Learning Agents Convergence in First-Price Auctions

Analyzing Gradient Dynamics in Coupled Optimization

Optimization Problem

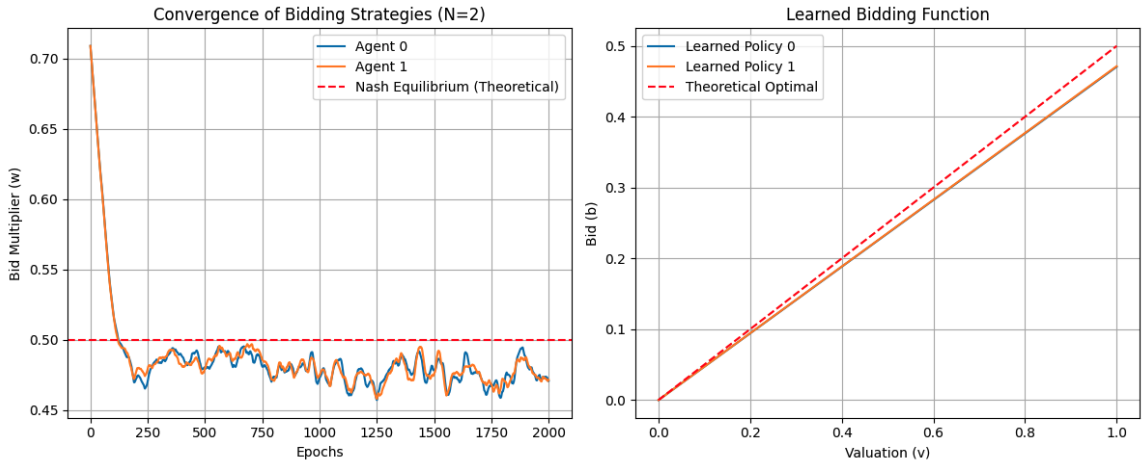
- We are investigating multi-agent reinforcement learning (MARL) in an economic setting, specifically first-price auctions (like online ad markets)
- **Question**: If everyone uses Gradient Ascent to bid, do we converge to a Nash Equilibrium?
- **Setup**:
  - First-Price Auctions: Highest bidder wins and pays their bid
  - Goal: Maximize surplus → $(Value - Bid) \times P(Win)$

Solution

- Differentiable Relaxation
  - **Agents**: PyTorch modules optimizing linear bid parameters (w)
  - **Sigmoid**: $P(Win) \approx \sigma\left(\beta\left(b_i - b_{max\_other}\right)\right)$
    - Simultaneous Gradient Ascent
  - **Parameters**:
    - $\beta = 50$ (Temperature/Steepness)
    - Batch Size = 128 valuations ($U[0,1]$)

Early Results (N = 2)

- Optimal $w = 0.5$ (for 2 agents)
  - Agents converged from random (0.71) to 0.48
- **Observations**:
  - Rapid initial learning (< 200 epochs)
  - Slight downward bias that we think is due to sigmoid smoothing
  - Stable oscillation instead of divergence



Next Steps

- **Sensitivity Analysis**: Does increasing β(steepness) remove the bias?
- **Scale**: Testing $N \in \{3, 5, 10\}$
- **Comparisons**: Benchmarking against multiplicative weights update (MWU)
- **Goal**: PettingZoo-compatible benchmark for auction dynamics.