

第8講 傾向スコアによる政策評価

1 はじめに

先に述べたように、政策を実施した処理群のデータはあるが、対照群のデータは不在であり、外部情報から対処群をみつけてマッチングさせる必要が出てくる場合には次のような手法を用いる。

実際に外部データが十分にあり、処理群に含まれる個別サンプルの全ての変数、属性にぴったり一致するような対照群サンプルを選ぶことが出来れば、これを完全一致マッチング (exact matching) と呼ぶが、分析に用いる変数が増えるに従って、個別変数をマッチングさせることは難しくなる。このような場合、変数一つ一つをマッチングさせるのではなく、ある程度変数を集約して表現した条件付き確率 (これを傾向スコア ; propensity score と呼ぶ) を処理群と対照群でマッチングさせるという方法が考えられる¹。

具体的な考え方は、処理群に選ばれる確率 $[\text{Pr}[d_i = 1 | x]]$ を全サンプルを用いたロジット推定によって求め、それを処理群と対照群に分け、さらに確率を均等な階層に分け、同じ階層に入るもの同士をマッチングさせ、処理効果の平均を求めるというものである。ここでマッチングをどうするかということが問題になる。すなわち、一度マッチングに使った対照サンプルを再び使うことを認めるかどうか、比較対照するために処理サンプルに対していくつの対照サンプルを割り当てるのか。最も propensity score が近いもの一つを選べばいいのか (caliper matching と言う)、それともその周辺の対照サンプルを複数割り当てるのがいいのか、また、具体的なマッチングの方法としてどのようなものを用いるのか、といった問題がある²。

マッチングの詳細な手法はかなり技術的に高度になるので、ここでは扱わないが、主要な手法の基本的な考え方を紹介しておきたい³。一般に処理効果

¹この方法は Rosenbaum and Rubin (1983) によって開発され、現在では以下で紹介するように沢山の拡張が行われている。

²実際にこれらの問題にどう対処するかということは研究者の判断にゆだねられている。逆に、決定的に正しい方法が知られているわけではなく、試行錯誤するしかない。

³最新のマッチング手法のアルゴリズムに関しては、Abadie et al (2004)、Becker and Ichino (2002)、Becker and Caliendo (2007) 等を参照されたい。

は次のように表すことが出来る。

$$\Delta^M = \frac{1}{N_T} \sum_{i \in \{d=1\}} [y_{i1} - \sum_j w(i, j) y_{j0}]$$

ここで $\sum w(i, j) = 1$ 、 $0 < w(i, j) \leq 1$ となるマッチング・ウェイトである。 N_T は処理群のサンプル数を表す。

対照群としてどのようなサンプルを処理群にマッチさせるかという事であるが次のようなマッチング手法が提案されている。(1) 最近傍マッチング (nearest-neighbor matching) の考え方では、全ての処理サンプル i に対して、次のような条件を満たす集合 $A_i(x) = \{j | \min_j \|x_i - x_j\|\}$ を対照群として選択する⁴。(2) カーネル・マッチング (kernel matching) では、ウェイトを次のように定義する $w(i, j) = K(x_j - x_i) / \sum_{j=1}^{N_{ic}} K(x_j - x_i)$ 、ここで K はカーネル関数を表す。(3) 層化マッチング (stratification matching) とは propensity score を均等に層化し、層内で処理群と対照群が同じスコアになるようにした後で、処理効果を推定する。同じスコアのペアが組めない場合には、その層内での処理効果は計算されない。(4) 半径マッチング (radius matching) では対照群集合を次のように定義する。 $A_i(p(x)) = \{p_j | \|p_i - p_j\| < r\}$ すなわち、 propensity score の差が半径 r 以内であればペアとしてマッチングするという方法である。

