Optimal weighting for estimating treatment effects

Michele Santacatterina
santam13@nyu.edu

New York University

Dec 15, 2022

Task

- Estimate the effect of a treatment, *e.g.*, ATE, on an outcome of interest using observational data
- Observational data are affected by confounding

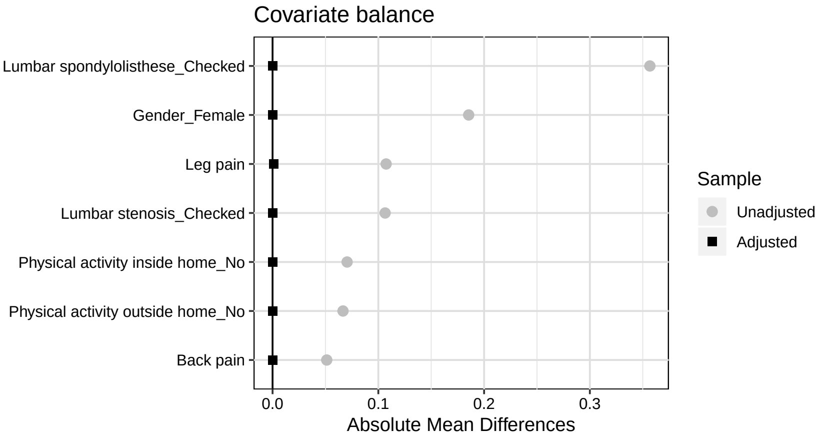
Covariate balance

How to control for confounding? *Covariate balance*

- the means of covariates between treatment groups should be the *same*
- the two treatment groups are similar except for the assignment of the treatment

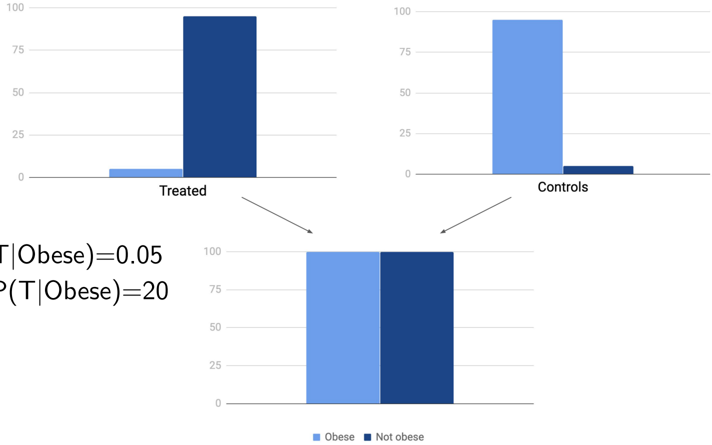
Consequently, potential differences can only be attributed to the treatment instead of other confounding factors.

Surgical Interventions on ODI



The absolute mean difference of covariates across treatment groups is null after weighting.

Inverse probability weighting



$P(T|Obese)=0.05$

$1/P(T|Obese)=20$

IPW is the unique set of weights that **balance the covariate distributions** of different treatment groups.

IPW issues

IPW can lead to poor covariate balance due to

- Extreme weights (lack of overlap)
- Model misspecification

IPW issues

Proposed solutions

- Extreme weights $\rightarrow$ post-hoc fixes, *e.g.*., truncation
- Model misspecification $\rightarrow$ flexible modeling, *e.g.*., machine learning

An alternative approach

Find weights that optimize covariate balance in the data at hand.

Generalized Average Treatment Effect

We consider GATE defined as

$$\tau_V = \sum_{i=1}^{n} V_i(g_1(X_i) - g_0(X_i)) \tag{1}$$

where V_i is chosen to target the estimand of interest. For instance, setting $V_i = \frac{1}{n}$ would target SATE. Let $X_i \in \mathcal{X}$ be the observed confounders

A weighted estimator for GATE

To estimate GATE in eq. (1) we propose to use the following weighted estimator

$$\hat{\tau}_W = \sum_{i \in \mathcal{T}} W_i Y_i - \sum_{i \in \mathcal{C}} W_i Y_i = \sum_{i=1}^{n} (-1)^{(T_i+1)} S_i W_i Y_i. \qquad (2)$$

where $T_i = \mathbb{I}\left[i \in \mathcal{T}\right]$ is the indicator of being treated with $t = 1$, and $S_i = \mathbb{I}\left[i \in \mathcal{S}\right]$ is the indicator of being in the labeled sample.

Error in estimation

$$\overbrace{E\left[\left(\hat{\tau}_{\mathsf{W}} - \tau_V\right)^2 \mid H_{1:n}\right]}^{\text{CMSE}} = \left[\underbrace{B_1(W_{1:n}, V_{1:n}, g_1)}_{\text{imbalance in } g_1} - \underbrace{B_0(W_{1:n}, V_{1:n}, g_0)}_{\text{imbalance in } g_0}\right]^2 + \underbrace{\sum_{i=1}^{n} S_i W_i^2 \sigma_i^2}_{\text{noise}}.$$

where $B_t(W_{1:n}, V_{1:n}, g_t) = \sum_{i=1}^{n}\left(S_i \mathbb{I}\left[T_i = t\right] W_i - V_i\right) g_t(X_i)$

Worst-case imbalance

By the representer theorem,

$$\begin{aligned}\Delta_t^2(W_{1:n}, V_{1:n}) &= \sup_{\|g_t\|_t \le 1} B_t(W_{1:n}, V_{1:n}, g_t) \\ &= \sup_{\|g\|_t^2 \le 1} \left(\sum_{i=1}^{n} \left(1[S_i = s] 1[T_i = t] W_i - V_i \right) g_t(X_i) \right)^2 \\ &= {W_{1:n}}^T \underbrace{I_S I_t K_t I_S I_t}_{Q} W_{1:n} - 2 {V_{1:n}}^T \underbrace{K_t I_S I_t}_{c} W_{1:n} \\ &+ {V_{1:n}}^T K_t V_{1:n}.\end{aligned}$$

