条件付き操作変数法の推定精度と 操作変数の選択

大阪大学 基礎工学研究科 黒 木 学

東京工業大学 社会理工学研究科 宮 川 雅 巳

京都大学 経済学研究科 川 田 亮 平

要 旨 線形回帰モデルにおいて、説明変数と誤差に相関があるとき、偏回帰係数の一致推定量を得る方法のひとつとして操作変数法がある。この操作変数法は、最近発展している因果ダイアグラムを用いた統計的因果推論において、その意義が再認識されている。しかし、一般に操作変数の条件を満たす変数が存在しない状況は少なくなく、適用範囲は限定されていた。これに対して Brito and Pearl (2002) が提案した条件付き操作変数法は、ある共変量を観測することで、操作変数の条件を緩和したものである。本論文では、線形構造方程式モデルのもとで、条件付き操作変数法を用いた総合効果の推定精度について調べた。その結果を利用して、条件付き操作変数の条件を満たす変数が複数あるときに、興味ある総合効果の推定精度という観点からの操作変数選択基準を与えた。

1. はじめに

操作変数法は、線形回帰モデルにおいて説明変数と誤差項に相関があるときに、偏回帰係数の一致推定量を導く発見的方法として計量経済学の分野などで古くから用いられており、代表的な著書として Bowden and Turkington (1984) がある。それによれば操作変数法の概要は次の通りである。いま、線形回帰モデル

$$y = X\beta + u \quad (1)$$

を考える。ここに、 n, p をそれぞれ標本サイズと説明変数の数としたとき、 y と u は $n \times 1$ の列ベクトルで、 X は $n \times p$ の行列、 β は $p \times 1$ の列ベクトルである。そして、目的変数 y だけでなく説明変数 X も観測変量の実現値であるという観察研究の場面を想定する。このとき、説明変数 X と誤差項 u に相関があるとする。

形式的に (1) 式の両辺に X' を左からかけることにより

$$X'y = X'X\beta + X'u$$

を得る。ここで、 $1/nX'u$ が $n \rightarrow \infty$ で 0 に確率収束しないときには、通常の最小 2 乗推定量

$$\hat{\beta} = (X'X)^{-1}X'y \quad (2)$$

は β の一致推定量にならない。しかし、 $n \rightarrow \infty$ で $1/nZ'u$ が 0 に確率収束するような都合の良い $n \times p$ 行列 Z があれば、 Z' を (1) 式の両辺に左からかけると

$$Z'y = Z'X\beta + Z'u \quad (3)$$

となるので、これより

$$\tilde{\beta} = (Z'X)^{-1}Z'y \quad (4)$$

を得ることができる。この推定量は β の一致推定量になる。このような方法が操作変数法で、このときの Z を操作変数という。

ところで、近年、因果ダイアグラムを用いた統計的因果推論の理論が発展し、因果的効果のいくつかの識別可能条件がグラフ用語によって表現されている。その流れの中で、操作変数法の意義が再確認されている(たとえば Pearl (2000) の 5.3 節)。その一方で、操作変数法の適用範囲が限定されていることも明らかになっている。それに対応すべく、Pearl (2001) と Brito and Pearl (2002) によって、操作変数以外の共変量を観測することで、操作変数の条件を緩和した条件付き操作変数法が提案された。本論文の主目的は、線形構造方程式モデルにおいて、条件付き操作変数法による総合効果の推定精度をしらべることである。この結果に基づいて、条件付き操作変数の条件を満たす変数が複数ある場合において、総合効果の推定精度という観点からの操作変数選択基準を与える。

本論文の構成は次のとおりである。第 2 節では、因果ダイアグラムに関する基本的な用語と概念を述べ、因果ダイアグラムの見地から従来の操作変数法を解釈する。続いて、条件付き操作変数法の紹介をする。第 3 節では、線形構造方程式モデルにおいて条件付き操作変数法で総合効果を推定したときのその漸近分散を与える。また、その結果を利用して、条件付き操作変数の条件を満たす変数が複数あるときに、興味ある総合効果の推定精度という観点からの操作変数選択基準を与える。第 4 節では、この操作変数選択基準の適用例と数値実験結果を与える。第 5 節ではまとめと残された課題を述べる。

