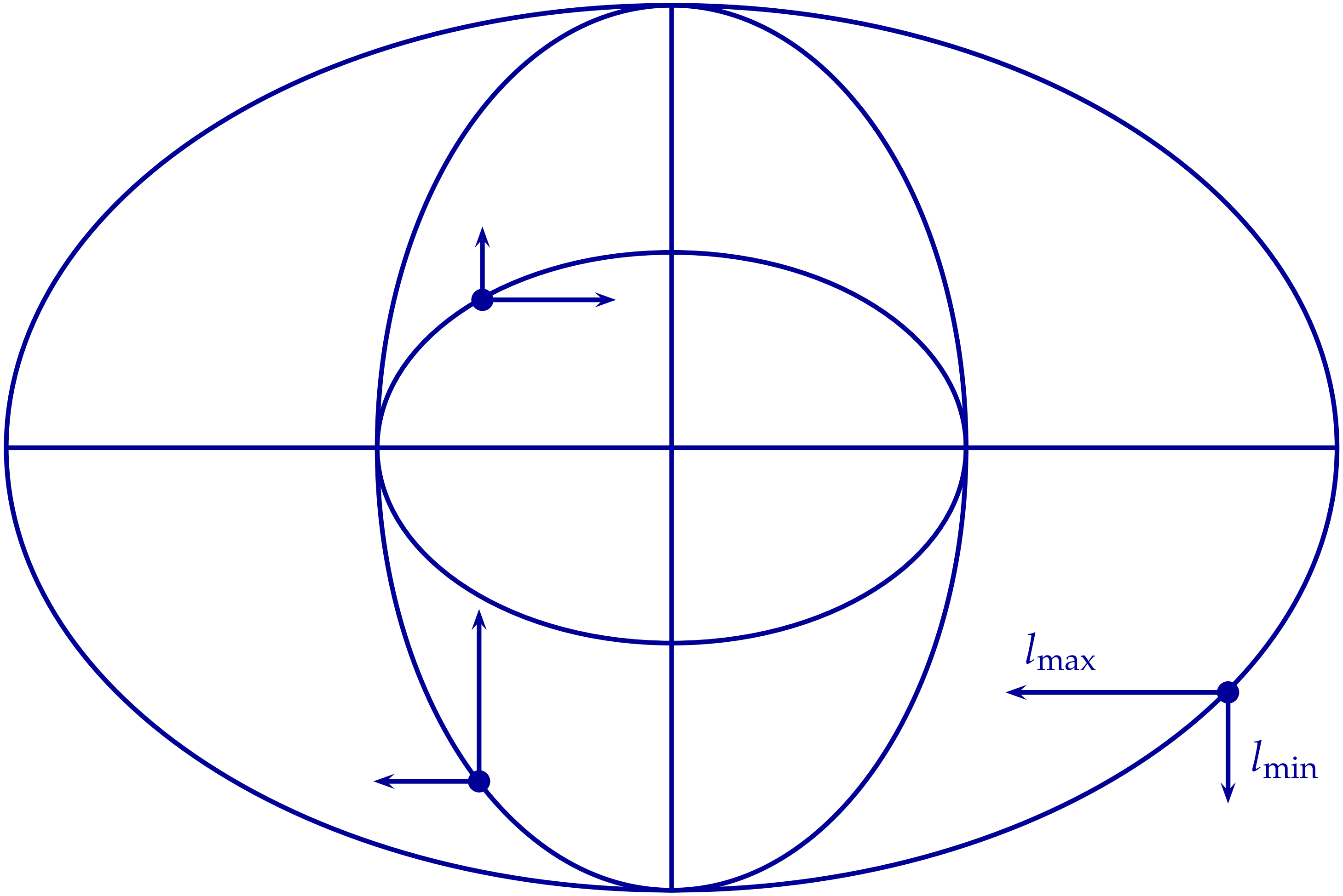
Trajectory in different quadrants (clockwise)



