

未観測交絡因子が存在する場合の 総合効果の識別可能条件

大阪大学大学院基礎工学研究科 黒 木 学

要 旨 本論文では、変数間の因果関係が非巡回的有向グラフと対応する線形構造方程式モデルで記述できると仮定する。このとき、処理変数と反応変数の間に未観測交絡因子が存在する場合において、処理変数に対して外的操作を行ったときの反応変数への総合効果を推測する問題を考える。この状況に対処するための総合効果の識別可能条件として、フロントドア基準 (Pearl, 2000) や条件付き操作変数法 (Brito and Pearl, 2002) が知られている。これらの識別可能条件は、総合効果を推定するために、未観測交絡因子とは直接的には関連しない第3の観測変数を利用するというアイデアに基づいている。しかし、現実の問題においては、未観測交絡因子は処理変数や反応変数だけでなく、さまざまな変数に影響を与えていることが多い。このような状況に対応するために、本論文では、因子モデルの識別可能条件を改良した新たな総合効果の識別可能条件を与える。本論文の結果は、総合効果の識別可能条件として有用であるだけでなく、因子モデルの識別可能条件の新たな見方与えている。

1. はじめに

線形構造方程式モデルを用いた統計的因果分析法は社会学、経済学、生物学等で幅広く活用されており、その起源は Wright (1923, 1934) のパス解析まで遡ることができる。その統計的因果分析法は、現在では因果ダイアグラムの理論として確率論的観点から整備されている (Pearl, 2000)。

因果関係を記述する統計モデルとして線形構造方程式モデルが与えられたとき、パス係数は単なる関連性の尺度ではなく「変数間の直接的な因果の強さ」を表した因果パラメータを意味する。そのとき、線形構造方程式モデルの重要な側面のひとつとして、総合効果の概念と定式化がある (Bollen, 1989)。統計的因果モデルとして線形構造方程式モデルが仮定できるとき、総合効果はある処理変数を外的操作により 1 単位変化させたときの反応変数の期待値の変化量を意味しており、その大きさは 2 つの変数を結ぶ有向道のそれぞれに対応するパス係数の積の総和で記述される。交絡因子を同定し解析に取り入れることが総合効果を推定する上で重要であることは、これまでもくりかえし指摘されてきた。

その一方で、交絡因子を観測しなければならないにもかかわらず、その作業を行うことが困難な場合がある。例えば、品質管理における工程解析では、解析に先立って、特性要因図を用いて品質特性に影響を与える要因を明らかにする活動が行われている。その活動をとおして多くの共

変量の存在が明らかにされる一方で、技術的あるいはコスト的な問題からそれらのいくつかについては観測することが困難な場合が多い。そのような共変量の中に交絡因子がある場合には、回帰モデルによる総合効果の推定は困難となる。このような問題を解決するために、Pearl (1995, 2000) は、総合効果を推定するための中間変数選択基準としてフロントドア基準を提案した。また、Brito and Pearl (2002) は、操作変数法 (Bowden and Turkington, 1984) を拡張した条件付き操作変数法を開発した。これらの識別可能条件は、総合効果を推定するために未観測交絡因子とは直接的には関連しない操作変数や中間変数といった第3変数を利用するというアイデアに基づいている。ところが、実際の問題においては、未観測交絡因子は処理変数や反応変数だけでなく、さまざまな要因とも関連していることが多い。Pearl (1995, 2000) や Brito and Pearl (2002) が与えた識別可能条件は、このような状況に対処することができない。

このような状況を踏まえて、本論文では、変数間の因果関係が非巡回的有向グラフと対応する線形構造方程式モデルによって記述できると仮定する。このとき、処理変数と反応変数の間に未観測交絡因子が存在する状況において、フロントドア基準 (Pearl, 1995, 2000) や条件付き操作変数法 (Brito and Pearl, 2002) とは異なる総合効果の識別可能条件を提案することが本論文の目的である。この識別可能条件は、Grzebyk et al. (2004), Stanghellini (1997), Vicard (2000) が提案した因子モデルの識別可能条件を改良することにより得られるものである。しかし、これらの識別可能条件とは異なり、本論文で提案する識別可能条件はモデル全体の識別可能性については問題としない。すなわち、現実の問題では統計的因果モデル自身が識別できないという状況がしばしば起こるが、このような場合においても本論文で与える条件を満たせば総合効果は識別可能となる。このようなことから、本論文の結果は、総合効果の識別可能条件として有用であるだけでなく、因子モデルの識別可能条件の新たな見方を与えるものとなっている。

2. パスダイアグラムと線形構造方程式モデル

2.1. グラフ

2.1.1. 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。 β の親全体からなる集合を $pa(\beta)$ と記す。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。 α の子孫全体からなる集合を $de(\alpha)$ と記すとき、 $V \setminus de(\alpha) \cup \{\alpha\}$ の要素を α の非子孫という。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。長さ n の有向道 $\alpha_0, \alpha_1, \dots, \alpha_n$ で、 $\alpha_0 = \alpha_n$ を許したものを長さ n の巡回閉路といい、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。

