# Partial identification via conditional linear programs

## with an application to learning individualized treatment rules

Eli Ben-Michael

Carnegie Mellon University

## Overview

Problem: estimate expectation of unknown linear functions

$$\theta = \mathbb{E}[\langle \boldsymbol{c}(X), \boldsymbol{p}^*(X)\rangle]$$ see also [1, 2]

- Know $J$ constraints on $K$ variables $\boldsymbol{p}^*(x)$: $A\boldsymbol{p}^*(x) = \boldsymbol{b}(x)$
- $\boldsymbol{b}(x)$ and $\boldsymbol{c}(x)$ are identifiable, and $\boldsymbol{p}^*(x) \geq 0$

For each value of covariates $x$, have pair of conditional LPs

$$\theta_L(x) = \min_{A\boldsymbol{p}=\boldsymbol{b}(x), \boldsymbol{p}\geq 0} \langle \boldsymbol{c}(x), \boldsymbol{p}\rangle \text{ and } \theta_U(x) = \max_{A\boldsymbol{p}=\boldsymbol{b}(x), \boldsymbol{p}\geq 0} \langle \boldsymbol{c}(x), \boldsymbol{p}\rangle$$

Sharp, covariate-assisted bounds: $\mathbb{E}[\theta_L(X)] \leq \theta \leq \mathbb{E}[\theta_U(X)]$

This work: estimation of (regularized) bounds + policy learning

## Application: collective utility functions

Setting: RCT or Obs. Study (i.e. ignorability + overlap)

- Covariates $X$, binary decision $D$
- Discrete potential outcomes $Y(0), Y(1)$; $\tau = Y(1) - Y(0)$
- Stochastic treatment policy $\pi(X)$

Individual expected utility under $\pi$ as a stochastic intervention

$$Y(\pi(X)) = Y(0) + \pi(X) \times \tau$$

Collective utility functions to aggregate individual utilities [3]

$$V^\lambda(\pi) = \frac{1}{\lambda}\mathbb{E}[Y(\pi(X))^\lambda - 1]$$

Egalitarian rule | 0 | Arithmetic mean utility | $\infty$

$-\infty$ | Geometric mean utility | 1 | Inequality-seeking

Nash cooperative bargaining value function

$$V^N(\pi) = \mathbb{E}[\{\tau > 0\} \log \pi(X) + \{\tau < 0\} \log(1 - \pi(X))]$$

Log-loss for positive/negative treatment effect

- Unconstrained opt: $\pi(x) = P(\tau > 0 \mid X = x, \tau \neq 0)$
- Bargaining + rationality conditions + affine invariance [4]

Goal: minimize regret relative to oracle w/knowledge of POs

- Oracle: $\pi^o = (\tau \geq 0)$; Regret $R(\pi) = V(\pi^o) - V(\pi)$
- Requires knowledge of joint distribution of $Y(0), Y(1)$
- Margins $A\boldsymbol{p}(x) = \boldsymbol{b}(x)$; regret for PO pairs $\boldsymbol{c}(\pi(x))$

Find minimax regret policy $\pi^* \in \text{argmin}_{\pi\in\Pi} \max V(\pi^o) - V(\pi)$

## References


[1] Vira Semenova. Aggregated Intersection Bounds and Aggregated Minimax Values. arXiv:2303.00982, 2024.

[2] Wenlong Ji, Lihua Lei, and Asher Spector. Model-Agnostic Covariate-Assisted Inference on Partially Identified Causal Effects. arXiv:2310.08115, 2024.

[3] Hervé Moulin. *Axioms of Cooperative Decision Making*. Cambridge, 1988.

[4] Mamoru Kaneko and Kenjiro Nakamura. The Nash Social Welfare Function. *Econometrica*, 47(2):423–435, 1979.

[5] Jonathan Weed. An explicit analysis of the entropic penalty in linear programming. In *Proceedings of the 31st Conference On Learning Theory*, pages 1841–1855, 2018.

[6] Amy Finkelstein, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group. The Oregon Health Insurance Experiment: Evidence from the First Year. *The Quarterly Journal of Economics*, 127(3):1057–1106, 2012.


## De-biased estimation of bounds from conditional linear programs

### Plugin basic feasible solutions

Solution in terms of optimal basis $\mathcal{B}^*_U(x) = \{i_1, \ldots, i_J\}$

$$\boldsymbol{p}_U(x) = A^{-1}_{\mathcal{B}^*_U(x)}\boldsymbol{b}(x) \Rightarrow \theta_U = \mathbb{E}\left[\langle \boldsymbol{c}(x), A^{-1}_{\mathcal{B}^*_U(x)}\boldsymbol{b}(x)\rangle\right]$$

Plugin optimal basic feasible solution:

$$\widehat{B}_U(x) \in \text{argmax}_{B\in\mathcal{B}}\langle \hat{\boldsymbol{c}}(x), A_B^{-1}\hat{\boldsymbol{b}}(x)\rangle$$

- Directly read off simplex algorithm (avg polynomial time)