2 傾向スコア法

balancing score (balancing score) とは次のように定義される。⁵

$$x \perp z | b(x) \quad (8.1)$$

それを条件付けすることにより、共変量と割り当てが独立になるような「共変量の関数」である。

すべてのbalancing scoreは $p(z = 1 | x) = g(b(x))$ と表現できることが必要。

$$\begin{aligned} p(z = 1 | b(x)) &= \int p(z = 1 | x \cdot b(x)) p(x | b(x)) dx \\ &= E_{x|b(x)} [p(z = 1 | x)] \\ &= E_{x|b(x) \cdot p(z=1|x)} [p(z = 1 | x)] \\ &= p(z = 1 | x) \\ &= p(z = 1 | x \cdot b(x)) \end{aligned} \quad (8.2)$$

⁴ $\|\cdot\|$ はベクトル間のユークリッド距離を表す。

⁵以下は、星野 (2009, pp.60-69) から引用。

(8.1) が成立するためには (8.2) の一段目と 2 段目の等号条件として $p(z = 1 | x \cdot b(x))$ が成り立つことが必要。強く無視できる割り当て条件 (strongly ignorable treatment assignment) [これは「割り当てはあくまで共変量にのみ依存し、結果変数には依存しない」ことを意味している。] が成立していると仮定すると、

$$\begin{aligned} p(z = 1 | y_1 \cdot y_0 \cdot b(x)) &= \int p(z = 1 | y_1 \cdot y_0 \cdot x \cdot b(x)) * p(x | y_1 \cdot y_0 \cdot b(x)) dx \\ &= E_{y_1 \cdot y_0 \cdot b(x)} [(p(z = 1 | y_1 \cdot y_0 \cdot x))] \\ &= p(z = 1 | x) \end{aligned}$$

$$\text{これを (8.2) とあわせて } P(z | y_1 \cdot y_0 \cdot b(x)) = P(z | b(x))$$

$$\implies (y_1 \cdot y_0) \Pi z | b(x)$$

(8.3)

バランススコアを条件付ければ、割り当て z と潜在的な結果変数 y_1, y_0 は独立になる。

傾向スコアの定義

第 i 対象者の共変量の値を x_i 割り当て変数の値を z_i とするとき、群 1 へ割り当てられる確率 e_i :

$$e_i = p(z_i = 1 | x_i)$$

を第 i 対象者の傾向スコアという ($0 \leq e \leq 1$)。

実証上はプロビットやロジットを用いて p を計算する。第 i 対象者の割り当て変数 z_i と共変量ベクトルを x_i とすると、 z_i が 1 となる確率をロジット推定する。

$$p(z_i = 1 | x_i) = e_i = \frac{1}{1 + \exp\{-\alpha^t x_i\}}$$

割り当てに関する尤度は

$$\prod_{i=1}^N \left(\frac{1}{1 + \exp\{-\alpha^t x_i\}} \right)^{z_i} \left(1 - \frac{1}{1 + \exp\{-\alpha^t x_i\}} \right)^{1 - z_i}$$

これを最大化する最尤推定値 $\hat{\alpha}$ を用いて、第 i 対象者の傾向スコアの推定値は

$$\hat{e} = \frac{1}{1 + \exp\{-\hat{\alpha}^t x_i\}} \text{ となる。}$$

傾向スコアを用いた共変量調整は、「強く無視できる割り当て」条件が満たされていれば、すべての共変量を用いて調整を行ったのと同じだけ偏りを減少させることができる。因果効果は、

$$E(y_1) - E(y_0) = E(y_1 - y_0) = E_e(E(y_1 - y_0 | e))$$

(8.3) より

$$E(y_1 | e) = E(y_1 | e, z = 1)$$

$$E(y_0 | e) = E(y_0 | e, z = 0)$$

結果として

$$E(y_1) - E(y_0) = E_e [E(y_1 | e, z = 1) - E(y_0 | e, z = 0)]$$

傾向スコアを所与とし (同じ傾向スコア値の元で) 各群から得られた y_1, y_0 から $y_1 - y_0$ を計算し、それを傾向スコアの分布で期待値をとると、もとも