2. 因果ダイアグラムからみた操作変数法

2.1. 非巡回的有向グラフ

グラフ G は頂点の集合 V と、その直積 $V \times V$ の部分集合である矢線の集合 E によって、 $G = (V, E)$ として表現される。2 つの頂点 $\alpha, \beta \in V$ に対して $(\alpha, \beta) \in E$ かつ $(\beta, \alpha) \notin E$ のとき、 α から β に向きのある有向の辺(矢線)を引く。矢線のみから構成されるグラフを有向グラフという。 α から β への矢線が存在するとき、 α は β の親であるといい、 β は α の子であるという。異なる頂点の列 $\alpha_0, \alpha_1, \dots, \alpha_n$ は、すべての $i = 1, \dots, n$ で $(\alpha_{i-1}, \alpha_i) \in E$ または $(\alpha_i, \alpha_{i-1}) \in E$ であるとき、長さ n の道という。特に、長さ n の道で、すべての $i = 1, \dots, n$ に対して $(\alpha_{i-1}, \alpha_i) \in E$ かつ $(\alpha_i, \alpha_{i-1}) \notin E$ であるとき、長さ n の有向道という。 α から β への有向道が存在するとき、 α は β の先祖であるといい、 β は α の子孫であるという。長さ n の道 $\alpha_0, \alpha_1, \dots, \alpha_n$ で、 $(\alpha_{i-1}, \alpha_i) \in E$ かつ $(\alpha_{i+1}, \alpha_i) \in E$ であるとき α_i を合流点といい、そうでないとき、 α_i を非合流点という。有向道 $\alpha_0, \alpha_1, \dots, \alpha_n$ で、 $\alpha_0 = \alpha_n$

を許したものを巡回閉路といい、巡回閉路の存在しない有向グラフを非巡回的有向グラフという。

2.2. 因果ダイアグラム

統計的因果分析において、変数間の因果関係を表現した非巡回的有向グラフを因果ダイアグラムと呼ぶ。以下にその数学的定義を与える。

【定義 1】 因果ダイアグラム

非巡回的有向グラフ G とその頂点に対応する確率変数の集合 $V = \{X_1, X_2, \dots, X_m\}$ が与えられている。グラフ G が確率変数間の因果的関係を

$$X_i = g_i(\text{pa}(X_i), \epsilon_i) \quad i = 1, \dots, m \quad (5)$$

なる形に規定し、確率変数がこの因果的関係にしたがって生成されるとき、 G を因果ダイアグラムという。ここに、 $\text{pa}(X_i)$ は X_i の親からなる集合であり、誤差変数 $\epsilon_1, \epsilon_2, \dots, \epsilon_m$ は互いに独立であるとする。□

$f(x_i|\text{pa}(x_i))$ と $f(x_i)$ をそれぞれ $\text{pa}(X_i)$ を与えたときの X_i の条件付き分布と X_i の周辺分布とする。また、 $\text{pa}(X_i)$ が空集合であるときには $f(x_i|\text{pa}(x_i)) = f(x_i)$ とする。他の周辺分布や条件付き分布についても同様に記述する。確率変数間の因果的関係が (5) 式によって規定されるとき、その同時分布 $f(x_1, x_2, \dots, x_m)$ は

$$f(x_1, x_2, \dots, x_m) = \prod_{i=1}^m f(x_i|\text{pa}(x_i))$$

と逐次的に因数分解された形で表現することができる (Pearl, 1995)。

因果ダイアグラム G が規定する条件付き独立性を記述するために以下の定義を用いる。

【定義 2】 有向分離

非巡回的有向グラフ G において、 X と Y を結ぶすべての道のそれぞれについて、 $\{X, Y\}$ と排反な頂点の集合 Z が次の条件のいずれかを満たすとき、 Z は X と Y を有向分離するという。

1. X と Y を結ぶ道に合流点で、その合流点とその子孫が Z に含まれないような頂点がある。
2. X と Y を結ぶ道に非合流点で、 Z に含まれるものがある。

特に、 X と Y を結ぶ道が存在しないとき、空集合は X と Y を有向分離するという。□

このとき、非巡回有向グラフ G において頂点集合 Z が X と Y を有向分離するならば、頂点に対応する確率変数において変数集合 Z を与えたとき X と Y は条件付き独立である (Geiger et al., 1990)。この性質は DAG の大域的マルコフ性と呼ばれる (たとえば, Lauritzen (1996))。

2.3. 線形構造方程式モデル

変数集合 V のすべての要素が連続変数の場合に、(5) 式に対応する基礎的モデルとして、次の線形構造方程式モデルが知られている。

$$X_i = \sum_{X_j \in \text{pa}(X_i)} \alpha_{x_i x_j} X_j + \epsilon_i \quad i = 1, \dots, m \quad (6)$$

