

博士論文

ノンパラメトリック操作変数法による

因果効果の統計的推測法の開発とその応用

Statistical inference of causal effects based on
nonparametric instrumental variable method and
its application

横浜国立大学大学院

理工学府

川上裕大

Yuta Kawakami

学位授与 西暦2023年9月

論文要旨

構造的因果モデルに基づく因果効果の統計的推測法は、統計学や機械学習などの情報学のみならず、経済学、政治学、医学やビジネス分野に至るまで広範な領域において基礎となる理論である。線形回帰分析における説明変数の内生性の解決法として、ふるくから知られる操作変数法は、構造的因果モデルに基づく因果推論においても重要な役割を果たす。とくに、操作変数法における構造的因果モデルの関数形に強い仮定を置かないモデルをノンパラメトリック操作変数モデルと呼び、本論文ではノンパラメトリック操作変数モデルを利用した新しい因果効果の統計的推測法を紹介する。

本論文は、5部構成となっており、第1部「はじめに」、第2部「共変量情報を用いた平均因果効果の識別性について」、第3部「共変量情報を用いた潜在反応タイプの確率の識別性および推定法の開発」、第4部「平均偏因果効果の推定法の開発」、第5部「まとめ」となっている。第2部と第3部は、2値の処置変数に関する研究である。第2部では、操作変数に加えて個体の共変量情報を利用することで、既存の研究より弱い仮定のもとで、平均因果効果が識別可能であることを示す。第3部では、個体の共変量情報を利用すると平均因果効果にとどまらず、潜在反応タイプの確率まで識別が可能であることを示し、さらに、潜在反応タイプの確率の有界制約付き拡張ラグランジュ法に基づく効率的な統計的推測法を提案する。第4部では、処置が連続変数の場合において、先行研究によって与えられている平均偏因果効果の識別可能条件のもとで、ピカールの反復法および線形基底モデルに基づいた平均偏因果効果の統計的推測法を提案する。

目次

1	はじめに	1
2	共変量情報を用いた平均因果効果の識別可能性について	12
2.1	準備	13
2.2	平均因果効果の識別可能性に関する定理	14
2.3	考察	20
3	共変量情報を用いた潜在反応タイプの確率の識別可能性および推定手法の開発	21
3.1	準備	22
3.2	潜在反応タイプの確率の識別可能性	24
3.2.1	$p(Y_{x_0} = y, Y_{x_1} = y')$ の識別可能性	25
3.2.2	$p(X_{z_0} = x, X_{z_1} = x')$ の識別可能性	27
3.2.3	$p(X_{z_0} = x, X_{z_1} = x', Y_{x_0} = y, Y_{x_1} = y')$ の識別可能性	28
3.3	潜在反応タイプの確率の推定法	30
3.3.1	$p(Y_{x_0} = y, Y_{x_1} = y')$ の推定法	30
3.3.2	$p(X_{z_0} = x, X_{z_1} = x')$ の推定法	34
3.3.3	$p(X_{z_0} = x, X_{z_1} = x', Y_{x_0} = y, Y_{x_1} = y')$ の推定法	37
3.4	数値実験	39
3.5	実データへの応用	41

3.6	考察	46
4	平均偏因果効果の推定法の開発	47
4.1	準備	48
4.2	ノンパラメトリック推定法 (N-APCE 推定量)	51
4.2.1	N-APCE 推定量のアルゴリズム	52
4.2.2	N-APCE 推定量の性質	57
4.3	パラメトリック推定法 (P-APCE 推定量)	60
4.3.1	P-APCE 推定量のアルゴリズム	60
4.3.2	P-APCE 推定量の性質	63
4.4	数値実験	64
4.5	実データへの応用	67
4.6	考察	68
5	まとめ	72
	参考文献	76
	付録	90

第 1 部

はじめに

統計的因果推論とは、事象の因果関係を収集したデータから統計的に評価することであり、統計学のみならず、情報学、経済学、政治学、医学やビジネス分野に至るまで広範な学問領域における中心的な試みである。研究に参加する被験者を2つの群にランダムに分け、原因となる処置の効果を検証する手法はランダム化試験 [Chalmers et al., 1981, Splawa-Neyman et al., 1990, Schulz et al., 2010, Athey and Imbens, 2017] とよばれ、ふるくから統計学や医学の分野において利用されており、「ゴールド・スタンダード」と呼ばれるほど理想的な試験方法であるとされている [Meldrum, 2000]。しかしながら、被験者がランダム化した割り当てに従わない状況（不遵守）などのランダム化試験が実行不可能な状況が多く存在し、ランダム化試験の限界が多くの研究者の間で認識されている。本論文では、ランダム化試験が実行不可能な際に、よく利用されるノンパラメトリック操作変数モデルにおける、新たな統計的推測法を提案する。

構造的因果モデルと因果推論。 Judea Pearl 教授は因果推論を行うための数学的な基盤を、構造的因果モデル、および、それによって定義される潜在反応変数によって与えた [Pearl, 2009, 黒木学, 2017]。はじめに、構造的因果モデル (Structural Causal Model; SCM) とは組

$$\mathcal{M} = \langle \mathbf{U}, \mathbf{V}, \mathcal{F}, \mathbb{P}_{\mathbf{U}} \rangle \quad (1.1)$$

に基づいて定義されるものである。変数 U は同時分布 \mathbb{P}_U に従う外生変数の集合であり、変数 V は構造関数族 $\mathcal{F} = \{f_{V_i}\}_{V_i \in V}$ によって値が決定される内生変数の集合である。ここで、各内生変数 V_i は構造関数 f_{V_i} をもちいて

$$V_i := f_{V_i}(\mathbf{PA}_{V_i}, U_{V_i}) \quad (1.2)$$

で与えられる。ただし、 $\mathbf{PA}_{V_i} \subseteq V$ かつ $U_{V_i} \subseteq U$ であり、各構造的因果モデル \mathcal{M} は外生変数の集合 V の観測可能な同時分布 \mathbb{P}_V を生成する。また、各構造的因果モデルには対応する V 上の因果グラフ $G(\mathcal{M})$ が存在しており、親 \mathbf{PA}_{V_i} に属する全ての変数から V_i への有向辺が引かれている。以上のように定義された構造的因果モデルが因果推論を数学的および統計的に議論するための基盤となっている [Pearl and Mackenzie, 2018]。つぎに、ある内生変数の部分集合 X をある定数 x に置き換える操作を介入とよび “ $do(x)$ ” と表記し、この介入によって得られる新たな構造的因果モデルを \mathcal{M}_x と表す。また、 \mathcal{M}_x におけるある外生変数を Y_x と表記し、これを潜在反応変数と呼ぶ。

因果効果の指標. つぎに、いくつかの因果効果の数学的な定義と解釈を紹介する。はじめに、2 値変数の処置 ($x \in \{x_0, x_1\}$) と 2 値変数のアウトカム ($y \in \{y_0, y_1\}$) を考えると、最もよく使われる因果効果の指標として、潜在反応変数の確率の差で定義される平均因果効果 [Holland, 1986]

$$p(Y_{x_1} = y_1) - p(Y_{x_0} = y_1) \quad (1.3)$$

が存在する。ランダム化試験においては、処置変数 X が 2 つの潜在反応変数の組 $\{Y_{x_0}, Y_{x_1}\}$ と独立であるため ($X \perp\!\!\!\perp \{Y_{x_0}, Y_{x_1}\}$)、平均因果効果は観測可能な条件付き期待値の差

$$p(Y = y_1 | X = x_1) - p(Y = y_1 | X = x_0) \quad (1.4)$$

を通して識別可能である。また、観察データにおいては “Do-calculus” の理論 [Pearl, 1995a, Tian and Pearl, 2002, Shpitser and Pearl, 2006] に基づいて、バックドア基

準 [Pearl, 1993] やフロントドア基準 [Pearl, 1995a] などを満たす共変量の集合を利用することで平均因果効果が識別可能となる。しかしながら、バックドア基準を満たす変数をすべて観測することや、妥当なフロントドア基準を満たす変数を観測することが困難な状況が多く存在している。

さらに、因果関係の反事実的な側面により踏み込んだ指標として、3つの原因の確率 [黒木学, 2017, Cai and Kuroki, 2005, Pearl, 2009] が知られている。3つの原因の確率の一つである必要十分性の確率は、潜在反応変数の組の同時確率

$$p(Y_{x_1} = y_1, Y_{x_0} = y_0) \quad (1.5)$$

で定義され、「もし $X = x_0$ であるなら $Y = y_0$ であり、かつ、もし $X = x_1$ であるなら $Y = y_1$ である」ような個体の確率を表している。この指標は因果効果の重要な指標ではあるが、たとえ完全なランダム化試験を行ったとしても、 X と Y の観測分布 $P(X, Y)$ からでは識別不可能なものとなっている。必要十分性の確率のほかにも、潜在反応変数の組 $(Y_{x_1} = y', Y_{x_0} = y'')$ ($y', y'' \in \{y_0, y_1\}$) の形式を取る指標は、全部で4種類考えることができ、Rothman et al. [2008] によれば、次の4種類

$$u_1: \text{“doomed”}, u_2: \text{“causative”}, u_3: \text{“preventive”}, u_4: \text{“immune”}$$

に分類される。ただし、

$$\begin{aligned} u_1 &= (Y_{x_0} = y_0, Y_{x_1} = y_0), & u_2 &= (Y_{x_0} = y_0, Y_{x_1} = y_1), \\ u_3 &= (Y_{x_0} = y_1, Y_{x_1} = y_0), & u_4 &= (Y_{x_0} = y_1, Y_{x_1} = y_1) \end{aligned} \quad (1.6)$$

であり、必要十分性の確率は“causative”の個体の確率に対応している。3つの原因の確率に関する先行研究として、Tian and Pearl [2000] は観測データの確率分布 $P(X, Y)$ から得られる3つの原因の確率の存在範囲をあたえている。また、強外生性 ($X \perp\!\!\!\perp \{Y_{x_0}, Y_{x_1}\}$) と単調性 II

$$p(u_3) = p(Y_{x_0} = y_1, Y_{x_1} = y_0) = 0 \quad (1.7)$$

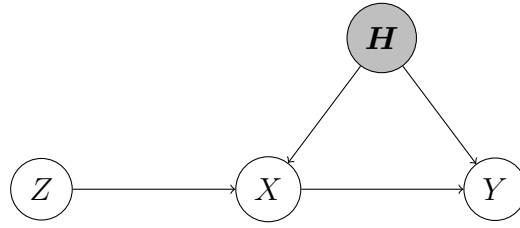


図 1.1: 操作変数モデルの因果グラフ

という比較的強い仮定のもとで，3つの原因の確率が識別可能となることを示している．さらに，Kuroki and Cai [2011] は個体の共変量情報を利用して3つの原因の確率のより狭い存在範囲を提供している．

操作変数モデルと操作変数法. 本論文では，特に，操作変数モデルと呼ばれる構造的因果モデル [Pearl, 1995b] を考える．操作変数モデルの構造的因果モデル M_{IV} は

$$Y := f_Y(X, \mathbf{H}, \epsilon_Y), X := f_X(Z, \mathbf{H}, \epsilon_X), Z := f_Z(\epsilon_Z) \quad (1.8)$$

によって定義され，操作変数モデルに基づく統計的手法群を操作変数法とよぶ．ここで，操作変数モデルに対応する因果グラフは図 1.1 によってあたえられており，任意の個数の非観測な離散変数および連続変数の集合 \mathbf{H} を許容している．とくに，構造関数族 \mathcal{F} に線形性などの強い制約を置かない操作変数モデルをノンパラメトリック操作変数モデルとよび，ノンパラメトリック操作変数モデルに基づく統計手法をノンパラメトリック操作変数法とよぶ．ここで，変数の集合 $\{\epsilon_X, \epsilon_Y, \epsilon_Z, \mathbf{H}\}$ が外生変数であり， $\{X, Y, Z\}$ が内生変数である．操作変数法による統計解析においては，内生変数の同時分布 $P(X, Y, Z)$ などからサンプリングされたデータを用いて因果効果の推定をおこなうこととなる．現実存在する多くの状況が操作変数モデルを用いて妥当にモデル化できるため，操作変数法は古くからさまざまな分野で研究されてきた．

操作変数法は，1920年代において，Wright [1928] によって線形回帰分析法における説明変数の内生性の解決法としてはじめて提案された．とくに，彼は次の

ような線形モデルを考えていた.

$$Y = \beta_0 + \beta_1 X + \epsilon \quad (1.9)$$

ここで, Y がアウトカム, X が処置変数, ϵ が誤差項である. ただし, 誤差項は説明変数と相関を持っているとし, これは内生性とよばれる [Wooldridge, 2015]. このとき, $X = \gamma_0 + \gamma_1 Z + \epsilon'$ かつ $Z \perp \epsilon'$ をみたす操作変数 Z を利用することで, 平均因果効果を表す係数である β_1 の識別可能とした. このような連続変数の処置変数に対する線形モデルに基づく操作変数法は, とくに, 経済学分野において広く普及していった. Angrist and Krueger [2001] や Imbens [2014] は操作変数法の経済学における歴史をまとめている.

1990年代において, 操作変数法は潜在反応変数を用いた議論においても重要な役割をはたすことが明らかになってきた. Imbens and Angrist [1994] は, ノンパラメトリック操作変数モデルが割り当ての不遵守が存在するランダム化試験を表現できることを明らかにし, 局所的平均因果効果 (Local Average Treatment Effect; LATE) と呼ばれる指標を導入した. また, Frangakis and Rubin [2002] は主要層別 (Principal Stratification) という枠組みによって, これらの議論を一般化しまとめている. この局所的平均因果効果が“局所的”であるのは, “compliance”と呼ばれる部分集団の平均因果効果のみを評価しているからである. この“compliance”と呼ばれる集団は潜在反応変数の組 ($X_{z_1} = x_1, X_{z_0} = x_0$) の形で定義されるものである. かわえて, 潜在反応変数の組 ($X_{z_1} = x', X_{z_0} = x''$) ($x', x'' \in \{x_0, x_1\}$) の形式をとる指標として, 以下の4種類が存在している.

$$v_1: \text{“never-take”}, v_2: \text{“compliance”}, v_3: \text{“defiance”}, v_4: \text{“always-take”},$$

ただし,

$$\begin{aligned} v_1 &= (X_{z_0} = x_0, X_{z_1} = x_0), v_2 = (X_{z_0} = x_0, X_{z_1} = x_1), \\ v_3 &= (X_{z_0} = x_1, X_{z_1} = x_0), v_4 = (X_{z_0} = x_1, X_{z_1} = x_1) \end{aligned} \quad (1.10)$$

である。局所的平均因果効果は単調性 I の仮定 “ $X_{z_1} \geq X_{z_0}$ ” が成り立つとき識別可能である [Imbens and Angrist, 1994]。この単調性 I の仮定は，“defiance” が存在していないと言い換えることができ、

$$p(v_3) = p(X_{z_0} = x_1, X_{z_1} = x_0) = 0 \quad (1.11)$$

と書くこともできる。また、Balke and Pearl [1994, 1997] はノンパラメトリック操作変数モデルのもとの平均因果効果 $p(Y_{x_1} = y_1) - p(Y_{x_0} = y_1)$ の存在範囲をあたえ、Cai et al. [2007] は共変量の集合には観測変数も含まれることに着目し、観測された共変量の情報を用いて Balke and Pearl [1997] の存在範囲よりも狭い範囲を導出した。

一方で、2000年代の経済学の分野において、Newey and Powell [2003] が連続変数の処置変数 X に対するノンパラメトリック操作変数法をはじめて提案した。この手法により、非線形な構造関数を解明することが可能となった。Newey and Powell [2003] は分離可能性の仮定

$$f_Y(X, \mathbf{H}, \epsilon_Y) = f_{Y_1}(X, \epsilon_Y) + f_{Y_2}(\mathbf{H}, \epsilon_Y) \quad (1.12)$$

と

$$\mathbb{E}[f_{Y_2}(\mathbf{H}, \epsilon_Y) | Z = z] = 0, \quad \forall z \in \Omega_Z \quad (1.13)$$

の仮定のもと、潜在反応変数の期待値（構造関数） $\mathbb{E}[Y_x] := \mathbb{E}_U[Y_x(\mathbf{U})]$ の識別可能性を、以下の積分方程式

$$\mathbb{E}[Y | Z = z] = \int_{\Omega_X} p(X = x | Z = z) \mathbb{E}[Y_x] dx \quad (1.14)$$

を通してあたえた。つまり、この積分方程式の解が一意であり、構造関数に一致するということであり、また、この積分方程式は第一種フレドホルム積分方程式と呼ばれるものである [Bôcher, 1926]。さらに、構造関数が識別可能であれば、任意の2つの処置 x' と x'' の平均因果効果は

$$\mathbb{E}[Y_{x'}] - \mathbb{E}[Y_{x''}] \quad (1.15)$$

で識別可能である。

Newey and Powell [2003] によって導入された積分方程式は、不良設定問題 [Tikhonov et al., 1995] と呼ばれるものであり、推定において難しい問題を抱えている。そこで、Newey and Powell [2003] は、推定におけるモデルの関数空間をコンパクトなものに限定する制約条件を導入することでこの問題を解決した。このコンパクトな空間への制約は、ノンパラメトリック操作変数法において標準的なものとなっている [Ai and Chen, 2003, Blundell et al., 2007]。また、処置変数が 2 値の研究では識別不可能とされていた平均因果効果が、彼らの研究で識別可能となるのは分離可能性の仮定のおかげであることに注意したい。本論文でも、[Newey and Powell, 2003] と同様に分離可能性の仮定をおくこととする。分離可能性の仮定の成り立たないモデルでの構造関数の推定は、Chernozhukov et al. [2007] や Torgovitsky [2015] などが研究しているが、彼らは分離可能性の仮定の代わりに非観測な変数 H に厳しい仮定を置いている。

さらに、機械学習の発展とともに、2010年代から2020年代にかけて、Deep IV [Hartford et al., 2017], Kernel IV [Singh et al., 2019], Deep GMM [Bennett et al., 2019], Dual IV [Muandet et al., 2020], DFIV [Xu et al., 2022] などの新たなノンパラメトリック操作変数法が開発されている。これらの手法はカーネル法や深層学習などの最先端の機械学習の手法を用いて推定を行なっている。また、Wu et al. [2022] は最近の機械学習分野におけるノンパラメトリック操作変数の研究をまとめており、機械学習におけるモデルの表現力の高さはノンパラメトリック操作変数法において重要な性質となっている。また、近年、用量反応関係 (Dose-response relationship) [Wong and Lachenbruch, 1996, Emilien et al., 2000] に代表されるように、医学、統計学や因果推論分野においても連続的な処置変数への関心が高まっている。

論文の構成. 本研究ではノンパラメトリック操作変数モデルを利用した因果効果の

新たな統計的推測法を紹介する。本論文は5部構成となっている。

- 第1部「はじめに」
- 第2部「共変量情報を用いた平均因果効果の識別可能性について」
- 第3部「共変量情報を用いた潜在反応タイプの確率の識別可能性および推定手法の開発」
- 第4部「平均偏因果効果の推定法の開発」
- 第5部「まとめ」

これらの内容は、[Kawakami, 2021], [Kawakami, Shingaki, and Kuroki, 2023b], [Kawakami, Kuroki, and Tian, 2023a]において公表されている。各部において、取り上げる因果効果の指標の定義や解釈、因果効果の識別可能性の定理の紹介、観測データを用いた因果効果の推定法の開発の順で解説をする。

第2部、第3部は、処置変数、アウトカムおよび操作変数が2値変数の場合のノンパラメトリック操作変数モデルに関する研究である。第2部では、操作変数に加えて個体の共変量情報を利用することで、既存の研究より弱い仮定のもとで平均因果効果の識別が可能であることを定理2.2.1を通して示す。ここで、個体の共変量情報は代替変数として操作変数モデルに組み込まれる。第3部では、個体の共変量情報を利用すると平均因果効果にとどまらず、より一般化された指標である潜在反応タイプの確率まで識別が可能であることを3つの定理(定理3.2.1, 3.2.2, 3.2.3)を通して示し、さらに、有界制約付き拡張ラグランジュ法[Birgin and Martínez, 2020]をもちいたより効率的な推定方法(アルゴリズム1とアルゴリズム2)を提案する。第4部では処置変数と操作変数が連続変数数量の場合のノンパラメトリック操作変数モデルにおいて、Wong [2022]によって与えられた平均偏因果効果の識別可能条件のもとで、ピカールの反復法[Diaz and Metcalf, 1970]を利用したN-APCE推定量(アルゴリズム3)と線形基底モデルを利用したP-APCE推定量(アルゴリズム4)を開

発する。この論文の定理などの証明はすべて付録に記載している。

第2部と第3部の概要. 第2部と第3部では、2値変数の処置変数(X)、アウトカム(Y)および操作変数(Z)にくわえて、代替変数(W)をともなうノンパラメトリック操作変数モデルを考える。ここで、多くのランダム化試験では、年齢、性別、人種などの個体の共変量の測定も可能であり [Lui, 2011], 個体の共変量を非観測な変数 \mathbf{H} の“代わり”に観測される代替変数とみなすことができる [Kuroki and Pearl, 2014, Miao et al., 2018, Tchetgen et al., 2020]. 想定する構造的因果モデルは次のような式で定義される。

$$\begin{aligned} Y &:= f_Y(X, \mathbf{H}, \epsilon_Y), \quad X := f_X(Z, \mathbf{H}, \epsilon_X), \\ Z &:= f_Z(\epsilon_Z), \quad W := f_W(\mathbf{H}, \epsilon_W) \end{aligned} \tag{1.16}$$

この構造的因果モデルに対応する因果グラフは図 1.2 で与えられる。この構造的因果モデルにおいても、非観測な変数 \mathbf{H} は任意の個数の連続および離散変数を含むことができるため、このモデルは多くのデータセットにおいて比較的妥当なものと考えられる。

つぎに、この操作変数モデルにおいてはランダム化後のふるまいを規定する情報として

$$(X_{z_0}, X_{z_1}, Y_{x_1}, Y_{x_0}) \tag{1.17}$$

による分類を考えることができる [Balke and Pearl, 1994]. 例えば、 $(X_{z_0} = x_0, X_{z_1} = x_1, Y_{x_1} = y_0, Y_{x_0} = y_1)$ に属する個体は、ランダム化によって z_1 に割り当てられると $(X, Y) = (x_1, y_1)$ に一意に決定される。この論文では、この分類を16値の離散値確率変数と見なし、「16種類の潜在反応タイプ」と呼ぶ。これは完全なランダム化試験であっても識別および推定することのできないものである。また、周辺化した2つの「4種類の潜在反応タイプ」

$$(X_{z_0}, X_{z_1}), (Y_{x_1}, Y_{x_0}) \tag{1.18}$$

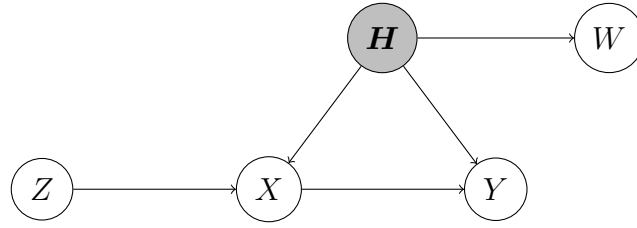


図 1.2: 共変量情報をともなう操作変数法の因果グラフ

も考えられることができる。これは、前述した二つの個体の分類方法 (1.6) と (1.10) に対応しており、この論文で対象にしている 16 種類の潜在反応タイプとは先行研究で考えられてきた必要十分性の確率や局所的平均因果効果における “compliance” などの議論を含むより一般化された概念となっている。第 2 部、第 3 部では操作変数にくわえて代替変数を利用することで、4 種類と 16 種類の潜在反応タイプの確率の識別可能性に関する定理をあたえる。また、有界制約付き拡張ラグランジュ法を利用したより効率的な推定法を紹介し、推定法の性質を数値実験および心理学分野での「賃金の公平さが仕事の成果に与える効果」に関する実験のデータへの適用例を通して紹介する。

第 4 部の概要. 第 4 部では連続変量の処置変数 (X)、アウトカム (Y)、および操作変数 (Z) で構成されるノンパラメトリック操作変数モデルに基づいて、ピカールの反復法と線形基底モデルを用いた平均偏因果効果の推定法を開発する。本論文の研究対象となる平均偏因果効果 (Average Partial Causal Effect; APCE) とは、潜在反応変数 Y_x の一次偏導関数の期待値 [Chamberlain, 1984, Wooldridge, 2005, Graham and Powell, 2012]

$$E[\partial_x Y_x] \tag{1.19}$$

で定義される。2 つの処置 x'' と x' に対する平均因果効果は、

$$E[Y_{x''}] - E[Y_{x'}] = \int_{x'}^{x''} E[\partial_x Y_x] dx \tag{1.20}$$

で表現できるため、平均因果効果は因果効果の評価において十分な指標である。こ

ここで、ピカールの反復法 [Picard, 1910, Diaz and Metcalf, 1970] は第一種フレドホルム積分方程式の数学的な解法ではあるが、構造関数 $E[Y_x]$ に関する積分方程式は不良設定問題であり、条件付き密度推定が必要となるため、ピカールの反復法をもちいて解くことは困難であった。

本研究では、近年、Wong [2022] によってあたえられた平均偏因果効果に関する積分方程式

$$\begin{aligned} & E[Y|Z = z_0] - E[Y|Z = z] \\ &= \int_{\Omega_x} \{P(X \leq x|Z = z) - P(X \leq x|Z = z_0)\} E[\partial_x Y_x] dx \end{aligned} \tag{1.21}$$

に着目して、ピカールの反復法を用いた平均因果効果の推定法を提案する。ここで、 $P(X \leq x|Z = z)$ は条件付き累積分布関数であり、 z_0 は任意の定数である。この積分方程式は、処置変数と操作変数の定義域が有界であるとき不良設定問題であり、積分核が有界な条件付き累積分布関数の差で構成されているという利点が存在している。また、構造関数 $E[Y_x]$ の推定に関する先行研究にならって線形基底モデルを利用した平均偏因果効果の推定方法も提案する。さいごに、提案法の一致性やバイアス、分散などの統計的な性質を紹介し、数値実験と経済学分野における「教育年数の賃金へ因果効果」の評価を目的としたデータセットへの適用例を紹介する。

第 2 部

共変量情報を用いた平均因果効果の識別可能性について

本論文では、不遵守の存在するランダム化試験における平均因果効果

$$p(Y_{x_1} = y_1) - p(Y_{x_0} = y_1) \quad (2.1)$$

の新しい識別可能性に関する定理を紹介する。不遵守の存在するランダム化試験はノンパラメトリック操作変数モデルで表現可能である [Imbens and Angrist, 1994]。しかしながら、この設定において平均因果効果が識別できないため、これまで、なにかしら厳しい仮定をおいたり [Mark and Robins, 1993, Angrist et al., 1996, Pearl, 2009, Robins and Richardson, 2011, Wang and Tchetgen, 2018, Li and Pearl, 2019]、平均因果効果の存在範囲 [Robins, 1989, Manski, 1990, Balke and Pearl, 1997] を通して平均因果効果が評価されてきた。また、Cai et al. [2007] は観測できる共変量の情報を用いて [Balke and Pearl, 1997] の存在範囲よりも狭い範囲を導出した。多くのランダム化試験では、割り当て (Z)、処置 (X)、アウトカム (Y) だけでなく、年齢、性別、人種などの共変量 (W) の測定も可能である [Lui, 2011]。

本論文では、代替変数として個体の共変量情報を利用して、新しい潜在反応変数の同時確率に関する識別可能条件を提案し、これを通して平均因果効果が先行研究に比べて弱い仮定のもとで識別可能であることを示す。本論文の結果は因果推論における解決可能な平均因果効果の識別問題の範囲を拡張するものである。第 2 部の

内容は [Kawakami, 2021] で公表されている。

2.1 準備

はじめに、平均因果効果の識別可能問題を議論するために必要な潜在反応変数の概念を紹介する [Pearl, 2009, Imbens and Rubin, 2015]. この部では、2 値の処置 ($x \in \{x_0, x_1\}$) と 2 値のアウトカム ($y \in \{y_0, y_1\}$) を考える. N 人の被験者のうち i 番目の被験者の処置 X が仮に x だった場合 ($X(i) = x$) に生じるであろうアウトカムを

$$Y_x(i) \tag{2.2}$$

と表記し、これを潜在反応変数とよぶ. 本研究では、因果推論において標準的に使われる SUTVA (Stable Unit Treatment Value Assumption) を仮定することとする. SUTVA は以下の二つの仮定から構成されている [Rubin, 2005].

1. 任意の被験者の処置が、他の被験者のアウトカムに影響を与えない.
2. すべての被験者の処置が均一であり、比較可能である.