との関心対象である因果効果になる。

傾向スコアと観測されている従属変数の情報を用いれば、因果効果を推定することができるということである。

「傾向スコアの値が等しい処置群と対照群のペアをマッチングすると、その差の平均は因果効果 ($E(y_1 - y_0)$) の不偏推定量となる」

傾向スコア 二段階推定法

(1) 傾向スコアの推定

割り当て変数 z を共変量 x によって説明するモデルを設定し、そのモデルの母数の推定を行う。母数の推定値を用いて、各対象者ごとに $z = 1$ に割り当てられる予測確率を計算し、これを傾向スコアの推定値とする。

(2) 傾向スコアを用いた調整

ローゼンバウム・ルービンの手法 マッチング、層別解析、共分散分析

マッチング

2つの群で傾向スコアの等しい対象者をペアにして、その差の平均をもって因果効果 $E(y_1) - E(y_0)$ の推定値とする。

最近傍マッチング、キャリパーマッチング (caliper matching)

傾向スコアという1次元の値でマッチングできることが利点。

層別解析

傾向スコアによっていくつかのサブクラス (通常は5つ) にわけ、その各クラスで処置群と対照群の平均の計算と、全体としての効果の推定量を計算する。

全体のサンプルサイズ N 、第 k サブクラスのサンプルサイズ N_k 、処置群の平均を \bar{y}_{1k} 、対照群の平均を \bar{y}_{0k} とすると、因果効果の推定量 \hat{d}_{01} は

$$\hat{d}_{01} = \sum_{k=1}^k \frac{N_k}{N} (\bar{y}_{1k} - \bar{y}_{0k}), \text{var}(\hat{d}_{01}) = \sum_{k=1}^k \left(\frac{N_k}{N}\right)^2 \text{var}(\bar{y}_{1k} - \bar{y}_{0k})$$

共分散分析

割り当て変数と傾向スコアを説明変数とした線形回帰分析を行う。

傾向スコアを用いた解析法の最大に利点は、「結果変数と共変量の回帰モデル (y_1, y_0 と x の関係)」を仮定する必要がないこと。

「割り当てと共変量の回帰モデル (z と x の関係)」のほうが「結果変数と共変量の回帰モデル」よりもモデルの語設定の可能性が低い。

傾向スコアの利点

- (1) 共変量を1変数に要約しているため、2つの群において局外要因・共変量の値に重なりがない (サポート問題) または少ない場合でも使える
- (2) 共変量と従属変数のモデルの設定を行わなくても良い (y と x)
- (3) モデルの語設定に強い

Drake (1993) は傾向スコア解析のほうが共分散分析よりバイアスが小さいことを示している。

傾向スコアの問題点

- (1) 2群ごとに傾向スコアを推定する必要がある。母集団が各2群ごとの解析ごとに異なる。
- (2) マッチング、層別解析では、因果効果の推定値は計算できるが、標準誤差は性格には計算できない。
- (3) マッチング層別解析ともに各周辺期待値 ($E(y_1), E(y_0)$) が推定できない
- (4) マッチングペアを選ぶ基準は恣意的。
- (5) マッチングをする場合、対象者の多い郡でのデータの多くが無駄になる。
- (6) 共分散分析のモデルで傾向スコア解析を行う (線形確率モデル) ための前提条件として確率が1以上、0以下の値はとらないが、推定値はそのような値をとるという矛盾がある。

参考文献

- [1] 星野高宏 (2009) 『調査観察データの統計科学』、岩波書店
- [2] 北村行伸 (2009) 『ミクロ計量経済学入門』、日本評論社
- [3] Anderson, T.W. (1984) *Introduction to Multivariate Statistical Analysis*, Wiley.
- [4] Anderson, T.W. and Rubin, H.(1949) “Estimators of the Parameters of a Single Equation in a Complete Set of Stochastic Equations”, *Annals of Mathematical Statistics*, 21, pp.570-82.
- [5] Andrew, Donald W.K. and Stock, James H.(2005) “Inference with Weak Instruments”, NBER Technical Working Paper 313.
- [6] Angrist, J.D. and Krueger, A.B.(1991) “Does Compulsory School Attendance Affect Schooling and Earnings?”, *Quarterly Journal of Economics*, 106, pp.979-1014.