De-biased estimator w/plugin basic feasible solution:

$$\hat{\theta}_U = \hat{\mathbb{E}}\left[\langle \hat{\boldsymbol{c}}(X) + \hat{\varphi}_c, \hat{\boldsymbol{p}}_U(X)\rangle + \langle \hat{\boldsymbol{c}}(X), A^{-1}_{\widehat{B}_U(X)}\hat{\varphi}_b\rangle\right]$$

Margin condition: $P(\text{best sol'n - 2nd best sol'n} \leq t) \leq t^\alpha$

$$\left|\mathbb{E}[\hat{\theta}_U - \theta_U]\right| \lesssim \left(\|\hat{\boldsymbol{b}} - \boldsymbol{b}\|_\infty + \|\hat{\boldsymbol{c}} - \boldsymbol{c}\|_\infty\right)^{1+\alpha} + o_p(n^{-1/2})$$

### Entropic regularization

Entropic-regularized solution:

$$\boldsymbol{p}^\eta_U(x) = \underset{A\boldsymbol{p}=\boldsymbol{b}(x)}{\text{argmax}} \langle \boldsymbol{c}(X), \boldsymbol{p}\rangle + \frac{1}{\eta}\text{Entropy}(\boldsymbol{p})$$

Plugin solution in terms of dual variables $\boldsymbol{d}(\boldsymbol{b}(x), \boldsymbol{c}(x))$

$$\hat{p}^\eta_U(x) = \exp\left(A'\boldsymbol{d}(\hat{\boldsymbol{b}}(x), \hat{\boldsymbol{c}}(x)) + \eta\hat{\boldsymbol{c}}(x)\right)$$

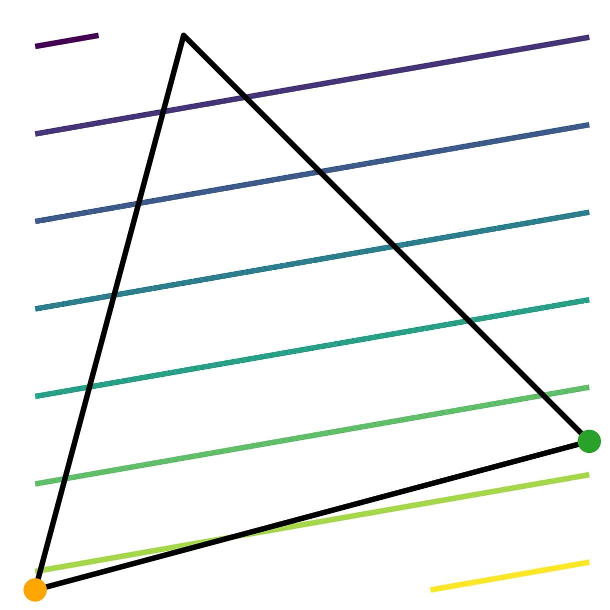
- Strongly convex, unconstrained, fast w/Sinkhorn algo

De-biased estimator w/entropic regularized solution:

$$\hat{\theta}^\eta_U = \hat{\mathbb{E}}[\langle \hat{\boldsymbol{c}}(X) + \hat{\varphi}_c, \hat{\boldsymbol{p}}^\eta_U(X)\rangle + \langle \hat{\boldsymbol{c}}(X), \nabla_b\hat{\boldsymbol{p}}^\eta_U(X)\hat{\varphi}_b + \nabla_c\hat{\boldsymbol{p}}^\eta_U(X)\hat{\varphi}_c\rangle]$$

If regularization penalty $\frac{1}{\eta}$ is small enough[5] relative to margin

$$\left|\mathbb{E}\left[\hat{\theta}^\eta_U - \theta_U\right]\right| \lesssim e^{-\eta} + o_p(n^{-1/2})$$



## Learning minimax regret policies

Unregularized minimax regret policy:

$$\hat{\pi} = \underset{\pi\in\Pi}{\text{argmin}}\ \hat{\mathbb{E}}\left[\langle \boldsymbol{c}(\pi(X)), A^{-1}_{\widehat{B}_U(X)}(\hat{\boldsymbol{b}}(X) + \hat{\varphi}_b)\rangle\right]$$

Excess regret $\lesssim$ Complexity$(\Pi)$ + $\|\hat{\boldsymbol{b}} - \boldsymbol{b}\|^{1+\alpha}_\infty$

Regularized minimax regret policy:

$$\hat{\pi}^\eta = \underset{\pi\in\Pi}{\text{argmin}}\ \hat{\mathbb{E}}[\langle \boldsymbol{c}(\pi(X)), \hat{\boldsymbol{p}}^\eta_U(X) + \nabla_b\hat{\boldsymbol{p}}^\eta_U(X)\hat{\varphi}_b\rangle]$$

Excess regret $\lesssim$ Complexity$(\Pi)$ + regularization bias

## Empirical illustration: Oregon health insurance experiment

Lottery for Medicaid enrollment[6]

- $X$: socioeconomic + health characteristics
- $D$: Medicaid offer; $U(d)$: -(# of ED visits)

Overall ITT: Medicaid offer ↑ ED visits by 12%

Q How suboptimal was random assignment?

Q What other targeting rules minimize maximum regret?