研究に参加する N 人の被験者を対象となる母集団からのランダムサンプルと考えると、 $X(i)$ と $Y_x(i)$ をそれぞれ確率変数 X と Y_x の実現値とみすことができる. また、平均因果効果 [Holland, 1986] は

$$p(Y_{x_1} = y_1) - p(Y_{x_0} = y_0) \tag{2.3}$$

で定義できる. 他の潜在反応変数についても同様の表記を用いることとする. ここで、潜在反応変数 Y_x は処置 X が x である場合にのみ観測され、因果推論の根本問題は、同一個体の潜在反応変数は片方しか観測不可能なことである. この性質は一致性 (consistency) と呼ばれ以下のように定式化される [Robins, 1989, Pearl, 2009].

$$X(i) = x \implies Y_x(i) = Y(i) \tag{2.4}$$

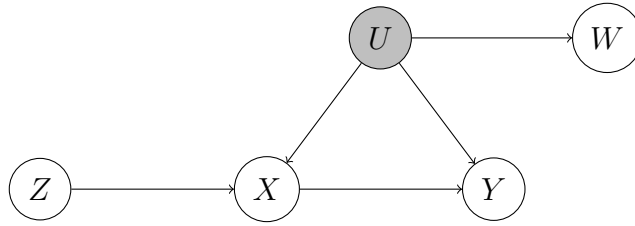


図 2.1: 共変量を伴う操作変数法の因果グラフ

ランダム化試験において被験者の遵守が完璧な場合、 X と潜在反応変数の組 (Y_{x_0}, Y_{x_1}) は独立であるので、平均因果効果は以下の式をとおして識別可能である。

$$p(Y_x = y) = p(y | x) \quad (2.5)$$

ここで、「識別可能である」とは、因果関係を評価するためのある数量が無限にデータが得られる状況において観測可能であることをいう。

一方、ランダム化試験の実行が困難であり、観察データしか得ることができない場合は、条件付き交換可能性の仮定 [Rosenbaum and Rubin, 1983] やバックドア基準 [Pearl, 1993, 2009] を満たす共変量の集合 S によって平均因果効果を評価できる。平均因果効果は共変量の集合 S を用いて以下の式を通して識別可能である。

$$p(Y_x = y) = E_s[p(y | x, S)] \quad (2.6)$$

ここで、 $E_s[p(y | x, S)]$ は $p(y | x, S)$ の S に関する期待値を表している。しかしながら、条件付き交換可能性の仮定やバックドア基準が成り立たない状況も多く存在している。

2.2 平均因果効果の識別可能性に関する定理

つぎに、操作変数に加えて、代替変数を利用する平均因果効果の識別可能性に関する定理を示す。図 1.2 は想定する構造的因果モデルの因果グラフであり、 W は H の代替変数として測定されているものとする。構造的因果モデルは以下の 5 つの式か

ら成り立っている。

$$Y := f_Y(X, \mathbf{H}, \epsilon_y), X := f_X(\mathbf{H}, Z, \epsilon_x), \quad (2.7)$$

$$W := f_W(\mathbf{H}, \epsilon_w), Z := f_Z(\epsilon_z), \mathbf{H} := \mathbf{f}_H(\epsilon_h)$$

ここで、変数 $\epsilon_x, \epsilon_y, \epsilon_z, \epsilon_w, \epsilon_h$ は互いに独立な外生変数であり $\mathbf{f}_H(\epsilon_h)$ は ϵ_h の関数の集合である。このような状況は [Pearl, 2010a] や [Kuroki and Pearl, 2014] でも述べられているが、[Kuroki and Pearl, 2014] とは異なり本論文では \mathbf{H} は任意の数の非観測な変数を含むことを許容している。したがって、多くの場面において \mathbf{H} が与えられたときに X, Y, Z から独立した代替変数 W の存在を仮定することは容易であり。条件付き交換可能性の仮定やバックドア基準と比べて、条件を満たすような代替変数を観測することはそれほど困難ではないと考えられる。

図 1.2 の状況において、外生変数の複雑さに関わらず外生変数がアウトカム Y に与える影響は 4 種類のタイプに分けることができる。ここで、 x_0 を「対照的な処置」、 x_1 を「実験的な処置」とし、 y_1 を「病気が発生すること」、 y_0 が「病気発生しないこと」を表しているとする、Rothman et al. [2008, p.59] によれば、被験者は以下の 4 種類に分類できると解釈できる。

“doomed”: $u_1 = (Y_{x_0} = y_0, Y_{x_1} = y_0)$ は実験的な処置でも対照的な処置によっても病気が発生する状況を表している。

“causative”: $u_2 = (Y_{x_0} = y_0, Y_{x_1} = y_1)$ は被験者が対照的な処置を受けた場合にのみ病気が発生する状況を表している。

“preventive”: $u_3 = (Y_{x_0} = y_1, Y_{x_1} = y_0)$ は被験者が実験的な処置を受けた場合にのみ病気が発生する状況を表している。

“immune”: $u_4 = (Y_{x_0} = y_1, Y_{x_1} = y_1)$ は実験的な処置でも対照的な処置でも病気が発生しない状況を表している。

従って、この外生変数の状態の分類に従って、外生変数は 4 つの値 $\{u_1, u_2, u_3, u_4\}$ をとる変数 U として再定義される。この論文では、この変数 U の確率の識別可能性を

通して平均因果効果が識別可能であることを示す。また、操作変数 Z は以下の操作変数法の仮定を満たしているとする。

1. X_z が z の非自明関数である。
2. $Y_{x,z} = Y_x$ が任意の x と z について成立する。
3. Z が $(X_{z_0}, X_{z_1}, Y_{x_0}, Y_{x_1})$ から独立している。

すると、Pearl [2009, ch.8] による不遵守問題の設定と同様に、任意の x, y, z, w に対して、図 1.2 は図 2.1 のように書き直すことができ、 X を与えられたもとでの Y, Z, W の条件付き同時確率 $p(y, z, w | x)$ は以下のように因子分解できる。

$$p(y, z, w | x) = \sum_{i=1}^4 p(y | x, u_i) p(z | x, u_i) p(w | u_i) p(u_i | x) \quad (2.8)$$

ここで、外生変数の状態は、Balke and Pearl [1994] では 16 種類の潜在結果タイプに分けて記述しており、 X と (Y_{x_0}, Y_{x_1}) の関係は (X_{z_0}, X_{z_1}) を通して生成されるが、図 2.1 において (X_{z_0}, X_{z_1}) は周辺化されているものと考えることができる。詳しくは、次の部を参照。つぎに、潜在反応変数の一致性の性質から次のような関係が得られる。

$$p(y_0, z, w, u_i | x_0) = p(z, w, u_i | x_0), \quad i = 1, 2, \quad (2.9)$$

$$p(y_0, z, w, u_i | x_1) = p(z, w, u_i | x_1), \quad i = 1, 3, \quad (2.10)$$

$$p(y_0, z, w, u_i | x_0) = 0, \quad i = 3, 4, \quad (2.11)$$

$$p(y_0, z, w, u_i | x_1) = 0, \quad i = 2, 4 \quad (2.12)$$

これらの式から、条件付き確率 $p(y, z, | x)$ と $p(z, | x)$ は次のように書き直すことができる。

$$p(y_0, z, w | x_0) = \sum_{i=1,2} p(w | u_i) p(z | x_0, u_i) p(u_i | x_0), \quad (2.13)$$

$$p(y_0, z, w | x_1) = \sum_{i=1,3} p(w | u_i) p(z | x_1, u_i) p(u_i | x_1), \quad (2.14)$$

$$p(z, w | x_0) = \sum_{i=1}^4 p(z | x_0, u_i) p(w | u_i) p(u_i | x_0), \quad (2.15)$$

$$p(z, w | x_1) = \sum_{i=1}^4 p(z | x_1, u_i) p(w | u_i) p(u_i | x_1) \quad (2.16)$$

ほかの変数 Y, Z, W の同時確率や周辺確率も同様に書き換えることができる。また、代替変数 W を k 値の変数とする ($k \geq 4$)。ここでは値 $\{w_1, w_2, w_3, w_4\}$ をとる変数とすると、任意の $x \in x_0, x_1$ に対して、次のように確率で構成される行列を定義する。

$$P_x = \begin{pmatrix} 1 & p(z_0 | x) \\ p(w_1 | x) & p(w_1, z_0 | x) \\ p(w_2 | x) & p(w_2, z_0 | x) \\ p(w_3 | x) & p(w_3, z_0 | x) \end{pmatrix}, \quad (2.17)$$

$$Q_x = \begin{pmatrix} p(y_0 | x) & p(y_0, z_0 | x) \\ p(y_0, w_1 | x) & p(y_0, w_1, z_0 | x) \\ p(y_0, w_2 | x) & p(y_0, w_2, z_0 | x) \\ p(y_0, w_3 | x) & p(y_0, w_3, z_0 | x) \end{pmatrix}, \quad (2.18)$$

$$A_{x_0} = \begin{pmatrix} 1 & p(w_1 | u_1) & p(w_2 | u_1) & p(w_3 | u_1) \\ 1 & p(w_1 | u_2) & p(w_2 | u_2) & p(w_3 | u_2) \\ 1 & p(w_1 | u_3) & p(w_2 | u_3) & p(w_3 | u_3) \\ 1 & p(w_1 | u_4) & p(w_2 | u_4) & p(w_3 | u_4) \end{pmatrix}, \quad (2.19)$$

$$A_{x_1} = \begin{pmatrix} 1 & p(w_1 | u_1) & p(w_2 | u_1) & p(w_3 | u_1) \\ 1 & p(w_1 | u_3) & p(w_2 | u_3) & p(w_3 | u_3) \\ 1 & p(w_1 | u_2) & p(w_2 | u_2) & p(w_3 | u_2) \\ 1 & p(w_1 | u_4) & p(w_2 | u_4) & p(w_3 | u_4) \end{pmatrix}, \quad (2.20)$$

$$B_{x_0} = \begin{pmatrix} 1 & p(z_0 | x_0, u_1) \\ 1 & p(z_0 | x_0, u_2) \\ 1 & p(z_0 | x_0, u_3) \\ 1 & p(z_0 | x_0, u_4) \end{pmatrix}, \quad B_{x_1} = \begin{pmatrix} 1 & p(z_0 | x_1, u_1) \\ 1 & p(z_0 | x_1, u_3) \\ 1 & p(z_0 | x_1, u_2) \\ 1 & p(z_0 | x_1, u_4) \end{pmatrix}, \quad (2.21)$$

$$\Delta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (2.22)$$

$$M_{x_0} = \begin{pmatrix} p(u_1 | x_0) & 0 & 0 & 0 \\ 0 & p(u_2 | x_0) & 0 & 0 \\ 0 & 0 & p(u_3 | x_0) & 0 \\ 0 & 0 & 0 & p(u_4 | x_0) \end{pmatrix}, \quad (2.23)$$

$$M_{x_1} = \begin{pmatrix} p(u_1 | x_1) & 0 & 0 & 0 \\ 0 & p(u_3 | x_1) & 0 & 0 \\ 0 & 0 & p(u_2 | x_1) & 0 \\ 0 & 0 & 0 & p(u_4 | x_1) \end{pmatrix}. \quad (2.24)$$

すると、先ほど示した同時確率の因数分解と潜在反応変数の一致性の性質から、任意の $x \in \{x_0, x_1\}$ に対して、

$$P_x = A_x^T M_x B_x, \quad Q_x = A_x^T \Delta M_x B_x \quad (2.25)$$

が成立する．ここで、“T”は転置ベクトル/転置行列を表している．また、任意の $x \in \{x_0, x_1\}$ に対して、 $P_{1,x}$ と $Q_{1,x}$ をそれぞれ P_x と Q_x の最初の2行から構成される 2×2 のブロック行列とする．くわえて、任意の $x \in \{x_0, x_1\}$ に対して、 $A_{11,x}$ を A_x^T の1行目と2列目で構成される 2×2 のブロック行列、 $A_{22,x}$ を A_x^T の3行目と4

列目で構成される 2×2 のブロック行列とし、 $A_{11,x}$ と $A_{22,x}$ と A_x はそれぞれ可逆であるとする。以上の関係式を利用して、以下の潜在反応変数の同時確率の識別可能性に関する定理を導出することができる。

定理 2.2.1 (U の識別可能性). Z が (X, Y) に対する操作変数であり、 X, Y, Z, W, U の因果グラフが図 2.1 で表すことができ、同時確率 $p(x, y, z, w, u)$ の因子分解がこれまで説明した通りである。くわえて、以下の 3 条件が成り立つとき、潜在反応変数の同時確率である $p(u_i)$ ($i = 1, 2, 3, 4$) は識別可能である。

1. 任意の $x \in \{x_0, x_1\}$ に対して、行列

$$A_{11,x}, A_{22,x}, A_x, Q_{1,x}, P_{1,x} - Q_{1,x} \quad (2.26)$$

が可逆である。

2. 任意の $x \in \{x_0, x_1\}$ に対して、行列

$$Q_{2,x} Q_{1,x}^{-1} \quad (2.27)$$

と

$$(P_{2,x} - Q_{2,x})(P_{1,x} - Q_{1,x})^{-1} \quad (2.28)$$

の第 2 列ベクトルがそれぞれ異なっている。

3. 任意の $x \in \{x_0, x_1\}$, $z \in \{z_0, z_1\}$, $w \in \{w_1, w_2, w_3, w_4\}$ に対して、 $Y_{x,z,w} = Y_x$ と $Y_{x,z,w} = Y_x$ が成立する。

定理 2.2.1 で得られた結果から、平均因果効果は識別可能であり、以下の式をとおして与えられる。

$$p(Y_{x_0} = y_1) = \sum_{i=3,4} p(u_i), \quad p(Y_{x_1} = y_1) = \sum_{i=2,4} p(u_i) \quad (2.29)$$

定理 2.2.1 の仮定は、既存の研究の単調性 II などの仮定より弱いものであると考えることができる。さらに、この定理はいままで平均因果効果が識別不可能であった

問題設定においても、平均因果効果の識別可能にするものである。また、定理 2.2.1 の仮定は潜在反応タイプ U の条件付き確率に依存しているが、変数 U に基づかない仮定のもとでの識別可能性の定理に関しては次の部で与える。

2.3 考察

第 2 部で示した定理 2.2.1 を通して、先行研究の単調性 II などの仮定よりも弱い条件において、平均因果効果が識別可能であることがわかった。これまで、ノンパラメトリック操作変数モデルにおいて、潜在反応変数の同時確率を評価するために代替変数を用いることは考えられてこなかったことである。最後に、平均因果効果が識別可能である場合、モーメント法 [Pearson, 1936] を利用して平均因果効果を推定することができる。しかしながら、推定問題は特異モデルに帰着するため、特異モデルに基づくより効率的な推定法の開発が必要となってくる。詳しくは、次部で解説する。また、定理 2.2.1 は 2 値以上の値をとる離散な処置変数にも容易に拡張することができる。より幅広い状況に適用することが可能となる。しかしながら、このような場合においてはスパース性により信頼できる統計量を得ることが困難となってくる。また、[Balke and Pearl, 1994] のように連続変数や 2 値以上の離散変数のアウトカムに関しても、ある閾値 y に基づく変数 $Y \leq y$ を考えることで、本論文の結果を適用することが可能である。

第 3 部

共変量情報を用いた潜在反応タイプの 確率の識別可能性および推定手法の 開発

第 3 部では、被験者の遵守が完全でないランダム化試験 [Imbens and Angrist, 1994] における、潜在反応タイプの確率の識別可能問題を考える。潜在反応タイプとは複数の潜在反応変数の同時分布で定義されるものであり、ランダム化後の個体の振る舞いを規定する因果効果の解明において重要な指標である。この論文では、はじめに、潜在反応タイプの定義と解釈を紹介し、個体の共変量情報を利用した潜在反応タイプの識別可能性に関する 3 つの定理をあたえる。

つぎに、収集したデータを用いて潜在反応タイプの確率の推定法を開発する。提案した識別可能性の定理によって潜在反応タイプの確率が識別可能であるとき、モーメント法 [Pearson, 1936] に基づいて潜在反応タイプの確率の一致推定量を得ることができるが、この潜在反応タイプの確率を推定する問題は特異モデルの問題に帰着されてしまい標準的な統計的尤度をもちいた手法では推定が困難である。そこで、制約付き最適化問題に対する求解法である有界制約付き拡張ラグランジュ法 [Birgin and Martínez, 2020] に基づく新しい統計的推定法によって、モーメント法による推定値よりも効率的な潜在反応タイプの確率の推定値を提案する。

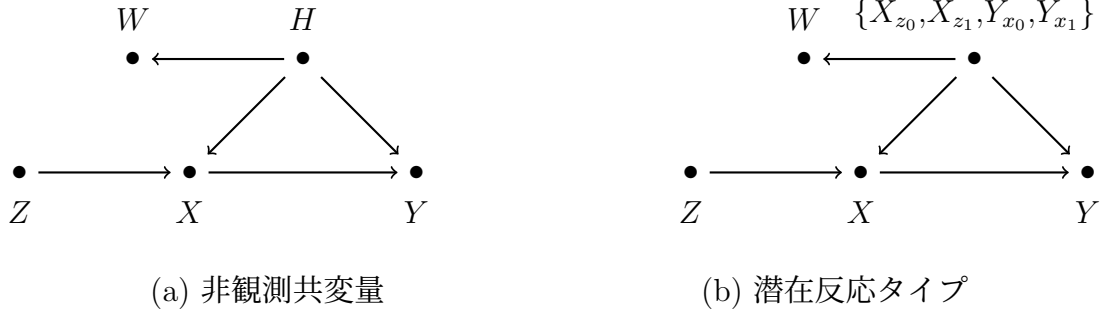


図 3.1: 共変量をともなう操作変数モデルの因果グラフ

最後に、数値実験と心理学における実験データ [Sajons, 2020] への適用例を通して提案法の有用性を検証する。第 3 部の内容は第 2 部の内容の拡張であり、[Kawakami, Shingaki, and Kuroki, 2023b] で公表されている。

3.1 準備

はじめに、潜在反応タイプの確率を識別可能性および推定に関する問題を議論するための準備を行う。表記は基本的に第 2 部で説明した通りとする。また、問題設定を以下のように定式化する。図 3.1(a) は想定する構造的因果モデルの対応する因果グラフである。図 3.1(a) に示す有向グラフに基づき潜在反応タイプの確率を識別可能性を考える。ここで、図 3.1(a) における共変量 W は非観測な共変量の集合 \mathbf{H} の代替変数 [Kuroki and Pearl, 2014, Miao et al., 2018, Tchetgen et al., 2020] として観測されているものである。なお、 \mathbf{H} は任意の個数の非観測な離散変数または連続変数の集合であることを許容している。

想定する構造的因果モデルは以下の 5 つの関係式で構成されている。

$$\begin{aligned}
 Y &:= f_Y(X, \mathbf{H}, \epsilon_y), & X &:= f_X(\mathbf{H}, Z, \epsilon_x), \\
 W &:= f_W(\mathbf{H}, \epsilon_w), & Z &:= f_Z(\epsilon_z), & \mathbf{H} &:= \mathbf{f}_H(\epsilon_h)
 \end{aligned}
 \tag{3.1}$$

ここで、変数 $\{\epsilon_x, \epsilon_y, \epsilon_z, \epsilon_w, \epsilon_h\}$ は互いに独立な外生変数であり、式 (3.1) のような構造方程式モデルを用いて構造的因果モデルを表現する場合、図 3.1(a) が対応する因果グラフとなっている。

図3.1(a)について、本論文では非観測な共変量である H が被験者の処置またはアウトカムに影響を与えるすべての離散および連続共変量を任意の個数含むことができると考えている。したがって、第一部と同様に、多くの状況において H が与えられたときに X, Y, Z から条件付きで独立した共変量 W の存在を仮定することは容易であり、上記の条件を満たすような代替変数を観測することは、条件付き交換可能性の仮定やバックドア基準と比べて、それほど困難ではないと考えられる。

つぎに、潜在反応タイプの定義をあたえる。まず、外生変数 $H \cup \{\epsilon_x, \epsilon_y, \epsilon_z, \epsilon_w\}$ の状態は、処置変数とアウトカムに関連して以下の4つの潜在反応タイプに分類できる。

$$u_1: \text{“doomed”}, u_2: \text{“causative”}, u_3: \text{“preventive”}, u_4: \text{“immune”}$$

これは、

$$\begin{aligned} u_1 &= (Y_{x_0} = y_0, Y_{x_1} = y_0), & u_2 &= (Y_{x_0} = y_0, Y_{x_1} = y_1), \\ u_3 &= (Y_{x_0} = y_1, Y_{x_1} = y_0), & u_4 &= (Y_{x_0} = y_1, Y_{x_1} = y_1) \end{aligned} \quad (3.2)$$

で定義される。また、外生変数がとる値はこのうち一つを選択することになり、外生変数は4つの値 $\{u_1, u_2, u_3, u_4\}$ をとる変数 U として再定義することができる。同様に、上記の4種類の潜在反応タイプにくわえて、操作変数と処置変数に関連して、外生変数は次の4種類の潜在反応タイプにも分けることが可能である [Frangakis and Rubin, 2002, Kowalski, 2020].

$$v_1: \text{“never take”}, v_2: \text{“compliance”}, v_3: \text{“defiance”}, v_4: \text{“always take”}$$

これは、

$$\begin{aligned} v_1 &= (X_{z_0} = x_0, X_{z_1} = x_0), & v_2 &= (X_{z_0} = x_0, X_{z_1} = x_1), \\ v_3 &= (X_{z_0} = x_1, X_{z_1} = x_0), & v_4 &= (X_{z_0} = x_1, X_{z_1} = x_1) \end{aligned} \quad (3.3)$$

で定義される。このとき、外生変数は4つの値 $\{v_1, v_2, v_3, v_4\}$ をとる変数 V と再定義できる。この潜在反応タイプは前の部では議論していなかったものである。

本研究では潜在反応タイプの確率，すなわち，

$$p(u_i), p(v_j), p(u_i, v_j), \quad \forall i, j = 1, 2, 3, 4 \quad (3.4)$$

の識別可能性および推定の問題を考える．とくに， U と V の同時確率は 4×4 つの組み合わせで，16種類の潜在反応タイプとなっている．16種類の潜在反応タイプが与えられるとランダム化後の全ても振る舞いが決定されることとなる [Balke and Pearl, 1997]．たとえば

$$(X_{z_0} = x_0, X_{z_1} = x_1, Y_{x_0} = y_0, Y_{x_1} = y_1) \quad (3.5)$$

である個体は， $Z = z_1$ に割り当てられると潜在反応変数の一致性よりランダム化後の変数は $(X, Y) = (x_1, y_1)$ に一意に決定される．逆に，2値の処置変数とアウトカムに関してはこの潜在反応タイプ以外の情報は冗長なものとなる．これらの確率は因果効果の評価のために重要な指標ではあるが，潜在反応変数の同時分布は各被験者について観測不可能であり，完全なランダム化試験でも識別不可能なものである．最後に，第2部と同様に，操作変数法の仮定は以下の通りである．

1. X_z は z の非自明関数である．
2. $Y_{x,z} = Y_x$ は任意の x と z について成り立つ．
3. Z は $(X_{z_0}, X_{z_1}, Y_{x_0}, Y_{x_1})$ から独立している．

この論文では，操作変数法にくわえて代替変数を利用し，潜在反応タイプの確率の識別可能性に関する3つの定理を示し，それらの推定法を開発する．

3.2 潜在反応タイプの確率の識別可能性

この節では，潜在反応タイプの確率の識別可能性に関する3つの定理を紹介する．任意の x, y, z, w について，因果グラフを図3.1(a)が図3.1(b)のように記述できる

と仮定すると、 X を与えられたときの Y, Z, W の条件付きの同時確率 $p(y, z, w|x)$ は次のような因子分解で与えられる。

$$p(y, z, w | x) = \sum_{i=1}^4 \sum_{j=1}^4 p(y | x, u_i, v_j) p(z | x, u_i, v_j) \times p(w | u_i, v_j) p(u_i, v_j | x) \quad (3.6)$$

また、任意の x, y, z, w に対して、確率の正值性 [Hernán and Robins, 2006]

$$p(y, z, w, u, v | x) > 0 \quad (3.7)$$

を仮定する。

3.2.1 $p(Y_{x_0} = y, Y_{x_1} = y')$ の識別可能性

はじめに、潜在反応タイプの確率 $p(Y_{x_0} = y, Y_{x_1} = y')$ ($y, y' \in \{y_0, y_1\}$) の識別可能性に関する定理をあたえる。代替変数 W を k 値 ($k \geq 4$) とする。ここでは w_1, w_2, w_3, w_4 の値をとる変数としたときに、任意の $x \in \{x_0, x_1\}$ に対して、次のように確率のブロック行列を定義する。

$$[P_x; Q_x] = \begin{pmatrix} P_{1,x} & Q_{1,x} \\ P_{2,x} & Q_{2,x} \end{pmatrix} \quad (3.8)$$

ここで、ブロック行列の各成分は

$$P_{1,x} = \begin{pmatrix} 1 & p(z_0 | x) \\ p(w_1 | x) & p(w_1, z_0 | x) \end{pmatrix}, \quad (3.9)$$

$$P_{2,x} = \begin{pmatrix} p(w_2 | x) & p(w_2, z_0 | x) \\ p(w_3 | x) & p(w_3, z_0 | x) \end{pmatrix}, \quad (3.10)$$

$$Q_{1,x} = \begin{pmatrix} p(y_0 | x) & p(y_0, z_0 | x) \\ p(y_0, w_1 | x) & p(y_0, w_1, z_0 | x) \end{pmatrix}, \quad (3.11)$$

$$Q_{2,x} = \begin{pmatrix} p(y_0, w_2 | x) & p(y_0, w_2, z_0 | x) \\ p(y_0, w_3 | x) & p(y_0, w_3, z_0 | x) \end{pmatrix} \quad (3.12)$$

である. 図 3.2(a) の状況において, 4 種類の潜在反応タイプ U の確率の識別可能性に関する以下の定理を導出できる.

定理 3.2.1 ($p(u)$ の識別可能性). W は k 値の変数とする ($k \geq 4$). ここでは, $w \in \{w_1, w_2, w_3, w_4\}$ とすると, 以下の 5 つの条件を満たすとき, 潜在反応タイプの確率 $p(u_1), p(u_2), p(u_3), p(u_4)$ は識別可能である.

1. 任意の $x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_0, z_1\}$ $w \in \{w_1, w_2, w_3, w_4\}$ に対して確率 $p(y, z, w|x)$ が識別可能である.
2. 任意の $x \in \{x_0, x_1\}$, $z \in \{z_0, z_1\}$, $w \in \{w_1, w_2, w_3, w_4\}$ に対して, 確率 $p(w|x, z, u)$ が 0 より大きく,

$$p(w|x, z, u) = p(w|u) \quad (3.13)$$

が成立する.

3. 4×4 の行列 $[P_x; Q_x]$ に対して, $P_{i,x}, Q_{i,x}$ を 2×2 ブロック行列とすると, 任意の $i = 1, 2$, $x \in \{x_0, x_1\}$ において,

$$P_{i,x} - Q_{i,x} \quad (3.14)$$

は可逆である.

4. 行列

$$Q_{2,x}Q_{1,x}^{-1} \text{ と } (P_{2,x} - Q_{2,x})(P_{1,x} - Q_{1,x})^{-1} \quad (3.15)$$

の二列目のベクトルが,

$$Q_{2,x'}Q_{1,x'}^{-1} \text{ と } (P_{2,x'} - Q_{2,x'})(P_{1,x'} - Q_{1,x'})^{-1} \quad (3.16)$$

の二列目のベクトルとはそれぞれ異なる. ただし, $x \neq x'$ である.

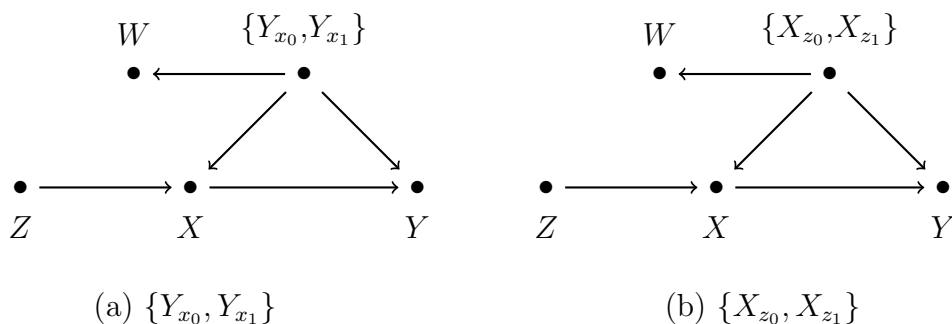


図 3.2: 共変量をともなう操作変数法の因果グラフ

5. Z は (X, Y) の操作変数である.

この定理は第 1 部での定理 2.2.1 の改良版である. 定理 2.2.1 とは違い, 定理の仮定が外生変数の分布で構成されており, 検証可能である. 第 2 部と同様に, 定理 3.2.1 は代表的な識別条件であるバックドア基準などが平均因果効果の推定に適用できない場合でも, 平均因果効果が識別できる場合があることを示しており, 平均因果効果は次の式で与えられる.

$$p(Y_{x_0} = y_1) = \sum_{i=3,4} p(u_i), \quad p(Y_{x_1} = y_1) = \sum_{i=2,4} p(u_i) \quad (3.17)$$

くわえて, 先行研究における単調性 II の仮定 ($p(u_3) = 0$) [Tian and Pearl, 2000] が検証可能なものとなった.

3.2.2 $p(X_{z_0} = x, X_{z_1} = x')$ の識別可能性

つぎに, 潜在反応タイプの確率 $p(X_{z_0} = x, X_{z_1} = x')$ ($x, x' \in \{x_0, x_1\}$) の識別可能性に関する定理をあたえる. 代替変数 W を k 値の変数 ($k \geq 3$) とする. ここでは, $\{w_1, w_2, w_3\}$ の値をとる変数とすると, 任意の $x \in \{x_0, x_1\}$, $z \in \{z_0, z_1\}$ に対して, 次のように確率で構成される行列を定義する.

$$P_{x,z}^w = \begin{pmatrix} 1 & p(y | x, z) \\ p(w | x, z) & p(y, w | x, z) \end{pmatrix} \quad (3.18)$$

すると、図3.2(b)に示すような状況において、4種類の潜在反応タイプ V の確率の識別可能性に関する以下の定理を導出できる。

定理 3.2.2 ($p(v)$ の識別可能性). W を k 値の変数とする($k \geq 3$). ここでは $w \in \{w_1, w_2, w_3\}$ とする. 以下の4つの条件を満たす場合に、潜在反応タイプの確率 $p(v_1)$, $p(v_2)$, $p(v_3)$, $p(v_4)$ は識別可能である.

1. 任意の $x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_0, z_1\}$, $w \in \{w_1, w_2, w_3\}$ に対して、確率 $p(y, w|x, z)$ が識別可能である.
2. 任意の $x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_0, z_1\}$, $w \in \{w_1, w_2, w_3\}$, $v \in \{v_1, v_2, v_3, v_4\}$ に対して、 $p(y, w|x, z, v)$ が0より大きく、

$$p(y|x, w, z, v) = p(y|x, v) \text{ と } p(w|x, z, v) = p(w|v) \quad (3.19)$$

が成立する.

3. 任意の x, y, z, w に対して、 $P_{x,z}^w$ が可逆であり、

$$\frac{\det(P_{x,z}^{w_1})}{\det(P_{x,z}^w)} \neq \frac{\det(P_{x',z'}^{w_1})}{\det(P_{x',z'}^w)} \quad (3.20)$$

を満たす. ただし、 $w \neq w_1$, $(x, z) \neq (x', z')$ ($x, x' \in \{x_0, x_1\}$, $z, z' \in \{z_0, z_1\}$, $w \in \{w_2, w_3\}$)である.

4. Z は (X, Y) に対する操作変数である.

この定理により、局所的平均因果効果[Imbens and Angrist, 1994]で重要になる“compliance”の確率は $p(v_2)$ が識別可能となり、さらに、先行研究における単調性Iの仮定($p(v_3) = 0$)が検証可能なものとなった.

3.2.3 $p(X_{z_0} = x, X_{z_1} = x', Y_{x_0} = y, Y_{x_1} = y')$ の識別可能性

さいごに、定理3.2.1と定理3.2.2と合わせて16種類の潜在反応タイプ (U, V) の確率 $p(X_{z_0} = x, X_{z_1} = x', Y_{x_0} = y, Y_{x_1} = y')$ ($x, x' \in \{x_0, x_1\}$, $y, y' \in \{y_0, y_1\}$)の識別可能性に関する定理をあたえる.

定理 3.2.3 ($p(u, v)$ の識別可能性). W は k 値の変数とする ($k \geq 4$). ここでは, $w \in \{w_1, w_2, w_3, w_4\}$ とすると, 以下の2つの条件を満たす場合, 潜在反応タイプの確率 $p(u_i, v_j)$ ($i, j = 1, 2, 3, 4$) は識別可能である.

1. 任意の $u \in \{u_1, u_2, u_3, u_4\}$, $v \in \{v_1, v_2, v_3, v_4\}$, $w \in \{w_1, w_2, w_3, w_4\}$ に対して, $p(v|w)$ と $p(u|w)$ の両方が識別可能である.

2. 以下のいずれか1つの条件を満たす.

(a) 任意の $u \in \{u_1, u_2, u_3, u_4\}$, $v \in \{v_1, v_2, v_3, v_4\}$, $w \in \{w_1, w_2, w_3, w_4\}$ に対して, 確率 $p(u, w|v)$ は0より大きく, $p(u|v, w) = p(u|v)$ が成り立ち, 4×4 行列

$$(p(v_i|w_j))_{i,j=1,\dots,4} \quad (3.21)$$

が可逆である.

(b) 任意の $u \in \{u_1, u_2, u_3, u_4\}$, $v \in \{v_1, v_2, v_3, v_4\}$, $w \in \{w_1, w_2, w_3, w_4\}$ に対して, 確率 $p(v, w|u)$ が0より大きく, $p(v|u, w) = p(v|u)$ が成立して, 4×4 行列

$$(p(u_i|w_j))_{i,j=1,\dots,4} \quad (3.22)$$

が可逆である.