ここに、 V の各要素は平均 0、分散 1 に標準化されているものとし、誤差変数 $\epsilon_1, \epsilon_2, \dots, \epsilon_m$ は互い

に独立に正規分布にしたがうとする。また、 $\alpha_{x_i x_j}$ は X_j から X_i への矢線に対応する係数でパス係数、あるいは直接効果と呼ばれるもので、 $X_j \in \text{pa}(X_i)$ に対して $\alpha_{x_i x_j} \neq 0$ とする。なお、 X_i から X_j への有向道のそれぞれについてパス係数すべての積を考えると、その総和は総合効果と呼ばれる。以下では、この線形構造方程式モデルのもとで議論を進める。

2.4. 操作変数法の解釈

X と Z 、さらに非観測変数 U もすべて 1 変数の場合について、操作変数法の仕組みを因果ダイアグラムで確認しよう。図 1 を見てほしい。図 1(a) が (1) 式に対応する困難な状況である。このとき、 U は X と Y に対する交絡因子であるから、 X と Y の相関関係のみでは、 X から Y への因果の効果である総合効果を識別できない。図 1(b) では、これに Z が加わっている。 Z はたしかに U と独立であり、かつ、 X と相関をもつ。ちなみに、(3) 式で行列 Z と行列 X の列ベクトルが直交していると、行列 $Z'X$ は 0 行列になってしまう。

図 1. 操作変数法の考え方

図 1(b) に対応する線形構造方程式モデルは

$$\begin{aligned} X &= \alpha_{xu}U + \alpha_{xz}Z + \epsilon_x \\ Y &= \alpha_{yu}U + \alpha_{yx}X + \epsilon_y \\ Z &= \epsilon_z \\ U &= \epsilon_u \end{aligned} \quad (7)$$

であるから、すべての変数の平均が 0 に基準化されているとき、(7) 式の第 2 式の両辺に Z をかけて期待値をとれば $\sigma_{yz} = \alpha_{yx}\sigma_{xz}$ を得る。ここに、 σ_{yz} は Y と Z の共分散、 σ_{xz} は X と Z の共分散である。これより、興味ある総合効果(この場合はパス係数 α_{yx})は 2 つの共分散 σ_{yz} と σ_{xz} を用いて、

$$\alpha_{yx} = \frac{\sigma_{yz}}{\sigma_{xz}} \quad (8)$$

と表現できることになる。これが操作変数法であり、 Z は操作変数と呼ばれる。操作変数 Z の要件を因果ダイアグラムで記述すると

- 1) Z は X および Y の子孫でない
- 2) 因果ダイアグラム G より X からでる矢線を除いたグラフで、空集合は Z と Y を有向分離する
- 3) 因果ダイアグラム G より X からでる矢線を除いたグラフで、空集合は Z と X を有向分離しない

となる。

2.5. 条件付き操作変数法

操作変数の要件はやや厳しいものといえる．その要件を緩和した方法が Brito and Pearl (2002) によって提案された条件付き操作変数法である．図 2 の状況を見てほしい．

図 2. 条件付き操作変数法の考え方

図 2 の状況で， U も W も観測しないとすると， W は Y への直接の矢線を持ち Z は W と独立でないので， Z は操作変数の要件を満たさない．しかし， W を観測し，これを条件付きにすることで，もうひとつの交絡因子 U を観測せずとも，総合効果(この場合はパス係数 α_{yx})は識別可能になる．このとき

$$\alpha_{yx} = \frac{\sigma_{yz:w}}{\sigma_{xz:w}} \quad (9)$$

という式が成り立つ(Pearl, 2001)．ここに， $\sigma_{yz:w}$ は W を与えたときの Y と Z の条件付き分布での共分散で， $\sigma_{xz:w}$ は W を与えたときの X と Z の条件付き分布での共分散である．これが条件付き操作変数法であり， Z を W を与えたときの条件付き操作変数と呼ぶ．一般に，条件付き操作変数法の要件を因果ダイアグラムで記述すると

- 1) Z と W に含まれる任意の頂点は X および Y の子孫でない
- 2) 因果ダイアグラム G より X からでる矢線を除いたグラフで， W は Z と Y を有向分離する
- 3) 因果ダイアグラム G より X からでる矢線を除いたグラフで， W は Z と X を有向分離しない

となる． W が空集合のときには，(9) 式は (8) 式に帰着できるという意味で，条件付き操作変数法は操作変数法を拡張したものになっている．

なお，総合効果の識別可能条件のひとつとしてバックドア基準がある(Pearl (1995))．因果ダイアグラム G において X から Y へ有向道があるとき， Z が (X, Y) に対してバックドア基準を満たす要件は，