2.1.2. 無向グラフ

頂点の集合 V と、その直積 $V \times V$ の部分集合である辺の集合 E の組 $G = (V, E)$ において、2つの頂点 $\alpha, \beta \in V$ に対して $(\alpha, \beta) \in E$ であるとき、 α と β を向きのない辺(無向辺)で結ぶ。このとき、 (α, β) と (β, α) は区別しない。無向辺のみから構成されるグラフ G を無向グラフという。異なる頂点の列 $\alpha_0, \alpha_1, \dots, \alpha_n$ は、すべての $i = 1, \dots, n$ で $(\alpha_{i-1}, \alpha_i) \in E$ であるとき、長さ n の道という。道で結ばれた頂点は連結しているといい、すべての頂点が連結している無向グラフを連結グラフという。極大である連結部分グラフを連結成分という。長さ n の道 $\alpha_0, \alpha_1, \dots, \alpha_n$ で、 $\alpha_0 = \alpha_n$ を許したものを長さ n の閉路という。 $\bar{E} = V \times V \setminus E$ とするとき、 $\bar{G} = (V, \bar{E})$ を G の補グラフという。

2.2. 線形構造方程式モデル

統計的因果分析では、確率変数間の因果関係を非巡回的有向グラフと対応する線形構造方程式モデルによって表現する。このグラフをパスダイアグラムという。本論文では、その数学的定義を以下のように与える。

定義 1(パスダイアグラム) 非巡回的有向グラフ G とその頂点に対応する連続型確率変数の集合 $V = \{V_1, V_2, \dots, V_m\}$ が与えられている。グラフ G が確率変数間の関数関係を線形構造方程式モデル

$$V_i = \mu_{v_i, pa(v_i)} + \sum_{v_j \in pa(v_i)} \alpha_{v_i v_j} V_j + \epsilon_i \quad (i = 1, 2, \dots, m) \quad (1)$$

なる形に規定し、データがこの関数関係にしたがって生成されるとき、グラフ G をパスダイアグラムという。ここに、誤差変数 $\epsilon_1, \epsilon_2, \dots, \epsilon_m$ は多変量正規分布にしたがい、その平均ベクトルは 0 とする。また、 $\mu_{v_i, pa(v_i)}$ は定数である。さらに、 $\alpha_{v_i v_j}$ は V_j から V_i への矢線に対応する係数でパス係数と呼ばれるものであり、 $V_j \in pa(V_i)$ に対して $\alpha_{ij} \neq 0$ とする。□

ここで、いくつか記号を定義する。 \mathbf{W} を与えたときの X の条件付き分散を $\sigma_{xx \cdot w}$ と記し、 \mathbf{W} を与えたときの X と Y の条件付き共分散を $\sigma_{xy \cdot w}$ と記す。 \mathbf{W} を与えたときの \mathbf{X} と \mathbf{Z} の共分散行列を $\Sigma_{xz \cdot w}$ と記し、 \mathbf{W} を与えたときの \mathbf{Z} の条件付き共分散行列を $\Sigma_{zz \cdot w}$ と記す。また、 Y を目的変数、 X と \mathbf{W} を説明変数とした回帰モデルでの X の偏回帰係数を $\beta_{y \cdot x \cdot w} = \sigma_{xy \cdot w} / \sigma_{xx \cdot w}$ と記す。さらに、 $\mathbf{X}, \mathbf{Z}, \mathbf{W}$ に対して、 $B_{xz \cdot w} = \Sigma_{xz \cdot w} \Sigma_{zz \cdot w}^{-1}$ とおき、 \mathbf{Z} が単一変数である場合にも記号の区別は行わないものとする。他の分散、共分散、回帰係数などについても同様に記し、 \mathbf{W} が空集合である場合には、上記の記号の添え字から \mathbf{W} を省略することとする。 V_j から V_i 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、その総和を V_j から V_i への総合効果と呼び、 $\tau_{v_i v_j}$ と記す。

線形構造方程式モデルが非巡回的有向グラフにしたがって与えられているとき、いくつかの条件付き独立性が成り立っている。これを記述するために、有向分離 (Pearl, 1988) の概念を用いる。

定義 2(有向分離) 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 \mathbf{Z} が次の条件のいずれかを満たすとき、 \mathbf{Z} は X と Y を有向分離するという。

1. X と Y を結ぶ道に合流点があるとき、その合流点とその子孫は \mathbf{Z} に含まれない。
2. X と Y を結ぶ道に非合流点があるとき、その非合流点の中で \mathbf{Z} に含まれるものがある。□

パスダイアグラム G において、 Z が X と Y を有向分離するとき、頂点に対応する確率変数において、 Z を与えたとき X と Y は条件付き独立である (Geiger et al., 1990).

なお、本論文では、変数集合 V の同時分布は対応するパスダイアグラムに忠実である (Spirtes, et al., 2001), すなわち、パスダイアグラムで規定される(条件付き)独立関係以外のいかなる独立性も付加的に成り立っていないと仮定する。

2.3. 総合効果の識別可能条件

一般に、 X から Y への総合効果を推定するためには、 X と Y 以外の変数の観測が必要になる。ある処理変数への外的操作を念頭において解析を進めるとき、総合効果を推定するのに観測すべき変数集合を認識することは重要である。その総合効果の識別可能条件として、バックドア基準、フロントドア基準 (Pearl, 1995, 2000), 操作変数法 (Bowden and Turkington, 1984; Brito and Pearl, 2002) が知られている。以下にこれらの識別可能条件を与える。ここに、総合効果やパス係数のようなパラメータが観測変数の分散および共分散によって記述されるとき、そのパラメータは識別可能であるという。

定義 3(バックドア基準) パスダイアグラム G において、 X は Y の非子孫であるとする。このとき、次の2条件を満たす変数集合 Z は順序対 (X, Y) についてバックドア基準を満たすという。