この定理の条件 (1) と (2-a) のもとで, 次のような関係式

$$\begin{pmatrix} p(u|v_1) \\ \vdots \\ p(u|v_4) \end{pmatrix} = \begin{pmatrix} p(v_1|w_1) & \cdots & p(v_4|w_1) \\ \vdots & \ddots & \vdots \\ p(v_1|w_4) & \cdots & p(v_4|w_4) \end{pmatrix}^{-1} \begin{pmatrix} p(u|w_1) \\ \vdots \\ p(u|w_4) \end{pmatrix} \quad (3.23)$$

が任意の $u \in \{u_1, u_2, u_3, u_4\}$ に対して成立する. ここで, 確率 $p(u|v)$ が識別可能なら

$$p(u, v) = \sum_{i=1}^4 p(u|v)p(v|w_i) \quad (3.24)$$

が、任意の $u \in \{u_1, u_2, u_3, u_4\}, v \in \{v_1, v_2, v_3, v_4\}$ に対して成立する。条件 (1) と (2-b) の元でも同様の式展開によって、16種類の潜在反応タイプ (U, V) の確率が識別可能となる。

3.3 潜在反応タイプの確率の推定法

定理 3.2.1 - 3.2.3 の証明を通して潜在反応タイプの確率は観測可能な確率で構成される明示的な式によって表現可能である。(付録を参照)。したがって、モーメント法 [Pearson, 1936] をもちいて正規条件下のデルタ法によって、これらの確率の漸近正規性をみたく一致推定量が得られる [Ferguson, 1996]。しかしながら、この推定問題は特異モデルに帰着されてしまううえ、モーメント法による推定量は複雑な式で構成されるため、非常に不安定な推定となってしまう。

そのため、本節ではより効率的な推定を行うために新たな潜在反応タイプの確率の統計的推定法を有界制約付き拡張ラグランジュ法 [Birgin and Martínez, 2020] に基づいて開発する。ここで、拡張ラグランジュ法とは、ラグランジュ法とペナルティ関数法を組み合わせた最適化問題に対する手法であり、ペナルティ関数法は制約付きの最適化問題を制約なしの最適化問題に変換してから勾配降下法などで解く手法である。

3.3.1 $p(Y_{x_0} = y, Y_{x_1} = y')$ の推定法

最初に、 $p(Y_{x_0} = y, Y_{x_1} = y')$ の推定法を説明する。図 3.1(a) の構造的因果モデルから得られる、サンプルサイズ N のデータセットを

$$(X_i, Y_i, Z_i, W_i)_{i=1}^N \quad (3.25)$$

とする。くわえて、任意の $x \in \{x_0, x_1\}, x' \in \{x_0, x_1\}, y \in \{y_0, y_1\}, z \in \{z_0, z_1\}, w \in \{w_1, w_2, w_3, w_4\}$ に対して、 $\hat{p}(y, w, z|x)$ をデータセットから得られる確率 $p(y, w, z|x)$ の最尤推定量とする。さらに、任意の $x \in \{x_0, x_1\}$ に対して、定理 3.2.1 における行

アルゴリズム 1 $p(Y_{x_0} = y, Y_{x_1} = y')$ の推定アルゴリズム

Input: $\widehat{P}_{x_i}, \widehat{Q}_{x_i}$ for $i = 0, 1, L$, α , $\{\delta^{(l)}\}_{l=0}^L$, $\{\mu^{(l)}\}_{l=0}^L$

Initialize: $\Theta^{(0)}$, $\lambda_{i,j}^{+(0)}$, $\lambda_{i,j}^{-(0)}$ for $i = 0, 1$ $j = 1, 2, 3, 4$

$(e_{i,1}^{(0)}, \dots, e_{i,4}^{(0)})^T \leftarrow \Theta_{x_i}^{(0)-T} \widehat{P}_{x_i}(1, 0)^T$ for $i = 0, 1$

for $l = 0, \dots, L - 1$ **do**

$t \leftarrow 0$; $\Theta^{[0]} \leftarrow \Theta^{(l)}$.

while $\left\| \nabla_{\Theta} F \left(\Theta^{[t]}, \left\{ \lambda_{i,j}^{+(l)}, \lambda_{i,j}^{-(l)} : i = 0, 1; j = 1, 2, 3, 4 \right\}, \mu^{(l)} \right) \right\|_F > \delta^{(l)}$ **do**

$\Theta^{[t+1]} \leftarrow P_{[0,1]} \left(\Theta^{[t]} - \alpha \nabla_{\Theta} F \left(\Theta^{[t]}, \left\{ \lambda_{i,j}^{+(l)}, \lambda_{i,j}^{-(l)} : i = 0, 1; j = 1, 2, 3, 4 \right\}, \mu^{(l)} \right) \right)$

$\Theta^{(l)} \leftarrow \Theta^{[t+1]}$

$t \leftarrow t + 1$

end while

$\Theta^{(l+1)} \leftarrow \Theta^{(l)}$

$(e_{i,1}^{(l+1)}, \dots, e_{i,4}^{(l+1)})^T \leftarrow \left(\Theta_{x_i}^{(l+1)} \right)^{-T} \widehat{P}_{x_i}(1, 0)^T$ for $i = 0, 1$

$\lambda_{i,j}^{+(l+1)} \leftarrow \lambda_{i,j}^{+(l)} + \mu^{(l)} \max \left\{ e_{i,j}^{(l+1)} - 1 + \frac{\lambda_{i,j}^{+(l)}}{\mu^{(l)}}, 0 \right\}$ for $i = 0, 1$ $j = 1, 2, 3, 4$

$\lambda_{i,j}^{-(l+1)} \leftarrow \lambda_{i,j}^{-(l)} + \mu^{(l)} \max \left\{ -e_{i,j}^{(l+1)} + \frac{\lambda_{i,j}^{-(l)}}{\mu^{(l)}}, 0 \right\}$ for $i = 0, 1$ $j = 1, 2, 3, 4$

end for

$(\widehat{p}(u_1|x_i), \dots, \widehat{p}(u_4|x_i))^T \leftarrow \left(\Theta_{x_i}^{(L)} \right)^{-T} \widehat{P}_{x_i}(1, 0)^T$ for $i = 0, 1$

$\widehat{p}(u_j) \leftarrow \sum_{i=0}^1 \widehat{p}(u_j|x_i) \widehat{p}(x_i)$ for $j = 1, 2, 3, 4$

列 P_x と Q_x のプラグイン推定量 [Vaart, 1998] を

$$\widehat{P}_x = \begin{pmatrix} 1 & \widehat{p}(z_0 | x) \\ \widehat{p}(w_1 | x) & \widehat{p}(w_1, z_0 | x) \\ \widehat{p}(w_2 | x) & \widehat{p}(w_2, z_0 | x) \\ \widehat{p}(w_3 | x) & \widehat{p}(w_3, z_0 | x) \end{pmatrix}, \quad (3.26)$$

$$\widehat{Q}_x = \begin{pmatrix} \widehat{p}(y_0 | x) & \widehat{p}(y_0, z_0 | x) \\ \widehat{p}(y_0, w_1 | x) & \widehat{p}(y_0, w_1, z_0 | x) \\ \widehat{p}(y_0, w_2 | x) & \widehat{p}(y_0, w_2, z_0 | x) \\ \widehat{p}(y_0, w_3 | x) & \widehat{p}(y_0, w_3, z_0 | x) \end{pmatrix} \quad (3.27)$$

とする。また、行列

$$A_{x_0} = \begin{pmatrix} 1 & p(w_1 | u_1) & p(w_2 | u_1) & p(w_3 | u_1) \\ 1 & p(w_1 | u_2) & p(w_2 | u_2) & p(w_3 | u_2) \\ 1 & p(w_1 | u_3) & p(w_2 | u_3) & p(w_3 | u_3) \\ 1 & p(w_1 | u_4) & p(w_2 | u_4) & p(w_3 | u_4) \end{pmatrix}, \quad (3.28)$$

$$A_{x_1} = \begin{pmatrix} 1 & p(w_1 | u_1) & p(w_2 | u_1) & p(w_3 | u_1) \\ 1 & p(w_1 | u_3) & p(w_2 | u_3) & p(w_3 | u_3) \\ 1 & p(w_1 | u_2) & p(w_2 | u_2) & p(w_3 | u_2) \\ 1 & p(w_1 | u_4) & p(w_2 | u_4) & p(w_3 | u_4) \end{pmatrix} \quad (3.29)$$

と定義すると、任意の $x \in \{x_0, x_1\}$ に対し、

$$P_x = A_x^T M_x B_x, \quad Q_x = A_x^T M_x \Delta B_x \quad (3.30)$$

の関係式が成立している。本論文では、行列 A_{x_0} と A_{x_1} を推定することを通して潜在反応タイプ U の確率を推定することを考える。

次に、上記の関係式を利用した行列 A_{x_0} と A_{x_1} の推定問題を紹介する。行列 A_{x_0} と A_{x_1} の推定のための行列を

$$\Theta_{x_0} = \begin{pmatrix} 1 & \theta_{11} & \theta_{12} & \theta_{13} \\ 1 & \theta_{21} & \theta_{22} & \theta_{23} \\ 1 & \theta_{31} & \theta_{32} & \theta_{33} \\ 1 & \theta_{41} & \theta_{42} & \theta_{43} \end{pmatrix}, \quad \Theta_{x_1} = \begin{pmatrix} 1 & \theta_{11} & \theta_{12} & \theta_{13} \\ 1 & \theta_{31} & \theta_{32} & \theta_{33} \\ 1 & \theta_{21} & \theta_{22} & \theta_{23} \\ 1 & \theta_{41} & \theta_{42} & \theta_{43} \end{pmatrix} \quad (3.31)$$

で用意すると、行列 A_{x_0} と A_{x_1} の推定問題は以下のような拡張ラグランジュ関数の最適化問題として定式化できる。

$$\begin{aligned}
& F\left(\Theta, \left\{\lambda_{i,j}^{+(l)}, \lambda_{i,j}^{-(l)} : i = 0, 1; j = 1, 2, 3, 4\right\}, \mu\right) \\
&= \sum_{i=0}^1 \|\Theta_{x_i}^T \Delta(\Theta_{x_i})^{-T} \widehat{P}_{x_i} - \widehat{Q}_{x_i}\|_F^2 \\
&+ \frac{\mu}{2} \left(\sum_{i=0}^1 \sum_{j=1}^4 \max \left\{ e_{i,j} - 1 + \frac{\lambda_{i,j}^+}{\mu}, 0 \right\} \right)^2 \\
&+ \frac{\mu}{2} \left(\sum_{i=0}^1 \sum_{j=1}^4 \max \left\{ -e_{i,j} + \frac{\lambda_{i,j}^-}{\mu}, 0 \right\} \right)^2 \\
&+ \sum_{i=0}^1 \sum_{j=1}^4 \lambda_{i,j}^+ \max \left\{ e_{i,j} - 1 + \frac{\lambda_{i,j}^+}{\mu}, 0 \right\} \\
&+ \sum_{i=0}^1 \sum_{j=1}^4 \lambda_{i,j}^- \max \left\{ -e_{i,j} + \frac{\lambda_{i,j}^-}{\mu}, 0 \right\} \tag{3.32}
\end{aligned}$$

このラグランジュ関数に対して勾配降下法 [Chong and Zak, 2013] を適用し行列 A_{x_0} と A_{x_1} の推定を行う。ここで、 $\|G\|_F$ は行列 G のフロベニウスノルムであり $G^{-T} = (G^T)^{-1}$ とである。推定のためのパラメータの集合

$$\Theta = \{\theta_{i,j} : i = 1, 2, 3, 4; j = 1, 2, 3\} \tag{3.33}$$

は Θ_{x_0} と Θ_{x_1} を構成する 12 個のパラメータから成り立っている。また、任意の $i = 0, 1$, $j = 1, 2, 3, 4$ に対し、

$$\mu, \lambda_{i,j}^+, \lambda_{i,j}^- \tag{3.34}$$

は各制約に対するラグランジュ乗数とよばれる係数である。 $m \times n$ 行列 $G = (g_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n}$ を入力とする関数 $P_{[0,1]}$ を

$$P_{[0,1]}(G) = \begin{cases} 0, & \text{if } g_{i,j} \leq 0 \\ g_{i,j}, & \text{if } g_{i,j} \in (0, 1) \quad \forall i = 1, \dots, m, \forall j = 1, \dots, n \\ 1, & \text{if } g_{i,j} \geq 1 \end{cases} \tag{3.35}$$

で定義する。これは、バリア関数と呼ばれるものである [Nocedal and Wright, 1999].

任意の $i = 0, 1$, $j = 1, 2, 3, 4$ に対し, 各 $e_{i,j}$ は行列

$$\Theta_{x_i}^{-T} \hat{P}_{x_i} (1, 0)^T \quad (3.36)$$

の j 番目の要素であり, これらは確率の推定量であるので, $0 \sim 1$ の範囲に収まるように上記のバリア関数をもちいて制約をかけている.

提案法の擬似コードをアルゴリズム 1 に記載している. α はステップサイズであり, L は反復回数, $\Theta^{(0)}$ は Θ の初期パラメータである. また, $\delta^{(l)}$ は l 回目の反復に関するパラメータである. アルゴリズム 1 により行列 A_{x_0} と A_{x_1} の推定量 $\hat{\Theta}$ を得る. 推定量 $\hat{p}(u_i|x_j)$ ($i = 1, 2, 3, 4, j = 0, 1$) は, 行列 A_{x_0} と A_{x_1} の要素の推定量 $\hat{\Theta}$ を用いて, 処置群と対照群における潜在反応タイプ U の条件付き確率の推定量は, 以下のよう求められる.

$$\hat{\Theta}_{x_0}^{-T} \hat{P}_{x_0} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \hat{p}(u_1|x_0) \\ \hat{p}(u_2|x_0) \\ \hat{p}(u_3|x_0) \\ \hat{p}(u_4|x_0) \end{pmatrix}, \quad \hat{\Theta}_{x_1}^{-T} \hat{P}_{x_1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \hat{p}(u_1|x_1) \\ \hat{p}(u_3|x_1) \\ \hat{p}(u_2|x_1) \\ \hat{p}(u_4|x_1) \end{pmatrix} \quad (3.37)$$

最後に, 以上の結果から潜在反応タイプ U の確率の推定量は次のように得られる.

$$\hat{p}(u_i) = \hat{p}(u_i|x_0)\hat{p}(x_0) + \hat{p}(u_i|x_1)\hat{p}(x_1), \quad i = 1, 2, 3, 4. \quad (3.38)$$

3.3.2 $p(X_{z_0} = x, X_{z_1} = x')$ の推定法

次に, $p(X_{z_0} = x, X_{z_1} = x')$ の推定法について説明する. 図 3.1(b) の構造的因果モデルから得られる, サンプルサイズ N のデータセットを

$$(X_i, Y_i, Z_i, W_i)_{i=1}^N \quad (3.39)$$

とする. また, 任意の $x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_0, z_1\}$, $w \in \{w_1, w_2, w_3, w_4\}$ に対して, $\hat{p}(y, w|x, z)$ をデータセットから得られる $p(y, w|x, z)$ の最尤推定量とする.

さらに、任意の $x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_0, z_1\}$, $w \in \{w_1, w_2, w_3\}$ に対して、行列 $P_{x,z}^w$ のプラグイン推定量を次のように計算する。

$$\hat{P}_{x,z}^w = \begin{pmatrix} 1 & \hat{p}(y|x, z) \\ \hat{p}(w|x, z) & \hat{p}(y, w|x, z) \end{pmatrix} \quad (3.40)$$

つぎに、任意の $w \in \{w_1, w_2, w_3\}$ に対して、確率で構成される行列

$$A_{x_0, z_0}^w = \begin{pmatrix} 1 & p(w | v_1) \\ 1 & p(w | v_2) \end{pmatrix}, \quad A_{x_1, z_0}^w = \begin{pmatrix} 1 & p(w | v_3) \\ 1 & p(w | v_4) \end{pmatrix}, \quad (3.41)$$

$$A_{x_0, z_1}^w = \begin{pmatrix} 1 & p(w | v_1) \\ 1 & p(w | v_3) \end{pmatrix}, \quad A_{x_1, z_1}^w = \begin{pmatrix} 1 & p(w | v_2) \\ 1 & p(w | v_4) \end{pmatrix} \quad (3.42)$$

を導入する。これらの行列は、任意の $x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_0, z_1\}$, $w, w' \in \{w_1, w_2, w_3\}$ に対し、

$$P_{x,z}^{w'} (P_{x,z}^w)^{-1} (A_{x,z}^w)^T = (A_{x,z}^{w'})^T \quad (3.43)$$

の関係式を満たしている。本論文では、行列 $A_{x,z}^w$ ($x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_0, z_1\}$, $w \in \{w_1, w_2, w_3\}$) を推定することを通して潜在反応タイプ V の確率を推定することを考える。

次に、行列 $A_{x,z}^w$ ($x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_0, z_1\}$, $w \in \{w_1, w_2, w_3\}$) の推定問題を紹介する。行列 $A_{x,z}^w$ ($x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_0, z_1\}$, $w \in \{w_1, w_2, w_3\}$) の推定のためのパラメータを、任意の $i = 1, 2, 3$ に対して、

$$\Theta_{x_0, z_0}^{w_i} = \begin{pmatrix} 1 & \theta_{i1} \\ 1 & \theta_{i2} \end{pmatrix}, \quad \Theta_{x_0, z_1}^{w_i} = \begin{pmatrix} 1 & \theta_{i1} \\ 1 & \theta_{i3} \end{pmatrix}, \quad (3.44)$$

$$\Theta_{x_1, z_0}^{w_i} = \begin{pmatrix} 1 & \theta_{i3} \\ 1 & \theta_{i4} \end{pmatrix}, \quad \Theta_{x_1, z_1}^{w_i} = \begin{pmatrix} 1 & \theta_{i2} \\ 1 & \theta_{i4} \end{pmatrix} \quad (3.45)$$

とする。行列 $A_{x,z}^w$ ($x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_0, z_1\}$, $w \in \{w_1, w_2, w_3\}$) の推定問題は、次のような拡張ラグランジュ関数の最適化問題に帰着する。

$$F(\Theta, \{\lambda_{i,j,k}^+, \lambda_{i,j,k}^- : i, j = 0, 1; k = 1, 2\}, \mu)$$

$$\begin{aligned}
&= \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=0}^1 \sum_{l=0}^1 \|\widehat{P}_{x_k, z_l}^{w_j} \widehat{P}_{x_k, z_l}^{w_i-1} (\Theta_{x_k, z_l}^{w_i})^T - (\Theta_{x_k, z_l}^{w_j})^T\|_F^2 \\
&\quad + \frac{\mu}{2} \left(\sum_{i=0}^1 \sum_{j=0}^1 \sum_{k=1}^2 \max \left\{ e_{i,j,k} - 1 + \frac{\lambda_{i,j,k}^+}{\mu}, 0 \right\} \right)^2 \\
&\quad + \frac{\mu}{2} \left(\sum_{i=0}^1 \sum_{j=0}^1 \sum_{k=1}^2 \max \left\{ -e_{i,j,k} + \frac{\lambda_{i,j,k}^-}{\mu}, 0 \right\} \right)^2 \\
&\quad + \sum_{i=0}^1 \sum_{j=0}^1 \sum_{k=1}^2 \lambda_{i,j,k}^+ \max \left\{ e_{i,j,k} - 1 + \frac{\lambda_{i,j,k}^+}{\mu}, 0 \right\} \\
&\quad + \sum_{i=0}^1 \sum_{j=0}^1 \sum_{k=1}^2 \lambda_{i,j,k}^- \max \left\{ -e_{i,j,k} + \frac{\lambda_{i,j,k}^-}{\mu}, 0 \right\} \tag{3.46}
\end{aligned}$$

パラメータの集合

$$\Theta = \{\theta_{i,j} : i = 1, 2, 3; j = 1, 2, 3, 4\} \tag{3.47}$$

は $\Theta_{x_i, z_j}^{w_k}$ ($i = 0, 1, j = 0, 1, k = 1, 2, 3$) を構成する 12 個のパラメータで構成されている。任意の $i = 0, 1, j = 0, 1, k = 1, 2$ に対し、パラメータ

$$\mu, \lambda_{i,j,k}^+, \lambda_{i,j,k}^- \tag{3.48}$$

は各制約に対するラグランジュ乗数である。くわえて、任意の $i, j = 0, 1, k = 1, 2, 3$ に対し、アルゴリズム 2 における $e_{i,j,k}$ は確率の推定量であり、これが 0 ~ 1 の範囲に収まるようにバリア関数を利用して制約をかけている。式

$$\sum_{k=1}^3 \beta_{i,j,k} \left(\widehat{\Theta}_{x_i, z_j}^{w_k} \right)^{-T} \widehat{P}_{x_i, z_j}^{w_k} (1, 0)^T \tag{3.49}$$

によって、 $(X, Z) = (x_i, z_j)$ ($i, j = 0, 1$) が与えられたもとでの、潜在反応タイプ V の確率の推定量がえられる。ここで、係数 $\beta_{i,j,k}$ ($i, j = 0, 1; k = 1, 2, 3$) は

$$\sum_{k=1}^3 \beta_{i,j,k} = 1, \forall i, j = 0, 1 \tag{3.50}$$

を満たす重みの集合である。

提案法の擬似コードをアルゴリズム 2 に記載している。 α はステップサイズであり、 L は反復回数であり、 $\Theta^{(0)}$ は Θ の初期値である。また、 $\delta^{(l)}$ は l 回目の反復に関する

るパラメータである。アルゴリズム 2 にて得られた推定量を利用して $\hat{p}(v_i|x_j, z_k)$ ($i = 1, 2, 3; j, k = 0, 1$) は以下のように求められる。

$$(\hat{\Theta}_{x,z}^w)^{-T} \hat{P}_{x,z}^w = \begin{pmatrix} \hat{p}(v_i|x, z) & \hat{p}(y, v_i|x, z) \\ \hat{p}(v_j|x, z) & \hat{p}(y, v_j|x, z) \end{pmatrix} \quad (3.51)$$

以上の議論より，潜在反応タイプ V の確率は次のように推定できる。

$$\hat{p}(v_i) = \sum_{j=1}^1 \sum_{k=0}^1 \hat{p}(v_i|x_j, z_k) \hat{p}(x_j, z_k), \quad i = 1, 2, 3, 4 \quad (3.52)$$

3.3.3 $p(X_{z_0} = x, X_{z_1} = x', Y_{x_0} = y, Y_{x_1} = y')$ の推定法

さいごに，定理 3.2.3 における関係式

$$\begin{pmatrix} p(u|v_1) \\ \vdots \\ p(u|v_4) \end{pmatrix} = \begin{pmatrix} p(v_1|w_1) & \cdots & p(v_4|w_1) \\ \vdots & \ddots & \vdots \\ p(v_1|w_4) & \cdots & p(v_4|w_4) \end{pmatrix}^{-1} \begin{pmatrix} p(u|w_1) \\ \vdots \\ p(u|w_4) \end{pmatrix} \quad (3.53)$$

に対して，アルゴリズム 1 と 2 で得られた推定量を式 (3.53) に代入することで条件付き確率 $p(u|v)$ が推定できる。16 種類の潜在反応タイプの確率 $p(u, v)$ は，さらに，アルゴリズム 2 で得られた $p(v)$ を利用することで推定できる。

アルゴリズム 2 $p(X_{z_0} = x, X_{z_1} = x')$ の推定アルゴリズム

Input: $\hat{R}_{x_i, z_j}^{w_k}$ for $i, j = 0, 1$ $k = 1, 2, 3, L$, α , $\{\delta^{(l)}\}_{l=0}^L$, $\{\mu^{(l)}\}_{l=0}^L$, $\{\beta_{i,j,k}^{(l)}\}_{l=0}^L$ for $i, j = 0, 1; k = 1, 2, 3$

Initialize: $\Theta^{(0)}$, $\lambda_{i,j,k}^{+(0)}$, $\lambda_{i,j,k}^{-(0)}$ for $i, j = 0, 1$ $k = 1, 2$

$(e_{i,j,1}^{(0)}, e_{i,j,2}^{(0)})^T \leftarrow \sum_{k=1}^3 \beta_{i,j,k}^{(0)} \left(\Theta_{x_i, z_j}^{w_k(0)} \right)^{-T} \hat{P}_{x_i, z_j}^{w_k} (1, 0)^T$ for $i, j = 0, 1$

for $l = 0, \dots, L - 1$ **do**

$t \leftarrow 0$; $\Theta^{[0]} \leftarrow \Theta^{(l)}$

while $\left\| \nabla_{\Theta} F \left(\Theta^{[t]}, \left\{ \lambda_{i,j,k}^{+(l)}, \lambda_{i,j,k}^{-(l)} : i, j = 1, 2; k = 1, 2 \right\}, \mu^{(l)} \right) \right\|_F > \delta^{(l)}$ **do**

$\Theta^{[t+1]} \leftarrow P_{[0,1]} \left(\Theta^{[t]} - \alpha \nabla_{\Theta} F \left(\Theta^{[t]}, \left\{ \lambda_{i,j,k}^{+(l)}, \lambda_{i,j,k}^{-(l)} : i, j = 1, 2; k = 1, 2 \right\}, \mu^{(l)} \right) \right)$

$\Theta^{(l)} \leftarrow \Theta^{[t+1]}$, $t \leftarrow t + 1$

end while

$\Theta^{(l+1)} \leftarrow \Theta^{(l)}$

$(e_{i,j,1}^{(l+1)}, e_{i,j,2}^{(l+1)})^T \leftarrow \sum_{k=1}^3 \beta_{i,j,k}^{(l+1)} \left(\Theta_{x_i, z_j}^{w_k(l+1)} \right)^{-T} \hat{P}_{x_i, z_j}^{w_k} (1, 0)^T$ for $i = 0, 1$ $j = 0, 1$

$\lambda_{i,j,k}^{+(l+1)} \leftarrow \lambda_{i,j,k}^{+(l)} + \mu^{(l)} \max \left\{ e_{i,j,k}^{(l+1)} - 1 + \frac{\lambda_{i,j,k}^{+(l)}}{\mu^{(l)}}, 0 \right\}$ for $i = 0, 1, j = 0, 1$ $k = 1, 2$

$\lambda_{i,j,k}^{-(l+1)} \leftarrow \lambda_{i,j,k}^{-(l)} + \mu^{(l)} \max \left\{ -e_{i,j,k}^{(l+1)} + \frac{\lambda_{i,j,k}^{-(l)}}{\mu^{(l)}}, 0 \right\}$ for $i = 0, 1, j = 0, 1$ $k = 1, 2$

end for

$(\hat{p}(v_1|x_0, z_0), \hat{p}(v_2|x_0, z_0))^T \leftarrow \sum_{k=1}^3 \beta_{0,0,k}^{(L)} \left(\Theta_{x_0, z_0}^{w_k(L)} \right)^{-T} \hat{P}_{x_0, z_0}^{w_k} (1, 0)^T$

$(\hat{p}(v_3|x_0, z_0), \hat{p}(v_4|x_0, z_0))^T \leftarrow (0, 0)^T$

$(\hat{p}(v_3|x_1, z_0), \hat{p}(v_4|x_1, z_0))^T \leftarrow \sum_{k=1}^3 \beta_{1,0,k}^{(L)} \left(\Theta_{x_1, z_0}^{w_k(L)} \right)^{-T} \hat{P}_{x_1, z_0}^{w_k} (1, 0)^T$

$(\hat{p}(v_1|x_1, z_0), \hat{p}(v_2|x_1, z_0))^T \leftarrow (0, 0)^T$

$(\hat{p}(v_1|x_0, z_1), \hat{p}(v_3|x_0, z_1))^T \leftarrow \sum_{k=1}^3 \beta_{0,1,k}^{(L)} \left(\Theta_{x_0, z_1}^{w_k(L)} \right)^{-T} \hat{P}_{x_0, z_1}^{w_k} (1, 0)^T$

$(\hat{p}(v_2|x_0, z_1), \hat{p}(v_4|x_0, z_1))^T \leftarrow (0, 0)^T$

$(\hat{p}(v_2|x_1, z_1), \hat{p}(v_4|x_1, z_1))^T \leftarrow \sum_{k=1}^3 \beta_{1,1,k}^{(L)} \left(\Theta_{x_1, z_1}^{w_k(L)} \right)^{-T} \hat{P}_{x_1, z_1}^{w_k} (1, 0)^T$

$(\hat{p}(v_1|x_1, z_1), \hat{p}(v_3|x_1, z_1))^T \leftarrow (0, 0)^T$

$\hat{p}(v_k) = \sum_{i=0}^1 \sum_{j=0}^1 \hat{p}(v_k|x_i, z_j) \hat{p}(x_i, z_j)$ for $k = 1, 2, 3, 4$

3.4 数値実験

この節では、有界制約付き拡張ラグランジュ法に基づく潜在反応タイプの確率の推定量の統計的な性質を調べるために数値実験を行う。

数値実験の設定. 変数 X, Y, Z, W, U, V は離散変数とし、提案法に基づく潜在反応タイプの確率

$$p(u_2), p(u_2, v_2), p(v_2) \quad (3.54)$$

の推定問題を考える。ただし、 $p(u_2)$ と $p(u_2, v_2)$ の推定に関しては図 3.2(a) の構造的因果モデルに、 $p(v_2)$ の推定に関しては図 3.2(b) の構造的因果モデルに従うものとする。本節のパラメータ設定は表 3.1 に示したものに加えて、操作変数は

$$p(z_1) = 0.5 \quad (3.55)$$

のベルヌーイ分布に従うものとする。ここで、図 3.2(a) においては、任意の $u \in \{u_1, u_2, u_3, u_4\}$, $v \in \{v_1, v_2, v_3, v_4\}$, $w \in \{w_1, w_2, w_3, w_4\}$ に対して、 $p(v|u, w) = p(v|u)$ が成り立つ。

また、確率 $p(v_2)$ を推定するためには、4つの値 w_1, w_2, w_3, w_4 から3つの値のみを使用することになっているが、今回は、4つの組み合わせ (${}_4C_3$) 全てにアルゴリズム 2 を適用し、その推定値の平均を $\hat{p}(v_2)$ とした。数値実験は、サンプルサイズ $N = 100, 200, 1000, 5000$ に対して、それぞれ1000回ずつ繰り返し行った。このとき、 $p(u_2, v_2)$ の真値は $5/32=0.156$ であり、 $p(u_2)$ と $p(v_2)$ の真値はともに $1/4=0.250$ である。

数値実験の結果. 次に、データセットに対して提案法を適用することで得られる推定量

$$\hat{p}(u_2), \hat{p}(v_2), \hat{p}(u_2, v_2) \quad (3.56)$$

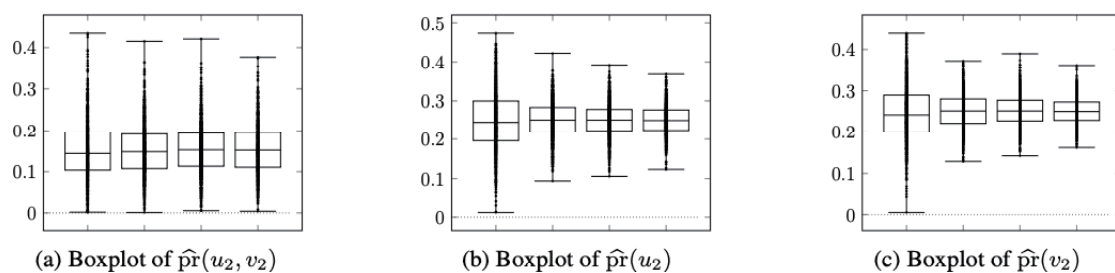


図 3.3: 潜在反応タイプの確率の推定量の箱ひげ図. サンプルサイズはそれぞれ左から $N = 100, 500, 1000, 5000$ である.

の統計的な性質を示す. 表 3.2 と図 3.3 はそれぞれ, 与えられたサンプルサイズ N のに対する $\hat{p}(u_2, v_2)$, $\hat{p}(u_2)$, $\hat{p}(v_2)$ の基本統計値と箱ひげ図である. 図 3.3 において, 横の実線はそれぞれ $\hat{p}(u_2, v_2)$, $\hat{p}(u_2)$, $\hat{p}(v_2)$ の最小値/第 1 分位数/中央値/第 3 分位数/最大値を表している.

表 3.2 より, 推定量 $\hat{p}(u_2, v_2)$, $\hat{p}(u_2)$, $\hat{p}(v_2)$ の平均は真値に近いことがわかる. また, 標準偏差はサンプルサイズが大きくなるにつれて小さくなっている. したがって, 提案法は $\hat{p}(u_2, v_2)$, $\hat{p}(u_2)$, $\hat{p}(v_2)$ の一致推定量となっていると考えられる. さらに, 図 3.3 より, $\hat{p}(u_2)$ と $\hat{p}(v_2)$ の四分位範囲が狭いことがわかり, サンプルサイズが大きくても真の値を含んでいる. 一方, $\hat{p}(u_2, v_2)$ の四分位範囲は, 真の値を含むがサンプルサイズが大きくなってもそれほど狭くならないが, これは定理 3.2.3 では潜在反応タイプの確率を推定するために, 小さな行列式を持つ確率で構成される行列の逆行列を計算する必要があることが理由かもしれない.

最後に, $\hat{p}(u_2)$ と $\hat{p}(v_2)$ は対称的に分布するのではっきりと漸近正規性が成り立つよう見え, 実際, 表 3.2 の $\hat{p}(u_2)$ と $\hat{p}(v_2)$ の推定値の分布の歪度と尖度は, サンプルサイズが大きくなるとともに正規分布の歪度と尖度に近くなっている. 一方で, $\hat{p}(u_2, v_2)$ は成り立たないようにみえるが, これは, 定理 3.2.3 は確率の逆行列の計算を, 定理 3.2.1 - 3.2.2 よりも必要とし, サンプルサイズが小さい場合に不安定になるためと考えられる.