- 1) X から Z の任意の頂点へ有向道がない
 - 2) G より X から出る矢線をすべて除いたグラフにおいて， Z は X と Y を有向分離する
- である．頂点集合 Z がバックドア基準を満たすとき，頂点に対応する確率変数において Y を目的変数， X と Z に含まれるすべての変数を説明変数とした回帰モデルでの X の偏回帰係数 β_{yxz} は， X から Y への総合効果と等しくなることが知られている(Pearl, 1998)．

このことから， $W \cup \{Z\}$ が (X, Y) についてバックドア基準を満たしていれば，これらの変数を X, Y とともに観測することで総合効果は識別可能になることがわかる．しかし，図 2 においては，交絡因子 U の存在によりバックドア基準は満たされていない．この点において，条件付き操作変数法には存在意義がある．

3. 総合効果の推定精度

3.1. 漸近分散

X の Y への総合効果 τ_{yx} を、 W を与えたときの Y と Z の標本共分散 $\hat{\sigma}_{yz \cdot w}$ と W を与えたときの X と Z の標本共分散 $\hat{\sigma}_{xz \cdot w}$ を用いて推定することを考える。

(6) 式の線形構造方程式モデルにおいて $\epsilon_1, \epsilon_2, \dots, \epsilon_m$ は互いに独立に正規分布に従うとする。このとき、 Z が W を与えたときの条件付き操作変数であれば、総合効果 τ_{yx} の推定量 $\tilde{\tau}_{yx}$ は

$$\tilde{\tau}_{yx} = \frac{\hat{\sigma}_{yz \cdot w}}{\hat{\sigma}_{xz \cdot w}} \quad (10)$$

と記述することができる。いま、(10) 式の厳密な分散を求めることは難しい。そこで、 $\sqrt{n}(\hat{\sigma}_{yz \cdot w} - \sigma_{yz \cdot w}, \hat{\sigma}_{xz \cdot w} - \sigma_{xz \cdot w})$ が平均 $\mathbf{0}$ 、分散共分散行列

$$\Sigma = \begin{pmatrix} \sigma_{zz \cdot w} \sigma_{yy \cdot w} + \sigma_{yz \cdot w}^2 & \sigma_{zz \cdot w} \sigma_{xy \cdot w} + \sigma_{xz \cdot w} \sigma_{yz \cdot w} \\ \sigma_{zz \cdot w} \sigma_{xy \cdot w} + \sigma_{xz \cdot w} \sigma_{yz \cdot w} & \sigma_{zz \cdot w} \sigma_{xx \cdot w} + \sigma_{xz \cdot w}^2 \end{pmatrix}$$

の正規分布に漸近的にしたがうことに注意すると、デルタ法 (Anderson, 1986) により、 $\tilde{\tau}_{yx}$ の漸近分散として

$$A.Var(\tilde{\tau}_{yx}) = \frac{1}{n} \frac{\sigma_{yy \cdot w} \sigma_{xz \cdot w}^2 - 2\sigma_{xy \cdot w} \sigma_{xz \cdot w} \sigma_{yz \cdot w} + \sigma_{xx \cdot w} \sigma_{yz \cdot w}^2}{\sigma_{xz \cdot w}^4 / \sigma_{zz \cdot w}} \quad (11)$$

を得る。ここに、 n は標本サイズ、 $\sigma_{yy \cdot w}$ は W を与えたときの Y の条件付き分布での分散、 $\sigma_{xx \cdot w}$ は W を与えたときの X での条件付き分布の分散である。また、 $\sigma_{zz \cdot w}$ は W を与えたときの Z の条件付き分布での分散、 $\sigma_{xy \cdot w}$ は W を与えたときの X と Y の条件付き分布での共分散である。

ここで、 $\beta_{yx \cdot w}$ を x と w を説明変数、 Y を目的変数とした回帰モデルでの x の偏回帰係数、 $\rho_{xz \cdot w}$ を W を与えたときの X と Z の偏相関係数とすると、 $\tau_{yx} = \sigma_{yz \cdot w} / \sigma_{xz \cdot w}$ であることより、(11) 式は

$$A.Var(\tilde{\tau}_{yx}) = \frac{1}{n} \frac{\sigma_{yy \cdot w} / \sigma_{xx \cdot w} - 2\beta_{yx \cdot w} \tau_{yx} + \tau_{yx}^2}{\rho_{xz \cdot w}^2} \quad (12)$$

