Assessing the Performance of Matching Methods in Observational Studies

Jack Potrykus

April 26, 2022

Abstract

Contents

1 Introduction

In the context of observational data, a bipartite matching problem is minimally defined by:

- X, an $n \times p$ matrix of feature values;
- $\mathbf{z}$, an n-vector of binary treatment assignments;
- $d : \mathbb{R}^p \times \mathbb{R}^p \mapsto \mathbb{R}^+$, a function which computes some "distance" or "cost of matching" between any two row vectors $\mathbf{x_i}$ and $\mathbf{x_j}$ of X.

$\mathcal{S}$, $\mathbf{b}$, A, β test Iacus, King, and Porro 2011

2 Literature Review

2.1 Measuring Similarity

Iacus, King, Porro 2011 - set the scene

2.1.1 Propensity Scoring

Rosenbaum and Rubin, 1983

- The *propensity score* is most often accredited to ...
- They offer a more general definition of a score: anything affording conditional independence
- Score seeks to replicated randomized trial
- Review of properties (particularly: when can we make unbiased estimates of ATE)
- Proposed use cases

Methods for calculating propensity scores: Garrido et al. 2014 Note: Blocking-based balance metrics

Extensions

- (Optimal) Caliper width: Austin 2011
- Prognostic score: Hansen 2008
- Miettinen score is the root of the above Miettinen 1976
- Joint use of Prognostic, Propensity, + Mahalanobis, Leacy and Stuart 2014
- Imai and Dyk 2004

2.1.2 (Coarsened) Exact Matching

- Iacus King Porro, 2011: MIB methods
- Iacus King Porro, 2012: Causal inference without balance checking

2.2 Balance Assessment

Garrido et al, 2014: don't use the outcome in the matching Garrido et al, 2014: balance in mean does not imply balance in scores

2.3 Matching Algorithms

(Rosenbaum 1989) (Ho et al. 2011) (Greedy) (Khan et al. 2016) (hungarian) (Munkres 1957) (hungarian) Hungarian – Munkres, Khan et al 2016

3 Python Package

4 Experiments

4.1 Data Generation

Papers with data generation:

- Austin 2011
- Stuart, Lee, and Leacy 2013

5 Results

6 Discussion

7 Conclusion

References

Austin, Peter C. (2011). "Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies". In: *Pharmaceutical Statistics* 10.2. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pst.433, pp. 150–161. ISSN: 1539-1612. DOI: 10.1002/pst.433. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/pst.433.

Garrido, Melissa M. et al. (Oct. 2014). "Methods for Constructing and Assessing Propensity Scores". In: *Health Services Research* 49.5, pp. 1701–1720. ISSN: 0017-9124, 1475-6773. DOI: 10.1111/1475-6773.12182. URL: https://onlinelibrary.wiley.com/doi/10.1111/1475-6773.12182.

Hansen, Ben B. (2008). "The Prognostic Analogue of the Propensity Score". In: *Biometrika* 95.2. Publisher: [Oxford University Press, Biometrika Trust], pp. 481–488. ISSN: 0006-3444. URL: https://www.jstor.org/stable/20441477.

Ho, Daniel et al. (June 14, 2011). "MatchIt: Nonparametric Preprocessing for Parametric Causal Inference". In: *Journal of Statistical Software* 42, pp. 1–28. ISSN: 1548-7660. DOI: 10.18637/jss.v042.i08. URL: https://doi.org/10.18637/jss.v042.i08.

Iacus, Stefano M., Gary King, and Giuseppe Porro (Mar. 2011). "Multivariate Matching Methods That Are Monotonic Imbalance Bounding". In: *Journal of the American Statistical Association* 106.493, pp. 345–361. ISSN: 0162-1459, 1537-274X. DOI: 10.1198/jasa.2011.tm09599. URL: http://www.tandfonline.com/doi/abs/10.1198/jasa.2011.tm09599.

Imai, Kosuke and David A van Dyk (Sept. 1, 2004). "Causal Inference With General Treatment Regimes". In: *Journal of the American Statistical Association* 99.467. Publisher: Taylor & Francis _eprint: https://doi.org/10.1198/016214504000001187, pp. 854–866. ISSN: 0162-1459. DOI: 10.1198/016214504000001187. URL: https://doi.org/10.1198/016214504000001187.

Khan, Arif et al. (Jan. 2016). "Efficient Approximation Algorithms for Weighted b-Matching". In: *SIAM Journal on Scientific Computing* 38.5, S593–S619. ISSN: 1064-8275, 1095-7197. DOI: 10.1137/15M1026304. URL: http://epubs.siam.org/doi/10.1137/15M1026304.

Leacy, Finbarr P. and Elizabeth A. Stuart (Sept. 10, 2014). "On the joint use of propensity and prognostic scores in estimation of the average treatment effect on the treated: a simulation study". In: *Statistics in Medicine* 33.20, pp. 3488–3508. ISSN: 02776715. DOI: 10.1002/sim.6030. URL: https://onlinelibrary.wiley.com/doi/10.1002/sim.6030.

Miettinen, O. S. (Dec. 1976). "Stratification by a multivariate confounder score". In: *American Journal of Epidemiology* 104.6, pp. 609–620. ISSN: 0002-9262. DOI: 10.1093/oxfordjournals.aje.a112339.

Munkres, James (Mar. 1957). "Algorithms for the Assignment and Transportation Problems". In: *Journal of the Society for Industrial and Applied Mathematics* 5.1, pp. 32–38. ISSN: 0368-4245, 2168-3484. DOI: 10.1137/0105003. URL: http://epubs.siam.org/doi/10.1137/0105003.

Rosenbaum, Paul R. (Dec. 1989). "Optimal Matching for Observational Studies". In: *Journal of the American Statistical Association* 84.408, pp. 1024–1032. ISSN: 0162-1459, 1537-274X. DOI: 10.1080/01621459.1989.10478868. URL: http://www.tandfonline.com/doi/abs/10.1080/01621459.1989.10478868.

Stuart, Elizabeth A., Brian K. Lee, and Finbarr P. Leacy (Aug. 2013). "Prognostic score–based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research". In: *Journal of Clinical Epidemiology* 66.8, S84–S90.e1. ISSN: 08954356. DOI: 10.1016/j.jclinepi.2013.01.013. URL: https://linkinghub.elsevier.com/retrieve/pii/S0895435613001625.