- [7] Angrist, Joshua, D., Imbens, Guido, W. and Rubin, Donald B.(1996) “Identification of Causal Effects Using Instrumental Variables”, *Journal of the American Statistical Association*, 91(434), pp.444-455.
- [8] Angrist, Joshua D. and Krueger, Alan B.(2001) “Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments”, *Journal of Economic Perspectives*, 15(4), pp.69-85.
- [9] Angrist, Joshua, D., Bettinger, Eric., Bloom, Erik., King, Elizabeth., and Kremer, Michael. (2002) “Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment”, *American Economic Review*, 92(5), pp.1535-1558.
- [10] Basman, R.L. (1960) “On Finite Sample Distributions of Generalized Classical Linear Identifiability Test Statistics”, *Journal of the American Statistical Association*, 55(292), pp.650-59.
- [11] Baum, Christopher. (2006) *An Introduction to Modern Econometrics Using Stata*, Stata Press.
- [12] Blackburn, McKinley and Neumark, David.(1992) “Unobserved Ability, Efficiency Wages, and Interindustry Wage Differentials”, *Quarterly Journal of Economics*, 107(4), pp.1421-1436.
- [13] Bound, John., Jaeger, David.A. and Baker, Regina. M.(1995) “Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak”, *Journal of the American Statistical Association*, 90(430), pp.443-50.
- [14] Bowden, R.J. and Turkington, D.A.(1984) *Instrumental Variables*, Cambridge University Press.
- [15] Breusch, Trevor., Qian, Hailong., Schmidt, Peter., and Wyhowski, Donald.(1999) “Redundancy of Moment Conditions”, *Journal of Econometrics*, 91, pp.89-111.
- [16] Cameron, A.C.and Trivedi, P.K.(1998) *Regression Analysis of Count Data*, Cambridge University Press.
- [17] Cameron, A.C. and Trivedi, P.K.(2005) *Microeconometrics: Methods and Applications*, Cambridge University Press.

- [18] Chao, John.C. and Swanson, Norman R.(2005) “Consistent Estimation with a Large Number of Weak Instruments”, *Econometrica*, 73(5), PP.1673-1692.
- [19] Cragg, John G.and Donald, Stephen G.(1993) “Testing Identifiability and Specification in Instrumental Variable Models”, *Econometric Theory*, 9, pp.222-40.
- [20] Davidson, Russell and MacKinnon, James G.(2004) *Econometric Theory and Methods*, Oxford University Press.
- [21] Durbin, J.(1954) “Errors in variables”, *Review of the International Statistical Institute*, 22, pp.23-32.
- [22] Griliches, Zvi.(1976) “Wages of Very Young Men”, *Journal of Political Economy*, 84(4. Part 2), pp. S69-S85.
- [23] Griliches, Zvi.(1977) “Estimating the Returns to Schooling: Some Econometric Problems”, *Econometrica*, 45(1), pp.1-22.
- [24] Griliches, Zvi., Hall, Bronwyn., and Hausman, Jerry.(1978) “Missing Data and self-Selection in Large Panels”, *Annales de L’INSEE*, XXX-XXXI, pp.137-76.
- [25] Hahn, Jinyoung and Hausman, Jerry. (2002a) “A New Specification Test for the Validity of Instrumental variables”, *Econometrica*, 70(1), pp.163-189.
- [26] Hahn, Jinyoung and Hausman, Jerry. (2002b) “Notes on Bias in Estimators for Simultaneous Equation Models”, *Economics Letters*, 75, pp.237-241.
- [27] Hahn, Jinyoung and Hausman, Jerry. (2003) “Weak Instruments: Diagnosis and Cures in Empirical Econometrics”, *American Economic Review*, 93(2), pp.118-125.
- [28] Hall, Alastair R., Rudebusch, Glenn D. and Wilcox, David W.(1996) “Judging Instrument Relevance in Instrumental Variables Estimation”, *International Economic Review*, 37(2), pp.283-298.
- [29] Hall, Alastair R. and Peixe, Fernanda P.M.(2000) “A Consistent Method for the Selection of Relevant Instruments”, A paper presented at Econometric Society World Congress 2000.
- [30] Hansen, Lars.P (1982) “Large Sample Properties of Generalized Method of Moments Estimators”, *Econometrica*, 50(4), pp.1029-1054.