と記述できる。この式の分子は W に依存し、分母は Z と W の両方に依存している。

3.2. 推定精度の観点からの操作変数選択基準の検討

まず、図 3(a) に示す因果ダイアグラムを例に議論を進めよう。図 3(a) では、 X の Y への総合効果の推定において、 Z_1 と Z_2 はそれぞれ W を与えたときの条件付き操作変数となっていることがわかる。また、 $\{W, Z_1\}$ は X と Z_2 を有向分離している。そこで、線形構造方程式モデルの枠組みにおいて、 X の Y への総合効果の推定精度という観点から、 Z_1 と Z_2 のどちらを条件付き操作変数として観測し、総合効果の推定に用いるべきか、という問題を考える。このとき、以下の定理が成り立つ。

【定理】

因果ダイアグラム G において、 Z_1 と Z_2 がそれぞれ、 (X, Y) に対して W を与えたときの条件付き操作変数であるとする。このとき、 $\{Z_1\} \cup W$ が Z_2 と X を有向分離するならば

$$A.Var\left(\frac{\hat{\sigma}_{y_{z_1 \cdot w}}}{\hat{\sigma}_{x_{z_1 \cdot w}}}\right) \leq A.Var\left(\frac{\hat{\sigma}_{y_{z_2 \cdot w}}}{\hat{\sigma}_{x_{z_2 \cdot w}}}\right) \quad (13)$$

が成り立つ。 □

この定理は、 $\sigma_{xz_2 \cdot w}^2 / \sigma_{zz_2 \cdot w} = \beta_{xz_2 \cdot w}^2 \sigma_{zz_2 \cdot w}$ であることおよび $\{Z_1\} \cup W$ が Z_2 と X を有向分離することから、

$$\begin{aligned} \beta_{xz_2 \cdot w}^2 \sigma_{zz_2 \cdot w} &= (\beta_{xz_2 \cdot z_1 w} + \beta_{xz_1 \cdot z_2 w} \beta_{z_1 z_2 \cdot w})^2 \sigma_{zz_2 \cdot w} = \beta_{xz_1 \cdot w}^2 \beta_{z_1 z_2 \cdot w}^2 \sigma_{zz_2 \cdot w} \\ &= \beta_{xz_1 \cdot w}^2 \sigma_{z_1 z_1 \cdot w} \rho_{z_1 z_2 \cdot w}^2 < \beta_{xz_1 \cdot w}^2 \sigma_{z_1 z_1 \cdot w} \end{aligned}$$

となるので、題意を得る。

定理で与えた仮定が成り立たない場合には、推定精度の観点から条件付き操作変数間に定性的な優劣関係を与えることは難しい。この場合には、(12) 式の分母が Z の選択に依存していることにより、 W を与えたときに X との偏相関係数が大きくなる変数を条件付き操作変数として選択することで、漸近分散を小さくすることができることがわかる。これに関する考察をおこなうために、図 3(b) に示す因果ダイアグラムを例に議論を進めよう。図 3(b) では、 X の Y への総合効果の推定において、 Z_1 と Z_2 はそれぞれ W を与えたときの条件付き操作変数となっている。この因果ダイアグラムは $\{W, Z_1\}$ は X と Z_2 を有向分離しないという点で、図 3(a) から得られる条件付き独立関係とは異なる。

図 3. 因果ダイアグラム

この因果ダイアグラムに対して、以下のような共分散行列の数値例を与える。

$$\begin{pmatrix} \sigma_{z_2 z_2} & & & & \\ \sigma_{z_2 z_1} & \sigma_{z_1 z_1} & & & \\ \sigma_{z_2 x} & \sigma_{z_1 x} & \sigma_{xx} & & \\ \sigma_{z_2 w} & \sigma_{z_1 w} & \sigma_{xw} & \sigma_{ww} & \\ \sigma_{z_2 y} & \sigma_{z_1 y} & \sigma_{xy} & \sigma_{wy} & \sigma_{yy} \end{pmatrix} = \begin{pmatrix} 1.000 & & & & \\ 0.447 & 1.000 & & & \\ 0.183 & 0.408 & 1.000 & & \\ 0.700 & 0.313 & 0.676 & 1.000 & \\ 0.320 & 0.227 & 0.592 & 0.654 & 1.000 \end{pmatrix} \quad (14)$$