1. X から Z の任意の要素へ有向道がない。
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。□

バックドア基準を満たす変数集合 Z が観測できるとき、 X から Y への総合効果 τ_{yx} は $\beta_{yx.z}$ により与えられる (Pearl, 2000)。

定義 4(フロントドア基準) パスダイアグラム G において、 X は Y の非子孫であるとする。このとき、次の3条件を満たす変数集合 Z は順序対 (X, Y) についてフロントドア基準を満たすという。

1. X から Y への任意の有向道に Z の要素がある。
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、空集合は X と Z の任意の要素を有向分離する。
3. G より Z の要素から出る矢線をすべて除いたグラフにおいて、 X は Y と Z の任意の要素を有向分離する。□

定義4の Z のように、処理変数から反応変数への有向道上にある変数を中間変数と呼ぶ。フロントドア基準を満たす変数集合 Z が観測できるとき、 X から Y への総合効果 τ_{yx} は $B_{yz.x}B_{zx}$ により与えられる (Pearl, 2000)。

定義 5(条件付き操作変数) パスダイアグラム G において、 X は Y の非子孫であるとする。このとき、次の2条件を満たす変数 Z は変数集合 W を与えたときの順序対 (X, Y) に対する条件付き操作変数であるという。

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. G より X からでる矢線を除いたグラフにおいて, \mathbf{W} は Z と Y を有向分離するが, Z と X を有向分離しない. \square

\mathbf{W} を与えたときの条件付き操作変数 Z が観測できるとき, 総合効果 τ_{yx} は $\sigma_{yz \cdot w} / \sigma_{xz \cdot w}$ により与えられる (Brito and Pearl, 2002). 特に, \mathbf{W} が空集合である場合には, 定義 5 は操作変数法 (Bowden and Turkington, 1984) のグラフィカル表現を与えている.

3. 未観測交絡因子が存在する場合の総合効果の識別可能条件

3.1. 準備

本節では, 総合効果の識別可能条件を与えるために必要となる基本定理を与える.

補題 1 $\{X, Y\} \cup S \cup T$ が多変量正規分布にしたがうとき,

$$\beta_{yx \cdot s} = \beta_{yx \cdot st} + B_{yt \cdot xs} B_{tx \cdot s} \quad (2)$$

$$\Sigma_{tt \cdot x} = \Sigma_{tt \cdot xs} + B_{ts \cdot x} \Sigma_{ss \cdot x} B'_{ts \cdot x} = \Sigma_{tt \cdot xs} + \Sigma_{ts \cdot x} \Sigma_{ss \cdot x}^{-1} \Sigma'_{ts \cdot x} \quad (3)$$

が成り立つ. \square

(2) 式と (3) 式は, それぞれ Cochran (1938) と Whittaker (1990) によるものである. ここに, (3) 式は T が単一変数である場合にも成り立つことに注意する.

また, Wermuth (1989) は次の補題を与えている.

補題 2 $\{X, Y\} \cup S \cup T$ が多変量正規分布にしたがうとき, S を与えたときに T と X が条件付き独立, あるいは $\{X\} \cup S$ を与えたときに Y と T が条件付き独立であるならば, $\beta_{yx \cdot st} = \beta_{yx \cdot s}$ が成り立つ. また, $S \cup \{X\}$ を与えたときに T が Y と条件付き独立であるならば, $\sigma_{yy \cdot xst} = \sigma_{yy \cdot xs}$ が成り立つ. \square

3.2. 共分散行列の表現

$\mathbf{X} = \{X_1, \dots, X_p\}$ を観測変数集合, $\mathbf{U} = \{U_1, \dots, U_k\}$ を非観測変数集合とした線形構造方程式モデル

$$X_i = \sum_{l=1}^k \alpha_{x_i u_l} U_l + \epsilon_{x_i}, \quad (i = 1, 2, \dots, p) \quad (4)$$

を考えよう ($k < p$). $k = 1$ のとき, この線形構造方程式モデルは 1 因子モデルと呼ばれ, $k \geq 2$ のとき多因子モデルと呼ばれる.

このとき, \mathbf{X} の共分散行列は

$$\Sigma_{xx} = \Sigma_{xx \cdot u} + \Sigma_{xu} \Sigma_{uu}^{-1} \Sigma_{ux} = \Sigma_{xx \cdot u} + \Omega(u) \Omega(u)' \quad (5)$$

なる構造で与えられる. ここに, $\Omega(u)$ は $p \times k$ 次行列である. また, (5) 式に対する逆行列は,

$$\Sigma_{xx}^{-1} = (\sigma^{x_i x_j}) = \Sigma_{xx \cdot u}^{-1} - \Lambda(u) \Lambda(u)' \quad (6)$$

と書くことができる (Rao, 1973). ここに, $\Lambda(u)$ は $p \times k$ 次行列である.

3.3. 1 因子モデルに基づく総合効果の識別可能条件

一般に, 第 2.3 節で与えた識別可能条件が総合効果を推定するための変数選択基準として有用であることはよく知られている (例えば, 黒木・宮川 (1999a, 1999b), 黒木他 (2003), 宮川・黒木 (1999), 宮川 (2004)). その一方で, Stanghellini (2004) は, 現実の問題では, これらの識別可能条件を満たす変数集合を観測できない場合があることを指摘している. このような問題に対処するためには, これまでに提案された識別可能条件とは異なる視点に基づいて, 新たな変数選択基準を開発することが必要となる. このような問題意識に基づいて, 本節では, 1 因子モデルの識別可能条件を利用した総合効果の識別可能条件を提案する.