- [31] Hausman, Jerry. (1978) "Specification tests in econometrics", *Econometrica*, 46, pp.1251-72.
- [32] Hausman, Jerry., Stock, James H. and Yogo, Motohiro.(2005) "Asymptotic Properties of the Hahn-Hausman Test for Weak-Instruments", *Economics Letters*, 89, pp.333-42.
- [33] Hayashi, Fumio.(2000) *Econometrics*, Princeton University Press.
- [34] Imbens, Guido W. and Angrist, Joshua D.(1994) "Identification and Estimation of Local Average Treatment Effects", *Econometrica*, 62(2). pp.467-475.
- [35] Koenker, Roger.(1981) "A Note on Studentizing a test for Heteroscedasticity", *Journal of Econometrics*, 17., pp.107-112.
- [36] Koenker, Roger. (2005) *Quantile Regression*, Cambridge University Press.
- [37] Nelson, Charles R. and Startz, Richard.(1990a) "The Distribution of the Instrumental Variables Estimator and Its t-Ratio When the Instrument is a Poor One", *Journal of Business*, 63(1, Part.2), pp. S125-S140.
- [38] Nelson, Charles R.and Startz, Richard.(1990) "Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator", *Econometrica*, 58(4), pp.967-76.
- [39] Pagan, A.R. and Hall, D. (1983) "Diagnostic Tests as Residual Analysis", *Econometric Reviews*, 2(2), pp.159-218.
- [40] Ruud, P.A. (2000) *An Introduction to Classical Econometric Theory*, Oxford University Press.
- [41] Sargan, J.D. (1958) "The Estimation of Economic Relationships Using Instrumental Variables", *Econometrica*, 26(3), pp.393-415.
- [42] Shea, John.(1997) "Instrument Relevance in Multivariate Linear Models: A Simple Measure", *Review of Economics and Statistics*, 79(2), pp.348-352.
- [43] Staiger, Douglas. and Stock, James.H. (1997) "Instrumental Variables Regression with Weak Instruments", *Econometrica*, 65(3), pp.557-86.
- [44] Stock, James H. and Wright Jonathan H. (2000) "GMM with Weak Identification", *Econometrica*, 68(5), pp.1055-96.

- [45] Stock, James H., Wright, Jonathan H. and Yogo, Motohiro. (2002) “A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments”, *Journal of Business and Economic Statistics*, 20(4), pp.518-29.
- [46] Stock, James H. and Yogo, Motohiro. (2005) “Testing for Weak Instruments in Linear IV Regression”, in Andrews, D.W.K. and Stock, J.H.(eds) *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge University Press. pp.80-108.
- [47] Winkleman, Rainer and Boes, Stefan. (2005) *Analysis of Microdata*, Springer.
- [48] White, Halbert. (1980) “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity”, *Econometrica*, 48(4), pp.817-838.
- [49] White, Halbert. (1982) “Instrumental Variables Regression with Independent Observations”, *Econometrica*, 50(2), pp.483-499.
- [50] Wooldridge, Jeffrey. M. (2002) *Econometric Analysis of Cross Section and Panel Data*, The MIT Press
- [51] Wu, D-M. (1973) “Alternative tests of independence between stochastic regressors and disturbances”, *Econometrica*, 41, pp.733-50.