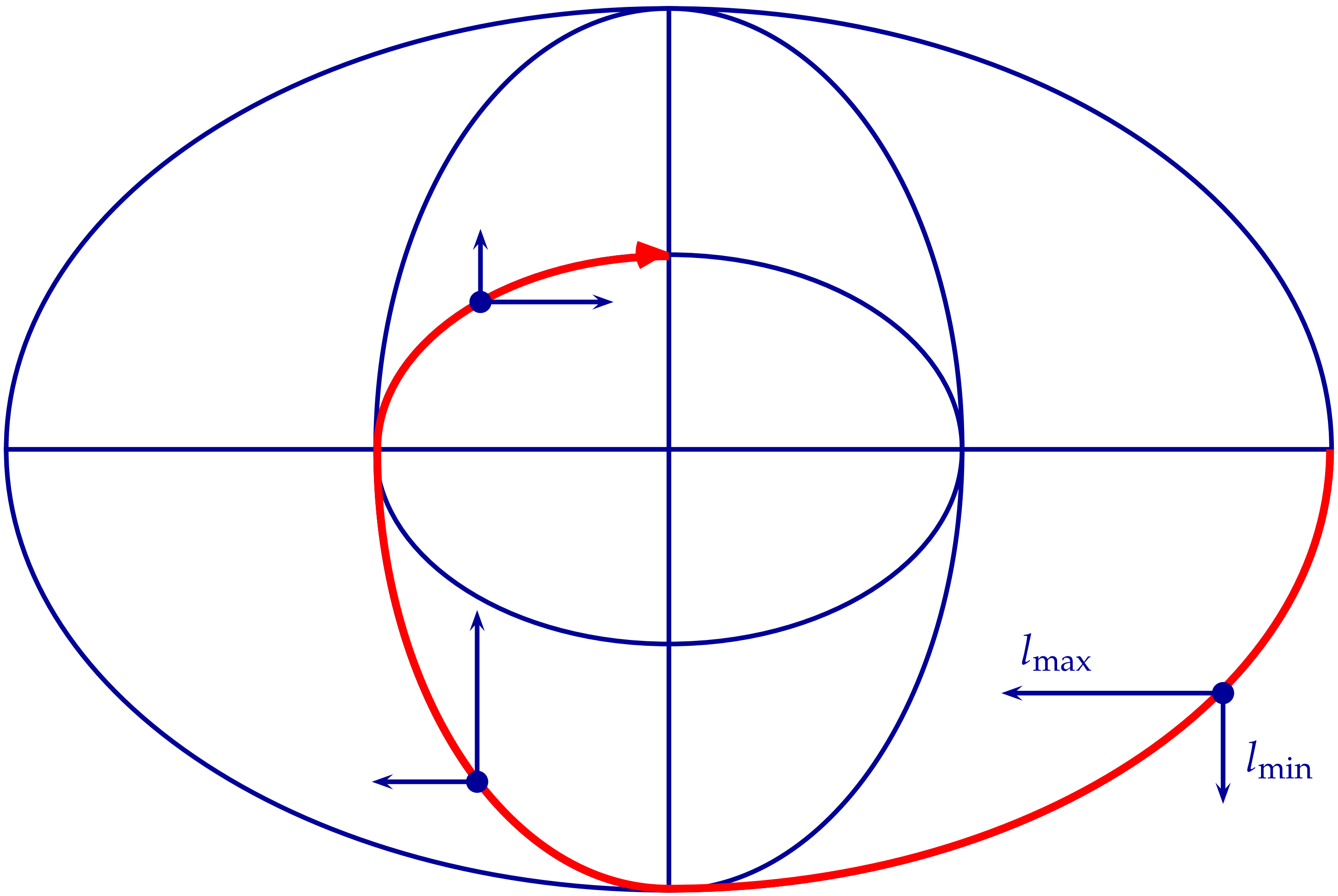
Trajectory in different quadrants (clockwise)



Trajectory in different quadrants (clockwise)



Compound trajectory



Trajectory in different quadrants (clockwise)

Claim. *The learning parameters* $l_{\min}$ / $l_{\max}$ *alternate in such a way that the ellipse that forms the trajectory in clockwise movement "lies flat" when* (α, β) *is in* $Q1$ *and* $Q3$ *of the ellipse and "stands" otherwise.*

Trajectory in different quadrants (clockwise)

Claim. *The learning parameters* $l_{\min}$ / $l_{\max}$ *alternate in such a way that the ellipse that forms the trajectory in clockwise movement "lies flat" when* (α, β) *is in* $Q1$ *and* $Q3$ *of the ellipse and "stands" otherwise.*

Proof.