まず, 第 2.3 節で与えた識別可能条件が成り立たない例として, 図 1 のパスダイアグラムにおいて X から Y への総合効果を推定する問題を考えよう. ここに, U は非観測変数, $\{X, Y, Z, W\}$ は観測変数集合である.

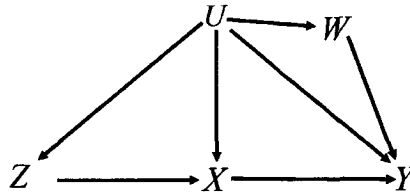


図 1. パスダイアグラム

図 1 のパスダイアグラムでは, U が非観測変数であるため, バックドア基準を満たす変数集合を観測することはできない. また, X と Y は矢線で結ばれているため, フロントドア基準を満たす変数集合も存在しない. さらに, U から観測変数のそれぞれへの矢線が存在しているため, 条件付き操作変数法のフレームワークも適用できない.

図 1 のパスダイアグラムより W から Y への矢線を取り除いたグラフに対応する線形構造方程式モデル (すなわち, $\alpha_{yw} = 0$ としたもの) を考えた場合, 対応する共分散構造は識別可能な 1 因子モデルのそれと一致する. したがって, X から Y へのパス係数 α_{yx} は識別可能である (Stanghellini, 2004). しかし, 図 1 では W から Y への矢線が存在しているため, 観測共分散の個数 (6 個) よりもパス係数の個数 (7 個) のほうが多くなる. このため, U の分散が既知であっても, 統計的因果モデル自身を識別することはできない. しかし, パス係数 α_{yx} は識別可能であり,

$$\alpha_{yx} = \frac{\sigma_{xw}\sigma_{yz} - \sigma_{zw}\sigma_{xy}}{\sigma_{xw}\sigma_{zx} - \sigma_{zw}\sigma_{xx}}$$

で与えられる. これを定理として与えると次のようになる.

定理 1 パスダイアグラム G において, $\{X, Y, W, Z, U\} \cup T$ が次の条件を満たすとする.

1. $\{X, U\} \cup T$ は $\{Y\}$ と $\{Z\}$ を有向分離する,
2. $\{U\} \cup T$ は $\{X, Z\}$ と $\{W\}$ を有向分離する,
3. $\{X\} \cup T$ は $\{Z\}$ と $\{W\}$ を有向分離しない.

X が Y の非子孫であるとき, $\{U\} \cup T$ が順序対 (X, Y) についてバックドア基準を満たすならば, X から Y への総合効果 τ_{yx} は $\{X, Y, W, Z\} \cup T$ を観測することによって識別可能となり,

$$\tau_{yx} = \frac{\sigma_{xw \cdot t} \sigma_{yz \cdot t} - \sigma_{zw \cdot t} \sigma_{xy \cdot t}}{\sigma_{xw \cdot t} \sigma_{zx \cdot t} - \sigma_{zw \cdot t} \sigma_{xx \cdot t}} \quad (7)$$

で与えられる. □

証明 $\{U\} \cup T$ が順序対 (X, Y) についてバックドア基準を満たすことから, $\tau_{yx} = \beta_{yx \cdot ut}$ が成り立つ. また, 補題 1(2) 式より,

$$\begin{aligned} \beta_{yz \cdot t} &= \beta_{yz \cdot xut} + \beta_{yx \cdot uzt} \beta_{xz \cdot t} + \beta_{yu \cdot xzt} \beta_{uz \cdot t}, & \beta_{yx \cdot t} &= \beta_{yx \cdot ut} + \beta_{yu \cdot xt} \beta_{ux \cdot t}, \\ \beta_{zx \cdot t} &= \beta_{zx \cdot wt} + \beta_{zw \cdot xt} \beta_{wx \cdot t}, & \beta_{zw \cdot t} &= \beta_{zw \cdot xt} + \beta_{zx \cdot tw} \beta_{xw \cdot t} \end{aligned}$$

が直ちに得られる. ここで, 条件 1 より $\{X, U\} \cup T$ を与えたときに Y と Z は条件付き独立になることから, $\beta_{yz \cdot xut} = 0$ が成り立つ. また, 補題 2 より, $\beta_{yx \cdot uzt} = \beta_{yx \cdot ut}$, $\beta_{yu \cdot xzt} = \beta_{yu \cdot xt}$ を得ることができる. これらのことから

$$\beta_{yz \cdot t} = \beta_{yx \cdot ut} \beta_{xz \cdot t} + \beta_{yu \cdot xt} \beta_{uz \cdot t}$$

が成り立つ. ここで, $\beta_{yz \cdot t} = \sigma_{yz \cdot t} / \sigma_{zz \cdot t}$, $\beta_{yx \cdot t} = \sigma_{xy \cdot t} / \sigma_{xx \cdot t}$, $\beta_{zx \cdot t} = \sigma_{xz \cdot t} / \sigma_{xx \cdot t}$, $\beta_{zw \cdot t} = \sigma_{zw \cdot t} / \sigma_{ww \cdot t}$ であることから, 上式は

$$\begin{aligned} \sigma_{yz \cdot t} &= \sigma_{xz \cdot t} \beta_{yx \cdot ut} + \sigma_{uz \cdot t} \beta_{yu \cdot xt}, & \sigma_{xy \cdot t} &= \beta_{yx \cdot ut} \sigma_{xx \cdot t} + \sigma_{ux \cdot t} \beta_{yu \cdot xt}, \\ \sigma_{xz \cdot t} &= \sigma_{xx \cdot t} \beta_{zx \cdot tw} + \beta_{zw \cdot tx} \sigma_{wx \cdot t}, & \sigma_{zw \cdot t} &= \sigma_{ww \cdot t} \beta_{zw \cdot xt} + \beta_{zx \cdot tw} \sigma_{xw \cdot t} \end{aligned}$$