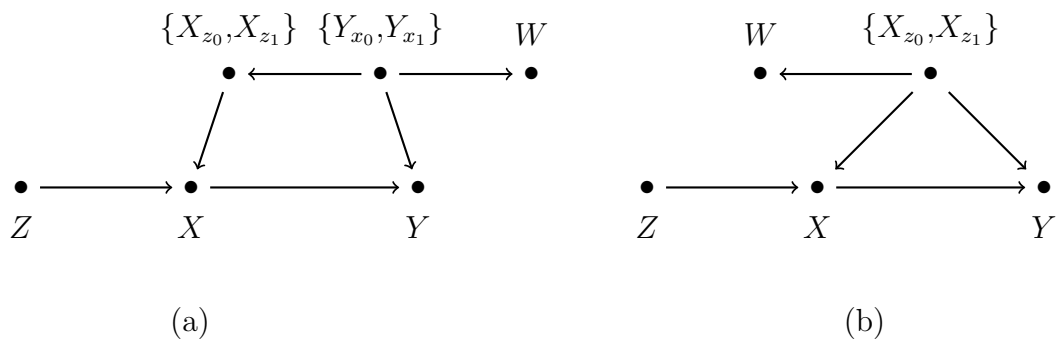


図 3.4: 数値実験の設定のグラフ表現

表 3.1: 数値実験のパラメータ設定

(a) $p(u, v)$					(b) $p(w u)$ と $p(w v)$				
	u_1	u_2	u_3	u_4		u_1	u_2	u_3	u_4
v_1	5/32	1/32	1/32	1/32		v_1	v_2	v_3	v_4
v_2	1/32	5/32	1/32	1/32	w_1	7/10	1/10	1/10	1/10
v_3	1/32	1/32	5/32	1/32	w_2	1/10	7/10	1/10	1/10
v_4	1/32	1/32	1/32	5/32	w_3	1/10	1/10	7/10	1/10
					w_4	1/10	1/10	1/10	7/10

3.5 実データへの応用

つぎに、実データへの適用例を紹介する。取り上げる事例は、賃金分配の公正さの認識が仕事の成果に与える因果効果を評価することを目的とした研究である。Sajons [2020] は賃金分配の公正さの認識と仕事の成果に関する因果効果を検証するために次のような実験室実験を設計、実行し、実験データを収集した。彼女はスペインの大学の異なる分野の学部生 200 人を集め、コンピュータを用いた実験をおこなった。

実験において、参加者は2つの期間（第1期、第2期）に与えられた業務をおこなった。具体的には、段落に含まれる特定の文字（r, s, n, i, o）を数え5分以内にできるだけ多くの段落について正しく数えるという課題をこなした。第1期の業

表 3.2: 数値実験による潜在反応タイプの確率の推定値の基本統計量

真値	$p(u_2, v_2) = 0.156$				$p(u_2) = 0.250$				$p(v_2) = 0.250$				
	N	100	500	1000	5000	100	500	1000	5000	100	500	1000	5000
最小値		0.002	0.001	0.005	0.004	0.012	0.094	0.106	0.123	0.005	0.129	0.142	0.162
第1分位数		0.104	0.107	0.113	0.111	0.196	0.219	0.224	0.225	0.199	0.222	0.228	0.230
中央値		0.143	0.147	0.152	0.151	0.245	0.251	0.251	0.250	0.242	0.252	0.252	0.251
第3分位数		0.195	0.190	0.193	0.194	0.300	0.283	0.278	0.277	0.290	0.280	0.277	0.273
最大値		0.435	0.415	0.421	0.377	0.474	0.422	0.392	0.368	0.439	0.372	0.390	0.361
平均		0.150	0.152	0.154	0.156	0.248	0.250	0.250	0.251	0.247	0.251	0.253	0.251
標準偏差		0.072	0.061	0.061	0.059	0.077	0.048	0.044	0.038	0.069	0.042	0.037	0.032
歪度		0.560	0.481	0.364	0.370	0.110	-0.124	-0.067	0.016	0.180	-0.069	0.087	0.019
尖度		3.590	3.299	3.341	3.114	2.897	2.950	3.115	3.045	3.033	2.869	3.092	2.981

務に対しては参加者全員に €6 の賃金が支払われた。次に、第2期の業務に対しては参加者の半数が €3 まで賃金を引き下げられた。このときで、各参加者はランダムに公平な上司と不公平な上司に割り当てられ、それぞれの上司によって賃金が引き下げられる人が決定した。本事例の詳細については Sajons [2020] を参照。本項で使用したデータセットはつぎのサイトで公開されている (<https://data.mendeley.com/datasets/rk9xb5c2y9/2>)。)

使用する変数. つぎに、データセットの中で使用する変数を紹介する。また、使用する変数は以下のように離散化している。

X : 処置変数「賃金分配の公正さに関する認識」,

1 (完全に不公平) から 7 (完全に公平) までのリッカート尺度 [Likert, 1932]

x_1 : “公平” (点数: $x \in (4, 7]$)

x_0 : “不公平” (点数: $x \in [1, 4]$)

Y : アウトカム「第2期における仕事の成果」

y_1 : “高い” (点数:16-30)

y_0 : “低い” (点数: 0-15)

Z : 操作変数「参加者がどの実験的処置に無作為に割り当てられたかを示すダミー変数」

z_1 : “公平な上司”

z_0 : “不公平な上司”

W_1 : 一つ目の代替変数「第1期の仕事の成果」

w_{11} : “低位” (点数: $w_1 \in [1, 10.75]$)

w_{12} : “中低位” (点数: $w_1 \in (10.75, 13.00]$)

w_{13} : “中高位” (点数: $w_1 \in (13.00, 16.00]$)

w_{14} : “高位” (点数: $w_1 \in (16.00, 25.00]$)

W_2 : 二つ目の代替変数「第1期における参加者の相対的な仕事の成果に対する自己評価」

w_{21} : “低位” (点数: $w_2 \in [0, 60]$)

w_{22} : “中位” (点数: $w_2 \in (60, 80]$)

w_{23} : “高位” (点数: $w_2 \in (80, 100]$)

変数 $\{X, Y, Z, W_1\}$ については構造的因果モデルが図 3.1(a) に、変数 $\{X, Y, Z, W_2\}$ については構造的因果モデルが図 3.1(b) に従うものと仮定する。

結果. 表 3.3 は、各潜在反応タイプの確率

“doomed” $\text{pr}(u_1)$, “causative” $\text{pr}(u_2)$,

“preventive” $\text{pr}(u_3)$, “immune” $\text{pr}(u_4)$,

“never take” $\text{pr}(v_1)$, “compliance” $\text{pr}(v_2)$,
“defiance” $\text{pr}(v_3)$, “always take” $\text{pr}(v_4)$

の 1000 回のブートストラップ法 [Tibshirani and Efron, 1993] による推定値と 95% 信頼区間を示している。95%信頼区間は、2.5%から 97.5%の区間を表している。表 3.3 によれば、潜在反応タイプ u_3 や v_3 の個体が存在していることがわかる。そのため、因果効果を評価するために先行研究のように単調性 II ($\text{pr}(u_3) = 0$) や単調性 I ($\text{pr}(v_3) = 0$) を仮定するのは妥当ではないことがわかる。また、“doomed”や“immune”の個体が多いため、賃金の分配が自分にとって公平か不公平かにかどうかにかかわらず、多くの参加者の仕事の成果は変化しないことが示唆される。

最後に、図 3.5 に箱ひげ図を示した $\hat{p}(u_2, v_2)$, $\hat{p}(u_2)$, $\hat{p}(v_2)$ のブートストラップ法による推定量のの基本統計量を表 3.3 に示した。表 3.3 の $\hat{p}(u_2)$ について、標準偏差は小さく尖度は 3.000 に近いことがわかる。 $\hat{p}(u_2)$ の分布はやや歪んでいるように見えるが、 $\hat{p}(v_2)$ は比較的小さなサンプルサイズ $N = 200$ でも正規分布に従うと判断できる。一方、 $\hat{p}(u_2, v_2)$ について、平均が 0.000 に近いが、推定値の外れ値が多いため、 $\hat{p}(u_2, v_2)$ の最小値は 0.000 以下になり標準偏差と尖度は $\hat{p}(u_2)$ よりもかなり大きくなっている。

表 3.3: ブートストラップ法による潜在反応タイプの確率の推定値および 95%信頼区間

(a) $\{Y_{x_0}, Y_{x_1}\}$

潜在反応タイプ	推定値	95%信頼区間
causative	0.169	[0.016,0.347]
preventive	0.181	[0.018,0.376]
doomed	0.376	[0.182,0.563]
immune	0.275	[0.049,0.512]

(b) $\{X_{z_0}, X_{z_1}\}$

潜在反応タイプ	推定値	95%信頼区間
compliance	0.318	[0.036,0.651]
defiance	0.191	[0.017,0.449]
never-take	0.232	[0.027,0.496]
always-take	0.259	[0.017,0.554]

表 3.4: ブートストラップ法による潜在反応タイプの確率の推定量の基本統計量

潜在反応タイプの確率の推定値	$\hat{p}(u_2, v_2)$	$\hat{p}(u_2)$	$\hat{p}(v_2)$
最小値	-1.488	0.001	0.001
第 1 分位数	-0.059	0.095	0.195
中央値	0.018	0.165	0.339
第 3 分位数	0.189	0.235	0.424
最大値	1.060	0.497	0.838
平均	0.045	0.169	0.318
標準偏差	0.261	0.086	0.162
歪度	-0.256	0.271	0.236
尖度	5.934	2.514	2.613

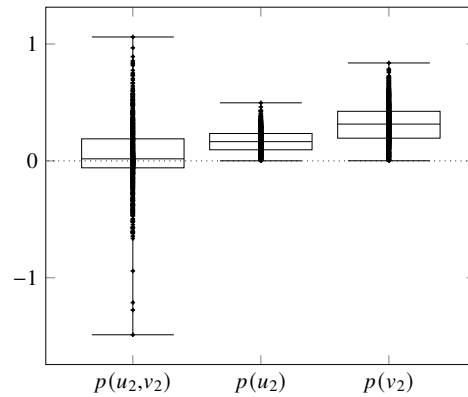


図 3.5: ブートストラップ法による各潜在反応タイプの確率の推定値の箱ひげ図

3.6 考察

本論文では，被験者の不遵守が存在するランダム化試験において，個体の共変量情報を利用した潜在反応タイプの確率の新しい識別可能性に関する定理を提案した．提案した定理によって潜在反応タイプの確率が識別可能である場合，モーメント法に基づく潜在反応タイプの確率の一致推定量を得ることができる．しかしながら，モーメント法に基づいた推定量は効率的でないため，有界制約付き拡張ラグランジュ法を用いたより効率的な推定法を提案した．拡張ラグランジュ法を用いた推定法は [Shingaki and Kuroki, 2021] でも議論されているものである．

この有界制約付き拡張ラグランジュ法を用いた推定法は非凸最適化問題に帰着するため，局所解に陥ってしまう可能性がありパラメータの初期値によって大きく真値と違う推定値を返すことがある．より安定した推定手法の開発は今後の課題とする．この論文で紹介した定理は2値以上の離散な処置変数への拡張が容易であり，多様な状況に適用可能である．ただし，第2部と同様に，スパース性が問題となってくる．また，[Balke and Pearl, 1994] のように連続変数や2値以上の離散変数のアウトカムに関しても，ある閾値 y に基づく変数 $Y \leq y$ を考えることで，本論文の結果を適用することが可能である．

第 4 部

平均偏因果効果の推定法の開発

連続変量の処置変数に対するノンパラメトリック操作変数法 [Newey and Powell, 2003, Ai and Chen, 2003, Blundell et al., 2007] は、非線形な因果効果を解明するために広く使われている手法である。本研究では連続変量である処置 X に対して、潜在反応変数の一次導関数の期待値で定義される平均偏因果効果 [Chamberlain, 1984, Wooldridge, 2005, Graham and Powell, 2012] を取りあげる。ノンパラメトリック操作変数モデルにおける平均偏因果効果の識別可能性に関する定理は、Wong [2022] によってある積分方程式を通して与えられているが、収集されたデータから平均偏因果効果を推定する手法に関してはこれまで議論されてこなかった。この論文では、Wong [2022] の定理に基づく新しい平均偏因果効果の推定法を 2 種類開発する。

1 つ目は、積分方程式、特に、第一種のフレドホルム積分方程式を解くための数学における一般的な解法であるピカールの反復法 [Picard, 1910, Diaz and Metcalf, 1970] を用いて平均偏因果効果の推定を行う。ピカールの反復法を用いた統計手法は、これまで考えられていないことであり、本論文ではピカールの反復法の統計的な性質についても議論する。2 つ目は、線形基底関数モデル [Bishop, 2006] を用いて平均偏因果効果をパラメータ化することで平均偏因果効果の推定を行う。線形基底関数モデルを用いた推定法は、操作変数法における標準的なアプローチである。

最後に、提案した手法群の統計的性質を紹介し、数値実験と経済学で大きな関心を集めている教育年数の賃金への因果効果 [Wooldridge, 2010, Griliches, 1977, Black-

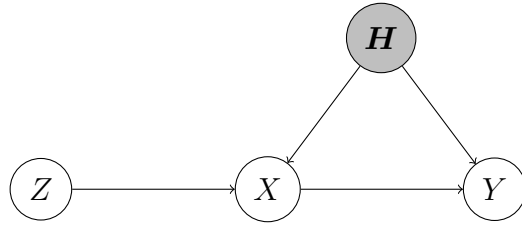


図 4.1: 操作変数モデルの因果グラフ

burn and Neumark, 1992, Card, 1999, Angrist and Krueger, 1991] に関する実データへの適用によってこれらの手法を説明する. 第 4 部の内容は [Kawakami, Kuroki, and Tian, 2023a] で公表されている.

4.1 準備

はじめに, 提案手法の説明のための準備を行う. 第 4 部では, 各変数を大文字の (X) , 実現値を小文字の (x) で表現する. また, $\mathbb{1}_\Omega(x)$ を指示関数とする. 指示関数は $x \in \Omega$ が成り立つとき 1 で, $x \notin \Omega$ が成り立つとき 0 をとる関数である. Ω_X を変数 X の定義域, $\mathbb{E}[Y]$ と $\mathbb{V}(Y)$ を変数 Y の期待値と分散, $\mathbb{P}_X[x] = \mathbb{P}(X \leq x)$ を X の累積分布関数であると定義する. また, $X = x$ の条件付き期待値を

$$\mathbb{E}[Y|X = x] \tag{4.1}$$

とし, $Z = z$ の条件付き累積分布関数を

$$\mathbb{P}_X[x|Z = z] = \mathbb{P}(X \leq x|Z = z) \tag{4.2}$$

とする. 任意の x に対して, $|g(x)| \leq Mh(x)$ となるような正の実数 M と実数 δ が存在すれば $g(x) = \mathcal{O}(h(x))$ ($x \rightarrow \infty$) と表記し, 任意の正の実数 M と, 任意の $|x| \leq \delta$ を満たす x に対して, $|g(x)| \leq Mh(x)$ を満たすような正の実数 δ が存在するとき, $g(x) = \mathcal{O}(h(x))$ ($|x| \rightarrow 0$) と書く.

関数解析. 本論文で使用する関数解析の基本的な表記を紹介する [Muscat, 2014]. ヒルベルト空間を \mathcal{H} とする. ただし, 任意の $a, b \in \mathcal{H}$ に対して内積は

$$\langle a, b \rangle = \int_{\Omega_x} a(x)b(x)dx \quad (4.3)$$

であり, ノルムは $\|a\| = \langle a, a \rangle^{\frac{1}{2}}$ で定義されるものとする. ある関数列 $\{a_n\}$ があり, $n \rightarrow \infty$ に対して $\|a_n - a\| \rightarrow 0$ を満たすとき関数列 a_n は関数 a に強収束するといふ. また, 作用素 \mathcal{L} を

$$(\mathcal{L}(a))(x) = \int_{\Omega_x} L(x', x)a(x')dx' \quad (4.4)$$

で定義される作用素であるとし, この作用素 \mathcal{L} のノルムを

$$\|\mathcal{L}\| = \sup\{\|\mathcal{L}(a)\| : \|a\| = 1, a \in \mathcal{H}\} \quad (4.5)$$

と定義する. \mathcal{L} が正の固有値の可算集合を持つときそれらを大きい順に

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq 0 \quad (4.6)$$

とし, $\mathcal{L}(v_i) = \lambda_i v_i$ のように対応する \mathcal{H} の固有ベクトル v_1, v_2, v_3, \dots と表す. ここで, 固有ベクトル $\{v_1, v_2, v_3, \dots\}$ は正規直交基底である.

平均偏因果効果. 介入 “do(x)” によって得られる Y に関する潜在反応変数を $Y_x(\mathbf{u})$ と表す. 本研究では連続変数である処置 X に対して平均偏因果効果 (Average Partial Causal Effect: APCE) を次のような x 上の実数値関数で定義する [Chamberlain, 1984, Wooldridge, 2005, Graham and Powell, 2012].

$$\mathbb{E}[\partial_x Y_x] := \mathbb{E}_{\mathcal{U}} \left[\frac{\partial}{\partial x} Y_x(\mathbf{U}) \right] \quad (4.7)$$

この論文では平均偏因果効果を推定することを目標とする. これは2値処置の平均因果効果 $\mathbb{E}[Y_1] - \mathbb{E}[Y_0]$ の自然な一般化と考えられる. 潜在反応変数の期待値 $\mathbb{E}[Y_x]$ が微分可能であるならば, 2つの処置 x'' と x' に対する平均因果効果は,

$$E[Y_{x''}] - E[Y_{x'}] = \int_{x'}^{x''} E[\partial_x Y_x] dx \quad (4.8)$$

で計算できる.

操作変数モデル. 平均偏因果効果の推定のために, 図 4.1 の因果グラフで表されるノンパラメトリック操作変数モデルを考える. これは以下の構造的因果モデル \mathcal{M}_{IV} で定義される.

$$\begin{cases} Y := f_Y(X, \mathbf{H}, \epsilon_Y) \\ X := f_X(Z, \mathbf{H}, \epsilon_X) \\ Z := f_Z(\epsilon_Z) \end{cases} \quad (4.9)$$

ただし, 第二部, 第三部とは異なり $\{X, Y, Z\}$ は連続変数変数とする. また, $\{\epsilon_X, \epsilon_Y, \epsilon_Z\}$ は外生ノイズである. ここで, 内生変数は $U = \{\mathbf{H}, \epsilon_X, \epsilon_Y, \epsilon_Z\}$ であり, 外生変数は $V = \{X, Y, Z\}$ である.

平均偏因果効果の識別可能性. つぎに, Wong [2022] によって与えられている分布 $\mathbb{P}(Y|Z)$ と $\mathbb{P}(X|Z)$ をもちいた平均偏因果効果の識別可能性に関する命題を紹介する. まずはじめに, 以下のような仮定をおく.

仮定 4.1.1. 構造的因果モデル \mathcal{M}_{IV} のもとで, 任意の $\mathbf{h} \in \Omega_H$ に対して

1. 潜在反応変数 Y_x は定義域 Ω_X において微分可能かつ有界である.
2. $\sup_{x,z} p(X_z = x) < \infty$ が成り立つ. ただし, p は密度関数である.
3. z を変化させて得られる分布の集合 $\mathbb{P}(X|Z = z)$ は完備である.

これらの仮定は, モデルを設定するために必要なものでありそれほど強い制限ではないと考えられている. 1つ目の仮定は, 任意の個体に対して平均偏因果効果が存在することを意味する. つぎに, 2つ目の仮定は, 潜在反応変数 X_z の密度関数がある有界であることを意味する. さいごに, 3つ目の仮定は $\mathbb{E}[h(X)|Z = z]$ が z に関してゼロ関数であるときに $h(X)$ はゼロ関数に限られることを意味している [Newey and

Powell, 2003]. これらの仮定に加えて、分離可能性の仮定をおく [Newey and Powell, 2003, Wong, 2022].

仮定 4.1.2 (分離可能性). 関数 $f_Y(X, \mathbf{H}, \epsilon_Y)$ が分離可能である. つまり,

$$f_Y(X, \mathbf{H}, \epsilon_Y) = f_1(X, \epsilon_Y) + f_2(\mathbf{H}, \epsilon_Y) \quad (4.10)$$

のように表現できる.

このとき、平均偏因果効果の識別可能性に関する以下の命題が成立する [Wong, 2022].

命題 4.1.1 (平均偏因果効果の識別可能性). 仮定 4.1.1, 4.1.2のもとで、平均偏因果効果 $\mathbb{E}[\partial_x Y_x]$ は以下の積分方程式を通して識別可能である.

$$\mu(z) = \int_{\Omega_X} k(x, z) \mathbb{E}[\partial_x Y_x] dx \quad (4.11)$$

ここで,

$$\begin{aligned} \mu(z) &= \mathbb{E}[Y|Z = z_0] - \mathbb{E}[Y|Z = z], \\ k(x, z) &= \mathbb{P}_X[x|Z = z] - \mathbb{P}_X[x|Z = z_0] \end{aligned} \quad (4.12)$$

であり、ここで z_0 は任意の定数である.

本論文では、積分方程式 (4.11) をもちいて観測したデータから平均偏因果効果を推定することを目的としている. 定数 z_0 の選択は識別可能性においては、結果に影響しないが、平均偏因果効果の推定問題においては推定量の分散などの側面において結果に影響してくると考えられる. ここで、積分核は条件付き累積分布関数の差で構成されていて、これは $-1 \sim 1$ の値を取る有界な関数であり、条件付き密度関数に比べて推定が容易なものである.

4.2 ノンパラメトリック推定法 (N-APCE 推定量)

本節では、積分方程式の数学的に標準的な解法であるピカールの反復法に基づいて平均偏因果効果のノンパラメトリック推定法を開発する. まずはじめに、つぎの仮定をおく.

仮定 4.2.1 (定義域の有界性). Ω_X と Ω_Z はどちらも有界である.

定義域の有界性は機械学習の分野において一般的な仮定である. つぎに, この仮定のもとで積分方程式 (4.11) を解く問題が良設定問題であることを示す. (良設定問題の定義は付録を参照.)

命題 4.2.1 (良設定問題). 仮定 4.1.1, 4.1.2, 4.2.1 のもとで, 積分方程式 (4.11) を解くことは良設定問題である. ここで, 良設定問題とは一意に解が存在し入力の変化に対して解が連続的に変化することをいう [Tikhonov et al., 1995].

既存のノンパラメトリック操作変数法における積分方程式は, 不良設定問題であったために, ピカールの反復法を利用した統計的な手法はこれまで考えられてこなかった. 一方で, 積分方程式 (4.11) は良設定問題であるためピカールの反復法を利用した統計的な手法を考えることが可能となる. 本研究は, 私の知る限り, ピカールの反復法を利用したはじめての統計的な手法である.

4.2.1 N-APCE 推定量のアルゴリズム

積分方程式 (4.11) は第一種のフレドホルム積分方程式と呼ばれるものであり [Bôcher, 1926], 関数 k に対応する部分を積分核と呼ぶ. ここで, 第一種のフレドホルム積分方程式とは任意の $a, b \in \mathcal{H}$ に対して,

$$b = \mathcal{L}(a) \tag{4.13}$$

という形をとる積分方程式である. この形の積分方程式に対して, 「ピカールの条件」という解の存在性に関する以下のような必要十分条件が Picard [1910] によって示されている.

ピカールの条件. 作用素 \mathcal{L} と関数 $b \in \mathcal{H}$ に対し, $\mathcal{L}(a) = b$ となる関数 a が存在し, かつ, その関数 a が $\mathcal{L}(a) = b$ のときのみ

$$\frac{\langle a, v_i \rangle^2}{\lambda_i^2} < \infty \tag{4.14}$$

を満たす。ここで、 v_i と λ_i は \mathcal{L} の i 番目の固有値と固有ベクトルである。

ピカールの反復法. つぎに、ピカールの反復法を説明する。最初に、次のような仮定をおく。

仮定 4.2.2 (定義域の包含関係). $\Omega_X \subseteq \Omega_Z$.

この仮定は Z の定義域が X の定義域を含むことを意味している。作用素 \mathcal{K} を次のように表す。

$$(\mathcal{K}(a))(x) = \int_{\Omega_X} k(x', x)a(x')dx', \quad \forall a \in \mathcal{H} \quad (4.15)$$

すると、積分方程式 (4.11) を解くためのピカールの反復法は、以下のように表現できる。

$$\theta_{t+1}(x) \leftarrow \theta_t(x) + \alpha \left(\mu(x) - \int_{\Omega_X} k(x', x)\theta_t(x')dx' \right) \quad (4.16)$$

ここで、ステップサイズ α は

$$0 < \alpha < \frac{2}{\|\mathcal{K}\|} \quad (4.17)$$

を満たす正の数である。つぎに、Diaz and Metcalf [1970] に関連する以下の補題が成立する。

補題 4.2.1. 仮定 4.1.1, 4.1.2, 4.2.1 のもとで、作用素 \mathcal{K} は以下の 3 つの性質を満たす。

1. 作用素 \mathcal{K} はコンパクトな作用素であり、有界集合を強収束の意味でコンパクトな集合に写す。
2. 作用素 \mathcal{K} は自己共役である。任意の $a, b \in \mathcal{H}$ に対して $\langle \mathcal{K}(a), b \rangle = \langle a, \mathcal{K}(b) \rangle$ が成立する。
3. 作用素 \mathcal{K} は半正定値である。任意の $a \in \mathcal{H}$ に対して $\langle \mathcal{K}(a), a \rangle \geq 0$ が成立する。

補題 4.2.1 から、 \mathcal{K} は正の固有値の可算集合を持ち、くわえて、以下の補題が成り立つ。

補題 4.2.2. 仮定 4.1.1 のもとで \mathcal{K} はピカールの条件を満たしている。

この補題から積分方程式 (4.11) に解が一意に存在する。最後に、以下の収束性に関する定理を得る。

定理 4.2.1 (推定値の収束性). 仮定 4.1.1, 4.1.2, 4.2.1, 4.2.2 のもとで、ピカールの反復法 (4.16) は平均偏因果効果に強収束し、

$$\lim_{t \rightarrow \infty} \theta_t(x) = \mathbb{E}[\partial_x Y_x] \quad (4.18)$$

がほとんど至る所で成立する。

この定理から、データが無限に存在するときにおいては、平均偏因果効果がピカールの反復法で求められることがわかった。

ノンパラメトリック平均偏因果効果 (N-APCE) 推定量. 次に、アルゴリズム 3 に示しているピカールの反復法 (4.16) に基づくノンパラメトリック平均偏因果効果推定法 (N-APCE 推定量) を説明する。N-APCE 推定量のアルゴリズムは、先行研究と同様に 2 段階推定法となっている。最初に、観測値の集合

$$\mathcal{D} = \{z^{(i)}, x^{(i)}, y^{(i)}\}_{i=1}^N \quad (4.22)$$

が利用可能であるとする。または、Angrist and Krueger [1995] や Inoue and Solon [2005] のように $\mathbb{P}(X, Z)$ と $\mathbb{P}(Y, Z)$ から別々にサンプリングしたデータセット

$$\mathcal{D}^{(1)} = \{z^{(i)}, x^{(i)}\}_{i=1}^{N_1}, \quad \mathcal{D}^{(2)} = \{z^{(i)}, y^{(i)}\}_{i=1}^{N_2} \quad (4.23)$$

のみでも問題ない。

まず、あらかじめアルゴリズムの停止閾値 ϵ 、ステップサイズ α 、初期関数 $\theta_1(x)$ を設定しておく。次に、第 1 段階として、 $Z = z$ が与えられたときの Y の期待値の

アルゴリズム 3 ノンパラメトリック平均偏因果効果 (N-APCE) 推定量

- 1: **入力:** データセット $\mathcal{D} = \{z^{(i)}, x^{(i)}, y^{(i)}\}_{i=1}^N$, 停止条件 ϵ , ステップサイズ α , X の値の集合 $\mathcal{X} = (x_0, x_1, \dots, x_R)$.
- 2: 二つの予測モデル $\hat{\mathbb{E}}[Y|Z = z]$ と $\hat{\mathbb{P}}_X[x|Z = z]$ を観測値 \mathcal{D} から任意の機械学習を利用し構築する.
- 3: $\theta_1(x)$ と $t \leftarrow 1$ を初期化する.
- 4: $R + R^2$ 個の値

$$\begin{aligned}\hat{\mu}(x_r) &= \hat{\mathbb{E}}[Y|Z = x_0] - \hat{\mathbb{E}}[Y|Z = x_r] \\ \hat{k}(x_q, x_r) &= \hat{\mathbb{P}}_X[x_q|Z = x_r] - \hat{\mathbb{P}}_X[x_q|Z = x_0]\end{aligned}\tag{4.19}$$

を計算する ($q, r = 1, \dots, R$).

- 5: 関数 $\theta_{t+1}(x_r)$ を

$$\{\mathcal{I}[(\hat{\mu}(x) - \mathcal{I}[\hat{k}(x', x)\hat{\theta}_t(x'); \mathcal{X}])^2; \mathcal{X}]\}^{1/2} < \epsilon\tag{4.20}$$

を満たすまで,

$$\hat{\theta}_{t+1}(x_r) \leftarrow \hat{\theta}_t(x_r) + \alpha \left(\hat{\mu}(x_r) - \mathcal{I}[\hat{k}(x', x_r)\hat{\theta}_t(x'); \mathcal{X}] \right)\tag{4.21}$$

で更新する ($r = 1, \dots, R$).

- 6: **出力:** 最後の更新でえられた関数 $\hat{\theta}_T(x_r)$ ($r = 1, \dots, R$) に対して数値補間を行い, N-APCE 推定量とし返す.
-

予測関数を $\hat{\mathbb{E}}[Y|Z = z]$ とし, $Z = z$ が与えられたときの X の累積分布関数の予測関数を $\hat{\mathbb{P}}_X[x|Z = z]$ とする. これらの予測関数は観測値 \mathcal{D} から任意の教師あり機械学習モデルを用いて学習される [Hastie et al., 2009]. 本論文で, この手法について特に指定はしない. ここで,

$$\hat{\mu}(x) = \hat{\mathbb{E}}[Y|Z = x_0] - \hat{\mathbb{E}}[Y|Z = x],\tag{4.24}$$

$$\hat{k}(x', x) = \hat{\mathbb{P}}_X[x'|Z = x] - \hat{\mathbb{P}}_X[x'|Z = x_0]\tag{4.25}$$

と表記し, 任意の $a \in \mathcal{H}$ に対し作用素 $\hat{\mathcal{K}}$ を

$$(\hat{\mathcal{K}}(a))(x) = \int_{\Omega_X} \hat{k}(x', x)a(x')dx \quad (4.26)$$

と表記する. つぎに, 数値積分を利用して積分を近似することを考える [Burden et al., 2015]. 数値積分のための有限個の値

$$\mathcal{X} = \{x_1, \dots, x_R\} \in \Omega_X \quad (4.27)$$

の集合を選ぶ. ここで, 任意の $r = 1, \dots, R-1$ に対して $x_r < x_{r+1}$ である. ただし, R が大きいほど精度はよくなるが, 計算量が多くなってしまふ. 例えば, 等距離な区間の分割

$$x_r = \frac{r\{\max(\Omega_X) - \min(\Omega_X)\}}{R} + \min(\Omega_X) \quad (4.28)$$

を用いることができる. なお, X と Z は \mathcal{X} という値を共有している点に注意. 次に, $\mathcal{I}[a(x); \mathcal{X}]$ は, \mathcal{X} が与えられたときの, 任意の関数 $a \in \mathcal{H}$ に対する積分 $\int_{\Omega_X} a(x)dx$ の数値積分を表しており, 次のような形で与えられる.

$$\mathcal{I}[a(x); \mathcal{X}] = \sum_{q=1}^R I(x_q, x_{q+1})(x_{q+1} - x_q) \quad (4.29)$$

ここで, $I(x_q, x_{q+1})$ は様々な形を取ることができる. 例えば, $I(x_q, x_{q+1}) = a(x_q)$ などである. 他の選択肢については, 付録5を参照.

さいごに, 停止基準として使用するための次のような経験損失関数を導入する.

$$J_N(\theta; \mathcal{D}) = \sqrt{\int_{\Omega_X} \left(\hat{\mu}(x) - \int_{\Omega_X} \hat{k}(x', x)\theta(x')dx' \right)^2 dx} \quad (4.30)$$

ここで, 経験損失関数には2つの積分が含まれており, これは数値積分で近似することとし, 数値的な近似経験損失関数は以下のように計算される.

$$\tilde{J}_N(\theta; \mathcal{D}) = \sqrt{\mathcal{I}[(\hat{\mu}(x) - \mathcal{I}[\hat{k}(x', x)\theta(x'); \mathcal{X}])^2; \mathcal{X}]} \quad (4.31)$$

第1段階でえられた予測モデルを利用して、ピカール反復をもちいて近似経験損失関数を計算するために、以下の $R + R^2$ 個の値 $(q, r = 1, \dots, R)$ を計算する.

$$\begin{aligned}\hat{\mu}(x_r) &= \hat{\mathbb{E}}[Y|Z = x_0] - \hat{\mathbb{E}}[Y|Z = x_r], \\ \hat{k}(x_q, x_r) &= \hat{\mathbb{P}}_X[x_q|Z = x_r] - \hat{\mathbb{P}}_X[x_q|Z = x_0]\end{aligned}\tag{4.32}$$

第2段階として、各 $t = 1, 2, \dots$ において $\hat{\theta}_t(x_r)$ ($r = 1, \dots, R$) を次のように更新する.

$$\hat{\theta}_{t+1}(x_r) \leftarrow \hat{\theta}_t(x_r) + \alpha \left(\hat{\mu}(x_r) - \mathcal{I}[\hat{k}(x', x_r)\hat{\theta}_t(x'); \mathcal{X}] \right)\tag{4.33}$$

ただし、初期関数 $\hat{\theta}_1(x) = \theta_1(x)$ から $\hat{\theta}_t$ が停止条件

$$\tilde{J}_N(\hat{\theta}_t; \mathcal{D}) \leq \epsilon\tag{4.34}$$

を満たすまで更新を繰り返す. ここで、 ϵ は任意の正の数である. 最後に、ピカールの反復法で $\hat{\theta}_T$ に収束した後、 $\hat{\theta}_T(x_r)$ ($r = 1, \dots, R$) を補間することで平均偏因果効果の推定量とする. 数値補間には、たとえば、ラグランジュ補間多項式 [Jeffreys and Jeffreys, 1988] などが使用できる. (詳しくは付録5を参照.)

4.2.2 N-APCE 推定量の性質

数値補間による N-APCE 推定量の誤差については数値解析の分野ではよく研究されている [Burden et al., 2015]. $\hat{\theta}_T(x_r)$ の誤差は数値積分の誤差と機械学習による予測関数の推定の誤差に依存し、前者は数値解析の分野でよく研究されている [Burden et al., 2015]. したがって、ここでは機械学習による予測関数の推定の誤差が N-APCE 推定量に与える影響にのみ注目し議論することとする.

一貫性と計算量. N-APCE 推定量の一貫性とアルゴリズムの計算量を調べる. まず、次のような仮定をおく.

仮定 4.2.3. 機械学習をもちいて学習した推定量 $\hat{\mu}$ と \hat{k} は式 (4.12) における μ と k の一致推定値となっている.

