Haiku Presentation
An Introduction to my Research Agenda

Carlos Gonzalez, Econometrics Lunch

November 26, 2024

Brief Oversight of my Work

- Interested in **Learning**, Dynamic Games, and its Policy Implications
- Looking at the intersection of ML Theory, Micro Theory and Applied Economics. Different approaches to similar problems
- **ML:** Result oriented. Heuristic approach to learning. Refined theory developed ex-post. Algorithms are very powerful, but usually a black box

Brief Oversight of my Work

- Interested in **Learning**, Dynamic Games, and its Policy Implications
- Looking at the intersection of ML Theory, Micro Theory and Applied Economics. Different approaches to similar problems
- **ML:** Result oriented. Heuristic approach to learning. Refined theory developed ex-post. Algorithms are very powerful, but usually a black box
- **MT:** Economically founded learning rules (positive and normative). Predom of Bayesian learning. Deliberately simple and analytically limited
- **Applications:** Adverse Selection, IC Recommendation systems, Monopsony markets, Labor markets, data rationalizability, recovery of utility functions ...

Three Lines of Action (WIP)

- Establish connections between Machine Learning and Econ Theory
  - Hannan Consistency and Convergence to Best Reply [Gonzalez, 2023]
  - An Epistemological Approach to Causality in ML [Gonzalez and Mazlish, 2024]
- Expand Economic Theory leveraging ML heuristics and Algorithms
  - A Prior-Free Theory of Adverse Selection [Gonzalez, 2023]
  - The Econometrics of Behavioral Parameters [Auer et al., 2024]
  - **Optimal Ordering Provisions in Sequential Search Problems**

Three Lines of Action (WIP)

- Establish connections between Machine Learning and Econ Theory
  - Hannan Consistency and Convergence to Best Reply [Gonzalez, 2023]
  - An Epistemological Approach to Causality in ML [Gonzalez and Mazlish, 2024]
- Expand Economic Theory leveraging ML heuristics and Algorithms
  - A Prior-Free Theory of Adverse Selection [Gonzalez, 2023]
  - The Econometrics of Behavioral Parameters [Auer et al., 2024]
  - **Optimal Ordering Provisions in Sequential Search Problems**
- The Economics of Machine Learning
  - Rationalizing Upper Confidence Bound Algorithms [Gonzalez, 2024]
  - Using Neural Networks to Recover Utility Functions [Gonzalez and Wu, 2024]

Stylized Problem

- There exists a long-lived welfare-maximizer **principal** who optimally trades-off exploration and exploitation across different options $j, h, \dots$

- And a short-lived **agent** who (i) conducts the actual exploration, but (ii) who just cares about exploitation. **Delegated Exploration Problem**

- P is the **platform owner**, but she can't force A to pick a particular option. Instead, she provides an **ordering of all options** $\{jhk, hjk, kjh, \dots\}$

Stylized Problem

- There exists a long-lived welfare-maximizer **principal** who optimally trades-off exploration and exploitation across different options $j, h, \ldots$

- And a short-lived **agent** who (i) conducts the actual exploration, but (ii) who just cares about exploitation. **Delegated Exploration Problem**

- P is the **platform owner**, but she can't force A to pick a particular option. Instead, she provides an **ordering of all options** $\{jhk, hjk, kjh, \ldots\}$

- Agent will query such options sequentially and select one

- Can we characterize a **(near)-optimal ordering policy** when only observing the actions of the agent and not the rewards?

Just another Bandit Paper?

- **Methodologically relevant:**
  - **Blurred trade-off** between exploration and exploitation (as every element can be observed in any sequence)
  - **Cross-exploration.** Need to query different arms to recover relevant statistics of desired arms
  - **New information structures**

Just another Bandit Paper?

- **Methodologically relevant:**
  - **Blurred trade-off** between exploration and exploitation (as every element can be observed in any sequence)
  - **Cross-exploration.** Need to query different arms to recover relevant statistics of desired arms
  - **New information structures**
- **Economically relvant:**
  - Agent's exploration is mediated by a **dynamic search problem**
  - Important economic applications like **job listings**

Stylized Model

- **Public Officer** wants to match **workers** $i = 1, \ldots, N$ and **firms** $J \in \{j, h\}$, where the quality of the worker-firm match is unknown

A	B	C	D	E
i arrives	**P**: Order $p^i \in \{jh, hj\}$	i plays "Worker's Problem"	Feedback is realized	**P** updates strategy

Workers' Problem

- When i visits firm J, he observes $m^i_J = \mu^J + \gamma_i + \varepsilon^J_i, \;\; \varepsilon^J_i \sim M^J$
- Outside option ϕ_i
- Expected value of exploration after observing first firm is $m^{J'}_{0i} = \mathbb{P}_{0i}(M^J_i \geq 0) \cdot \mathbb{E}_{0i}[M^J_i \mid M^J_i \geq \phi_i] + \mathbb{P}_{0i}(M^J_i < 0) \cdot \phi_i.$