と変形できる. 以上のことから

$$\sigma_{xw \cdot t} \sigma_{yz \cdot t} - \sigma_{zw \cdot t} \sigma_{xy \cdot t} = \beta_{yx \cdot ut} (\sigma_{xw \cdot t} \sigma_{xz \cdot t} - \sigma_{zw \cdot t} \sigma_{xx \cdot t}) + \beta_{yu \cdot xt} (\sigma_{xw \cdot t} \sigma_{uz \cdot t} - \sigma_{zw \cdot t} \sigma_{ux \cdot t})$$

を得ることができる. ここで, 条件 3 と忠実性より $\beta_{zw \cdot tx} \neq 0$ が成り立つことに注意すると, 補題 1(3) 式より, 第 1 項について,

$$\sigma_{xw \cdot t} \sigma_{xz \cdot t} - \sigma_{zw \cdot t} \sigma_{xx \cdot t} = -\beta_{zw \cdot tx} \sigma_{ww \cdot t} (\sigma_{xx \cdot t} - \beta_{xw \cdot t} \sigma_{wx \cdot t}) = -\beta_{zw \cdot tx} \sigma_{ww \cdot t} \sigma_{xx \cdot tw} \neq 0$$

を得る. 一方, 補題 1(2) 式より,

$$\beta_{xw \cdot t} = \beta_{xw \cdot ut} + \beta_{xu \cdot tw} \beta_{uw \cdot t}, \quad \beta_{zw \cdot t} = \beta_{zw \cdot ut} + \beta_{zu \cdot tw} \beta_{uw \cdot t}$$

が得られる. ここで, 条件 2 より $\beta_{xw \cdot ut} = \beta_{zw \cdot ut} = 0$ を得ることができることに注意すると, $\sigma_{xw \cdot t} = \beta_{xu \cdot t} \sigma_{uw \cdot t}$ と $\sigma_{zw \cdot t} = \beta_{zu \cdot t} \sigma_{uw \cdot t}$ が得られることから

$$\sigma_{xw \cdot t} \sigma_{uz \cdot t} - \sigma_{zw \cdot t} \sigma_{ux \cdot t} = 0$$

が成り立つ. これらをまとめることによって, 定理を得ることができる. □

定理 1 では, U の分散は既知でなくてもよいことに注意する.

3.4. 多因子モデルに基づく総合効果の識別可能条件

前節では、未観測交絡因子が1つである状況を考え、その下での総合効果の識別可能条件を与えた。しかし、一般には、未観測交絡因子が数多く存在するという状況は少なくない。そこで、本節では、複数の未観測交絡因子が存在する場合の総合効果の識別可能条件を与える。

ここで記号の定義を行う。 $X_i, X_j \in \mathbf{X}$ に対して、 \mathbf{U} を与えたときに X_i と X_j が条件付き独立である場合には、 X_i と X_j を無向辺で結ばないという手続きによって得られる \mathbf{X} を頂点集合とする無向グラフを $G_{cov}^{x \cdot u}$ と表記する。また、 $\mathbf{U} \cup \mathbf{X} \setminus \{X_i, X_j\}$ を与えたときに X_i と X_j が条件付き独立である場合には、 X_i と X_j を無向辺で結ばないという手続きによって得られる \mathbf{X} を頂点集合とする無向グラフを $G_{con}^{x \cdot u}$ と表記する。以上の準備の下で、Stanghellini and Werumth (2005) によって与えられた1因子モデルの識別可能条件を紹介する。

定理 2 1因子モデルによって規定される共分散行列において、 $\Omega(u)\Omega(u)'$ (あるいは $\Lambda(u)\Lambda(u)'$) と $\Sigma_{xx \cdot u}$ が唯一の解を持つための必要十分条件は、次のいずれかが成り立つことである。

1. (5) 式において $\Omega(u) \neq 0$ であって、 $G_{cov}^{x \cdot u}$ の補グラフに含まれる連結成分すべてが長さが奇数の閉路を持つ。
2. (6) 式において $\Lambda(u) \neq 0$ であって、 $G_{con}^{x \cdot u}$ の補グラフに含まれる連結成分すべてが長さが奇数の閉路を持つ。 □

因子モデルの識別可能条件でしばしば使われる σ_{uu} の値が既知であるという条件は、 Σ_{uu} を識別するために使われるものであり、 $\Sigma_{xx \cdot u}$ を識別するには必要ではない (Stanghellini and Werumth, 2005)。このことに注意することにより、次の定理が成り立つ。

定理 3 パスダイアグラム G において、 $\mathbf{X} = \{X_1, \dots, X_p\}$ を観測変数集合とし、 $\mathbf{U} = \{U_1, \dots, U_k\}$ を非観測変数集合とする。このとき、 $\{U_1, \dots, U_{i-1}\}$ を条件付け $\{U_{i+1}, \dots, U_k\}$ で周辺を取るによって構成された新たな線形構造方程式モデルを非観測変数 U_i に関する1因子モデルとみなしたとき、任意の i に対して定理2が成り立つならば、 $\Sigma_{xx \cdot u}$ は識別可能である。 □