Trajectory in different quadrants (clockwise)

Claim. *The learning parameters $l_{\min}$ / $l_{\max}$ alternate in such a way that the ellipse that forms the trajectory in clockwise movement "lies flat" when (α, β) is in $Q1$ and $Q3$ of the ellipse and "stands" otherwise.*

Proof.
Suppose movement is clockwise.

Trajectory in different quadrants (clockwise)

Claim. *The learning parameters l_{min} / l_{max} alternate in such a way that the ellipse that forms the trajectory in clockwise movement "lies flat" when (α, β) is in $Q1$ and $Q3$ of the ellipse and "stands" otherwise.*

Proof.
Suppose movement is clockwise.

1. Suppose (α, β) is in $Q1$ (upper right) of the ellipse. Then row "wins" and col "loses". By Lemma 2 horizontal velocity $<$ vertical velocity, and the ellipse "stands".

Trajectory in different quadrants (clockwise)

Claim. *The learning parameters $l_{\min}$ / $l_{\max}$ alternate in such a way that the ellipse that forms the trajectory in clockwise movement "lies flat" when (α, β) is in $Q1$ and $Q3$ of the ellipse and "stands" otherwise.*

Proof.
Suppose movement is clockwise.

1. Suppose (α, β) is in $Q1$ (upper right) of the ellipse. Then row "wins" and col "loses". By Lemma 2 horizontal velocity $<$ vertical velocity, and the ellipse "stands".

2. Suppose (α, β) is in $Q2$ (lower right) of the ellipse. Then row "loses" and col "wins". By Lemma 2 horizontal velocity $>$ vertical velocity, and the ellipse "lies flat".

Trajectory in different quadrants (clockwise)

Claim. *The learning parameters* $l_{\min}$ / $l_{\max}$ *alternate in such a way that the ellipse that forms the trajectory in clockwise movement "lies flat" when* (α, β) *is in* $Q1$ *and* $Q3$ *of the ellipse and "stands" otherwise.*

Proof.
Suppose movement is clockwise.

1. Suppose (α, β) is in $Q1$ (upper right) of the ellipse. Then row "wins" and col "loses". By Lemma 2 horizontal velocity $<$ vertical velocity, and the ellipse "stands".

2. Suppose (α, β) is in $Q2$ (lower right) of the ellipse. Then row "loses" and col "wins". By Lemma 2 horizontal velocity $>$ vertical velocity, and the ellipse "lies flat".

Clearly, the reasoning is similar when
– the strategy pair (α, β) is in the other two quadrants

Trajectory in different quadrants (clockwise)

Claim. *The learning parameters $l_{\min}$ / $l_{\max}$ alternate in such a way that the ellipse that forms the trajectory in clockwise movement "lies flat" when (α, β) is in $Q1$ and $Q3$ of the ellipse and "stands" otherwise.*

Proof.
Suppose movement is clockwise.

1. Suppose (α, β) is in $Q1$ (upper right) of the ellipse. Then row "wins" and col "loses". By Lemma 2 horizontal velocity $<$ vertical velocity, and the ellipse "stands".

2. Suppose (α, β) is in $Q2$ (lower right) of the ellipse. Then row "loses" and col "wins". By Lemma 2 horizontal velocity $>$ vertical velocity, and the ellipse "lies flat".

Clearly, the reasoning is similar when
– the strategy pair (α, β) is in the other two quadrants, or
– when movement is counter-clockwise. $\square$

Part 4: Another solution

Why not utilise Singh *et al.*'s result on emp. frequencies

Why not utilise Singh *et al.*'s result on emp. frequencies

- **Theorem** (Singh, Kearns and Mansour, 2000) *If players follow IGA, where* $\eta \to 0$*, then their strategies will converge to a Nash equilibrium. If not, then at least their average payoffs will converge to the expected payoffs of a Nash equilibrium.*

Why not utilise Singh *et al.*'s result on emp. frequencies

- **Theorem** (Singh, Kearns and Mansour, 2000) *If players follow IGA, where* $\eta \to 0$, *then their strategies will converge to a Nash equilibrium. If not, then at least their average payoffs will converge to the expected payoffs of a Nash equilibrium.*
- Idea: use average payoffs to correct the gradient.