ここで、 \mathcal{X} の値は $[x_0, x_R] = [\min(\Omega_X), \max(\Omega_X)]$ と $\lim_{R \rightarrow \infty} |x_{r+1} - x_r| = 0$ ($r = 1, \dots, R$) を満たすものとする。すると、次のような N-APCE 推定量の一致性に関する定理が得られる。

定理 4.2.2 (一致性). 仮定 4.1.1, 4.1.2, 4.2.1, 4.2.2, 4.2.3 のもとで、極限 $N \rightarrow \infty$, $R \rightarrow \infty$, $t \rightarrow \infty$ を取ると、N-APCE 推定量 $\hat{\theta}_t(x)$ はほとんどすべての $x \in \Omega_X$ において平均偏因果効果 $\mathbb{E}[\partial_x Y_x]$ の一致推定量である。

次に、アルゴリズム 3 の有限回反復での停止可能性を示す。まず、以下のように仮定する。

仮定 4.2.4. $\hat{\mathcal{K}}$ はピカールの条件を満たしている。

すると、次のアルゴリズムの停止性に関する定理が成立する。

定理 4.2.3 (停止性). 仮定 4.1.1, 4.1.2, 4.2.1, 4.2.2, 4.2.4 のもとで、 $R \rightarrow \infty$ の極限をとることで、アルゴリズム 3 は任意の正の数 ϵ に対して、有限回の繰り返して停止する。

ここで、 $\lim_{t \rightarrow \infty} \hat{\theta}_t(x) = \hat{\theta}_\infty(x)$ と表記すれば、次の系が成立する。

系 4.2.1 (収束レート). 仮定 4.1.1, 4.1.2, 4.2.1, 4.2.2, 4.2.4 のもとで、関数列 $\hat{\theta}_t$ は $\hat{\theta}_\infty$ に線形的に収束する。

予測モデルの構築や $R + R^2$ の予測値の算出の複雑さは使用する機械学習に依存する。アルゴリズム 3 が T 回の繰り返して停止すると仮定すると、数値積分は合計 $\mathcal{O}(T \times R^2)$ 時間かかり最後の補間は $\mathcal{O}(N \times R^2)$ 時間がかかる。よって、長い繰り返し回数 T は計算に重大な問題を引き起こしてしまう。

バイアスと分散. つぎに、N-APCE 推定量のバイアスと分散を調べる。この推定量には $\hat{\mathcal{K}}$ の誤差に起因する減衰バイアスが含まれている [Wooldridge, 2010]。減衰バ

イアスとは、説明変数にノイズが存在することで、推定量が過小評価される傾向になるようなバイアスのことである。最初に、次のような仮定をおく。

仮定 4.2.5. 機械学習をもちいて学習した推定量 $\hat{\mu}$ と \hat{k} は式 (4.12) における μ と k の不偏推定値となっている。

すると、次のような N-APCE 推定値のバイアスに関する定理が得られる。

定理 4.2.4 (推定値のバイアス). 仮定 4.1.1, 4.1.2, 4.2.1, 4.2.2, 4.2.4, 4.2.5 のもとで、逆作用素 $\hat{\mathcal{K}}^{-1} = \alpha \sum_{t=0}^{\infty} (I - \alpha \hat{\mathcal{K}})^t$ のノルム $\|\hat{\mathcal{K}}^{-1}\|$ が M 以下であるとき、バイアスのノルムの期待値は

$$\mathbb{E}[\|\hat{\theta}_{\infty} - \theta_{\infty}\|] \leq M(A + \|\theta_{\infty}\|B) \quad (4.35)$$

を満たす。ここで、

$$A = \sqrt{\int_{\Omega_X} \mathbb{V}(\hat{\mu}(x)) dx}, \quad B = \sqrt{\int_{\Omega_X} \mathbb{V}(\hat{k}(x, x)) dx} \quad (4.36)$$

である。

A と B の条件付き分散関数は [Fan and Yao, 1998] の方法で計算できる。Hable [2012] Shen and Bellec [2020] の機械学習のように条件付き分散関数が $\mathcal{O}(N^{-1})$ に従って減少するのであれば、 $\mathcal{O}(N^{-1/2})$ に従ってバイアスの期待値のノルムは減少する ($N \rightarrow \infty$)。さらに、N-APCE 推定量には数値積分や数値補間による他のバイアスも存在している。最後に、 $X = x$ における N-APCE 推定量の分散を以下の定理を通して評価する。

定理 4.2.5 (推定値の分散). 仮定 4.1.1, 4.1.2, 4.2.1, 4.2.2, 4.2.5 のもと、アルゴリズム 3 が $t = T$ で停止したとき、N-APCE 推定量 $\hat{\theta}_T(x)$ の分散の上界は

$$\alpha^2(T-1)^2\nu(x) + \mathcal{O}(\alpha^3) \quad (4.37)$$

である。ただし、 $\alpha \rightarrow 0$, $\forall x \in \Omega_X$ であり、関数 $\nu(x)$ は

$$\begin{aligned} & \mathbb{V}(\hat{\mu}(x)) + \left(\int_{\Omega_X} \sqrt{\mathbb{V}(\hat{k}(x', x))} |\theta_T(x')| dx' \right)^2 \\ & + 2\sqrt{\mathbb{V}(\hat{\mu}(x))} \left(\int_{\Omega_X} \sqrt{\mathbb{V}(\hat{k}(x', x))} |\theta_T(x')| dx' \right) \end{aligned} \quad (4.38)$$

である.

さらに, 定理 4.2.5 から次の系が成立する.

系 4.2.2 (推定値の分散). 仮定 4.1.1, 4.1.2, 4.2.1, 4.2.2, 4.2.5 のもと, アルゴリズム 3 が $t = T$ で停止したとき, $\hat{\theta}_T(x)$ の分散は任意の $x \in \Omega_X$ に対して

$$\alpha^2(T-1)^2 \mathbb{V}(\hat{\mu}(x)) + \mathcal{O}(\alpha^3) + \mathcal{O}(\{\max_{x'} \{\mathbb{V}(\hat{k}(x', x))\}\}^{1/2}) \quad (4.39)$$

である. ここで,

$$\alpha \rightarrow 0 \text{ かつ } \{\max_{x'} \{\mathbb{V}(\hat{k}(x', x))\}\}^{1/2} \rightarrow 0$$

である.

ここで, 極限 $\{\max_{x'} \{\mathbb{V}(\hat{k}(x', x))\}\}^{1/2} \rightarrow 0$ は機械学習の推定量 $\hat{\mathbb{E}}[Y|Z = z]$ が小さな条件つき分散を持っていることを意味している. この定理や系により, N-APCE 推定量の分散の評価が可能となった.

4.3 パラメトリック推定法 (P-APCE 推定量)

本節では平均偏因果効果を推定するための線形基底関数モデル [Bishop, 2006] を用いたパラメトリック推定法を紹介する. 基底関数には, 多項式などの非線形関数を想定している. 線形基底関数モデルは, カーネル法や深層学習などさまざまな機械学習モデルの基礎となるモデルである.

4.3.1 P-APCE 推定量のアルゴリズム

線形基底関数モデル. 積分方程式 (4.11) を解くために, 平均偏因果効果を線形基底関数モデルでパラメトリック化する.

$$\mathbb{E}[\partial_x Y_x] = \sum_{p=1}^P \theta_p \phi_p(x) \quad (4.40)$$

基底関数の集合を $\{\phi_p(x)\}_{p=1,\dots,P}$ とし [Bishop, 2006], データから推定するモデルパラメータは $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_P\}$ を使用する. 基底関数には非線形な関数を利用することを想定している. すると, 積分方程式 (4.11) は次のようになる.

$$\mu(z) = \sum_{p=1}^P \theta_p \int_{\Omega_X} k(x, z) \phi_p(x) dx \quad (4.41)$$

任意の $p = 1, \dots, P$ に対し, 基底関数の不定積分を

$$\varphi_p(x) = \int \phi_p(x) dx \quad (4.42)$$

とすると, 積分方程式 (4.11) は以下の等式となる.

$$\mu(z) = \sum_{p=1}^P \theta_p \{\mathbb{E}[\varphi_p(X)|Z = z] - \mathbb{E}[\varphi_p(X)|Z = z_0]\} \quad (4.43)$$

次に, 推定問題が連立線形方程式に帰着されることを示す. まず, $r = 1, \dots, R-1$ に対して $z_r < z_{r+1}$ となる値の集合 $\{z_1, \dots, z_R\} \in \Omega_Z$ を選ぶ. 通常, 先行研究と同様に観測データ $z^{(i)}$ ($i = 1, \dots, N$) を利用することを想定している. 任意の $r = 1, \dots, R$ $p = 1, \dots, P$ に対して,

$$\begin{aligned} u_r &= \mathbb{E}[Y|Z = z_r] - \mathbb{E}[Y|Z = z_0] \\ d_r^p &= \mathbb{E}[\varphi_p(X)|Z = z_r] - \mathbb{E}[\varphi_p(X)|Z = z_0] \end{aligned} \quad (4.44)$$

とする. さらに, 任意の $p = 1, \dots, P$ に対して, $\mathbf{d}^p = (d_1^p, \dots, d_R^p)^T$ とし, $\mathbf{u} = (u_1, \dots, u_R)^T$, $\mathbf{D} = (\mathbf{d}^1, \dots, \mathbf{d}^P)$ と表記する. すると, パラメータ $\boldsymbol{\theta}$ は連立線形方程式

$$\mathbf{u} = \mathbf{D}^T \boldsymbol{\theta} \quad (4.45)$$

を解くことで得られる. ただし, \mathbf{D}^T は \mathbf{D} の転置行列を表している.

パラメトリック平均偏因果効果推定量. 次に, パラメトリック平均偏因果効果推定量 (P-APCE 推定量) を紹介する. P-APCE 推定量のアルゴリズムは, N-APCE 推

アルゴリズム 4 パラメトリック平均偏因果効果 (P-APCE) 推定量

1: **入力:** データセット $\mathcal{D} = \{z^{(i)}, x^{(i)}, y^{(i)}\}_{i=1}^N$, 基底関数 $\{\phi_p(x)\}_{p=1, \dots, P}$, 集合 $\{z_1, \dots, z_R\} \in \Omega_Z$.

2: 予測モデル $\hat{\mathbb{E}}[Y|Z = z]$ と $\hat{\mathbb{E}}[\varphi_p(X)|Z = z] (p = 1, \dots, P)$ を観測値 \mathcal{D} から, 任意の機械学習を用いて構築する.

3: $R + R \times P$ 個の値

$$\begin{aligned}\hat{u}_r &= \hat{\mathbb{E}}[Y|Z = z_r] - \hat{\mathbb{E}}[Y|Z = z_0] \\ \hat{d}_r^p &= \hat{\mathbb{E}}[\varphi_p(X)|Z = z_r] - \hat{\mathbb{E}}[\varphi_p(X)|Z = z_0]\end{aligned}\tag{4.46}$$

を計算する ($r = 1, \dots, R; p = 1, \dots, P$).

4: $\hat{\mathbf{u}} = (\hat{u}_1, \dots, \hat{u}_R)^T$, $\hat{\mathbf{d}}^p = (\hat{d}_1^p, \dots, \hat{d}_R^p)^T (p = 1, \dots, P)$, $\hat{\mathbf{D}} = (\hat{\mathbf{d}}^1, \dots, \hat{\mathbf{d}}^P)$ とし, 最適化問題

$$\hat{\boldsymbol{\theta}} = \operatorname{argmin}_{\boldsymbol{\theta}} \|\hat{\mathbf{u}} - \hat{\mathbf{D}}\boldsymbol{\theta}\|^2\tag{4.47}$$

を解く.

5: **出力:** $\sum_{p=1}^P \hat{\theta}_p \phi_p(x)$ を P-APCE 推定量として返す.

定量と同様に, 2段階推定法となっている. P-APCE 推定量のアルゴリズムはアルゴリズム 4 に示している. 操作変数モデルから観測データ

$$\mathcal{D} = \{z^{(i)}, x^{(i)}, y^{(i)}\}_{i=1}^N\tag{4.48}$$

が得られているものとする. または, Angrist and Krueger [1995] や Inoue and Solon [2005] のように $\mathbb{P}(X, Z)$ と $\mathbb{P}(Y, Z)$ から別々にサンプリングしたデータセット

$$\mathcal{D}^{(1)} = \{z^{(i)}, x^{(i)}\}_{i=1}^{N_1}, \quad \mathcal{D}^{(2)} = \{z^{(i)}, y^{(i)}\}_{i=1}^{N_2}\tag{4.49}$$

のみでも問題ない.

第1段階として, 任意の $p = 1, \dots, P$ に対して, $\hat{\mathbb{E}}[Y|Z = z]$ を $Z = z$ が与えられたときの Y の期待値の予測値, $\hat{\mathbb{E}}[\varphi_p(X)|Z = z]$ を $Z = z$ が与えられたときの変

数 $\varphi_p(X)$ の期待値の予測値とする。これらの期待値の予測値は、観測値 \mathcal{D} から機械学習を用いて学習される。次に、以下の $R + R \times P$ の値を計算する。

$$\begin{aligned}\hat{u}_r &= \hat{\mathbb{E}}[Y|Z = z_r] - \hat{\mathbb{E}}[Y|Z = z_0] \\ \hat{d}_r^p &= \hat{\mathbb{E}}[\varphi_p(X)|Z = z_r] - \hat{\mathbb{E}}[\varphi_p(X)|Z = z_0]\end{aligned}\tag{4.50}$$

ただし、 $r = 1, \dots, R$, $p = 1, \dots, P$ である。任意の $p = 1, \dots, P$ に対して、 $\hat{\mathbf{d}}^p = (\hat{d}_1^p, \dots, \hat{d}_R^p)^T$ とし、 $\hat{\mathbf{u}} = (\hat{u}_1, \dots, \hat{u}_R)^T$, $\hat{\mathbf{D}} = (\hat{\mathbf{d}}^1, \dots, \hat{\mathbf{d}}^P)$ と表記する。

第2段階として、以下の経験損失関数を最小化することで推定量 $\hat{\boldsymbol{\theta}}$ を求めることができる [Olive, 2017]。

$$J_P(\boldsymbol{\theta}; \mathcal{D}) = \|\hat{\mathbf{u}} - \hat{\mathbf{D}}\boldsymbol{\theta}\|^2.\tag{4.51}$$

解は行列 $\hat{\mathbf{D}}^T \hat{\mathbf{D}}$ が可逆であれば、最小二乗法を用いて、

$$(\hat{\mathbf{D}}^T \hat{\mathbf{D}})^{-1} \hat{\mathbf{D}}^T \hat{\mathbf{u}}\tag{4.52}$$

によって与えられ、そうでない場合は特異値分解 [Mandel, 1982] または正則化 [Hilt et al.] などで解くところができる。

4.3.2 P-APCE 推定量の性質

次に一致性や計算量に関する P-APCE 推定量の性質を示す。まず、以下の仮定をする。

仮定 4.3.1. 機械学習で学習した $\hat{\mathbf{u}}$ と $\hat{\mathbf{D}}$ は \mathbf{u} と \mathbf{D} の一致推定量となっている。

すると、次の P-APCE 推定量の一致性に関する定理が成り立つ。

定理 4.3.1 (一致性). 仮定 4.1.1, 4.1.2, 4.2.1, 4.3.1 のもとで、極限 $N \rightarrow \infty$ に対して $\hat{\boldsymbol{\theta}}$ は $\boldsymbol{\theta}$ の一致推定値である。

このモデルは、変数誤差モデル [Söderström, 2007] や測定誤差モデル [Fuller, 2009] と考えられるため、推定量 $\hat{\boldsymbol{\theta}}$ はバイアスがかかる。説明変数にノイズが存在することで、

係数が過小評価される傾向になる。変数誤差モデルのバイアスのノルムは $\hat{\mathbf{D}}^T \hat{\mathbf{D}}$ のノルムの逆数に従って減少するので、仮に予測値の条件付き分散が $\mathcal{O}(N^{-1/2})$ ($N \rightarrow \infty$) に従って減少するなら [Greene, 1997], バイアスの期待値ノルムは $\mathcal{O}(N^{-1})$ ($N \rightarrow \infty$) に従って減少する。また, P-APCE 推定量の分散は変数誤差モデルの理論に帰着する。計算量については, 予測モデルの構築と $R + R \times P$ の予測値の計算の複雑さは機械学習に依存し, 逆行列の計算は通常 $\mathcal{O}(R^3)$ 時間で計算される。

モデル選択について. 式 (4.51) の経験損失関数は, 別の検証データ

$$\mathcal{D}' = \{z^{(i)}, x^{(i)}, y^{(i)}\}_{i=1}^{N'} \quad (4.53)$$

を用いて学習したモデルの検証誤差として使うことができる。 $\hat{\mathbf{u}}'$ と $\hat{\mathbf{D}}'$ はデータセット \mathcal{D}' を使って計算されたとする。そして, 次のような検証誤差によって学習済みモデルの性能を評価することができる。

$$J_P^{test}(\hat{\boldsymbol{\theta}}; \mathcal{D}') = \|\hat{\mathbf{u}}' - \hat{\mathbf{D}}' \hat{\boldsymbol{\theta}}\|^2 \quad (4.54)$$

この検証誤差は第2段階のモデルの性能のみを評価しているが, 第1段階の推定は標準的な回帰問題であるため, 第1段階は標準的な方法でモデルの評価および選択が行われていることを想定している。この検証誤差は様々な基底関数候補の中からモデルを選択したり, 基底の数 P の選択に利用可能である。また, 交差検証 [Kohavi, 1995] に応用することも可能である。

4.4 数値実験

この節では, P-APCE 推定量と N-APCE 推定量の性能を調べるために数値実験を行う。とくに, 先行研究である二段階残差投入法 (Two-stage predictor substitution; TSPS) [Terza et al., 2008] やノンパラメトリック二段階推定法 (Nonparametric two-stage least squares; NPTLS) [Newey and Powell, 2003] と比較することにする。

表 4.1: 100 回のシミュレーションによってえられる P-APCE 推定量と TSPS の推定量の平均と標準偏差. ただし, サンプルサイズはそれぞれ $N = 100$, $N = 1000$ である.

係数の真値 (1, 2, 3)		$D = 0$	$D = 1$	$D = 2$
P-APCE	平均	7.879	-11.128	8.525
$N = 100$	標準偏差	10.516	20.884	9.214
TSPS	平均	-9.749	29.489	-11.571
$N = 100$	標準偏差	48.039	118.938	69.845
P-APCE	平均	0.995	2.132	2.878
$N = 1000$	標準偏差	4.951	9.477	3.997
TSPS	平均	-1.590	8.403	1.531
$N = 1000$	標準偏差	15.110	37.188	21.470

数値実験の設定. 数値実験の設定として以下のような構造的因果モデルを考える.

$$\begin{cases} X := \frac{1}{25}Z^2 + \frac{1}{5}Z + 0.5 + (\frac{Z}{3} + 0.1)U \\ Y := X^3 + X^2 + X + U + E \end{cases} \quad (4.55)$$

この構造的因果モデルは f_X と f_Y の両方が非線形関数であり, 分離可能性と仮定 4.1.1 を満たしている. 外生変数 U と E の各実現値は一様分布 $U[-1, 1]$ の独立同一分布からサンプリングされる. また, 11 個の異なる Z の値 $(0, 0.3, \dots, 2.7, 3)$ において, 式 (4.55) の構造的因果モデルを用いて人工データセットを生成した. 各 Z の値におけるサンプルサイズはそれぞれ 10 と 100 であり, 合計サンプルサイズは $N = 100$ と $N = 1000$ である. 式 (4.12) の $\hat{\mu}$ と \hat{k} は条件付き標本平均

$$\hat{\mathbb{E}}[Y|Z = z] = \frac{\sum_{i=1}^N y^{(i)} \mathbb{1}_{z^{(i)}=z}}{\sum_{i=1}^N \mathbb{1}_{z^{(i)}=z}} \quad (4.56)$$

で計算することとする. また, 条件付き標本平均をもちいて $\hat{\mathbf{u}}$ と $\hat{\mathbf{D}}$ を計算する. 条件付き標本平均は, 第 1 段階の一致性と不偏性に関する仮定 4.2.3 と 4.2.5 を満たし

ている。各シミュレーションは100回ずつ行うこととする。

N-APCE 推定量の設定 (アルゴリズム 3) つぎに、N-APCE 推定量の設定を説明する。 $\mathcal{X} = \{0, 0.3, \dots, 2.7, 3\}$ とし、定数 x_0 および z_0 は0とする。ただし、 $X = 0$ での N-APCE 推定量は定義されないことに注意されたい。初期関数 $\hat{\theta}_1$ を零関数、停止閾値 ϵ を10とする。ステップサイズ α は、アルゴリズムが100回反復する前に停止する、 $(1, 0.5, 0.1, \dots)$ の中から最小のものを選ぶこととし、結果としてステップサイズ α は0.5となった。

P-APCE 推定量の設定 (アルゴリズム 4) つぎに、P-APCE 推定量の設定を説明する。基底関数には多項式基底関数

$$\phi_p(x) = x^{p-1}, \quad p = 1, 2, \dots \quad (4.57)$$

を用い、 $(\hat{\mathbf{D}}^T \hat{\mathbf{D}})^{-1} \hat{\mathbf{D}}^T \hat{\mathbf{u}}$ により式(4.51)の解を算出する。このモデルの最適な次数を決定するために、データセットを訓練データ \mathcal{D} と検証データ \mathcal{D}' に分け、訓練データで $\hat{\theta}$ を推定し、検証誤差(4.54)を用いて訓練済みモデルを評価した。さらに、100回シミュレーションを行い検証誤差を算出した。その結果、検証誤差の平均が最も小さくなるのはP-APCE 推定量の多項式関数の次数が3のときであった。くわえて、TSPS と NPTSLS の次数も同様に3とする。

実験結果. 数値実験を100回実行したときのN-APCE 推定量の平均と標準偏差、および、式(4.39)による近似標準偏差を表4.3に示す。また、P-APCE 推定量、TSPS、NPTSLS の平均と標準偏差を表4.1に示す。また、 $X = (0.3, 0.6, \dots, 2.7, 3)$ の各点における推定値の箱ひげ図を図4.2に示す。数値実験の結果から、P-APCE 推定量はTSPS や NPTSLS よりも優れた性能を示した。P-APCE 推定量のバイアスと標準偏差は $N = 100$ のときに比較的大きく、 $N = 1000$ のときに比較的小きなバイアスと標準偏差を持つ。一方、TSPS 推定量と NPTSLS 推定量は $N = 100$ と $N = 1000$

の両方で相対的に大きなバイアスと標準偏差を持つ。N-APCE 推定量も同様に標準偏差が比較的小さく、平均は真の平均偏因果効果の値に近い。また、式 (4.39) で計算した近似標準偏差は N-APCE 推定値の実際の標準偏差に近い値になっている。

4.5 実データへの応用

本節では、経済学における実データに対する提案法の適用例を紹介する。

実データセット. この節では実データへの応用例として R パッケージ “wooldridge” (<https://cran.r-project.org/package=wooldridge>) にあるオープンデータセットを取りあげる。データソースは “National Longitudinal Survey of Young Men” と呼ばれる研究の一部であり、サンプルサイズは 935 である。このデータセットは、Griliches [1977] と Blackburn and Neumark [1992] などによって分析されており、経済学で関心の高い「月給に対する教育年数の因果効果」の評価を目標としている [Card, 1999, Angrist and Krueger, 1991]。ここで、研究者は学生を強制的に学校に通わせたり退学させたりすることは不可能なので、先行研究と同様に、母親の教育年数を操作変数として利用することとする。本論文で使用する変数を以下に列挙する。

処置変数 (X): 教育年数

アウトカム (Y): 月給

操作変数 (Z): 母親の教育年数

ここで、処置変数 X と操作変数 Z は離散化された連続変数であるとみなす。変数 X と Z のとる値はどちらも $\{9, 10, \dots, 18\}$ であり、これは高校 1 年から修士 2 年までに対応している。3 つの変数 X, Y, Z のうち 1 つでも利用不可であるサンプルは除外し、条件付き平均を用いて条件付き期待値を推定する。P-APCE 推定量の多項式の次数を検証誤差から求め、平均偏因果効果の候補には一次関数を選んだ。平均偏因果効果は 1000 回のブートストラップ法で評価する。推定量の分散を小さくし安定化

させるために、行列 $\hat{\mathbf{D}}^T \hat{\mathbf{D}}$ に $0.1 \times \mathbf{I}$ を加えて正則化する。ここで、行列 \mathbf{I} は $R \times R$ の単位行列である。

適用結果. ブートストラップ法による P-APCE 推定量と TSPS の基本的な統計的な性質を表 4.2 に示す。ただし、N-APCE 推定量は収束しなかった。P-APCE 推定量の定数項の平均は 192.491、1 次項の係数の平均は -10.267 であり、TSPS の定数項の平均は 108.484、1 次項の係数の平均は 0.073 であった。P-APCE 推定量、TSPS とともに教育年数が賃金を増加させることを示し、これは先行研究の結果とも一致する [Blackburn and Neumark, 1992, Wooldridge, 2010]。一方、TSPS の結果は教育年数の賃金への効果がほぼ一定であることを意味しているが、P-APCE 推定量の推定結果は教育年数の賃金への効果が年ごとに弱くなることを意味している。

P-APCE 推定量の結果は義務教育レベルでは教育が賃金に大きく影響することを示唆しており、この点については [Angrist and Krueger, 1991] と一致している。一方で、大学レベルで高い教育を受けることで賃金上がるのは、教育による効果ではなく同じスキルを持っていても学歴がある人の方が学歴がない人よりも高い所得を得るという Spence [1973] や Caplan [2018] が述べた現象によるものかもしれない。この現象は、「シープスキン効果」と呼ばれているものである [Jaeger and Page, 1996]。

4.6 考察

本論文では、2つの新しい平均偏因果効果を推定法を開発した。提案した2つの推定量（P-APCE 推定量と N-APCE 推定量）のそれぞれの統計的な性質を紹介し、数値実験と実データへの応用で提案法の有用性を実証した。P-APCE 推定量の性能は基底関数の種類や次数の選択に依存することが欠点であり、N-APCE 推定量はモデルの形式に関する仮定をする必要がなくハイパーパラメータが存在しないが、計算量が多いことが欠点となっている。

表 4.2: 実データに対する P-APCE 推定量と TSPS の適用結果

	P-APCE 推定量		TSPS	
	$D = 0$	$D = 1$	$D = 0$	$D = 1$
最小値	-664.970	-55.854	-990.480	-82.835
第一分位数	73.724	-18.613	-106.334	-15.593
中央値	186.969	-9.842	118.563	-0.657
第三分位数	312.781	-1.939	313.215	16.273
最大値	834.128	51.523	1216.730	82.870
平均	192.491	-10.267	108.484	0.073
標準偏差	182.698	13.0290	325.414	24.646

P-APCE 推定量や N-APCE 推定量は構造関数 $\mathbb{E}[Y_x]$ を推定するのではなく、平均偏因果効果 $\mathbb{E}[\partial_x Y_x]$ を“直接”推定している。しかし、因果推論の主な関心は変化する条件もとでの処置の効果を推定することであるため、因果効果を明らかにするには平均偏因果効果で十分であることが多い [Pearl, 2010b]。特に、平均偏因果効果を用いるとよく用いられる指標である処置を x' から x'' に処置を変えたときの平均因果効果を

$$\mathbb{E}[Y_{x'}] - \mathbb{E}[Y_{x''}] = \int_{x''}^{x'} \mathbb{E}[\partial_x Y_x] dx \quad (4.58)$$

によって評価することができる。

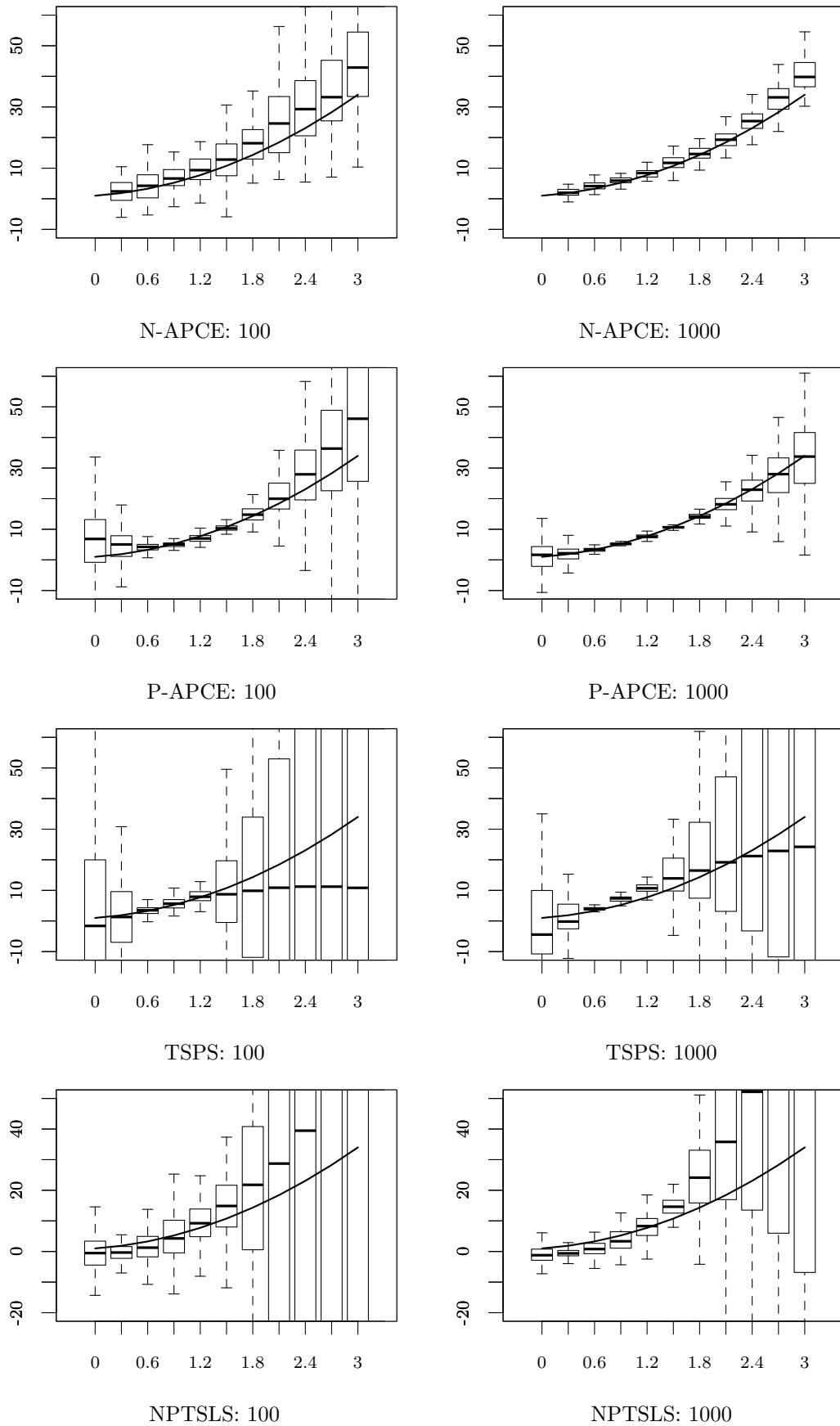


図 4.2: それぞれの推定量による平均偏因果効果の値の推定値の箱ひげ図. X 軸は処置変数 X の値, Y 軸は $X = x$ のときの平均偏因果効果の値であり, 黒い曲線は真の平均偏因果効果である.

表 4.3: N-APCE を 100 回実行したときの (0.3, 0.6, ..., 2.7, 3.0) における推定量の平均と標準偏差. サンプルサイズは, $N = 100$ と $N = 1000$ であり, 式 (4.39) による近似標準偏差の平均も記載している.