証明 まず、 $\{U_2, \dots, U_k\}$ で周辺を取った線形構造方程式モデルに対応する共分散行列は、

$$\Sigma_{xx} = \Sigma_{xx \cdot u_1} + \frac{1}{\sigma_{u_1 u_1}} \Sigma_{xu_1} \Sigma'_{xu_1}$$

で与えられる。このとき、仮定より $\Sigma_{xx \cdot u_1}$ は識別可能である。

ここで、 $i(\geq 2)$ に対して、 $\Sigma_{xx \cdot u_1 \dots u_{i-1}}$ が識別可能であると仮定する。このとき、補題1(3)式より、 $\{U_1, \dots, U_{i-1}\}$ を条件付け、 $\{U_{i+1}, \dots, U_k\}$ で周辺を取った線形構造方程式モデルに対応する共分散行列は、

$$\Sigma_{xx \cdot u_1 \dots u_{i-1}} = \Sigma_{xx \cdot u_1 \dots u_i} + \frac{1}{\sigma_{u_i u_i \cdot u_1 \dots u_{i-1}}} \Sigma_{xu_i \cdot u_1 \dots u_{i-1}} \Sigma'_{xu_i \cdot u_1 \dots u_{i-1}}$$

で与えられる。仮定より、 $\Sigma_{xx \cdot u_1 \dots u_i}$ は識別可能である。この手順を繰り返すことにより、定理を得ることができる。 □

定理 3 が成り立つとき, $\{X, Y\} \cup Z \subset X$ に対して, $Z \cup U$ が順序対 (X, Y) についてバックドア基準を満たすならば, 総合効果 τ_{yx} は識別可能となる. このとき, $\{U_1, \dots, U_k\}$ に含まれる要素の分散は既知でなくてもよいことに注意する. また, $\{U_1, \dots, U_k\}$ の分散が既知で周辺独立となっている場合には, 多因子モデルの識別可能条件 (Grzebyk, et al., 2004) としても利用できる.

例として, 図 2(a) で与えられるパスダイアグラムを考えよう. ここに, $X = \{X, Y, Z_1, Z_2, W_1, W_2\}$, $U = \{U_1, U_2\}$ とする.

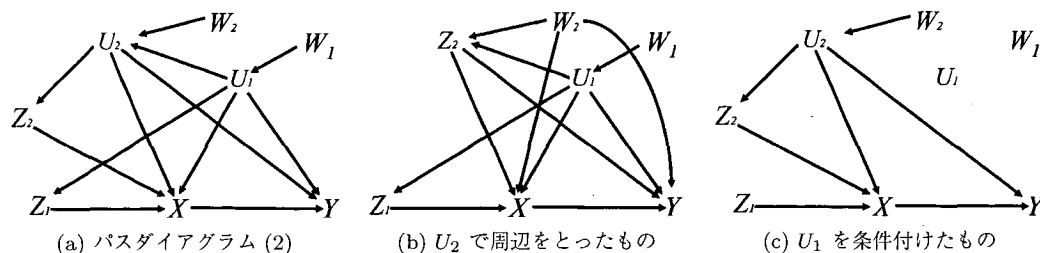


図 2. 定理 3 の適用過程

まず, U_2 で周辺をとった線形構造方程式モデルに対応するグラフを図 2(b) に与える. このグラフに基づいて構成される無向グラフ $G_{con}^{x:u_1}$ について, その補グラフに含まれる連結成分は 1 つであり, 長さが奇数の閉路 ($W_1 - Z_1 - Y - W_1$) が存在することがわかる.

次に, U_1 で条件付けた線形構造方程式モデルに対応するグラフを図 2(c) に与える. このグラフに基づいて構成される無向グラフ $G_{con}^{x:u_1 u_2}$ についてもその補グラフに含まれる連結成分は一つであり, 長さが奇数の閉路 ($W_2 - Z_1 - Y - W_2$) が存在することがわかる. $\{U_1, U_2\}$ が順序対 (X, Y) についてバックドア基準を満たしていることに注意すると, 以上のことから, 定理 3 より X から Y への総合効果は識別可能となる. その明示表現は, (5) 式の右辺をパス係数で記述し, それを α_{yx} について解くことによって

$$\tau_{yx} = \frac{\sigma_{yz_1} \sigma_{z_2 w_1} - \sigma_{y w_1} \sigma_{z_1 z_2}}{\sigma_{x z_1} \sigma_{z_2 w_1} - \sigma_{x w_1} \sigma_{z_1 z_2}}$$

で与えられる.