Why not utilise Singh *et al.*'s result on emp. frequencies

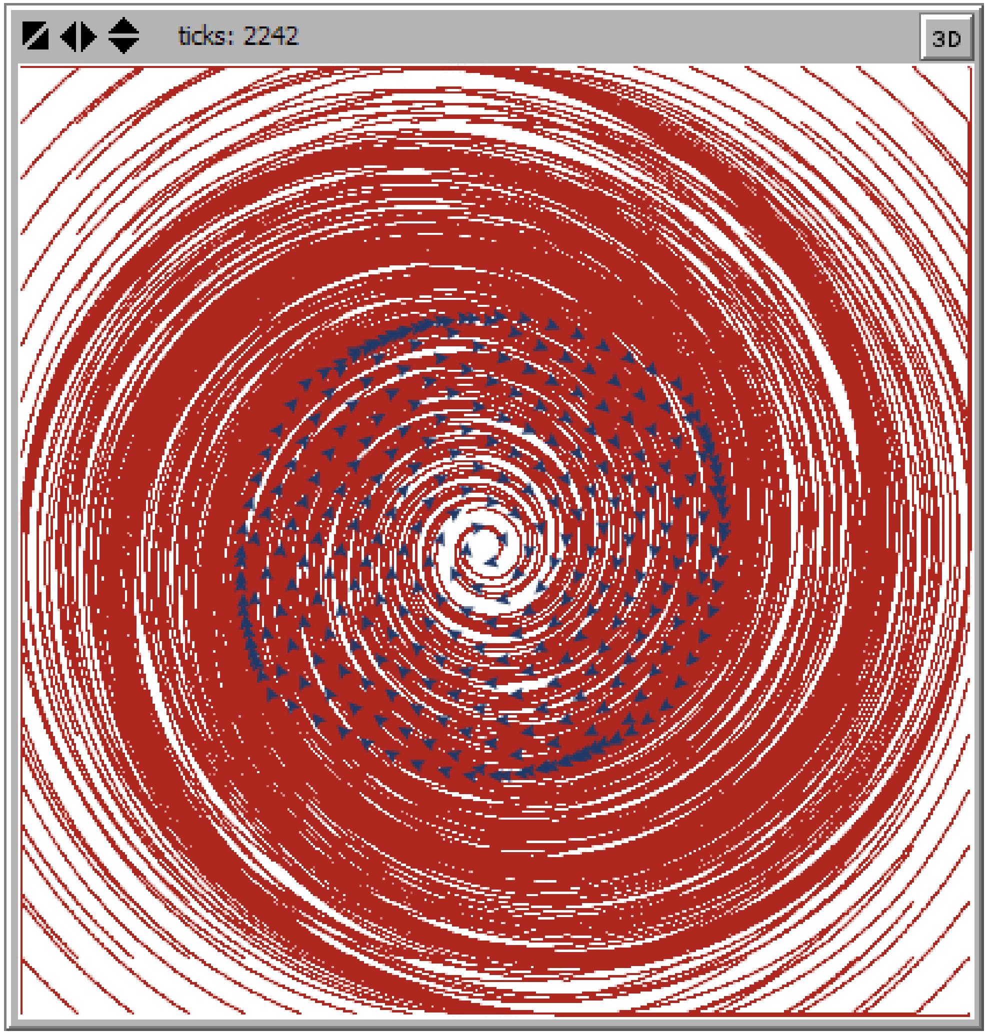
- **Theorem** (Singh, Kearns and Mansour, 2000) *If players follow IGA, where $\eta \to 0$, then their strategies will converge to a Nash equilibrium. If not, then at least their average payoffs will converge to the expected payoffs of a Nash equilibrium.*
- Idea: use average payoffs to correct the gradient.
- So gradient points slight more in direction of average payoffs.