$X = x$	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3
平均偏因果効果の真の値	1.87	3.28	5.23	7.72	10.75	14.32	18.43	23.08	28.27	34
($N = 100$)	2.519	4.416	6.907	10.001	13.588	18.721	24.918	30.223	35.955	44.400
標準偏差	4.386	4.577	5.387	6.335	8.269	8.145	11.114	12.389	14.135	16.371
近似標準偏差	4.578	5.481	5.408	5.478	7.472	11.881	15.197	14.683	19.716	24.110
($N = 1000$)	1.822	3.647	5.397	7.514	10.561	14.394	18.601	24.017	30.521	37.536
標準偏差	1.365	1.180	1.161	1.415	1.971	2.373	3.094	3.813	4.472	5.004
近似標準偏差	1.246	1.414	1.655	2.010	2.468	2.987	3.817	4.789	5.843	6.824

第 5 部

まとめ

さいごに、本論文ではノンパラメトリック操作変数モデルに基づいた新たな因果効果の統計的推測法を提供した。どれもいままで考えられてこなかったものであり、既存の推定方法より優れている点が存在していることを、この論文を通して示した。また、数値実験や実データへの応用を通して提案法の性質を実証した。これらの内容は、[Kawakami, 2021], [Kawakami, Shingaki, and Kuroki, 2023b], [Kawakami, Kuroki, and Tian, 2023a] において公表されている。

第 2 部では、操作変数に加えて代替変数として個体の共変量情報を利用することで、既存の手法より弱い仮定のもとで平均因果効果が識別可能であることを示した。さらに、第 3 部において、個体の共変量情報を利用することで平均因果効果にとどまらず、より一般化された指標である潜在反応タイプの確率まで識別可能であることを示した。さらに、有界制約付き拡張ラグランジュ法に基づく潜在反応タイプの確率のより効率的な推定法を提案した。16 種類の潜在反応タイプの確率は、

$$p(X_{z_0}, X_{z_1}, Y_{x_0}, Y_{x_1}) \quad (5.1)$$

で表され、これは以下の因果推論における 3 つの重要な指標を評価することを可能にしている。

1 つ目は、処置がアウトカムを引き起こすのにどの程度必要、十分、あるいは必要十分であったかを評価するための指標である 3 つの原因の確率 [黒木学, 2017, Cai

and Kuroki, 2005, Kuroki and Cai, 2011, Pearl, 2009] である。3つの原因の確率とは、必要十分性の確率

$$p(Y_{x_0} = y_0, Y_{x_1} = y_1), \quad (5.2)$$

必要性の確率

$$p(Y_{x_0} = y_0 | X = x_1, Y = y_1), \quad (5.3)$$

十分性の確率

$$p(Y_{x_1} = y_1 | X = x_0, Y = y_0) \quad (5.4)$$

の3つである。この3つの原因の確率は、単調性IIなどの強い仮定を置かなければ、操作変数 (Z) と処置変数 (X) とアウトカム (Y) のデータのみでは識別不可能であり、確率の存在範囲が与えられるだけであった [Tian and Pearl, 2000, Kuroki and Cai, 2011]。くわえて、先行研究における単調性IIの仮定

$$p(u_3) = p(Y_{x_0} = y_1, Y_{x_1} = y_0) = 0 \quad (5.5)$$

が検証可能なものとなった。

2つ目は、処置群や対照群における平均因果効果である [Hirano et al., 2003, Angrist and Pischke, 2009, Cunningham, 2021]。これは、ランダム化試験において、実験的処置（対照的処置）を受けた部分集団における平均因果効果であり

$$p(Y_{x_1} = y_1 | X = x_1) - p(Y_{x_0} = y_1 | X = x_1) \quad (5.6)$$

$$p(Y_{x_1} = y_1 | X = x_0) - p(Y_{x_0} = y_1 | X = x_0) \quad (5.7)$$

で定義される。これらの指標は、式(3.37)における $p(u_2|x_1) - p(u_3|x_1)$ や $p(u_2|x_0) - p(u_3|x_0)$ を通して識別および推定可能である。

3つ目は、“defiance”の確率である。局所平均因果効果 [Imbens and Angrist, 1994] は、“defiance”の個体が存在しないという単調性Iの仮定

$$p(v_3) = p(X_{z_0} = x_1, X_{z_1} = x_0) = 0 \quad (5.8)$$

のもとで識別および推定が可能となるが、この単調性Iの仮定はいままで検証不可能な仮定であった。しかしながら、本論文の結果によってこの“defiance”の確率を識別および推定可能とし、先行研究における単調性Iの仮定は検証可能なものとなった。以上のように、本論文の結果はいままで不可能であった指標を識別可能にし、より踏み込んだ因果効果の評価を可能にするものである。

最後に、[Balke and Pearl, 1994]のように連続変数や2値以上の離散変数のアウトカムに関しても、ある閾値 y に基づく変数 $Y \leq y$ を考えることで、本論文の結果を適用することが可能である。また、2値以上の値をとる離散な処置変数にも拡張可能であるが、スパース性が問題となる。さらに、この有界制約付き拡張ラグランジュ法を用いた推定法は非凸最適化問題であるため、局所解に陥ってしまう可能性がありパラメータの初期値によっては大きく真値から離れた推定値を返すことがある。これらの問題は今後の課題とする。

第4部では、操作変数と処置変数が連続変数の場合において、[Wong, 2022]によって与えられた平均偏因果効果の識別可能条件に基づいて、ピカールの反復法と線形基底モデルに基づいた2つの平均偏因果効果の推定法(N-APCE推定量とP-APCE推定量)を紹介した。とくに、N-APCE推定量は、ピカールの反復法を利用した統計的な手法という点において既存の手法とは全く異なる新たな推定手法となっている。平均偏因果効果は潜在反応変数 Y_x の一次偏導関数の期待値

$$E[\partial_x Y_x] \tag{5.9}$$

で定義されており、二つの処置 x'' と x' に対する平均因果効果は、

$$E[Y_{x''}] - E[Y_{x'}] = \int_{x'}^{x''} E[\partial_x Y_x] dx \tag{5.10}$$

で表現できる。因果推論の主な関心は変化する処置の効果の推定することであるため [Pearl, 2010b], 平均偏因果効果は因果効果の評価において十分な指標である。

構造関数 $E[Y_x]$ の推定と平均偏因果効果の推定は、[Chamberlain, 1984, Wooldridge,

2005, Graham and Powell, 2012] を除いて、これまであまり区別されて議論が行われることはなかった。これは、構造関数をとおして平均偏因果効果が簡単に求められるからである。しかしながら、本論文では、操作変数モデルにおいては平均偏因果効果を“直接”推定することは、構造関数の推定に比べておおきく2つの長所が存在することを示した。1つ目は、積分核が条件付き累積分布関数の差で構成されていることである。条件付き累積分布関数の差は $-1 \sim 1$ の値を取る有界な関数であり、条件付き密度関数に比べて推定が容易なものである。2つ目は、定義域が有界であれば積分方程式が良設定問題となることである。これらの性質によって、ピカールの反復法が適用可能となっている。また、

$$\mathbb{E}[f_{Y_2}(\mathbf{H}, \epsilon_Y) | Z = z] = 0 \quad \forall z \in \Omega_Z \quad (5.11)$$

の仮定が必要ないことも利点である。

さいごに、共変量を伴うノンパラメトリック操作変数モデルへの拡張や、分離可能性の仮定が成り立たない場合の因果効果の統計解析の手法は今後の課題とする。また、P-APCE 推定量の発展形として、機械学習の分野で盛んに研究されている操作変数モデルに基づく構造関数の推定法、たとえば、Deep IV [Hartford et al., 2017], Kernel IV [Singh et al., 2019], Deep GMM [Bennett et al., 2019], Dual IV [Muandet et al., 2020], DFIV [Xu et al., 2022] などが平均偏因果効果の推定に応用可能かどうかは今後の課題とする。

謝辞

まず、指導教官の黒木学先生に深く感謝の意を示します。黒木先生には、2018年4月から2023年9月の5年半にわたり、統計学および因果推論の基礎から、研究活動の基本までご指導いただきました。学部生の頃より多くの国内外での学会参加や発表の機会をいただき、私にとって貴重な経験となりました。また、2022年10月より開始したアイオワ州立大学のJin Tian教授のもとでの留学の際にもお力添えを頂きました。

学位論文の博論審査等にあたり、梶原先生、植木先生、竹居先生、本田先生には大変お世話になりました。発表および博士論文に関する多大なるご指導をいただき深く感謝申し上げます。また、共著者の新垣隆生様にも深く感謝申し上げます。

I also would like to express my sincere gratitude to my advisor at Iowa State University (ISU), Professor Jin Tian, for his thoughtful guidance. I started my research with Prof. Tian at ISU in October 2022, and Chapter 4 is based on my work with him.

参考文献

- Chunrong Ai and Xiaohong Chen. Efficient estimation of models with conditional moment restrictions containing unknown functions. *Econometrica*, 71(6):1795–1843, 2003.
- Alen Alexanderian. On compact operators. *Journal of Mathematical Analysis and Applications*, 18(2):5–36, 2013.
- Joshua Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. 01 2009. ISBN 9780691120348 (hardcover : alk. paper).
- Joshua D. Angrist and Alan B. Krueger. Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics*, 106(4):979–1014, 1991.
- Joshua D. Angrist and Alan B. Krueger. Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives*, 15(4):69–85, December 2001.
- Joshua D. Angrist, Guido W. Imbens, and Donald B. Rubin. Identification of causal effects using instrumental variables. *J. Amer. Statist. Assoc.*, 91(434):444–455, 1996.
- Joshua David Angrist and Alan B. Krueger. Split-sample instrumental variables es-

- imates of the return to schooling. *Journal of Business & Economic Statistics*, 13: 225–235, 1995. URL <https://api.semanticscholar.org/CorpusID:56415374>.
- S. Athey and G.W. Imbens. Chapter 3 - the econometrics of randomized experiments. In Abhijit Vinayak Banerjee and Esther Duflo, editors, *Handbook of Field Experiments*, volume 1 of *Handbook of Economic Field Experiments*, pages 73–140. North-Holland, 2017.
- Alexander Balke and Judea Pearl. Counterfactual probabilities: Computational methods, bounds and applications. In *Proceedings of the Tenth International Conference on Uncertainty in Artificial Intelligence, UAI'94*, page 46–54, San Francisco, CA, USA, 1994. Morgan Kaufmann Publishers Inc. ISBN 1558603328.
- Alexander Balke and Judea Pearl. Bounds on treatment effects from studies with imperfect compliance. *Journal of the American Statistical Association*, 92(439): 1171–1176, 1997.
- Andrew Bennett, Nathan Kallus, and Tobias Schnabel. Deep generalized method of moments for instrumental variable analysis. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.
- Ernesto G Birgin and José Mario Martínez. Complexity and performance of an augmented lagrangian algorithm. *Optimization Methods and Software*, 35(5):885–920, 2020.
- Christopher Bishop. *Pattern Recognition and Machine Learning*. Springer, January 2006.
- McKinley Blackburn and David Neumark. Unobserved ability, efficiency wages, and interindustry wage differentials. *The Quarterly Journal of Economics*, 107

(4):1421–1436, 1992. ISSN 00335533, 15314650. URL <http://www.jstor.org/stable/2118394>.

Richard Blundell, Xiaohong Chen, and Dennis Kristensen. Semi-nonparametric iv estimation of shape-invariant engel curves. *Econometrica*, 75(6):1613–1669, 2007.

Maxime Bôcher. *An introduction to the study of integral equations*. Number 10. University Press, 1926.

Richard L Burden, J Douglas Faires, and Annette M Burden. *Numerical analysis*. Cengage learning, 2015.

Zhihong Cai and Manabu Kuroki. Variance estimators for three “probabilities of causation”. *Risk Analysis*, 25(6):1611–1620, 2005.

Zhihong Cai, Manabu Kuroki, and Tosiya Sato. Non-parametric bounds on treatment effects with non-compliance by covariate adjustment. *Statist. Med.*, 26(16):3188–3204, 2007.

Bryan Caplan. *The Case against Education: Why the Education System Is a Waste of Time and Money*. Princeton University Press, 2018. ISBN 9780691174655. URL <http://www.jstor.org/stable/j.ctvc772xh>.

David Card. Chapter 30 - the causal effect of education on earnings. volume 3 of *Handbook of Labor Economics*, pages 1801–1863. Elsevier, 1999.

T C Chalmers, H Jr Smith, B Blackburn, B Silverman, B Schroeder, D Reitman, and A Ambroz. A method for assessing the quality of a randomized control trial. *Control Clin Trials*, 2(1):31–49, May 1981.

Gary Chamberlain. Panel data. In Z. Griliches † and M. D. Intriligator, editors,

- Handbook of Econometrics*, volume 2, chapter 22, pages 1247–1318. Elsevier, 1 edition, 1984. URL <https://EconPapers.repec.org/RePEc:eee:ecochp:2-22>.
- Victor Chernozhukov, Guido W. Imbens, and Whitney K. Newey. Instrumental variable estimation of nonseparable models. *Journal of Econometrics*, 139(1): 4–14, 2007.
- Edwin KP Chong and Stanislaw H Żak. *An introduction to optimization*, volume 75. John Wiley & Sons, 2013.
- Scott Cunningham. *Causal Inference: The Mixtape*. Yale University Press, 2021. ISBN 9780300251685. URL <http://www.jstor.org/stable/j.ctv1c29t27>.
- J. B. Diaz and F. T. Metcalf. On iteration procedures for equations of the first kind, $ax = y$, and picard’s criterion for the existence of a solution. *Mathematics of Computation*, 24(112):923–935, 1970.
- G Emilien, W van Meurs, and J M Maloteaux. The dose-response relationship in phase i clinical trials and beyond: use, meaning, and assessment. *Pharmacol Ther*, 88(1):33–58, Oct 2000.
- Jianqing Fan and Qiwei Yao. Efficient estimation of conditional variance functions in stochastic regression. *Biometrika*, 85(3):645–660, 1998.
- Thomas S Ferguson. *A Course in Large Sample Theory*. Routledge, 1996.
- Constantine E. Frangakis and Donald B. Rubin. Principal stratification in causal inference. *Biometrics*, 58(1):21–29, 2002.
- Wayne A Fuller. *Measurement error models*. John Wiley & Sons, 2009.

Bryan S. Graham and James L. Powell. Identification and estimation of average partial effects in "irregular" correlated random coefficient panel data models. *Econometrica*, 80(5):2105–2152, 2012. ISSN 00129682, 14680262. URL <http://www.jstor.org/stable/23271443>.

W.H. Greene. *Econometric Analysis*. Prentice-Hall international editions. Prentice Hall, 1997.

Zvi Griliches. Estimating the returns to schooling: Some econometric problems. *Econometrica*, 45(1):1–22, 1977. ISSN 00129682, 14680262.

Robert Hable. Asymptotic normality of support vector machine variants and other regularized kernel methods. *Journal of Multivariate Analysis*, 106:92–117, 2012.

Jason Hartford, Greg Lewis, Kevin Leyton-Brown, and Matt Taddy. Deep IV: A flexible approach for counterfactual prediction. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 1414–1423. PMLR, 06–11 Aug 2017.

T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*. Springer Series in Statistics. Springer New York, 2009.

Miguel A Hernán and James M Robins. Estimating causal effects from epidemiological data. *J Epidemiol Community Health*, 60(7):578–586, Jul 2006.

Nicholas J Higham. *Accuracy and stability of numerical algorithms*. SIAM, 2002.

Donald E. Hilt, Donald W. Seegrist, United States. Forest Service., and Pa.) North-eastern Forest Experiment Station (Radnor. *Ridge, a computer program for cal-*

culating ridge regression estimates, volume no.236. Upper Darby, Pa, Dept. of Agriculture, Forest Service, Northeastern Forest Experiment Station, 1977.

Keisuke Hirano, Guido W. Imbens, and Geert Ridder. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189, 2003.

Paul W. Holland. Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960, 1986. ISSN 01621459.

Guido W. Imbens. Instrumental variables: An econometrician’s perspective. *Statistical Science*, 29(3):323–358, 2014. ISSN 08834237, 21688745. URL <http://www.jstor.org/stable/43288511>.

Guido W. Imbens and Joshua D. Angrist. Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475, 1994.

Guido W Imbens and Donald B Rubin. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press, 2015.

Atsushi Inoue and Gary Solon. Two-sample instrumental variables estimators. Working Paper 311, National Bureau of Economic Research, June 2005. URL <http://www.nber.org/papers/t0311>.

David Jaeger and Marianne Page. Degrees matter: New evidence on sheepskin effects in the returns to education. *The Review of Economics and Statistics*, 78(4):733–40, 1996.

H Jeffreys and BS Jeffreys. Lagrange ’ s interpolation formula. *Methods of mathematical physics*, 3:260, 1988.

- Yuta Kawakami. Instrumental variable-based identification for causal effects using covariate information. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(13):12131–12138, May 2021.
- Yuta Kawakami, Manabu Kuroki, and Jin Tian. Instrumental variable estimation of average partial causal effects. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 16097–16130. PMLR, 23–29 Jul 2023a. URL <https://proceedings.mlr.press/v202/kawakami23a.html>.
- Yuta Kawakami, Ryusei Shingaki, and Manabu Kuroki. Identification and estimation of the probabilities of potential outcome types using covariate information in studies with non-compliance. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(10):12234–12242, Jun. 2023b.
- Ron Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, IJCAI’95, page 1137–1143, San Francisco, CA, USA, 1995. Morgan Kaufmann Publishers Inc. ISBN 1558603638.
- Amanda Kowalski. Counting defiers: Examples from health care. *arXiv preprint arXiv:1912.06739*, 2020.
- Manabu Kuroki and Zhihong Cai. Statistical analysis of ‘probabilities of causation’ using covariate information. *Scand. J. Statist.*, 38(3):564–577, 2011.
- Manabu Kuroki and Judea Pearl. Measurement bias and effect restoration in causal inference. *Biometrika*, 101(2):423–437, 2014.
- Ang Li and Judea Pearl. Unit selection based on counterfactual logic. In *Proceedings*

of the *Twenty-Eighth International Joint Conference on Artificial Intelligence*, pages 1793–1799, 2019.

Rensis Likert. A technique for the measurement of attitudes. *Archives of psychology*, 1932.

Kung-Jong Lui. *Binary Data Analysis of Randomized Clinical Trials with Noncompliance*. John Wiley & Sons, 2011.

John Mandel. Use of the singular value decomposition in regression analysis. *The American Statistician*, 36(1):15–24, 1982.

Charles F Manski. Nonparametric bounds on treatment effects. *American Economic Review, Papers and Proceedings*, 80(2):319–323, Oct 1990.

Steven D Mark and James M Robins. A method for the analysis of randomized trials with compliance information; an application to the multiple risk factor intervention trial. *Controlled Clinical Trials*, 14(1):79–97, Oct 1993.

Marcia L. Meldrum. A brief history of the randomized controlled trial, August 2000. URL [https://doi.org/10.1016/s0889-8588\(05\)70309-9](https://doi.org/10.1016/s0889-8588(05)70309-9).

Wang Miao, Zhi Geng, and Eric Tchetgen Tchetgen. Identifying causal effects with proxy variables of an unmeasured confounder. *Biometrika*, 105(4):987–993, Dec 2018.

Krikamol Muandet, Arash Mehrjou, Si Kai Lee, and Anant Raj. Dual instrumental variable regression. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS’20, Red Hook, NY, USA, 2020. Curran Associates Inc.

- J. Muscat. *Functional Analysis: An Introduction to Metric Spaces, Hilbert Spaces, and Banach Algebras*. Springer International Publishing, 2014.
- Whitney K. Newey and James L. Powell. Instrumental variable estimation of non-parametric models. *Econometrica*, 71(5):1565–1578, 2003.
- Jorge Nocedal and Stephen J Wright. *Numerical optimization*. Springer, 1999.
- David J Olive. Multiple linear regression. In *Linear regression*, pages 17–83. Springer, 2017.
- Judea Pearl. [bayesian analysis in expert systems]: Comment: Graphical models, causality and intervention. *Statistical Science*, 8(3):266–269, 1993. ISSN 08834237. URL <http://www.jstor.org/stable/2245965>.
- Judea Pearl. On the testability of causal models with latent and instrumental variables. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence, UAI'95*, page 435–443, San Francisco, CA, USA, 1995a. Morgan Kaufmann Publishers Inc.
- Judea Pearl. On the testability of causal models with latent and instrumental variables. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence, UAI'95*, page 435–443. Morgan Kaufmann Publishers Inc., 1995b. ISBN 1558603859.
- Judea Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, 2nd edition, 2009.
- Judea Pearl. On measurement bias in causal inference. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, pages 425–432, Arlington, Virginia, USA, 2010a. AUAI Press.

- Judea Pearl. An introduction to causal inference. *Int J Biostat*, 6(2):Article 7, Feb 2010b.
- Judea Pearl and Dana Mackenzie. *The book of why: the new science of cause and effect*. Basic books, 2018.
- Karl Pearson. Method of moments and method of maximum likelihood. *Biometrika*, 28(1/2):34–59, 1936. ISSN 00063444. URL <http://www.jstor.org/stable/2334123>.
- Émile Picard. Sur un théorème général relatif aux équations intégrales de première espèce et sur quelques problèmes de physique mathématique. *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, 29:79–97, 1910.
- James Robins. The analysis of randomized and non-randomized aids treatment trials using a new approach to causal inference in longitudinal studies. In *Health Service Research Methodology: A Focus on AIDS*, pages 113–159. U.S. Public Health Service, Washington, DC., 1989.
- James Robins and Thomas S. Richardson. Alternative graphical causal models and the identification of direct effects. In *Causality and psychopathology: Finding the determinants of disorders and their cures*, pages 103–158. Oxford University Press, 2011.
- Paul R. Rosenbaum and Donald B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.
- K.J. Rothman, S. Greenland, and T.L. Lash. *Modern Epidemiology*. Lippincott Williams & Wilkins, 3rd edition, 2008.

- Donald B. Rubin. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, 100(469):322–331, 2005. ISSN 01621459. URL <http://www.jstor.org/stable/27590541>.
- Gwendolin B Sajons. Estimating the causal effect of measured endogenous variables: A tutorial on experimentally randomized instrumental variables. *The Leadership Quarterly*, 31(5):101348, 2020.
- Kenneth F Schulz, Douglas G Altman, and David Moher. Consort 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*, 340:c332, Mar 2010.
- Yiwei Shen and Pierre C Bellec. Asymptotic normality and confidence intervals for derivatives of 2-layers neural network in the random features model. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 18625–18636. Curran Associates, Inc., 2020.
- Ryusei Shingaki and Manabu Kuroki. Identification and estimation of joint probabilities of potential outcomes in observational studies with covariate information. In *Proceedings of the Thirty-Fifth Conference on Neural Information Processing Systems*, pages 26475–26486, 2021.
- Ilya Shpitser and Judea Pearl. Identification of joint interventional distributions in recursive semi-markovian causal models. In *Proceedings of the 21st National Conference on Artificial Intelligence - Volume 2, AAAI'06*, page 1219–1226. AAAI Press, 2006. ISBN 9781577352815.
- Rahul Singh, Maneesh Sahani, and Arthur Gretton. Kernel instrumental variable regression. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox,

- and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.
- Torsten Söderström. Errors-in-variables methods in system identification. *Automatica*, 43(6):939–958, 2007.
- Michael Spence. Job market signaling. *The Quarterly Journal of Economics*, 87(3): 355–374, 1973. ISSN 00335533, 15314650. URL <http://www.jstor.org/stable/1882010>.
- Jerzy Splawa-Neyman, D. M. Dabrowska, and T. P. Speed. On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. *Statistical Science*, 5(4):465 – 472, 1990.
- Eric J Tchetgen Tchetgen, Andrew Ying, Yifan Cui, Xu Shi, and Wang Miao. An introduction to proximal causal learning, 2020.
- Joseph V. Terza, Anirban Basu, and Paul J. Rathouz. Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*, 27(3):531–543, 2008.
- Jin Tian and Judea Pearl. Probabilities of causation: Bounds and identification. *Annals of Mathematics and Artificial Intelligence*, 28(1):287–313, 2000.
- Jin Tian and Judea Pearl. A general identification condition for causal effects. In *Eighteenth National Conference on Artificial Intelligence*, page 567–573, USA, 2002. American Association for Artificial Intelligence. ISBN 0262511290.
- Robert J Tibshirani and Bradley Efron. An introduction to the bootstrap. *Monographs on statistics and applied probability*, 57(1), 1993.

- Andrei Nikolaevich Tikhonov, AV Goncharsky, VV Stepanov, and Anatoly G Yagola. *Numerical methods for the solution of ill-posed problems*, volume 328. Springer Science & Business Media, 1995.
- Alexander Torgovitsky. Identification of nonseparable models using instruments with small support. *Econometrica*, 83(3):1185–1197, 2015.
- A. W. van der Vaart. *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998. doi: 10.1017/CBO9780511802256.
- Linbo Wang and Eric Tchetgen Tchetgen. Bounded, efficient and multiply robust estimation of average treatment effects using instrumental variables. *Journal of the Royal Statistical Society. Series B, Statistical methodology*, 80(3):531, 2018.
- W K Wong and P A Lachenbruch. Tutorial in biostatistics. designing studies for dose response. *Stat Med*, 15(4):343–359, Feb 1996.
- Wing Hung Wong. An equation for the identification of average causal effect in nonlinear models. *Statistica Sinica*, 32:539–545, 2022.
- Jeffrey M Wooldridge. Unobserved heterogeneity and estimation of average partial effects. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, pages 27–55, 2005.
- Jeffrey M Wooldridge. *Econometric analysis of cross section and panel data*. MIT press, 2010.
- Jeffrey M Wooldridge. *Introductory econometrics: A modern approach*. Cengage learning, 2015.

P.G. Wright. *The Tariff on Animal and Vegetable Oils*. Investigations in international commercial policies. Macmillan, 1928.

Anpeng Wu, Kun Kuang, Ruoxuan Xiong, and Fei Wu. Instrumental variables in causal inference and machine learning: A survey, 12 2022.

Sascha Xu, Osman A Mian, Alexander Marx, and Jilles Vreeken. Inferring cause and effect in the presence of heteroscedastic noise. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato, editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 24615–24630. PMLR, 17–23 Jul 2022.

黒木学. **構造的因果モデルの基礎**. 共立出版, 2017.

付録

この付録では本文で述べている定理などの証明を与える。

第1部の付録

定理 2.2.1 の証明

証明. まず, 任意の $x \in \{x_0, x_1\}$ に対して,

$$Q_x = A_x^T \Delta A_x^{T-1} P_x. \quad (1)$$

が成り立っている. また, 任意の $x \in \{x_0, x_1\}$ に対して,

$$A_x^T = \begin{pmatrix} A_{11,x} & A_{12,x} \\ A_{21,x} & A_{22,x} \end{pmatrix}, \quad (2)$$

$$P_x = \begin{pmatrix} P_{1,x} \\ P_{2,x} \end{pmatrix}, \quad Q_x = \begin{pmatrix} Q_{1,x} \\ Q_{2,x} \end{pmatrix} \quad (3)$$

と表記する. ここで,

$$P_{1,x} = \begin{pmatrix} 1 & \text{pr}(z_0 | x) \\ \text{pr}(w_1 | x) & \text{pr}(w_1, z_0 | x) \end{pmatrix}, \quad (4)$$

$$P_{2,x} = \begin{pmatrix} \text{pr}(w_2 | x) & \text{pr}(w_2, z_0 | x) \\ \text{pr}(w_3 | x) & \text{pr}(w_3, z_0 | x) \end{pmatrix}, \quad (5)$$

$$Q_{1,x} = \begin{pmatrix} \text{pr}(y_0 | x) & \text{pr}(y_0, z_0 | x) \\ \text{pr}(y_0, w_1 | x) & \text{pr}(y_0, w_1, z_0 | x) \end{pmatrix}, \quad (6)$$

$$Q_{2,x} = \begin{pmatrix} \text{pr}(y_0, w_2 | x) & \text{pr}(y_0, w_2, z_0 | x) \\ \text{pr}(y_0, w_3 | x) & \text{pr}(y_0, w_3, z_0 | x) \end{pmatrix} \quad (7)$$

である。くわえて、

$$A_{11,x_0} = \begin{pmatrix} 1 & 1 \\ \text{pr}(w_1 | u_1) & \text{pr}(w_1 | u_2) \end{pmatrix}, \quad (8)$$

$$A_{12,x_0} = \begin{pmatrix} 1 & 1 \\ \text{pr}(w_1 | u_3) & \text{pr}(w_1 | u_4) \end{pmatrix}, \quad (9)$$

$$A_{21,x_0} = \begin{pmatrix} \text{pr}(w_2 | u_1) & \text{pr}(w_2 | u_2) \\ \text{pr}(w_3 | u_1) & \text{pr}(w_3 | u_2) \end{pmatrix}, \quad (10)$$

$$A_{22,x_0} = \begin{pmatrix} \text{pr}(w_2 | u_3) & \text{pr}(w_2 | u_4) \\ \text{pr}(w_3 | u_3) & \text{pr}(w_3 | u_4) \end{pmatrix}, \quad (11)$$

$$A_{11,x_1} = \begin{pmatrix} 1 & 1 \\ \text{pr}(w_1 | u_1) & \text{pr}(w_1 | u_3) \end{pmatrix}, \quad (12)$$

$$A_{12,x_1} = \begin{pmatrix} 1 & 1 \\ \text{pr}(w_1 | u_2) & \text{pr}(w_1 | u_4) \end{pmatrix}, \quad (13)$$

$$A_{21,x_1} = \begin{pmatrix} \text{pr}(w_2 | u_1) & \text{pr}(w_2 | u_3) \\ \text{pr}(w_3 | u_1) & \text{pr}(w_3 | u_3) \end{pmatrix}, \quad (14)$$

$$A_{22,x_1} = \begin{pmatrix} \text{pr}(w_2 | u_2) & \text{pr}(w_2 | u_4) \\ \text{pr}(w_3 | u_2) & \text{pr}(w_3 | u_4) \end{pmatrix} \quad (15)$$

と表記する. すると, 任意の $x \in \{x_0, x_1\}$ に対して, $A_x^{T^{-1}}$ は

$$A_x^{T^{-1}} = \begin{pmatrix} A_x^{11} & -A_x^{11} A_{12,x} A_{22,x}^{-1} \\ -A_{22,x}^{-1} A_{21,x} A_x^{11} & A_x^{22} \end{pmatrix} \quad (16)$$

と表現できる. ここで,

$$A_x^{11} = (A_{11,x} - A_{12,x} A_{22,x}^{-1} A_{21,x})^{-1}, \quad (17)$$

$$A_x^{22} = (A_{22,x} - A_{21,x} A_{11,x}^{-1} A_{12,x})^{-1} \quad (18)$$

であり, また,

$$A_x^T \Delta A_x^{T^{-1}} = \begin{pmatrix} A_{11,x} A_x^{11} \\ A_{21,x} A_x^{11} \end{pmatrix} (I_{2,2}; -A_{12,x} A_{22,x}^{-1}) \quad (19)$$

が成り立つ. ここで, $I_{2,2}$ は 2×2 の単位行列である. すると,

$$Q_{1,x} = A_{11,x} A_x^{11} (P_{1,x} - A_{12,x} A_{22,x}^{-1} P_{2,x}), \quad (20)$$

$$Q_{2,x} = A_{21,x} A_x^{11} (P_{1,x} - A_{12,x} A_{22,x}^{-1} P_{2,x}) \quad (21)$$

が成立する. これらの式から, $Q_{2,x} = A_{21,x} A_{11,x}^{-1} \times Q_{1,x}$ と

$$A_{21,x} = Q_{2,x} Q_{1,x}^{-1} A_{11,x} \quad (22)$$

と

$$A_x^{11,-1} A_{11,x}^{-1} Q_{1,x} \quad (23)$$

$$= Q_{1,x} - A_{12,x} A_{22,x}^{-1} A_{21,x} A_{11,x}^{-1} Q_{1,x} \quad (24)$$

$$= P_{1,x} - A_{12,x} A_{22,x}^{-1} P_{2,x} \quad (25)$$

が成立する。また、関係式

$$Q_{1,x} - P_{1,x} \tag{26}$$

$$= A_{12,x} A_{22,x}^{-1} (A_{21,x} A_{11,x}^{-1} Q_{1,x} - P_{2,x}) \tag{27}$$

$$= A_{12,x} A_{22,x}^{-1} (Q_{2,x} - P_{2,x}) \tag{28}$$

と

$$A_{22,x} = (Q_{2,x} - P_{2,x})(Q_{1,x} - P_{1,x})^{-1} A_{12,x} \tag{29}$$

も成立する。ここで、任意の $x \in \{x_0, x_1\}$ に対して、 a_{ij}^x と b_{ij}^x を行列 $Q_{2,x} Q_{1,x}^{-1}$ と $(Q_{2,x} - P_{2,x})(Q_{1,x} - P_{1,x})^{-1}$ の第 (i, j) 成分とする ($i, j = 1, 2$)。すると、

$$Q_{2,x_0} Q_{1,x_0}^{-1} A_{11,x_0} \tag{30}$$

$$= \begin{pmatrix} a_{11}^{x_0} + a_{12}^{x_0} \text{pr}(w_1 | u_1) & a_{11}^{x_0} + a_{12}^{x_0} \text{pr}(w_1 | u_2) \\ a_{21}^{x_0} + a_{22}^{x_0} \text{pr}(w_1 | u_1) & a_{21}^{x_0} + a_{22}^{x_0} \text{pr}(w_1 | u_2) \end{pmatrix}$$

$$= A_{21,x_0} = \begin{pmatrix} \text{pr}(w_2 | u_1) & \text{pr}(w_2 | u_2) \\ \text{pr}(w_3 | u_1) & \text{pr}(w_3 | u_2) \end{pmatrix}, \tag{31}$$

$$Q_{2,x_1} Q_{1,x_1}^{-1} A_{11,x_1} \tag{32}$$

$$= \begin{pmatrix} a_{11}^{x_1} + a_{12}^{x_1} \text{pr}(w_1 | u_1) & a_{11}^{x_1} + a_{12}^{x_1} \text{pr}(w_1 | u_3) \\ a_{21}^{x_1} + a_{22}^{x_1} \text{pr}(w_1 | u_1) & a_{21}^{x_1} + a_{22}^{x_1} \text{pr}(w_1 | u_3) \end{pmatrix}$$

$$= A_{21,x_1} = \begin{pmatrix} \text{pr}(w_2 | u_1) & \text{pr}(w_2 | u_3) \\ \text{pr}(w_3 | u_1) & \text{pr}(w_3 | u_3) \end{pmatrix}, \tag{33}$$

$$A_{22,x_0} = \begin{pmatrix} \text{pr}(w_2 | u_3) & \text{pr}(w_2 | u_4) \\ \text{pr}(w_3 | u_3) & \text{pr}(w_3 | u_4) \end{pmatrix} \tag{34}$$

$$= (Q_{2,x_0} - P_{2,x_0})(Q_{1,x_0} - P_{1,x_0})^{-1} A_{12,x_0} \tag{35}$$

$$= \begin{pmatrix} b_{11}^{x_0} + b_{12}^{x_0} \text{pr}(w_1 | u_3) & b_{11}^{x_0} + b_{12}^{x_0} \text{pr}(w_1 | u_4) \\ b_{21}^{x_0} + b_{22}^{x_0} \text{pr}(w_1 | u_3) & b_{21}^{x_0} + b_{22}^{x_0} \text{pr}(w_1 | u_4) \end{pmatrix} \tag{36}$$

$$A_{22,x_1} = \begin{pmatrix} \text{pr}(w_2 | u_2) & \text{pr}(w_2 | u_4) \\ \text{pr}(w_3 | u_2) & \text{pr}(w_3 | u_4) \end{pmatrix} \quad (37)$$

$$= (Q_{2,x_1} - P_{2,x_1})(Q_{1,x_1} - P_{1,x_1})^{-1} A_{12,x_1} \quad (38)$$

$$= \begin{pmatrix} b_{11}^{x_1} + b_{12}^{x_1} \text{pr}(w_1 | u_2) & b_{11}^{x_1} + b_{12}^{x_1} \text{pr}(w_1 | u_4) \\ b_{21}^{x_1} + b_{22}^{x_1} \text{pr}(w_1 | u_2) & b_{21}^{x_1} + b_{22}^{x_1} \text{pr}(w_1 | u_4) \end{pmatrix} \quad (39)$$