4. 適用例

前節で与えた識別可能条件を奥野他 (1986) の “ボディ塗装条件の設定” データに適用する. 自動車のボディ塗装は, 防錆性能と美観性能を付与するために行われるものであるが, このとき塗装膜厚を確保することが不可欠である. 塗装膜厚は, 下塗り・中塗り・上塗りという 3 層から形成されるが, 中塗り・上塗りについては塗料を作業者が吹き付けるため, 塗着率が悪いという問題が生じていた. そこで, 奥野他 (1986) 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ために, スプレーガンに対してさまざまな条件を設定した実験を行い, 38 組のデータを採取している. ここに, 奥野他 (1986) が与えた標本相関行列および記述より, この工程解析事例では処理条件の無作為割り付けが実施されていないと判断できることから, 処理変数の存在する観察研究であるとみなすことができる (黒木・宮川 (1999a), 宮川

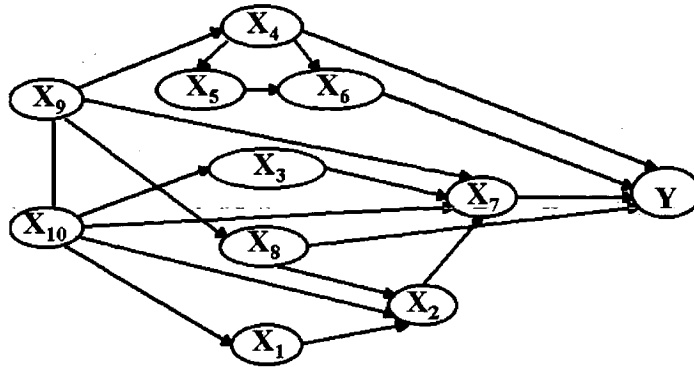


図 3. “ボディ塗装条件の設定” データによるパスダイアグラム(黒木・宮川, 1999a; 宮川, 2004)

表 1. 図 3 の因果モデルに基づく相関行列の推定値(黒木・宮川, 1999b; 宮川, 2004)

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	Y
X_1	1.000	-0.736	-0.152	0.148	0.028	-0.043	0.324	0.216	0.286	-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	-0.043	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	0.286	-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	-0.091	0.326	0.134	-0.614	-0.277	-0.250	-0.044	-0.493	-0.475	0.283	1.000

(2004)).

データ収集に先立って、塗着率 Y に影響を与えるさまざまな要因が特性要因図を用いて吟味されており、その結果、要因系変量として

“希釈率” X_1 , “粘度” X_2 , “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 , “塗料温度” X_8 , “温度” X_9 , “湿度” X_{10}

が選択されている。奥野他 (1986) の記述より X_1, \dots, X_6 が制御可能な処理変数であると推察されるが、本論文では X_2 と X_6 を処理変数として取り上げることとする。これらの処理変数に対して適切な工程条件の設定を行うためには、「処理変数を外的操作により一定の値に固定したときの塗着率の期待値」を評価することが不可欠であり、これは総合効果を評価することと等価である。このことから、 X_2 から Y への総合効果と X_6 から Y への総合効果を評価するために、どのような共変量集合を観測すればよいのかを明らかにすることが重要となる。

この工程解析事例は黒木・宮川 (1999a, 1999b) と宮川 (2004) でも議論されており、そこではパスダイアグラムとそれに基づいて推定された相関行列が与えられている。このパスダイアグラムと相関行列をそれぞれ図 3 と表 1 に与える。本節では、図 3 と表 1 に基づいて X_2 と X_6 のそれぞれから Y への総合効果を評価することを考えよう。

X_2 と X_6 のそれぞれに対して、定理 1 を利用したときの総合効果の推定値を表 2 に与える。

表 2. 総合効果の推定値

処理変数	共変量	総合効果
X_2	$Z = X_1, W = X_9, T = X_8$	-0.116
X_6	$Z = X_5, W = X_9, T = \varphi$	-0.465

表 2 では、第 1 列には処理変数、第 2 列には総合効果を識別するために用いる共変量集合、第 3 列には総合効果の推定値を記している。ここに、第 2 列の各要素に記述されている変数は、順に定理 1 の Z, W, T に対応している。

まず、 X_{10} を用いずに処理変数 X_2 から Y への総合効果を推定することを考えよう。このとき、バックドア基準や条件付き操作変数法に基づいて総合効果を推定することはできないことが図 3 より確認できる。また、中間変数との間に交絡因子が存在しているため、フロントドア基準に基づいて総合効果を推定することはできないことも容易にわかる。しかし、第 2 列に与えた共変量集合は順序対 (X_2, Y) について定理 1 の要件を満たすことから、(7) 式を用いて推定することができる。

同様に、 X_4 を用いずに処理変数 X_6 から Y への総合効果を推定することを考えよう。このとき、バックドア基準に基づいて総合効果を推定することはできないことが容易に確認できる。また、条件付き操作変数法の要件やフロントドア基準を満たす変数集合も存在しないため、これらの枠組みに基づいて総合効果を推定することはできないこともわかる。しかし、図 3 より第 2 列に与えた共変量集合が定理 1 の要件を満たすことが確認できるので、(7) 式を用いて総合効果を推定することができる。

5. 数値実験

第 4 節より、バックドア基準、条件付き操作変数法、本論文で提案した識別可能条件のいずれを用いても、総合効果が識別可能であると判断できる場合があることがわかる。そこで、本節では、図 4 に与えるパスダイアグラムに基づいて、総合効果の推定精度の観点からこれらの識別可能条件の優劣関係について考察しよう。図 4 では、 U は順序対 (X, Y) についてバックドア基準を満たしており、 Z は U を与えたときの条件付き操作変数となっている。また、 $\{X, Y, Z, W\}$ は定理 1 の要件を満たしている。これらのことから、いずれの識別可能条件を用いても総合効果を推定することができることがわかる。

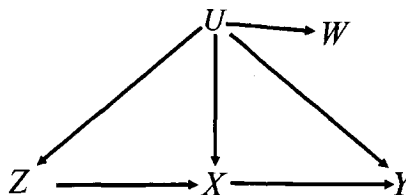


図 4. パスダイアグラム

ここで、図 4 のパスダイアグラムにおいて、識別可能条件のそれぞれに対して、推定量の分散が小さくなる状況を定性的に考察しよう。まず、バックドア基準を用いて総合効果を推定する場