Workers' Problem

- When i visits firm J, he observes $m^i_J = \mu^J + \gamma_i + \varepsilon^J_i, \;\; \varepsilon^J_i \sim M^J$

- Outside option ϕ_i

- Expected value of exploration after observing first firm is $m^{J'}_{0i} = \mathbb{P}_{0i}(M^J_i \geq 0) \cdot \mathbb{E}_{0i}[M^J_i \mid M^J_i \geq \phi_i] + \mathbb{P}_{0i}(M^J_i < 0) \cdot \phi_i.$

- Under two "small" **simplifying conditions** we may assume that $m^J_{0i} = m^J_0, m^J_i = \mu^J + \varepsilon^J_i, \phi_i = 0$

(Simplified) Worker's Problem

$$a_i^{jh} = \begin{cases} T & \text{if } m_i^j \geq m_0^h \\ \{C, T\} & \text{if } m_i^j < m_0^h \ \& \ m_i^h \geq 0 \\ \{C, C\} & \text{if } m_i^j < m_0^h \ \& \ m_i^h < 0 \end{cases} \tag{1}$$

- Rewards given by

$$r_i^{jh} = \mathbb{1}(m_i^j \geq m_0^h) \cdot m_i^j + \mathbb{1}(m_i^j < m_0^h, m_i^h \geq 0) \cdot m_i^h \tag{2}$$

- **Other assumptions**: No participation cost (no IR), no discounting, risk neutrality, **agents can't go back**, they only get to play once, present bias if indifferent

Principal's Problem

- Maximize expected welfare = Minimize Stochastic Regret

$$
\begin{aligned}
\arg\max_{\pi} \mathbb{E}\left[\sum_{i}^{N} r_i^{\pi(i)}\right] &= \arg\min_{\pi} N \cdot \mathscr{r}^{p^*} - \mathbb{E}\left[\sum_{i}^{N} r_i^{\pi(i)}\right] \\
&= \arg\min_{\pi} \mathscr{R}_N(\pi) \qquad (3)
\end{aligned}
$$

Intuitions

- In a nutshell, P needs to learn $V = \{\mu^j, \mu^h, m_0^j, m_0^h\}$ from agents decisions

- Parametric assumptions st V is identified from these decisions

$$q^1 = \mathbb{P}(a_{1i}^{jh} = T), \;\; \hat{q}^1 = \frac{\sum_i \mathbb{1}(a_{1i}^{jh} = T, p^i = jh)}{\sum_i \mathbb{1}(p^i = jh)} \tag{4}$$

Intuitions

- In a nutshell, P needs to learn $V = \{\mu^j, \mu^h, m_0^j, m_0^h\}$ from agents decisions

- Parametric assumptions st V is identified from these decisions

$$q^1 = \mathbb{P}(a_{1i}^{jh} = T), \;\; \hat{q}^1 = \frac{\sum_i \mathbb{1}(a_{1i}^{jh} = T, p^i = jh)}{\sum_i \mathbb{1}(p^i = jh)} \tag{4}$$

- Assuming $\varepsilon^J \sim Log(0, 1) \implies$

$$\mu^J = \ln\left(\frac{q^k}{1 - q^k}\right), \;\; m_0^J = \mu^{J'} - \ln\left(\frac{q^l}{1 - q^l}\right) \tag{5}$$

- This allows us to rewrite $\mathbb{E}[r^p] = f(q^1, q^2, q^3, q^4)$

UNIVERSITY OF OXFORD

Challenges

- If $\mathbb{E}[r^p] = f(q^1, q^2, q^3, q^4)$ this means that we need to sample p' to recover consistent estimates of $\hat{r}^p \implies$ **Cross-exploration**
- "Q-space" is non-monotonic (as opposed to the reward space). Adjust classic algorithms in non-trivial ways
- $f(Q)$ is very much not well-behaved (non-Lipschitz)

Some Results - To be verified

Proposition 1: Near Optimality under Partial Feedback

Let $k = \mid Q \mid$, then a version of UCB yields $\mathcal{R}_N = \mathcal{O}(2^k \cdot \ln(N))$

- **Same regret as under full-feedback (?!)** with worse constant $(2^k, k = 3)$ and at the expense of parametric assumptions. Still not bad

Moving Forward

- Exploit internal searching structure. Some initial results
- Extend analysis to $J > 2$ arms
- Refine bounds
- Data Application
- Interplay between Information and Sequencing strategies

Bibliography I

- Auer, J., Ballester, M., and Gonzalez, C. (2024).
  The econometrics of behavioral parameters.
- Gonzalez, C. (2023).
  Adaptive wage setting: A prior-free theory of adverse selection and monopsony markets.
- Gonzalez, C. (2024).
  Rationalizing upper confidence bound algorithms.
- Gonzalez, C. and Mazlish, J. Z. (2024).
  An epistemological approach to causility in machine learning.
- Gonzalez, C. and Wu, H. (2024).
  Using neural networks to recover utility functions.