が成立する．これらの関係式から確率 $\text{pr}(w|u)$ は識別可能である．くわえて，任意の $x \in \{x_0, x_1\}$ に対して， A_x が識別可能であるため，

$$A_x^{T-1} P_x = M_x B_x$$

の第一列を比較することで行列 M_{x_0} と M_{x_1} 各成分は識別可能である．また，確率 $\text{pr}(x_0)$ と $\text{pr}(x_1)$ は観測分布から推定可能であり潜在反応変数の同時確率は識別可能であるため因果効果の識別も可能となる． \square

第2部の付録

定理 3.2.1 の証明

証明. 潜在反応変数の一致性の性質から，

$$p(y_0, z, w, u_i | x_0) = p(z, w, u_i | x_0), \quad i = 1, 2, \quad (40)$$

$$p(y_0, z, w, u_i | x_0) = 0, \quad i = 3, 4, \quad (41)$$

$$p(y_0, z, w, u_i | x_1) = p(z, w, u_i | x_1), \quad i = 1, 3, \quad (42)$$

$$p(y_0, z, w, u_i | x_1) = 0, \quad i = 2, 4 \quad (43)$$

が任意の $z \in \{z_0, z_1\}$ と $w \in \{w_1, w_2, w_3, w_4\}$ に対して成立する．これらの式と条件 (2) から条件付き確率 $p(y, z, w | x)$ と $p(z, w | x)$ は

$$p(y_0, z, w | x_0) = \sum_{i=1,2} p(w | u_i) p(z | x_0, u_i) p(u_i | x_0), \quad (44)$$

$$p(y_0, z, w | x_1) = \sum_{i=1,3} p(w | u_i) p(z | x_1, u_i) p(u_i | x_1), \quad (45)$$

$$p(z, w | x_0) = \sum_{i=1}^4 p(z | x_0, u_i) p(w | u_i) p(u_i | x_0), \quad (46)$$

$$p(z, w | x_1) = \sum_{i=1}^4 p(z | x_1, u_i) p(w | u_i) p(u_i | x_1) \quad (47)$$

とかける。また、変数 Y, Z, W に関する他の同時および周辺化確率も同様に書き下すことができる。また、ブロック行列を

$$A_x^T = \begin{pmatrix} A_{11,x} & A_{12,x} \\ A_{21,x} & A_{22,x} \end{pmatrix}, P_x = \begin{pmatrix} P_{1,x} \\ P_{2,x} \end{pmatrix}, Q_x = \begin{pmatrix} Q_{1,x} \\ Q_{2,x} \end{pmatrix} \quad (48)$$

と定義する。ここで、任意の $x \in \{x_0, x_1\}$ に対して、各成分は

$$P_{1,x} = \begin{pmatrix} 1 & p(z_0 | x) \\ p(w_1 | x) & p(w_1, z_0 | x) \end{pmatrix}, \quad (49)$$

$$P_{2,x} = \begin{pmatrix} p(w_2 | x) & p(w_2, z_0 | x) \\ p(w_3 | x) & p(w_3, z_0 | x) \end{pmatrix}, \quad (50)$$

$$Q_{1,x} = \begin{pmatrix} p(y_0 | x) & p(y_0, z_0 | x) \\ p(y_0, w_1 | x) & p(y_0, w_1, z_0 | x) \end{pmatrix}, \quad (51)$$

$$Q_{2,x} = \begin{pmatrix} p(y_0, w_2 | x) & p(y_0, w_2, z_0 | x) \\ p(y_0, w_3 | x) & p(y_0, w_3, z_0 | x) \end{pmatrix}, \quad (52)$$

$$A_{11,x_0} = \begin{pmatrix} 1 & 1 \\ p(w_1 | u_1) & p(w_1 | u_2) \end{pmatrix}, \quad (53)$$

$$A_{12,x_0} = \begin{pmatrix} 1 & 1 \\ p(w_1 | u_3) & p(w_1 | u_4) \end{pmatrix}, \quad (54)$$

$$A_{21,x_0} = \begin{pmatrix} p(w_2 | u_1) & p(w_2 | u_2) \\ p(w_3 | u_1) & p(w_3 | u_2) \end{pmatrix}, \quad (55)$$

$$A_{22,x_0} = \begin{pmatrix} p(w_2 | u_3) & p(w_2 | u_4) \\ p(w_3 | u_3) & p(w_3 | u_4) \end{pmatrix}, \quad (56)$$

$$A_{11,x_1} = \begin{pmatrix} 1 & 1 \\ p(w_1 | u_1) & p(w_1 | u_3) \end{pmatrix}, \quad (57)$$

$$A_{12,x_1} = \begin{pmatrix} 1 & 1 \\ p(w_1 | u_2) & p(w_1 | u_4) \end{pmatrix}, \quad (58)$$

$$A_{21,x_1} = \begin{pmatrix} p(w_2 | u_1) & p(w_2 | u_3) \\ p(w_3 | u_1) & p(w_3 | u_3) \end{pmatrix}, \quad (59)$$

$$A_{22,x_1} = \begin{pmatrix} p(w_2 | u_2) & p(w_2 | u_4) \\ p(w_3 | u_2) & p(w_3 | u_4) \end{pmatrix} \quad (60)$$

である。くわえて、

$$B_{x_0} = \begin{pmatrix} 1 & p(z_0 | x_0, u_1) \\ 1 & p(z_0 | x_0, u_2) \\ 1 & p(z_0 | x_0, u_3) \\ 1 & p(z_0 | x_0, u_4) \end{pmatrix}, B_{x_1} = \begin{pmatrix} 1 & p(z_0 | x_1, u_1) \\ 1 & p(z_0 | x_1, u_3) \\ 1 & p(z_0 | x_1, u_2) \\ 1 & p(z_0 | x_1, u_4) \end{pmatrix}, \quad (61)$$

$$\Delta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (62)$$

$$M_{x_0} = \begin{pmatrix} p(u_1 | x_0) & 0 & 0 & 0 \\ 0 & p(u_2 | x_0) & 0 & 0 \\ 0 & 0 & p(u_3 | x_0) & 0 \\ 0 & 0 & 0 & p(u_4 | x_0) \end{pmatrix}, \quad (63)$$

$$M_{x_1} = \begin{pmatrix} p(u_1 | x_1) & 0 & 0 & 0 \\ 0 & p(u_3 | x_1) & 0 & 0 \\ 0 & 0 & p(u_2 | x_1) & 0 \\ 0 & 0 & 0 & p(u_4 | x_1) \end{pmatrix} \quad (64)$$

とすると,

$$P_x = A_x^T M_x B_x, \quad Q_x = A_x^T M_x \Delta B_x \quad (65)$$

が任意の $x \in \{x_0, x_1\}$ に対して成立する. ここで, “T” は転置行列/ベクトルを表している. 4×4 のブロック行列 $[P_x; Q_x]$ が条件 (3) より可逆であるので A_x も任意の $x \in \{x_0, x_1\}$ に対して可逆である. ここで

$$[P_x; Q_x] = A_x^T M_x [B_x; M_x B_x] \quad (66)$$

であり, 結合行列 $[B_x; M_x B_x]$ は (4×4) のブロック行列である. くわえて,

$$P_{2,x} - Q_{2,x} = A_{22,x} M_{22,x} B_{2,x}, \quad Q_{1,x} = A_{11,x} M_{11,x} B_{1,x} \quad (67)$$

が条件 (3) より可逆なので $A_{11,x}$ と $A_{22,x}$ のどちらも任意の $x \in \{x_0, x_1\}$ に対して可逆である. したがって, 任意の $x \in \{x_0, x_1\}$ に対して仮定より

$$Q_x = A_x^T \Delta A_x^{-T} P_x \quad (68)$$

が成立し, 任意の $x \in \{x_0, x_1\}$ に対して $A_x^{-T} = (A_x^T)^{-1}$ は

$$A_x^{-T} = \begin{pmatrix} A_x^{11} & -A_x^{11} A_{12,x} A_{22,x}^{-1} \\ -A_{22,x}^{-1} A_{21,x} A_x^{11} & A_x^{22} \end{pmatrix} \quad (69)$$

とかける。ただし,

$$A_x^{11} = (A_{11,x} - A_{12,x}A_{22,x}^{-1}A_{21,x})^{-1}, \quad A_x^{22} = (A_{22,x} - A_{21,x}A_{11,x}^{-1}A_{12,x})^{-1} \quad (70)$$

である。くわえて,

$$A_x^T \Delta A_x^{-T} = \begin{pmatrix} A_{11,x} A_x^{11} \\ A_{21,x} A_x^{11} \end{pmatrix} [I_{2,2}; -A_{12,x} A_{22,x}^{-1}] \quad (71)$$

が成立する。ここで, $I_{2,2}$ は 2×2 の単位行列である。また, $[I_{2,2}; -A_{12,x} A_{22,x}^{-1}]$ は (2×4) の結合行列である。すると,

$$Q_{1,x} = A_{11,x} A_x^{11} (P_{1,x} - A_{12,x} A_{22,x}^{-1} P_{2,x}), \quad (72)$$

$$Q_{2,x} = A_{21,x} A_x^{11} (P_{1,x} - A_{12,x} A_{22,x}^{-1} P_{2,x}) \quad (73)$$

が成立する。これらの式から $Q_{2,x} = A_{21,x} A_{11,x}^{-1} Q_{1,x}$ と

$$A_{21,x} = Q_{2,x} Q_{1,x}^{-1} A_{11,x} \quad (74)$$

と

$$A_x^{11,-1} A_{11,x}^{-1} Q_{1,x} = Q_{1,x} - A_{12,x} A_{22,x}^{-1} A_{21,x} A_{11,x}^{-1} Q_{1,x} \quad (75)$$

$$= P_{1,x} - A_{12,x} A_{22,x}^{-1} P_{2,x} \quad (76)$$

が成立し。くわえて, 等式

$$Q_{1,x} - P_{1,x} = A_{12,x} A_{22,x}^{-1} (A_{21,x} A_{11,x}^{-1} Q_{1,x} - P_{2,x}) \quad (77)$$

$$= A_{12,x} A_{22,x}^{-1} (Q_{2,x} - P_{2,x}) \quad (78)$$

と

$$A_{22,x} = (P_{2,x} - Q_{2,x})(P_{1,x} - Q_{1,x})^{-1} A_{12,x} \quad (79)$$

が成立する。ここで, 任意の $x \in \{x_0, x_1\}$ と $i = 1, 2$ と $j = 1, 2$ に対して, $a_{i,j}^x$ と $b_{i,j}^x$ を観測可能な行列 $Q_{2,x} Q_{1,x}^{-1}$ と $(P_{2,x} - Q_{2,x})(P_{1,x} - Q_{1,x})^{-1}$ の第 (i, j) 成分とすると,

$$a_{11}^x = \frac{1}{\det Q_{1,x}} \{p(y_0, w_2|x)p(y_0, w_1, z_0|x) - p(y_0, w_2, z_0|x)p(y_0, z_0|x)\}, \quad (80)$$

$$a_{12}^x = \frac{1}{\det Q_{1,x}} \{p(y_0, w_2, z_0|x)p(y_0|x) - p(y_0, w_2|x)p(y_0, w_1|x)\}, \quad (81)$$

$$a_{21}^x = \frac{1}{\det Q_{1,x}} \{p(y_0, w_3|x)p(y_0, w_1, z_0|x) - p(y_0, w_3, z_0|x)p(y_0, z_0|x)\}, \quad (82)$$

$$a_{22}^x = \frac{1}{\det Q_{1,x}} \{p(y_0, w_3, z_0|x)p(y_0|x) - p(y_0, w_3|x)p(y_0, w_1|x)\}, \quad (83)$$

$$b_{11}^x = \frac{1}{\det(P_{1,x} - Q_{1,x})} \{p(y_1, w_2|x)p(y_1, w_1, z_0|x) - p(y_1, w_2, z_0|x)p(y_1, z_0|x)\}, \quad (84)$$

$$b_{12}^x = \frac{1}{\det(P_{1,x} - Q_{1,x})} \{p(y_1, w_2, z_0|x)p(y_1|x) - p(y_1, w_2|x)p(y_1, w_1|x)\}, \quad (85)$$

$$b_{21}^x = \frac{1}{\det(P_{1,x} - Q_{1,x})} \{p(y_1, w_3|x)p(y_1, w_1, z_0|x) - p(y_1, w_3, z_0|x)p(y_1, z_0|x)\}, \quad (86)$$

$$b_{22}^x = \frac{1}{\det(P_{1,x} - Q_{1,x})} \{p(y_1, w_3, z_0|x)p(y_1|x) - p(y_1, w_3|x)p(y_1, w_1|x)\} \quad (87)$$

とかけ,

$$\begin{aligned} Q_{2,x_0} Q_{1,x_0}^{-1} A_{11,x_0} &= \begin{pmatrix} a_{11}^{x_0} + a_{12}^{x_0}p(w_1 | u_1) & a_{11}^{x_0} + a_{12}^{x_0}p(w_1 | u_2) \\ a_{21}^{x_0} + a_{22}^{x_0}p(w_1 | u_1) & a_{21}^{x_0} + a_{22}^{x_0}p(w_1 | u_2) \end{pmatrix} \\ &= A_{21,x_0} = \begin{pmatrix} p(w_2 | u_1) & p(w_2 | u_2) \\ p(w_3 | u_1) & p(w_3 | u_2) \end{pmatrix}, \end{aligned} \quad (88)$$

$$\begin{aligned} Q_{2,x_1} Q_{1,x_1}^{-1} A_{11,x_1} &= \begin{pmatrix} a_{11}^{x_1} + a_{12}^{x_1}p(w_1 | u_1) & a_{11}^{x_1} + a_{12}^{x_1}p(w_1 | u_3) \\ a_{21}^{x_1} + a_{22}^{x_1}p(w_1 | u_1) & a_{21}^{x_1} + a_{22}^{x_1}p(w_1 | u_3) \end{pmatrix} \\ &= A_{21,x_1} = \begin{pmatrix} p(w_2 | u_1) & p(w_2 | u_3) \\ p(w_3 | u_1) & p(w_3 | u_3) \end{pmatrix}, \end{aligned} \quad (89)$$

$$\begin{aligned} A_{22,x_0} &= \begin{pmatrix} p(w_2 | u_3) & p(w_2 | u_4) \\ p(w_3 | u_3) & p(w_3 | u_4) \end{pmatrix} = (P_{2,x_0} - Q_{2,x_0})(P_{1,x_0} - Q_{1,x_0})^{-1} A_{12,x_0} \\ &= \begin{pmatrix} b_{11}^{x_0} + b_{12}^{x_0}p(w_1 | u_3) & b_{11}^{x_0} + b_{12}^{x_0}p(w_1 | u_4) \\ b_{21}^{x_0} + b_{22}^{x_0}p(w_1 | u_3) & b_{21}^{x_0} + b_{22}^{x_0}p(w_1 | u_4) \end{pmatrix}, \end{aligned} \quad (90)$$

$$A_{22,x_1} = \begin{pmatrix} p(w_2 | u_2) & p(w_2 | u_4) \\ p(w_3 | u_2) & p(w_3 | u_4) \end{pmatrix} = (P_{2,x_1} - Q_{2,x_1})(P_{1,x_1} - Q_{1,x_1})^{-1} A_{12,x_1}$$

$$= \begin{pmatrix} b_{11}^{x_1} + b_{12}^{x_1}p(w_1 | u_2) & b_{11}^{x_1} + b_{12}^{x_1}p(w_1 | u_4) \\ b_{21}^{x_1} + b_{22}^{x_1}p(w_1 | u_2) & b_{21}^{x_1} + b_{22}^{x_1}p(w_1 | u_4) \end{pmatrix} \quad (91)$$

が成立する。これらの式から $p(w_1|u)$ は以下のように与えられる。

$$p(w_1|u_1) = -\frac{a_{11}^{x_0} - a_{11}^{x_1}}{a_{12}^{x_0} - a_{12}^{x_1}}, \quad p(w_1|u_1) = -\frac{a_{21}^{x_0} - a_{21}^{x_1}}{a_{22}^{x_0} - a_{22}^{x_1}}, \quad (92)$$

$$p(w_1|u_2) = -\frac{a_{11}^{x_0} - b_{11}^{x_1}}{a_{12}^{x_0} - b_{12}^{x_1}}, \quad p(w_1|u_2) = -\frac{a_{21}^{x_0} - b_{21}^{x_1}}{a_{22}^{x_0} - b_{22}^{x_1}}, \quad (93)$$

$$p(w_1|u_3) = -\frac{a_{11}^{x_1} - b_{11}^{x_0}}{a_{12}^{x_1} - b_{12}^{x_0}}, \quad p(w_1|u_3) = -\frac{a_{21}^{x_1} - b_{21}^{x_0}}{a_{22}^{x_1} - b_{22}^{x_0}}, \quad (94)$$

$$p(w_1|u_4) = -\frac{b_{11}^{x_0} - b_{11}^{x_1}}{b_{12}^{x_0} - b_{12}^{x_1}}, \quad p(w_1|u_4) = -\frac{b_{21}^{x_0} - b_{21}^{x_1}}{b_{22}^{x_0} - b_{22}^{x_1}} \quad (95)$$

さらに、 $p(w_2|u)$ と $p(w_3|u)$ も

$$a_{11}^{x_0} + a_{12}^{x_0}p(w_1|u_1) = p(w_2|u_1), \quad a_{11}^{x_0} + a_{12}^{x_0}p(w_1|u_2) = p(w_2|u_2), \quad (96)$$

$$a_{21}^{x_0} + a_{22}^{x_0}p(w_1|u_1) = p(w_3|u_1), \quad a_{21}^{x_0} + a_{22}^{x_0}p(w_1|u_2) = p(w_3|u_2), \quad (97)$$

$$a_{11}^{x_1} + a_{12}^{x_1}p(w_1|u_1) = p(w_2|u_1), \quad a_{11}^{x_1} + a_{12}^{x_1}p(w_1|u_3) = p(w_2|u_3), \quad (98)$$

$$a_{21}^{x_1} + a_{22}^{x_1}p(w_1|u_1) = p(w_3|u_1), \quad a_{21}^{x_1} + a_{22}^{x_1}p(w_1|u_3) = p(w_3|u_3), \quad (99)$$

$$b_{11}^{x_0} + b_{12}^{x_0}p(w_1|u_3) = p(w_2|u_3), \quad b_{11}^{x_0} + b_{12}^{x_0}p(w_1|u_4) = p(w_2|u_4), \quad (100)$$

$$b_{21}^{x_0} + b_{22}^{x_0}p(w_1|u_3) = p(w_3|u_3), \quad b_{21}^{x_0} + b_{22}^{x_0}p(w_1|u_4) = p(w_3|u_4), \quad (101)$$

$$b_{11}^{x_1} + b_{12}^{x_1}p(w_1|u_2) = p(w_2|u_2), \quad b_{11}^{x_1} + b_{12}^{x_1}p(w_1|u_4) = p(w_2|u_4), \quad (102)$$

$$b_{21}^{x_1} + b_{22}^{x_1}p(w_1|u_2) = p(w_3|u_2), \quad b_{21}^{x_1} + b_{22}^{x_1}p(w_1|u_4) = p(w_3|u_4) \quad (103)$$

のように与えられる。以上の議論により、条件付き確率 $p(w|u)$ は識別可能である。

くわえて、 A_x は任意の $x \in \{x_0, x_1\}$ に対して識別可能であるため、

$$A_x^{-T} P_x = M_x B_x \quad (104)$$

の第一列を比べることで M_{x_0} と M_{x_1} も識別可能である。さらに、 $p(x_0)$ と $p(x_1)$ のどちらも観測分布から推定可能なので、潜在反応変数の同時分布は任意の $u \in \{u_1, u_2, u_3, u_4\}$ に対して、

$$\sum_x p(u|x, z)p(x) = p(u) \quad (105)$$

をとおして識別可能である。

□

定理 3.2.2 の証明

証明. 潜在反応変数の一致性の性質から,

$$p(y, w, v_i | x_0, z_0) = 0, \quad i = 3, 4, \quad (106)$$

$$p(y, w, v_i | x_1, z_0) = 0, \quad i = 1, 2, \quad (107)$$

$$p(y, w, v_i | x_0, z_1) = 0, \quad i = 2, 4, \quad (108)$$

$$p(y, w, v_i | x_1, z_1) = 0, \quad i = 1, 3 \quad (109)$$

が任意の $y \in \{y_0, y_1\}$ と $w \in \{w_1, w_2, w_3\}$ に対して成立する. これらの等式から, 任意の $y \in \{y_0, y_1\}$ と $w \in \{w_1, w_2, w_3\}$ に対して, $p(y, w | x, z)$ は

$$p(y, w | x_0, z_0) = \sum_{i=1,2} p(y | x_0, v_i)p(w | v_i)p(v_i | x_0, z_0), \quad (110)$$

$$p(y, w | x_1, z_0) = \sum_{i=3,4} p(y | x_1, v_i)p(w | v_i)p(v_i | x_1, z_0), \quad (111)$$

$$p(y, w | x_0, z_1) = \sum_{i=1,3} p(y | x_0, v_i)p(w | v_i)p(v_i | x_0, z_1), \quad (112)$$

$$p(y, w | x_1, z_1) = \sum_{i=2,4} p(y | x_1, v_i)p(w | v_i)p(v_i | x_1, z_1), \quad (113)$$

と書ける. ここで,

$$P_{x,z}^w = \begin{pmatrix} 1 & p(y | x, z) \\ p(w | x, z) & p(y, w | x, z) \end{pmatrix}, \quad (114)$$

$$A_{x_0,z_0}^w = \begin{pmatrix} 1 & p(w | v_1) \\ 1 & p(w | v_2) \end{pmatrix}, \quad A_{x_1,z_0}^w = \begin{pmatrix} 1 & p(w | v_3) \\ 1 & p(w | v_4) \end{pmatrix}, \quad (115)$$

$$A_{x_0,z_1}^w = \begin{pmatrix} 1 & p(w | v_1) \\ 1 & p(w | v_3) \end{pmatrix}, \quad A_{x_1,z_1}^w = \begin{pmatrix} 1 & p(w | v_2) \\ 1 & p(w | v_4) \end{pmatrix}, \quad (116)$$

$$B_{x_0, z_0} = \begin{pmatrix} 1 & p(y | x_0, v_1) \\ 1 & p(y | x_0, v_2) \end{pmatrix}, \quad B_{x_1, z_0} = \begin{pmatrix} 1 & p(y | x_1, v_3) \\ 1 & p(y | x_1, v_4) \end{pmatrix}, \quad (117)$$

$$B_{x_0, z_1} = \begin{pmatrix} 1 & p(y | x_0, v_1) \\ 1 & p(y | x_0, v_3) \end{pmatrix}, \quad B_{x_1, z_1} = \begin{pmatrix} 1 & p(y | x_1, v_2) \\ 1 & p(y | x_1, v_4) \end{pmatrix}, \quad (118)$$

$$M_{x_0, z_0} = \begin{pmatrix} p(v_1 | x_0, z_0) & 0 \\ 0 & p(v_2 | x_0, z_0) \end{pmatrix}, \quad (119)$$

$$M_{x_1, z_0} = \begin{pmatrix} p(v_3 | x_1, z_0) & 0 \\ 0 & p(v_4 | x_1, z_0) \end{pmatrix}, \quad (120)$$

$$M_{x_0, z_1} = \begin{pmatrix} p(v_1 | x_0, z_1) & 0 \\ 0 & p(v_3 | x_0, z_1) \end{pmatrix}, \quad (121)$$

$$M_{x_1, z_1} = \begin{pmatrix} p(v_2 | x_1, z_1) & 0 \\ 0 & p(v_4 | x_1, z_1) \end{pmatrix} \quad (122)$$

と定義する. すると,

$$P_{x,z}^w = (A_{x,z}^w)^T M_{x,z} B_{x,z} \quad (123)$$

が任意の $x \in \{x_0, x_1\}$, $z \in \{z_0, z_1\}$, $w \in \{w_1, w_2, w_3\}$ に対して成立する. すると, 任意の $w, w' \in \{w_1, w_2, w_3\}$ に対して, 行列 $P_{x,z}^w$ と $P_{x,z}^{w'}$ は可逆である. また,

$$P_{x,z}^{w'} (P_{x,z}^w)^{-1} (A_{x,z}^w)^T = (A_{x,z}^{w'})^T \quad (124)$$

が成立する. 式を書き下すと,

$$\begin{aligned} & \frac{1}{\det P_{x,z}^w} \begin{pmatrix} 1 & p(y|x, z) \\ p(w'|x, z) & p(y, w'|x, z) \end{pmatrix} \\ & \times \begin{pmatrix} p(y, w|x, z) & -p(y|x, z) \\ -p(w|x, z) & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ p(w|v) & p(w|v') \end{pmatrix} \end{aligned} \quad (125)$$

$$= \begin{pmatrix} 1 & 1 \\ p(w'|v) & p(w'|v') \end{pmatrix} \quad (126)$$

と

$$\det(P_{x,z}^w)p(w'|x, z) - \det(P_{x,z}^{w'})p(w|x, z) + \det(P_{x,z}^{w'})p(w|v) \quad (127)$$

$$= \det(P_{x,z}^w)p(w'|v) \quad (128)$$

が成立する。これは

$$\det(P_{x,z}^{w'})(p(w|v) - p(w|x, z)) = \det(P_{x,z}^w)(p(w'|v) - p(w'|x, z)) \quad (129)$$

と等しい。よって、 $x = x_0$ と $z = z_0$ に対して、

$$\begin{aligned} p(w_1|v_1) &= \frac{\det(P_{x_0, z_0}^{w_1})}{\det(P_{x_0, z_0}^{w_2})}(p(w_2|v_1) - p(w_2|x_0, z_0)) + p(w_1|x_0, z_0) \\ &= \frac{\det(P_{x_0, z_0}^{w_1})}{\det(P_{x_0, z_0}^{w_3})}(p(w_3|v_1) - p(w_3|x_0, z_0)) + p(w_1|x_0, z_0), \end{aligned} \quad (130)$$

$$\begin{aligned} p(w_1|v_2) &= \frac{\det(P_{x_0, z_0}^{w_1})}{\det(P_{x_0, z_0}^{w_2})}(p(w_2|v_2) - p(w_2|x_0, z_0)) + p(w_1|x_0, z_0) \\ &= \frac{\det(P_{x_0, z_0}^{w_1})}{\det(P_{x_0, z_0}^{w_3})}(p(w_3|v_2) - p(w_3|x_0, z_0)) + p(w_1|x_0, z_0) \end{aligned} \quad (131)$$

が、 $x = x_0$ と $z = z_1$ に対して、

$$\begin{aligned} p(w_1|v_1) &= \frac{\det(P_{x_0, z_1}^{w_1})}{\det(P_{x_0, z_1}^{w_2})}(p(w_2|v_1) - p(w_2|x_0, z_1)) + p(w_1|x_0, z_1) \\ &= \frac{\det(P_{x_0, z_1}^{w_1})}{\det(P_{x_0, z_1}^{w_3})}(p(w_3|v_1) - p(w_3|x_0, z_1)) + p(w_1|x_0, z_1), \end{aligned} \quad (132)$$

$$\begin{aligned} p(w_1|v_3) &= \frac{\det(P_{x_0, z_1}^{w_1})}{\det(P_{x_0, z_1}^{w_2})}(p(w_2|v_3) - p(w_2|x_0, z_1)) + p(w_1|x_0, z_1) \\ &= \frac{\det(P_{x_0, z_1}^{w_1})}{\det(P_{x_0, z_1}^{w_3})}(p(w_3|v_3) - p(w_3|x_0, z_1)) + p(w_1|x_0, z_1) \end{aligned} \quad (133)$$

が、 $x = x_1$ と $z = z_0$ に対して、

$$\begin{aligned} p(w_1|v_3) &= \frac{\det(P_{x_1, z_0}^{w_1})}{\det(P_{x_1, z_0}^{w_2})}(p(w_2|v_3) - p(w_2|x_1, z_0)) + p(w_1|x_1, z_0) \\ &= \frac{\det(P_{x_1, z_0}^{w_1})}{\det(P_{x_1, z_0}^{w_3})}(p(w_3|v_3) - p(w_3|x_0, z_1)) + p(w_1|x_1, z_0), \end{aligned} \quad (134)$$

$$\begin{aligned}
p(w_1|v_4) &= \frac{\det(P_{x_1, z_0}^{w_1})}{\det(P_{x_1, z_0}^{w_2})} (p(w_2|v_4) - p(w_2|x_1, z_0)) + p(w_1|x_1, z_0) \\
&= \frac{\det(P_{x_1, z_0}^{w_1})}{\det(P_{x_1, z_0}^{w_3})} (p(w_3|v_4) - p(w_3|x_1, z_0)) + p(w_1|x_1, z_0) \quad (135)
\end{aligned}$$

が、 $x = x_1$ と $z = z_1$ に対して、

$$\begin{aligned}
p(w_1|v_2) &= \frac{\det(P_{x_1, z_1}^{w_1})}{\det(P_{x_1, z_1}^{w_2})} (p(w_2|v_2) - p(w_2|x_1, z_1)) + p(w_1|x_1, z_1) \\
&= \frac{\det(P_{x_1, z_1}^{w_1})}{\det(P_{x_1, z_1}^{w_3})} (p(w_3|v_2) - p(w_3|x_1, z_1)) + p(w_1|x_1, z_1), \quad (136)
\end{aligned}$$

$$\begin{aligned}
p(w_1|v_4) &= \frac{\det(P_{x_1, z_1}^{w_1})}{\det(P_{x_1, z_1}^{w_2})} (p(w_2|v_4) - p(w_2|x_1, z_1)) + p(w_1|x_1, z_1) \\
&= \frac{\det(P_{x_1, z_1}^{w_1})}{\det(P_{x_1, z_1}^{w_3})} (p(w_3|v_4) - p(w_3|x_1, z_1)) + p(w_1|x_1, z_1) \quad (137)
\end{aligned}$$

がそれぞれ成立する。すると、任意の $w \neq w_1$, $(x, z) \neq (x', z')$ ($x, x' \in \{x_0, x_1\}$, $z, z' \in \{z_0, z_1\}$, $w \in \{w_2, w_3\}$) に対して、

$$\frac{\det(P_{x, z}^{w_1})}{\det(P_{x, z}^w)} \neq \frac{\det(P_{x', z'}^{w_1})}{\det(P_{x', z'}^w)} \quad (138)$$

が成り立つため、任意の $w \in \{w_1, w_2, w_3\}$ と $v \in \{v_1, v_2, v_3, v_4\}$ に対して、確率 $p(w|v)$ は識別可能である。次に、

$$(A_{x, z}^w)^{-T} P_{x, z}^w = M_{x, z} B_{x, z} = \begin{pmatrix} p(v_i|x, z) & p(y, v_i|x, z) \\ p(v_j|x, z) & p(y, v_j|x, z) \end{pmatrix} \quad (139)$$

が成立する。 $P_{x, z}^w$ と $A_{x, z}^w$ は観測分布から識別可能なので式の両辺の第一列を比較することで $p(v_i|x, z)$ と $p(v_j|x, z)$ は識別可能である。さらに、 $p(x_0, z_0)$, $p(x_0, z_1)$, $p(x_1, z_0)$, $p(x_1, z_1)$ は観測分布から識別可能なので、任意の $v \in \{v_1, v_2, v_3, v_4\}$ に対して、潜在反応変数の同時分布は

$$\sum_{x, z} p(v|x, z)p(x, z) = p(v) \quad (140)$$

によって計算できる。

□

第3部の付録

命題 4.2.1 の証明

証明. はじめに, 問題 $\mathcal{K}f = g$ が良設定問題であるとは以下の3つの性質を満たす場合のことをいう.