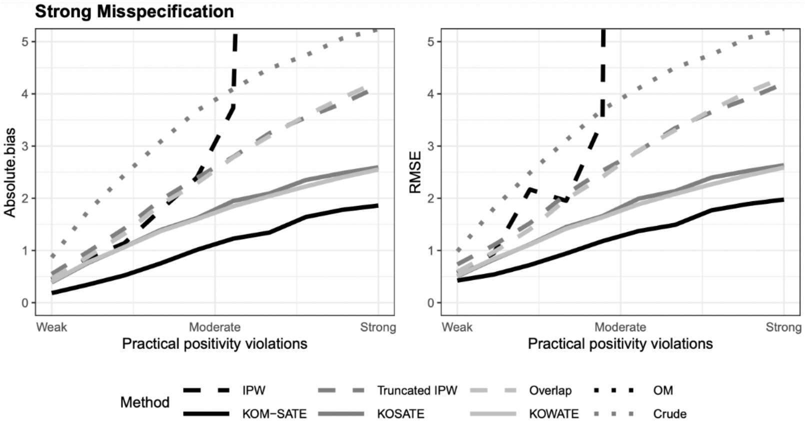
where the matrix $K_t \in \mathbb{R}^{n \times n}$ is defined as $K_{tij} = \mathcal{K}_t(X_i, X_j)$.

Kernel Optimal Matching (KOM)

Linearly-constrained convex-quadratic optimization problem,

$$\min_{\substack{W_{1:n}\geq 0,\\ W_{1:n}^T I_S I_1 e_n = n,\\ W_{1:n}^T I_S I_0 e_n = n}} \frac{1}{n^2}\left({W_{1:n}}^T Q W_{1:n} - 2{V_{1:n}}^T c W_{1:n}\right)$$

Simulations



Surgical Interventions on ODI

We applied KOM in the evaluation of two spine surgical interventions on the Oswestry Disability Index (ODI).

Table: The effect of fusion-plus-laminectomy on ODI

	ATE	Unadjusted
$\hat{\tau}_W$ (SE)	1.33 (3.98)	5.09* (2.31)

* indicates statistical significance at the 0.05 level.

Time-dependent confounding

An MSM is a model for the marginal causal effect of a *time-varying* treatment regime on the mean of Y, that is,

$$E\left[Y(\overline{a})\right] = g(\overline{a}, \beta), \tag{3}$$

where $g(\overline{a}, \beta)$ is some known function class parametrized by β.

- Q: <u>How do we balance time-dependent confounders?</u>

Kernel Optimal Weighting

$$E\left[W1[\overline{A}=\overline{a}]Y\right]-E\left[Y(\overline{a})\right]=\sum_{t=1}^{T}\delta_{a_t}^{(t)}(W,g_{\overline{a}}^{(t)}).$$

Imbalances for any time $t\geq 3$ as

$$\begin{aligned}\delta_{a_t}^{(t)}(W,h^{(t)})&=E\left[W1[A_t=a_t]h^{(t)}(\overline{A}_{t-1},\overline{X}_t)\right]\\&-E\left[Wh^{(t)}(\overline{A}_{t-1},\overline{X}_t)\right]\end{aligned}$$

KOW

$$\min_{W_{1:n} \in \mathcal{W}} \quad \frac{1}{2} W_{1:n}^T K_\lambda^\circ W_{1:n} - e^T K_\lambda W_{1:n} \tag{4}$$

KOW

Table: Effect of HIV treatment on time to death.

	KOW		Logistic	
	$\mathcal{K}_1$	$\mathcal{K}_2$	IPTCW	sIPTCW
$\hat{HR}$	0.40*	0.48*	0.14	1.25
SE	(0.30)	(0.28)	(1.15)	(0.30)

* indicates statistical significance at the 0.05 level.

Publications

- Kallus, N., **Santacatterina, M.**, *Optimal Weighting for Estimating Generalized Average Treatment Effects*, 2022, Journal of Causal Inference, 10(1), pp.123-140.
- Kallus, N., Pennicooke, B., **Santacatterina, M.**, *More robust estimation of sample average treatment effects using Kernel Optimal Matching in an observational study of spine surgical interventions*, 2021, Statistics in medicine, 40(10), pp.2305-2320.
- Kallus, N. and **Santacatterina, M.**, 2021. *Optimal balancing of time-dependent confounders for marginal structural models*. Journal of Causal Inference, 9(1), pp.345-369.

Publications

Controlling for extreme weights.

- **Santacatterina, M.**, Bottai, M., 2018. *Optimal probability weights for inference with constrained precision*. JASA, 113(523), pp.983-991.
- **Santacatterina, M.**, et.al., 2019. *Optimal probability weights for estimating causal effects of time-varying treatments with marginal structural Cox models*. Statistics in medicine, 38(10), pp.1891-1902.

Publications

Dealing with continuous treatments and time-to-event data

- Kallus, N. and **Santacatterina, M.**, 2020. *Kernel Optimal Orthogonality Weighting: A Balancing Approach to Estimating Effects of Continuous Treatments*. arXiv preprint arXiv:1910.11972.
- **Santacatterina, M.**, 2022. *Robust weights that optimally balance confounders for estimating marginal hazard ratios*, SMMR.

Questions?

Contact me at **santam13@nyu.edu**

Visit my website **https://michelesantacatterina.github.io/**