ECMA31100 Introduction to Empirical Analysis II

Winter 2022, Week 9: Discussion Session

Conroy Lau

The University of Chicago

March 10, 2022

Welcome to week 9 of the winter quarter!

Topics for this TA session

- Similar to last week, I will cover two topics:
  - Synthetic control
  - Wild bootstrap
- I will be skipping technical details due to time constraints
- These two topics were mentioned only very briefly in lectures
- The goal here is to provide an overview of the methods
- I also provide some remarks on treatment choice based on last week's discussion

Contents

Some relevant papers on treatment choice

Extensions?

- Sun (2021): extends to the case with unknown budget/cost
- Kitagawa and Tetenov (2021): considers an objective that is rank-dependent

Some other papers

- Armstrong and Shen (2015): inference on first-best rules
- Athey and Wager (2021): doubly-robust estimator and asymptotic bounds
  - The two papers we saw in week 8 derived finite-sample regret bounds

Contents

Synthetic control

Set-up

- Data: $j = 1, \ldots, J+1$ ($j = 1$ is treated, and remaining are untreated)
- Intervention period: T_0
- Pre-period outcome and covariates for each unit j: $X_j \equiv (X_{1j}, \ldots, X_{kj})'$
  - X_1 is the vector of pre-period of outcome and covariates for the treated unit
  - $X_0 \equiv (X_2, \ldots, X_{J+1})$ is the corresponding $k \times J$ matrix for the untreated units
- Potential outcomes without intervention for unit j in period T: Y_{jt}^N
- Potential outcome in post-intervention period $t > T_0$: Y_{1t}^I

Goal

- We want to estimate the effect of intervention on unit 1: $\tau_{1t} = Y_{1t}^I - Y_{1t}^N$

Synthetic control

Optimization problem

- Let $w_j \geq 0$ be a potential synthetic control for unit $j = 2, \ldots, J+1$ s.t. $\sum_{j=2}^{J+1} w_j = 1$
- Stack them together $w \equiv (w_2, \ldots, w_{J+1})'$
- We obtain the optimal weighting by solving the optimization problem below:

$$
\begin{aligned}
w^\star \equiv \underset{w \in \mathbb{R}^J}{\arg\min} \quad & \|X_1 - X_0 w\|_V \\
\text{s.t.} \quad & \sum_{j=2}^{J+1} w_j = 1 \\
& w_2, \ldots, w_{J+1} \geq 0
\end{aligned}
$$

Synthetic control

More about the objective function

- The objective function can be written explicitly as

$$\|X_1 - X_0 w\|_V = \sqrt{(X_1 - X_0 w)' V (X_1 - X_0 w)} = \sqrt{\sum_{h=1}^{k} v_h \left(X_{h1} - \sum_{j=2}^{J+1} w_j X_{hj} \right)^2}$$

Treatment effect estimate

$$\widehat{\tau}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j^\star Y_{jt}$$

Synthetic control

Weighting matrix V

- The weight $w^\star$ is dependent on the choice of V
- That is, each potential choice of V produce a set of synthetic controls

$$w(V) \equiv (w_2(V), \ldots, w_{J+1}(V))$$

Abadie and Gardeazabal (2003) and Abadie et al. (2010)

- Let $\mathcal{T}_0 \subseteq \{1, 2, \ldots, T_0\}$ be the set of preintervention periods
- They choose V such that $w(V)$ minimize the

$$\sum_{t \in \mathcal{T}_0} \left(Y_{1t} - w_2(V)Y_{2t} - \cdots - w_{J+1}(V)Y_{J+1,t}\right)^2,$$

i.e., minimize the mean squared prediction error of synthetic control w.r.t. Y_{1t}^N

Penalized synthetic control (Abadie and L'Hour, 2021)

Optimization problem

$$
\begin{aligned}
w^\star \equiv \underset{w \in \mathbb{R}^J}{\arg\min} \quad & \left\| X_1 - \sum_{j=2}^{J} w_j X_j \right\|^2 + \lambda \sum_{j=2}^{J} w_j \|X_i - X_j\|^2 \\
\text{s.t.} \quad & \sum_{j=2}^{J+1} w_j = 1 \\
& w_2, \ldots, w_{J+1} \geq 0
\end{aligned}
$$

Interpretation

- $\lambda \longrightarrow 0$: pure synthetic control that min. pairwise matching discrepancies
- $\lambda \longrightarrow \infty$: nearest neighbor matching
- $\lambda > 0$: trade-off between fitting well and minimizing sum of pairwise distance

Any questions?

Contents

The bootstrap

Idea

- The bootstrap was introduced by Efron (1979)
- We want to estimate the distribution of estimator by resampling the data
- Some of the discussion here follows Horowitz (2019) (highly recommended)

Notations

- Random variable and data: X and $\{X_i\}_{i=1}^n$
- Distribution function of X: F_0
- Empirical distribution using the sample: $F_n(x) \equiv \frac{1}{n}\sum_{i=1}^n \mathbb{1}[X_i \leq x]$
- Object of interest: $T_n(X_1, \ldots, X_n)$
- Distribution of T_n using data sampled from distribution F: $G_n(\cdot, F)$

Nonparametric bootstrap

Idea

1. Do the followings for B times:
   (a) Sample the original data with replacement to get $\{X_i^\star\}_{i=1}^n$
   (b) Compute $T_n^\star = T_n(X_1^\star, \ldots, X_n^\star)$
2. Compute the probability of $T_n^\star \leq \tau$ using these B draws for each τ