Why not utilise Singh *et al.*'s result on emp. frequencies

- **Theorem** (Singh, Kearns and Mansour, 2000) *If players follow IGA, where* $\eta \to 0$*, then their strategies will converge to a Nash equilibrium. If not, then at least their average payoffs will converge to the expected payoffs of a Nash equilibrium.*
- Idea: use average payoffs to correct the gradient.
- So gradient points slight more in direction of average payoffs.
- At least works empirically :-)

Why not utilise Singh *et al.*'s result on emp. frequencies

- **Theorem** (Singh, Kearns and Mansour, 2000) *If players follow IGA, where $\eta \to 0$, then their strategies will converge to a Nash equilibrium. If not, then at least their average payoffs will converge to the expected payoffs of a Nash equilibrium.*
- Idea: use average payoffs to correct the gradient.
- So gradient points slight more in direction of average payoffs.
- At least works empirically :-)



Literature

Literature

- Original work on gradient ascent in general-sum games:

Literature

- Original work on gradient ascent in general-sum games:

Singh, Kearns, and Mansour (2000). "Nash convergence of gradient ascent in general-sum games". In: *Proc. of the Sixteenth Conf. on Uncertainty in Artificial Intelligence* (pp. 541-548).

Literature

- Original work on gradient ascent in general-sum games:

  Singh, Kearns, and Mansour (2000). "Nash convergence of gradient ascent in general-sum games". In: *Proc. of the Sixteenth Conf. on Uncertainty in Artificial Intelligence* (pp. 541-548).

- Today's presentation was mainly based on this conference publication:

Literature

- Original work on gradient ascent in general-sum games:

  Singh, Kearns, and Mansour (2000). "Nash convergence of gradient ascent in general-sum games". In: *Proc. of the Sixteenth Conf. on Uncertainty in Artificial Intelligence* (pp. 541-548).

- Today's presentation was mainly based on this conference publication:

  Bowling and Veloso (2001). "Convergence of gradient ascent with a Variable Learning Rate". In *Proc. of the Eighteenth Int. Conf. on Machine Learning (ICML)*, pp. 27-34, June 2001.

Literature

- Original work on gradient ascent in general-sum games:

Singh, Kearns, and Mansour (2000). "Nash convergence of gradient ascent in general-sum games". In: *Proc. of the Sixteenth Conf. on Uncertainty in Artificial Intelligence* (pp. 541-548).

- Today's presentation was mainly based on this conference publication:

Bowling and Veloso (2001). "Convergence of gradient ascent with a Variable Learning Rate". In *Proc. of the Eighteenth Int. Conf. on Machine Learning (ICML)*, pp. 27-34, June 2001.

Conference publication was elaborated, and published as a journal article:

Literature

- Original work on gradient ascent in general-sum games:

Singh, Kearns, and Mansour (2000). "Nash convergence of gradient ascent in general-sum games". In: *Proc. of the Sixteenth Conf. on Uncertainty in Artificial Intelligence* (pp. 541-548).

- Today's presentation was mainly based on this conference publication:

Bowling and Veloso (2001). "Convergence of gradient ascent with a Variable Learning Rate". In *Proc. of the Eighteenth Int. Conf. on Machine Learning (ICML)*, pp. 27-34, June 2001.

Conference publication was elaborated, and published as a journal article:

Bowling and Veloso (2002). "Multiagent Learning Using a Variable Learning Rate". In: *Artificial Intelligence* **136**, pp. 215-250, 2002.

What next?

What next?

- With fictitious play, or gradient ascent, opponents are modelled by a single mixed strategy.

What next?

- With fictitious play, or gradient ascent, opponents are modelled by a single mixed strategy.
- With Bayesian play, opponents are modelled by a probability distribution over all opponent strategies

$$\Delta \left[\Pi_{j \neq i} \, \Delta(X_j)^H \right].$$

  - $\Delta(A)$ denotes the set of all probability distributions over A.
  - B^A denotes the set of all functions from A to B.
  - $\Pi_{j \neq i} A_j$ denotes the Cartesian product of $\{A_j\}_{j \neq i}$. In case of a finite product, this can be written as

$$\Pi_{j \neq i} A_j \; = \; A_1 \times A_2 \times \cdots \times A_{i-1} \times A_{i+1} \times \cdots \times A_n.$$