このとき、(12) 式の分母のみが条件付き操作変数に依存する。したがって、 $\rho_{xz_1 \cdot w}^2 = 0.078$ 、 $\rho_{xz_2 \cdot w}^2 = 0.304$ であることから、 $A.Var(\hat{\sigma}_{y_{z_1 \cdot w}} / \hat{\sigma}_{x_{z_1 \cdot w}}) / A.Var(\hat{\sigma}_{y_{z_2 \cdot w}} / \hat{\sigma}_{x_{z_2 \cdot w}}) = 0.198 / 0.051 = 3.882$ となり、 Z_1 を条件付き操作変数として用いた場合のほうが Z_2 を用いた場合よりも漸近分散が大きいことがわかる。一般に、従来の操作変数法においては、 X の直接的原因となる変量を操作変数として選択することで常に漸近分散を小さくすることができる。しかし、この例は条件付き操

作変数法の枠組みにおいては、 X の直接的原因となる変量を条件付き操作変数として選択しても、必ずしも推定精度の向上には結びつかないことを意味している。

4. 応 用

4.1. 適用例

前節までに述べた方法論を、奥野他(1986)の事例1「ボディ塗装条件の設定」に報告されているデータに適用する。このデータは自動車のボディ塗装工程において塗着率を高め、安定させるような工程条件を設定することを目的として収集されたものである。

データ数は38で、要因系変量として

- 塗料条件：希釈率 (X_1)、粘度 (X_2)、塗料温度 (X_8)
- 吹付条件：ガンスピード (X_3)、吹付距離 (X_4)、霧化エア圧 (X_5)、パターン幅 (X_6)、吐出量 (X_7)
- 環境条件：温度 (X_9)、湿度 (X_{10})

が測定されている。また、反応変量として

- 塗着率 (Y_1)、中心膜圧 (Y_2)、膜圧幅 (Y_3)

が測定されている。以下の解析では、反応変量としては Y_1 のみを対象にする。また、奥野他(1986)の記述より、 $X_1, X_2, X_3, X_4, X_5, X_6$ が制御可能な処理変量であると推察できる。黒木・宮川(1999a)は、奥野他(1986)の記述および相関行列より、この工程に対する因果ダイアグラムとして、図4を与えている。また、図4の因果ダイアグラムに基づいて推定された相関行列が黒木・宮川(1999b)により与えられている。これを表1に示す。

図4. 「ボディ塗装条件の設定」データによる因果ダイアグラム(黒木・宮川, 1999a)

この因果ダイアグラムに基づく統計的因果分析に関する議論については、黒木・宮川(1999a, 1999b, 2002)を参照されたい。

さて、図4から上述の各処理変数に対して操作変数は存在しないことが容易にわかる。そこで、表1の相関行列に対して条件付き操作変数法を適用することにより総合効果を推定することを考える。表2は、そのときの各処理変量の総合効果の推定値を示したものである。表2では、第一列目に条件付き操作変数法が利用可能な処理変量、第二列目に条件付き操作変数の候補、第三列

表 1. 図 4 の因果モデルに基づく相関行列の推定値(黒木・宮川, 1999b)

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	Y_1
X_1	1.000	-0.736	-0.152	0.148	0.028	-0.042	0.324	0.216	0.283	-0.496	-0.091
X_2	-0.736	1.000	0.210	-0.331	-0.063	0.095	-0.479	-0.684	-0.635	0.684	0.326
X_3	-0.152	0.210	1.000	-0.091	-0.017	0.026	0.195	-0.134	-0.175	0.307	0.134
X_4	0.148	-0.331	-0.091	1.000	0.191	-0.286	0.184	0.397	0.521	-0.298	-0.614
X_5	0.028	-0.063	-0.017	0.191	1.000	0.291	0.035	0.076	0.099	-0.057	-0.277
X_6	-0.042	0.095	0.026	-0.286	0.291	1.000	-0.053	-0.114	-0.149	0.085	-0.250
X_7	0.324	-0.479	0.195	0.184	0.035	-0.053	1.000	0.396	0.353	-0.146	-0.044
X_8	0.216	-0.684	-0.134	0.397	0.076	-0.114	0.396	1.000	0.761	-0.435	-0.493
X_9	0.283	-0.635	-0.175	0.521	0.099	-0.149	0.353	0.761	1.000	-0.571	-0.475
X_{10}	-0.496	0.684	0.307	-0.298	-0.057	0.085	-0.146	-0.435	-0.571	1.000	0.283
Y_1	-0.091	0.326	0.134	-0.614	-0.277	-0.250	-0.044	-0.493	-0.475	0.283	1.000

表 2. 総合効果の推定値

処理変量	条件付き操作変数	共変量集合	総合効果
X_2	$\{X_1\}$	$\{X_{10}\}$	-0.124
X_4	$\{X_1\}, \{X_2\}, \{X_3\}, \{X_9\}, \{X_{10}\}$	$\{X_7, X_8\}$	-0.502
X_6	$\{X_5\}$	$\{X_4\}$	-0.464

目に条件付き操作変数法で用いられる共変量集合の候補, 第四列目には総合効果の推定値をそれぞれ配している.