Consistency

- Based on the above procedure, we have $\mathbb{P}[T_n \leq \tau] = G_n(\tau, F_0)$
- The bootstrap approximates $G_n(\tau, F_0)$ by $G_n(\tau, F_n)$
- Let P_n be the probability distribution based on the sample $\{X_i\}_{i=1}^n$
- The bootstrap estimator $G_n(\cdot, F_n)$ is consistent if for all $\epsilon > 0$ and F_0, we have

$$\lim_{n\to\infty} \mathbb{P}_n \left[\sup_\tau |G_n(\tau, F_n) - G_\infty(\tau, F_0)| > \epsilon \right] = 0$$

(Nonreplacement) subsampling

Idea

- Proposed by Politis and Romano (1994)
- They show it consistently est. the distribution of statistic under weaker conditions

Some details

- Let $t_n \equiv t_n(X_1, \ldots, X_n)$ be an estimator of the population parameter θ
- Let ρ_n be a normalizing factor s.t. $G_n(\tau, F_0) \equiv \mathbb{P}[\rho_n(t_n - \theta) \leq \tau] \longrightarrow G_\infty(\tau, F_0)$
  - Requires this to hold at nondegenerate $G_\infty(\tau, F_0)$ at points τ where G_∞ is continuous
- Let $N_{nm} = \binom{n}{m}$ to be the total number of size $m < n$ subsets that can be formed
- We estimate $G_n(\tau, F_0)$ by (t_{mk} means the estimator t_m at the kth subset)

$$G_{nm}(\tau) = \frac{1}{N_{nm}} \sum_{k=1}^{N_{nm}} \mathbb{1}[\rho_m(t_{mk} - t_n) \leq t)]$$

The Wild bootstrap

Idea

- Originally introduced by Wu (1986)
- I focus on the clustered linear regression model: $Y_{ij} = X_{ij}\beta + U_{ij}$
- Assume that there are J clusters in total

Rademacher distribution

- $R_i \in \{-1, 1\}$
- $\mathbb{P}[R_i = 1] = \mathbb{P}[R_i = -1] = \frac{1}{2}$

Further readings

- Canay et al. (2021): Justify wild bootstrap with small number of large clusters

The Wild bootstrap

Null

- Given a null hypothesis H_0, let $\widehat{\beta}$ be the constrained OLS estimator under the null
- Let $\widehat{w}$ be the cluster-robust Wald statistic from constrained regression

Procedure

1. Compute the residuals $\widehat{U}_{ij} \equiv Y_{ij} - X_{ij}'\widehat{\beta}$ from the constrained regression
2. For each bootstrap replication $b = 1, \ldots, B$, do the followings:
   (a) Draw one independent Rademacher draws $R_j \in \{-1, 1\}$ for each cluster $j = 1, \ldots, J$
   (b) Create $U_{ij}^\star = R_j \widehat{U}_{ij}$ based on the J draws
   (c) Regress $Y_{ij}^\star \equiv X_{ij}\widehat{\beta} + U_{ij}^\star$ on X_{ij} to obtain the cluster-robust Wald statistic $\widehat{w}_b^\star$
3. Reject null at level α if $|\widehat{w}| > (1-\alpha)$-quantile of $\{|\widehat{w}_b|\}_{b=1}^B$

Any questions?

Contents

Final remarks

- I hope you have learned useful econometric tools from lectures and TA sessions
- Don't forget to submit the final problem set and the final project
- Feel free to email me if you have any questions

- Good luck with the spring quarter!
- Good luck with your pre-doc/Ph.D./industry job applications!!
- Good luck with your research!!!
- Good luck with everything!!!!

Thank you!

References 1

ABADIE, A., A. DIAMOND, AND J. HAINMUELLER (2010): "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 105, 493–505.

ABADIE, A. AND J. GARDEAZABAL (2003): "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review*, 93, 113–132.

ABADIE, A. AND J. L'HOUR (2021): "A Penalized Synthetic Control Estimator for Disaggregated Data," *Journal of the American Statistical Association*, 116, 1817–1834.

ARMSTRONG, T. AND S. SHEN (2015): "Inference on Optimal Treatment Assignments," *Working paper*.

References 2

ATHEY, S. AND S. WAGER (2021): "Policy Learning With Observational Data," *Econometrica*, 89, 133–161.

CANAY, I. A., A. SANTOS, AND A. M. SHAIKH (2021): "The Wild Bootstrap with a "Small" Number of "Large" Clusters," *The Review of Economics and Statistics*, 103, 346–363.

EFRON, B. (1979): "Bootstrap Methods: Another Look at the Jackknife," *The Annals of Statistics*, 7, 1 – 26.

HOROWITZ, J. L. (2019): "Bootstrap Methods in Econometrics," *Annual Review of Economics*, 11, 193–224.

KITAGAWA, T. AND A. TETENOV (2021): "Equality-Minded Treatment Choice," *Journal of Business & Economic Statistics*, 39, 561–574.

References 3

Politis, D. N. and J. P. Romano (1994): “The Stationary Bootstrap,” *Journal of the American Statistical Association*, 89, 1303–1313.

Sun, L. (2021): “Empirical Welfare Maximization with Constraints,” *Working Paper*.

Wu, C. F. J. (1986): “Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis,” *The Annals of Statistics*, 14, 1261 – 1295.