1. 解が存在している.
2. その解が一意である.
3. 入力 \mathcal{K} と g の変化に対して解が連続的に変化する.

これらのうち1つでも条件が成立しないような問題は不良設定問題と呼ばれる [Tikhonov et al., 1995]. 最初に, この積分方程式は Wong [2022] の結果から唯一解を持ち1つ目と2つ目の条件を満たす. 仮定 4.1.1, 4.1.2, 4.2.1 より, 作用素 \mathcal{K} は有界作用素である. つぎに, 有界作用素は連続性を満たすことを示す. 任意の関数 $f, f^* \in \mathcal{H}$ ($f \neq f^*$) に対してコーシー・シュワルツの不等式から,

$$\|\mathcal{K}(f) - \mathcal{K}(f^*)\|^2 \quad (141)$$

$$= \|\mathcal{K}(f - f^*)\|^2 \quad (142)$$

$$= \int_{\Omega_Z} \left(\int_{\Omega_X} k(x', x) \{f(x') - f^*(x')\} dx' \right)^2 dx \quad (143)$$

が成立し,

$$\leq \int_{\Omega_Z} \left(\int_{\Omega_X} k(x', x)^2 dx' \right) \left(\int_{\Omega_X} \{f(x') - f^*(x')\}^2 dx' \right) dx \quad (144)$$

$$= \left(\int_{\Omega_Z} \int_{\Omega_X} k(x', x)^2 dx' dx \right) \left(\int_{\Omega_X} \{f(x') - f^*(x')\}^2 dx' \right) \quad (145)$$

$$\leq \left(\int_{\Omega_Z} \int_{\Omega_X} k(x', x)^2 dx' dx \right) \|f - f^*\|^2 < \infty \quad (146)$$

となる. また,

$$\|\mathcal{K}(f) - \mathcal{K}(f^*)\| \leq \sqrt{\left(\int_{\Omega_Z} \int_{\Omega_X} k(x', x)^2 dx' dx \right)} \|f - f^*\| \quad (147)$$

なので \mathcal{K} は連続作用素である。また、開写像定理から逆作用素 \mathcal{K}^{-1} もまた連続作用素である。ここで、 $\mathcal{K}^{-1} = \alpha \sum_{t=1}^{\infty} (\mathcal{I} - \alpha \mathcal{K})^t$ である。ただし、 \mathcal{I} は単位作用素であり $0 < \alpha < 2/\|\mathcal{K}\|$ を満たしている。さらに、任意の $\mathcal{K}, \mathcal{K}^*$ ($\mathcal{K} \neq \mathcal{K}^*$) と g, g^* ($g \neq g^*$) に対して、

$$\|\mathcal{K}^{-1}(g) - \mathcal{K}^{*-1}(g^*)\| \quad (148)$$

$$\leq \|\mathcal{K}^{-1}(g) - \mathcal{K}^{*-1}(g)\| + \|\mathcal{K}^{*-1}(g) - \mathcal{K}^{*-1}(g^*)\| \quad (149)$$

$$\leq \|\mathcal{K}^{-1}(g) - \mathcal{K}^{*-1}(g)\| + \|\mathcal{K}^{*-1}\| \|g - g^*\| \quad (150)$$

が成立する。よって、関数 k は連続なので入力関数 \mathcal{K} と g の変化に応じて解が連続的に変化する。 \square

数値積分の詳細

$\mathcal{I}[a(x); \mathcal{X}]$ を \mathcal{X} が与えられたときの積分 $\int_{\Omega_X} a(x) dx$ の数値積分を表している。これは

$$\mathcal{I}[a(x); \mathcal{X}] = \sum_{q=1}^R I(x_q, x_{q+1})(x_{q+1} - x_q) \quad (151)$$

の形で表すことができる。たとえば、左点則

$$\mathcal{I}[a(x); \mathcal{X}] = \sum_{q=1}^R a(x_q)(x_{q+1} - x_q) \quad (152)$$

や、中点則

$$\mathcal{I}[a(x); \mathcal{X}] = \sum_{q=1}^R a\left(\frac{x_q + x_{q+1}}{2}\right)(x_{q+1} - x_q) \quad (153)$$

や、台形則

$$\mathcal{I}[a(x); \mathcal{X}] = \sum_{q=1}^R \frac{a(x_q) + a(x_{q+1})}{2}(x_{q+1} - x_q) \quad (154)$$

が存在している。中点則の誤差はせいぜい

$$(x_R - x_0)^3 B / 24R^2 \quad (155)$$

である。ここで、 B は関数 $a(x)$ の二階微分の上界であり、 $x_{r+1} - x_r$ は任意の $r = 1, \dots, R$ に対して等間隔である。また、数値積分の誤差は $R \rightarrow \infty$ としたときに 0 に収束し、収束のオーダーは $\mathcal{O}(R^{-2})$ である。

数値補間の詳細

つぎに、数値補間について説明する。集合 \mathcal{X} が与えられたときの関数 $\hat{\theta}$ の数値補間は次のような線形和で表現できる。

$$\hat{\theta}(x) = \sum_{r=1}^R w_r l_r(x) \quad (156)$$

たとえば、ラグランジュ多項式補間

$$l_r(x) = \frac{x - x_0}{x_r - x_0} \cdots \frac{x - x_{r-1}}{x_r - x_{r-1}} \frac{x - x_{r+1}}{x_r - x_{r+1}} \cdots \frac{x - x_R}{x_r - x_R} \quad (157)$$

$$= \prod_{0 \leq m \leq R, m \neq r} \frac{x - x_m}{x_r - x_m} \quad (158)$$

がある。係数 w_1, \dots, w_R は

$$\hat{\theta}(x_r) = \sum_{r=1}^R w_r l_r(x_r), \quad r = 1, \dots, R \quad (159)$$

を解くことで求めることができる。関数 $\theta(x)$ をラグランジュ多項式補間によって補間するときの誤差の上界は

$$|\theta(x) - \hat{\theta}(x)| \leq \frac{Ch^R}{4R} \quad (160)$$

で抑えられる。ここで、

$$C = \max_{x \in [x_0, x_R]} \left| \frac{d^R}{d^R x} \theta(x) \right| \quad (161)$$

と $h = \max_{r=0, \dots, R-1} |x_{r+1} - x_r|$ が成立する。誤差は $R \rightarrow \infty$ と

$$\lim_{R \rightarrow \infty} |x_{r+1} - x_r| = 0 \quad (162)$$

($r = 1, \dots, R$) によって 0 に収束する。また、 h は w/R とかけるので誤差は $\mathcal{O}(R^{-R})$ のオーダーで収束する。

補題 4.2.1 の証明

証明. はじめに, 積分方程式 (4.11)

$$\begin{aligned} & \mathbb{E}[Y|Z = z_0] - \mathbb{E}[Y|Z = z] \\ &= \int_{\Omega_X} \{\mathbb{P}_X[x|Z = z] - \mathbb{P}_X[x|Z = z_0]\} \mathbb{E}[\partial_x Y_x] dx \end{aligned} \quad (163)$$

を考える. 作用素 \mathcal{K} は任意の $f \in \mathcal{H}$ に対して,

$$\mathcal{K}(f) = \int_{\Omega_X} \{\mathbb{P}_X[x|Z = z] - \mathbb{P}_X[x|Z = z_0]\} f(x) dx \quad (164)$$

であり,

$$\int_{\Omega_Z} \int_{\Omega_X} |\{\mathbb{P}_X[x|Z = z] - \mathbb{P}_X[x|Z = z_0]\}|^2 dx dz < \infty \quad (165)$$

を満たしている. したがって, \mathcal{K} はコンパクトな積分核である [Alexanderian, 2013].

つぎに,

$$\begin{aligned} & \langle \mathcal{K}(f), g \rangle \\ &= \int_{\Omega_Z} \left(\int_{\Omega_X} \{\mathbb{P}_X[x|Z = z] - \mathbb{P}_X[x|Z = z_0]\} f(x) dx \right) g(z) dz \quad (166) \\ &= \int_{\Omega_Z} f(x) \left(\int_{\Omega_X} \{\mathbb{P}_X[x|Z = z] - \mathbb{P}_X[x|Z = z_0]\} g(z) dz \right) dx \\ &= \langle f, \mathcal{K}(g) \rangle \end{aligned} \quad (167)$$

であるので作用素 \mathcal{K} は自己共役である. さいごに, 任意の $x \in X$ に対して $\{\mathbb{P}_X[x|Z = z] - \mathbb{P}_X[x|Z = z_0]\} > 0$ が成り立つので,

$$\begin{aligned} & \langle \mathcal{K}(f), f \rangle \\ &= \int_{\Omega_Z} \left(\int_{\Omega_X} \{\mathbb{P}_X[x|Z = z] - \mathbb{P}_X[x|Z = z_0]\} f(x) dx \right) f(z) dz \end{aligned}$$

となり. 積分作用素 \mathcal{K} は補題 4.2.1 の 3 つの性質を満たしている. \square

補題 4.2.2 の証明

証明. 仮定 4.1.1 から積分方程式 (4.11) の解として平均偏因果効果は存在している.

ピカールの条件は解の存在の必要十分条件なので \mathcal{K} はピカールの条件を満たす. \square

定理 4.2.1 の証明

証明. はじめに, Diaz and Metcalf [1970] によって与えられている定理を以下に示す.

Diaz and Metcalf [1970] の定理 (c_1). 作用素 \mathcal{L} はコンパクト, 自己共役, かつ半正定値とする. \mathcal{L} がピカールの条件を満たすことは, ピカールの反復法が強収束することの必要十分条件である.

補題 4.2.1 によって作用素 \mathcal{K} がコンパクト, 自己共役, かつ, 半正定値であるので補題 4.2.2 から作用素 \mathcal{K} はピカールの条件を満たす. したがって, ピカールの反復法 $\{\theta_t\}_{t=0}^\infty$ は強収束する. また, 解の一意性より収束先は平均偏因果効果となる.

□

定理 4.2.2 の証明

証明. 関数 $\hat{\mu}(x)$ と $\hat{k}(x', x)$ は

$$\begin{cases} \hat{\mu}(x) = \mu(x) + e(x) \\ \hat{k}(x', x) = k(x', x) + \epsilon(x', x) \end{cases} \quad (168)$$

のようにかける. ここで, 関数 $e(x)$ と $\epsilon(x', x)$ は誤差項である. 簡潔に,

$$\begin{cases} \hat{\mu} = \mu + e \\ \hat{k} = k + \epsilon \end{cases} \quad (169)$$

のように表現できる. 数値積分のために, 任意の $q = 1, \dots, Q$ に対して, $[x_q, x_{q+1}]$ が $[x_0, x_Q] = [\min(\Omega_X), \max(\Omega_X)]$ と $\lim_{Q \rightarrow \infty} |x_{q+1} - x_q| = 0$ を満たすように選ぶ. かわえて, 数値補間のために, 任意の $r = 1, \dots, R$ に対して, $[x_r, x_{r+1}]$ が $[x_0, x_R] = [\min(\Omega_X), \max(\Omega_X)]$ と $\lim_{R \rightarrow \infty} |x_{r+1} - x_r| = 0$ を満たすように選ぶ. すると, ピカールの反復法は機械学習の誤差を伴って以下のように書き下せる. 関数 $\hat{\theta}$ は $X = x_r$ において,

$$\hat{\theta}_{t+1}(x_r) = \hat{\theta}_t(x_r) + \alpha \left(\hat{\mu}(x_r) - \mathcal{I}[\hat{k}(x', x) \hat{\theta}_t(x'); \mathcal{X}] \right) \quad (170)$$

$$\begin{aligned}\hat{\theta}_{t+1}(x_r) &= \hat{\theta}_t(x_r) + \alpha \left(\mu(x_r) + e(x_r) - \mathcal{I}[\{k(x', x_r) + \epsilon(x', x_r)\} \hat{\theta}_t(x'); \mathcal{X}] \right) \\ \hat{\theta}_{t+1}(x_r) &= \\ & \hat{\theta}_t(x_r) + \alpha \left(\mu(x_r) + e(x_r) - \mathcal{I}[k(x', x_r) \hat{\theta}_t(x'); \mathcal{X}] - \mathcal{I}[\epsilon(x', x_r) \hat{\theta}_t(x'); \mathcal{X}] \right)\end{aligned}$$

で更新する. ここで, \mathcal{I} は数値積分を意味しており,

$$\begin{aligned}\hat{\theta}_{t+1}(x_r) &= \hat{\theta}_t(x_r) \\ &+ \alpha \left(\mu(x_r) + e(x_r) - \sum_{q=0}^Q A(k(x_q, x_r) \hat{\theta}_t(x_q), k(x_{q+1}, x_r) \hat{\theta}_t(x_{q+1})) (x_{q+1} - x_q) \right. \\ &\quad \left. - \sum_{q=0}^Q A(\epsilon(x_q, x_r) \hat{\theta}_t(x_q), \epsilon(x_{q+1}, x_r) \hat{\theta}_t(x_{q+1})) (x_{q+1} - x_q) \right) \quad (171)\end{aligned}$$

のように書ける. すると, 極限 $Q \rightarrow \infty$ をとると数値積分は積分に収束する. すると,

$$\begin{aligned}\lim_{Q \rightarrow \infty} \hat{\theta}_{t+1}(x_r) &= \lim_{Q \rightarrow \infty} \hat{\theta}_t(x_r) \quad (172) \\ &+ \alpha \left(\mu(x_r) + e(x_r) - \int_{\Omega_X} k(x', x_r) \lim_{Q \rightarrow \infty} \hat{\theta}_t(x') dx' - \int_{\Omega_X} \epsilon(x', x_r) \lim_{Q \rightarrow \infty} \hat{\theta}_t(x') dx' \right)\end{aligned}$$

が成立する. 極限 $N \rightarrow \infty$ をとると, $e(x_r)$ は 0 に確率収束し $\epsilon(x', x_r)$ は任意の $r = 1, \dots, R$ に対して 0 に確率収束するため,

$$\begin{aligned}\lim_{Q \rightarrow \infty, N \rightarrow \infty} \hat{\theta}_{t+1}(x_r) &= \lim_{Q \rightarrow \infty, N \rightarrow \infty} \hat{\theta}_t(x_r) \\ &+ \alpha \left(\mu(x_r) - \int_{\Omega_X} k(x', x_r) \left\{ \lim_{Q \rightarrow \infty, N \rightarrow \infty} \hat{\theta}_t(x') \right\} dx' \right) \quad (173)\end{aligned}$$

が成立し, 推定値 $\hat{\theta}_t(x_r)$ は任意の $r = 1, \dots, R$ に対して $\theta_t(x_r)$ の一致推定量である. さらに, 極限 $t \rightarrow \infty$ をとると, $\lim_{t \rightarrow \infty} \hat{\theta}_t(x_r)$ は, θ_t が $X = x_r$ において平均偏因果効果に収束するため, $X = x_r$ において平均偏因果効果の一致推定量である. 数値補間の性質から, 極限 $R \rightarrow \infty$ をとると, 関数 $\hat{\theta}(x)$ は任意の $x \in \Omega_X$ に対して平均偏因果効果の一致推定量である. \square

定理 4.2.3 の証明

証明. 定理 4.2.1 と同様に, ピカールの反復法 (4.33) は仮定 4.2.4 のもとで解に強収束する. したがって, アルゴリズム 3 は任意の $\epsilon > 0$ に対して有限回数で停止する. \square

となり,

$$\begin{aligned}
& \hat{\theta}_{t+1}(x) - \theta_{t+1}(x) \\
&= \hat{\theta}_t(x) - \theta_t(x) - \alpha \int_{\Omega_X} k(x', x) \{\hat{\theta}_t(x') - \theta_t(x')\} dx' \\
&\quad + \alpha e(x) - \alpha \int_{\Omega_X} \epsilon(x', x) \{\hat{\theta}_t(x') - \theta_t(x')\} dx' \\
&\quad + \alpha \int_{\Omega_X} \epsilon(x', x) \theta_t(x') dx'
\end{aligned} \tag{178}$$

が成立する. 簡潔に, 作用素 $I - \alpha\mathcal{K} - \alpha\mathcal{E}$ を

$$(I - \alpha\mathcal{K} - \alpha\mathcal{E})(\hat{\theta}_t - \theta_t)(x) \tag{179}$$

$$\begin{aligned}
&:= \hat{\theta}_t(x) - \theta_t(x) - \alpha \int_{\Omega_X} k(x', x) \{\hat{\theta}_t(x') - \theta_t(x')\} dx' \\
&\quad - \alpha \int_{\Omega_X} \epsilon(x', x) \{\hat{\theta}_t(x') - \theta_t(x')\} dx'
\end{aligned} \tag{180}$$

と定義すると,

$$\hat{\theta}_{t+1} - \theta_{t+1} = (I - \alpha\mathcal{K} - \alpha\mathcal{E})(\hat{\theta}_t - \theta_t) + \alpha e + \alpha\mathcal{E}\theta_t \tag{181}$$

となる. $\theta_1 = \hat{\theta}_1$ が成り立つので, 機械学習の誤差をともなうピカールの反復法は

$$\begin{aligned}
\hat{\theta}_2 - \theta_2 &= \alpha e + \alpha\mathcal{E}\theta_2 \\
\hat{\theta}_3 - \theta_3 &= (I - \alpha\mathcal{K} - \alpha\mathcal{E})(\alpha e + \alpha\mathcal{E}\theta_2) + \alpha e + \alpha\mathcal{E}\theta_3 \\
\hat{\theta}_4 - \theta_4 &= (I - \alpha\mathcal{K} - \alpha\mathcal{E})^2(\alpha e + \alpha\mathcal{E}\theta_2) + (I - \alpha\mathcal{K} - \alpha\mathcal{E})(\alpha e + \alpha\mathcal{E}\theta_3) \\
&\quad + \alpha e + \alpha\mathcal{E}\theta_4 \\
\hat{\theta}_5 - \theta_5 &= (I - \alpha\mathcal{K} - \alpha\mathcal{E})^4(\alpha e + \alpha\mathcal{E}\theta_2) + (I - \alpha\mathcal{K} - \alpha\mathcal{E})^2(\alpha e + \alpha\mathcal{E}\theta_3) \\
&\quad + (I - \alpha\mathcal{K} - \alpha\mathcal{E})(\alpha e + \alpha\mathcal{E}\theta_4) + \alpha e + \alpha\mathcal{E}\theta_4 \\
&\vdots
\end{aligned} \tag{182}$$

となる. したがって, t 回反復後の誤差は

$$\hat{\theta}_t - \theta_t = \sum_{t'=0}^{t-2} (I - \alpha\mathcal{K} - \alpha\mathcal{E})^{t'} (\alpha e + \alpha\mathcal{E}\theta_{t'})$$

$$= \sum_{t'=0}^{t-2} (I - \alpha(\mathcal{K} + \mathcal{E}))^{t'} (\alpha e + \alpha \mathcal{E} \theta_{t'}) \quad (183)$$

となる。ここで、作用素 $(\mathcal{K} + \mathcal{E})$ が連続であり、

$$\begin{aligned} & \left\| \sum_{t'=0}^{\infty} (I - \alpha(\mathcal{K} + \mathcal{E}))^{t'} (\alpha e + \alpha \mathcal{E} \theta_{\infty}) - \sum_{t'=0}^{t-2} (I - \alpha(\mathcal{K} + \mathcal{E}))^{t'} (\alpha e + \alpha \mathcal{E} \theta_{t'}) \right\| \\ & \leq \left\| \sum_{t'=0}^{\infty} (I - \alpha(\mathcal{K} + \mathcal{E}))^{t'} (\alpha e + \alpha \mathcal{E} \theta_{\infty}) - \sum_{t'=0}^{t-2} (I - \alpha(\mathcal{K} + \mathcal{E}))^{t'} (\alpha e + \alpha \mathcal{E} \theta_{\infty}) \right\| \\ & + \left\| \sum_{t'=0}^{t-2} (I - \alpha(\mathcal{K} + \mathcal{E}))^{t'} (\alpha e + \alpha \mathcal{E} \theta_{\infty}) - \sum_{t'=0}^{t-2} (I - \alpha(\mathcal{K} + \mathcal{E}))^{t'} (\alpha e + \alpha \mathcal{E} \theta_{t'}) \right\| \end{aligned}$$

が成り立つので、作用素 $(\mathcal{K} + \mathcal{E})$ がピカールの条件を満たすなら

$$\sum_{t'=0}^{t-2} (I - \alpha(\mathcal{K} + \mathcal{E}))^{t'} (\alpha e + \alpha \mathcal{E} \theta_{t'}) \quad (184)$$

は

$$\sum_{t'=0}^{\infty} (I - \alpha(\mathcal{K} + \mathcal{E}))^{t'} (\alpha e + \alpha \mathcal{E} \theta_{\infty}) \quad (185)$$

に強収束する。 $\hat{\theta}_t - \theta_t$ は積分方程式の解 σ に強収束し、

$$\alpha e + \alpha \mathcal{E} \theta_{\infty} = \alpha(\mathcal{K} + \mathcal{E})\sigma \Leftrightarrow \sigma = (\mathcal{K} + \mathcal{E})^{-1}(e + \mathcal{E} \theta_{\infty}) \quad (186)$$

とかける。ここで、 $(\mathcal{K} + \mathcal{E})^{-1}(e + \mathcal{E} \theta_{\infty}) = \alpha \sum_{t=0}^{\infty} (I - \alpha(\mathcal{K} + \mathcal{E}))^t (e + \mathcal{E} \theta_{\infty})$ [Diaz and Metcalf, 1970] である。誤差のノルムは

$$\|\sigma\| \leq \|(\mathcal{K} + \mathcal{E})^{-1}\| \|e + \mathcal{E} \theta_{\infty}\| \leq \|(\mathcal{K} + \mathcal{E})^{-1}\| \{\|e\| + \|\mathcal{E}\| \|\theta_{\infty}\|\} \quad (187)$$

で抑えられる。これは

$$\|\hat{\theta}_{\infty} - \theta_{\infty}\| \leq \|\hat{\mathcal{K}}^{-1}\| \{\|\hat{\mu} - \mu\| + \|\hat{\mathcal{K}} - \mathcal{K}\| \|\theta_{\infty}\|\} \quad (188)$$

を意味している。もし $\|\hat{\mathcal{K}}^{-1}\|$ が M で抑えられているなら、

$$\|\hat{\theta}_{\infty} - \theta_{\infty}\| \leq M \{\|\hat{\mu} - \mu\| + \|\hat{\mathcal{K}} - \mathcal{K}\| \|\theta_{\infty}\|\} \quad (189)$$

が成立する。仮に $\hat{\mu}$ が μ に等しくさらに $\hat{\mathcal{K}}$ が \mathcal{K} に等しい場合 $\hat{\theta}_{\infty}$ は θ_{∞} に一致する。

□

定理 4.2.5 の証明

証明. 各 x における誤差に関して t 回反復を行った後の x における誤差は

$$\hat{\theta}_t(x) - \theta_t(x) = \sum_{t'=0}^{t-2} (I - \alpha\mathcal{K} - \alpha\mathcal{E})^{t'} (\alpha e + \alpha\mathcal{E}\theta_t)(x) \quad (190)$$

$$= \alpha(t-1)(e + \mathcal{E}\theta_t)(x) + \mathcal{O}(\alpha^2) \quad (191)$$

となり, 絶対値誤差は

$$|\hat{\theta}_t(x) - \theta_t(x)| = |(t-1)\alpha(e + \mathcal{E}\theta_t)(x)| + \mathcal{O}(\alpha^2) \quad (192)$$

となり, 2乗誤差は

$$(\hat{\theta}_t(x) - \theta_t(x))^2 = (t-1)^2\alpha^2(e + \mathcal{E}\theta_t)(x)^2 + \mathcal{O}(\alpha^3) \quad (193)$$

となる. したがって, 分散は

$$\mathbb{V}(\hat{\theta}_t(x)) = \mathbb{E}\{(\hat{\theta}_t(x) - \theta_t(x))^2\} = (t-1)^2\alpha^2\mathbb{E}[(e + \mathcal{E}\theta_t)(x)^2] + \mathcal{O}(\alpha^3) \quad (194)$$

となる. ここで,

$$\mathbb{E}[(e + \mathcal{E}\theta_t)(x)^2] \leq \mathbb{E}[e(x)^2] + 2\mathbb{E}[e(x)\mathcal{E}(\theta_t)(x)] + \mathbb{E}[\mathcal{E}(\theta_t)(x)^2] \quad (195)$$

$$\begin{aligned} &= \mathbb{V}(\hat{\mu}(x)) + 2\mathbb{E}\{(\hat{\mu}(x) - \mu(x))\{\hat{\mathcal{K}}(\theta_t)(x) - \mathcal{K}(\theta_t)(x)\}\} \\ &\quad + \mathbb{V}(\hat{\mathcal{K}}(\theta_t)(x)) \end{aligned} \quad (196)$$

が成立する. 機械学習による予測モデルの不偏性の仮定とコーシーシュワルツの不等式から $\mathbb{E}[XY] \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]}$ がなりたつので,

$$\mathbb{E}[(e + \mathcal{E}\theta_t)(x)^2] \leq \mathbb{V}(\hat{\mu}(x)) + 2\sqrt{\mathbb{V}(\hat{\mu}(x))\mathbb{V}(\hat{\mathcal{K}}(\theta_t)(x))} + \mathbb{V}(\hat{\mathcal{K}}(\theta_t)(x)) \quad (197)$$

が成立する. したがって, $\mathbb{V}(\hat{\mathcal{K}}(\theta_t)(x))$ は Hölder の不等式から

$$\mathbb{V}(\hat{\mathcal{K}}(\theta_t)(x)) = E\left[\left(\int_{\Omega_X} \{\hat{k}(x', x) - k(x', x)\}\theta_t(x')dx'\right)^2\right] \quad (198)$$

$$\leq E\left[\left|\int_{\Omega_X} \{\hat{k}(x', x) - k(x', x)\}\theta_t(x')dx'\right|^2\right] \quad (199)$$

$$\leq \left(\int_{\Omega_X} \mathbb{E}[|\hat{k}(x', x) - k(x', x)|\theta_t(x')|dx'] \right)^2 \quad (200)$$

$$\leq \left(\int_{\Omega_X} \sqrt{\mathbb{V}(\hat{k}(x', x))|\theta_t(x')|dx'} \right)^2 \quad (201)$$

でおさえられる。さいごに、

$$\mathbb{V}(\hat{\theta}_t(x)) \leq (t-1)^2 \alpha^2 \nu(x) + \mathcal{O}(\alpha^3) \quad (202)$$

が成立する。ここで、

$$\begin{aligned} \nu(x) = & \mathbb{V}(\hat{\mu}(x)) + 2\sqrt{\mathbb{V}(\hat{\mu}(x)) \left(\int_{\Omega_X} \sqrt{\mathbb{V}(\hat{k}(x', x))|\theta_t(x')|dx'} \right)^2} \\ & + \left(\int_{\Omega_X} \sqrt{\mathbb{V}(\hat{k}(x', x))|\theta_t(x')|dx'} \right)^2 \end{aligned} \quad (203)$$

である。

□

系 4.2.2 の証明

証明. はじめに、不等式

$$\left(\int_{\Omega_X} \sqrt{\mathbb{V}(\hat{k}(x', x))|\theta_t(x')|dx'} \right)^2 \leq \max_{x'} \{\mathbb{V}(\hat{k}(x', x))\} \left(\int_{\Omega_X} |\theta_t(x')|dx' \right)^2 \quad (204)$$

が成立するため、定理 4.2.5 から

$$\begin{aligned} \mathbb{E}[\{\hat{\theta}_t(x) - \theta_t(x)\}^2] = & (t-1)^2 \alpha^2 \mathbb{V}(\hat{\mu}(x)) (\alpha^3) \\ & + \mathcal{O}(\alpha^3) + \mathcal{O}\left(\sqrt{\max_{x'} \{\mathbb{V}(\hat{k}(x', x))\}}\right) \end{aligned} \quad (205)$$

が成立する。

□

定理 4.3.1 の証明

証明. はじめに、 \hat{u} と \hat{D} は

$$\begin{cases} u_r = \theta_P d_r^P + \dots + \theta_1 d_r^1 \\ \hat{u}_r = u_r + e_r, \quad \forall r = 1, \dots, R \\ \hat{d}_r^p = d_r^p + \epsilon_r^p, \quad \forall r = 1, \dots, R, \quad \forall p = 1, \dots, P \end{cases} \quad (206)$$

とかける. ここで, $e_r(r = 1, \dots, R)$ と $\epsilon_r^p(r = 1, \dots, R, p = 1, \dots, P)$ は誤差項である. $\mathbf{e} = (e_1, \dots, e_R)$ と $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_R)$ と書く. すると, 推定量 $\hat{\boldsymbol{\theta}}$ は

$$\hat{\boldsymbol{\theta}} = (\hat{\mathbf{D}}^T \hat{\mathbf{D}})^{-1} \hat{\mathbf{D}}^T \hat{\mathbf{u}} = \{(\mathbf{D} + \boldsymbol{\epsilon})^T (\mathbf{D} + \boldsymbol{\epsilon})\}^{-1} (\mathbf{D} + \boldsymbol{\epsilon})^T (\mathbf{u} + \mathbf{e}) \quad (207)$$

となる. $\mathbf{D} \perp \boldsymbol{\epsilon}$ なので,

$$= (\mathbf{D}^T \mathbf{D} + \boldsymbol{\epsilon}^T \boldsymbol{\epsilon})^{-1} (\mathbf{D} + \boldsymbol{\epsilon})^T (\mathbf{u} + \mathbf{e}) \quad (208)$$

が成立し, 逆行列の補助定理 [Higham, 2002] から,

$$= (\mathbf{D}^T \mathbf{D} + \boldsymbol{\epsilon}^T \boldsymbol{\epsilon})^{-1} (\mathbf{D} + \boldsymbol{\epsilon})^T (\mathbf{u} + \mathbf{e}) \quad (209)$$

$$= [(\mathbf{D}^T \mathbf{D})^{-1} - (\mathbf{D}^T \mathbf{D})^{-1} \boldsymbol{\epsilon}^T (I + \boldsymbol{\epsilon} (\mathbf{D}^T \mathbf{D})^{-1} \boldsymbol{\epsilon}^T)^{-1} \boldsymbol{\epsilon} (\mathbf{D}^T \mathbf{D})^{-1}] (\mathbf{D} + \boldsymbol{\epsilon})^T (\mathbf{u} + \mathbf{e})$$

$$= (\mathbf{D}^T \mathbf{D})^{-1} (\mathbf{D} + \boldsymbol{\epsilon})^T (\mathbf{u} + \mathbf{e})$$

$$- (\mathbf{D}^T \mathbf{D})^{-1} \boldsymbol{\epsilon}^T (I + \boldsymbol{\epsilon} (\mathbf{D}^T \mathbf{D})^{-1} \boldsymbol{\epsilon}^T)^{-1} \boldsymbol{\epsilon} (\mathbf{D}^T \mathbf{D})^{-1} (\mathbf{D} + \boldsymbol{\epsilon})^T (\mathbf{u} + \mathbf{e}) \quad (210)$$

$$= \boldsymbol{\theta} + (\mathbf{D}^T \mathbf{D})^{-1} \boldsymbol{\epsilon}^T \mathbf{u} + (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{e} + (\mathbf{D}^T \mathbf{D})^{-1} \boldsymbol{\epsilon}^T \mathbf{e}$$

$$- (\mathbf{D}^T \mathbf{D})^{-1} \boldsymbol{\epsilon}^T (I + \boldsymbol{\epsilon} (\mathbf{D}^T \mathbf{D})^{-1} \boldsymbol{\epsilon}^T)^{-1} \boldsymbol{\epsilon} (\mathbf{D}^T \mathbf{D})^{-1} (\mathbf{D} + \boldsymbol{\epsilon})^T (\mathbf{u} + \mathbf{e}) \quad (211)$$

となる. すると,

$$\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} = (\mathbf{D}^T \mathbf{D})^{-1} \boldsymbol{\epsilon}^T \mathbf{u} + (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{e} + (\mathbf{D}^T \mathbf{D})^{-1} \boldsymbol{\epsilon}^T \mathbf{e} \quad (212)$$

$$- (\mathbf{D}^T \mathbf{D})^{-1} \boldsymbol{\epsilon}^T (I + \boldsymbol{\epsilon} (\mathbf{D}^T \mathbf{D})^{-1} \boldsymbol{\epsilon}^T)^{-1} \boldsymbol{\epsilon} (\mathbf{D}^T \mathbf{D})^{-1} (\mathbf{D} + \boldsymbol{\epsilon})^T (\mathbf{u} + \mathbf{e})$$

となり, 極限 $N \rightarrow \infty$ をとると $\boldsymbol{\epsilon}$ は零行列に確率収束し, \mathbf{e} は零ベクトルに確率収束する. したがって, 推定量 $\hat{\boldsymbol{\theta}}$ は一致推定量である. \square