ここに, X_1, X_3, X_5 に関しては総合効果の推定をおこなう上で, 条件付き操作変数法を利用できないため, 表 2 にはこれらに関する項目は記載されていない.

まず, 図 4 および表 2 から, X_2 の総合効果の推定においては, 条件付き操作変数法に用いられる共変量集合はバックドア基準を満たしていないことが確認できる. 一方, X_4 と X_6 のそれぞれの総合効果の推定においては, 条件付き操作変数法に用いられる共変量集合はバックドア基準を満たしていることが確認できる. また, 表 2 から, X_4 の総合効果の推定に条件付き操作変数法を用いるとき, 共変量集合 $\{X_7, X_8\}$ を与えたときの条件付き操作変数が複数存在していることがわかる. そこで, 定理 2 を利用して, 条件付き操作変数間の優劣関係について検討してみよう. このとき, 図 4 から $\{X_7, X_8, X_9\}$ は X_4 と $\{X_1, X_2, X_3, X_{10}\}$ を有向分離していることが確認できる. したがって, 定理 2 より X_9 を条件付き操作変数として選択した場合には他の変数を条件付き操作変数として用いた場合よりも漸近分散が小さいことがわかる. X_1, X_2, X_3, X_{10} については, 総合効果の推定精度間の優劣関係をグラフから読み取ることは難しい.

4.2. 数値実験

この節では, 条件付き操作変数法による総合効果の推定量の分散を実験的に求め, 3.1 節に与えた漸近分散との比較をおこなう. 対象する因果ダイアグラムとして図 3(b) で与えたものを, 対応する母相関行列として (14) 式で与えたものを用いる. これは 3 節の考察から, Z_1 を条件付き操作変数として用いた場合のほうが Z_2 を用いた場合よりも分散が大きいと推察される状況である.

標本サイズを $n = 50, 100, 500, 1000$ の 4 水準として, (14) 式の相関行列に従う 5 次元正規乱数を発生させた. この条件の下で, X, Y , 条件付き操作変数 Z_1, Z_2 および共変量 W を観測し, 総合効果を (10) 式に基づいて推定する. これを 1000 回繰り返し, 推定量の分散値を実験的に求め

表 3. 数値実験結果

n	50	100	500	1000
$Var(\hat{\tau}_1)$	7.130	0.659	0.044	0.020
$A.Var(\hat{\tau}_1)$	0.388	0.194	0.039	0.019
$Var(\hat{\tau}_2)$	0.127	0.057	0.011	0.005
$A.Var(\hat{\tau}_2)$	0.101	0.050	0.010	0.005

た. その結果を表 3 に示す. ここに, 表 3 の $\hat{\tau}_1$ および $\hat{\tau}_2$ は, それぞれ条件付き操作変数として Z_1 を用いた場合の総合効果の推定量と Z_2 を用いた場合の総合効果の推定量である.

表 3 を観察して, 次の知見を得る.

- 1) Z_1 を条件付き操作変数として用いたときの総合効果の推定量の漸近分散と分散の比率は, $n \leq 100$ の場合において 3 を超えているが, $n \geq 500$ の場合には漸近分散と分散の比率は 1 に近い値となっている. 一方, Z_2 を条件付き操作変数として用いたときの総合効果の推定量の漸近分散と分散との比率は, $n \geq 50$ の場合において 1 に近い値となっている. このことから, Z_1 を用いた場合には小標本における漸近分散による分散の近似は十分でないが, Z_2 を用いた場合には漸近分散による分散の近似は比較的良好いと推察できる
- 2) $n \geq 50$ の場合において, $Var(\hat{\tau}_2) < Var(\hat{\tau}_1)$ となっており, 3.2 節の考察と一致している. このことから, 条件付き操作変数の選択による分散の差異が現れていることがわかる
- 3) $n = 50$ の場合には, $Var(\hat{\tau}_1)/Var(\hat{\tau}_2)$ は 50 を超えており, 小標本による影響がでていることがわかる

5. ま と め

本論文では, 変数間の因果関係が因果ダイアグラムと対応する線形構造方程式モデルにより記述できるとき, 条件付き操作変数法による総合効果の推定量の漸近分散を明示した. また, ある共変量集合を与えたときに条件付き操作変数となる変数が複数あり, それらがグラフ上で特定の関係にあるときに, 総合効果の推定精度という点から, 変数間の優劣関係について考察した. これは, 条件付き操作変数法を用いて総合効果を推定する際の操作変数選択基準として有用と考える.

本論文では, ある変数が条件付き操作変数となるような共変量が複数ある場合において, それら共変量間の優劣関係に関する考察をおこなわなかった. (12) 式の分子および分母がともに共変量に依存することから, この問題について定量的な優劣関係を考察することは可能であるが, 定性的な優劣関係を与えることは容易ではない. これについては今後の課題である. また, 推定精度に基づく条件付き操作変数間の優劣関係は, 小標本においても (12) 式から導かれる結果と一致していることが数値実験より確認できた. しかし, (12) 式は分散の十分な近似となっていない. より正確な近似法を確立することも今後の課題としたい.

謝 辞 有益な助言をくださった査読者に感謝の意を表します.

参 考 文 献

- Anderson, T.W. (1986): *Introduction to Multivariate Statistical Analysis*, Second edition. John Wiley & Sons.
- Bowden, R.J., and Turkington, D.A. (1984): *Instrumental Variables*, Cambridge University Press.
- Brito, C. and Pearl, J. (2002): Generalized instrumental variables. *Uncertainty in Artificial Intelligence* **18**, 85–93.
- Geiger, D., Verma, T.S., and Pearl, J. (1990): Identifying independence in Bayesian networks. *Networks* **20**, 507–534.
- 黒木学・宮川雅巳 (1999a): 因果ダイアグラムにおける介入効果の推定と工程解析への応用, 品質 **29**, 237–247.
- 黒木学・宮川雅巳 (1999b): 適応制御における条件付き介入効果の定式化とその推定, 品質 **29**, 476–486.
- 黒木学・宮川雅巳 (2002): 線形構造方程式モデルにおける同時介入効果の線形回帰母数による表現, 応用統計学 **31**, 107–121.
- Lauritzen, S.L. (1996): *Graphical models*, Clarendon Press, Oxford.
- 奥野忠一・片山善三郎・上郡長昭・伊東哲二・入倉則夫・藤原信夫 (1986): 工業における多変量データの解析, 日科技連, 東京.
- Pearl, J. (1995): Causal Diagrams for Empirical Research. *Biometrika* **82**, 669–709.
- Pearl, J. (1998): Graphs, causality and structural equation models, *Sociological Methods and Research* **27**, 226–284.
- Pearl, J. (2000): *Causality: Models, Reasoning, and Inference*. Cambridge University Press.
- Pearl, J. (2001): Parameter Identification: A New Perspective. UCLA Cognitive Systems Laboratory, Technical Report R-276.

(2003 年 5 月 28 日受付 8 月 27 日最終修正 9 月 1 日採択)

著者連絡先: 黒木 学

〒 560-8531 大阪府豊中市待兼山町 1-3 大阪大学基礎工学研究科

Instrumental Variable (IV) Selection for Estimating Total Effects in Conditional IV Method

Manabu KUROKI^{1,*}, Masami MIYAKAWA² and Ryouhei KAWATA³

¹ Graduate School of Engineering Science, Osaka University

1–3, Machikaneyama-cho, Toyonaka, Osaka, 560–8531, Japan

² Department of Industrial Engineering and Management, Tokyo Institute of Technology,

2–12–1, Oh-okayama, Meguro-ku, Tokyo, 152–8552, Japan

³ Graduate School of Economics, Kyoto University,

Yoshida-Honmachi, Sakyo-ku, Kyoto, 606–8501 Japan

Abstract

Suppose that causal relationships among variables can be described by a causal diagram and the corresponding linear structural equation model. In order to identify total effects, Brito and Pearl (2002) have suggested the conditional instrumental variable method, which is provided as a generalization of the instrumental variable method. However, the selection of conditional instrumental variable given a set of covariates is not unique. In this paper, when a total effect can be estimated through the conditional instrumental variable method, we formulate the asymptotic variance of the total effect. In addition, when there are some conditional instrumental variables given a set of covariates, the difference between them is investigated based on the formulation.

Key words: back door criterion, conditional independencies, confounding factor, path coefficient.

* Corresponding author.

████████████████████ (Manabu Kuroki).

Received May 28, 2003; Received in final form August 27, 2003; Accepted September 1, 2003.