合, α_{xu} の値が小さくなると推定量の分散も小さくなると推察される. また, 条件付き操作変数法を用いて総合効果を推定する場合, α_{xz} の値が大きくなると推定量の分散は小さくなると推察される. この考察に基づいて, α_{xu} と α_{xz} 以外のパス係数を 0.3 に固定した上で,

ケース 1: $\alpha_{xu} = 0.1, \alpha_{xz} = 0.1$

ケース 2: $\alpha_{xu} = 0.1, \alpha_{xz} = 0.5$

ケース 3: $\alpha_{xu} = 0.5, \alpha_{xz} = 0.1$

ケース 4: $\alpha_{xu} = 0.5, \alpha_{xz} = 0.5$

の 4 通りのケースを数値実験の対象とする.

標本サイズを $n = 500, 1000, 1500, 2000$ の 4 水準として, 各ケースにおいて上述のパス係数から決定される相関行列を求め, それに基づく 5 次元正規乱数を発生させた. この条件の下で, パスダイアグラム上のすべての変数を観測し, そこから得られる標本共分散を用いて総合効果を推定する. これを 3000 回繰り返し, 3 つの識別可能条件のそれぞれに基づいて総合効果の推定量の分散値を実験的に求めた. その結果を表 3 に示す.

表 3. 数値実験結果

	$(\alpha_{xu}, \alpha_{xz})$	$n = 500$	$n = 1000$	$n = 1500$	$n = 2000$
バックドア基準	(0.1, 0.1)	0.0018	0.0009	0.0006	0.0005
	(0.1, 0.5)	0.0019	0.0010	0.0006	0.0005
	(0.5, 0.1)	0.0022	0.0011	0.0007	0.0005
	(0.5, 0.5)	0.0025	0.0012	0.0008	0.0006
条件付き操作変数	(0.1, 0.1)	3.6228	0.2800	0.0761	0.0519
	(0.1, 0.5)	0.0104	0.0051	0.0035	0.0025
	(0.5, 0.1)	3.1562	0.1154	0.0641	0.0477
	(0.5, 0.5)	0.0056	0.0025	0.0018	0.0013
定理 1	(0.1, 0.1)	0.0119	0.0014	0.0009	0.0006
	(0.1, 0.5)	0.1919	0.0155	0.0027	0.0016
	(0.5, 0.1)	4.5311	0.0993	0.0118	0.0075
	(0.5, 0.5)	6.3753	1.1822	0.0347	0.0125

表 3 を観察することにより, 次の知見を得る. なお, 総合効果の推定精度に基づくバックドア基準と条件付き操作変数法の比較については Kuroki and Cai (2004) を参照されたい.

- すべてのケースにおいて, バックドア基準に基づいて推定された総合効果の分散よりも, 定理 1 を用いた場合の分散のほうが大きくなっており, 識別可能条件の選択による差異が現れている.
- $\alpha_{xz} = 0.1$ のとき, ケース 1 ではすべての標本サイズにおいて, ケース 3 では $n \geq 1000$ において, 定理 1 を用いた場合の分散のほうが条件付き操作変数法を用いた場合の分散よりも小さくなっている. このことから, α_{xz} の値が小さい場合には条件付き操作変数法を用いることによって推定精度が下がると推察される.
- $\alpha_{xu} = 0.1$ のとき, ケース 1 ではすべての標本サイズにおいて, ケース 2 では $n \geq 1500$ において, 定理 1 を用いた場合の分散のほうが条件付き操作変数法を用いた場合の分散よりも小さくなっている. このことから, α_{xu} の値が小さい場合には, 定理 1 に基づいて総合効果を

推定することで推定精度が上がると推察される。

6. ま と め

本論文では、変数間の因果関係が線形構造方程式モデルにより記述できるとき、未観測交絡因子が存在する場合の総合効果の識別可能条件を与えるとともに、奥野他 (1986) の工程解析事例をとおして、実質科学への適用可能性について検討した。これによって、本論文で与えた識別可能条件は、単なる Stanghelli (2004) の結果の拡張というだけでなく、Pearl (2000) や Brito and Pearl (2002) によって提案された識別可能条件が適用できない状況においても適用できるという点で有用であることが明らかとなった。また、総合効果の推定精度の観点からバックドア基準、条件付き操作変数法、本論文で提案した識別可能条件の優劣関係を調べるために数値実験を行った。その結果、本論文で提案した識別可能条件はバックドア基準を用いた場合よりも総合効果の推定精度は劣るものの、いくつかのケースでは条件付き操作変数法を用いた場合よりも優れていることが明らかとなった。

実質科学においては、因果関係が線形構造方程式モデルで記述できない状況も少なくない。より現実的な問題に適用するためには、ノンパラメトリック構造方程式モデルに基づく総合効果の識別可能条件の開発が必要であると考えられる。未観測交絡因子が存在する場合の総合効果の識別可能問題については、Shpitser and Pearl (2006a, 2006b) によって一般的な識別可能条件が与えられている。しかし、選択バイアス問題を解決するための統計的因果推論技術の開発については、Stanghellini and Wermuth (2005) や Kuroki and Cai (2006) で線形構造方程式モデルに基づく議論が行われているものの、ノンパラメトリック構造方程式モデルに基づくフレームワークは構築されていない。これについては今後の課題としたい。

謝 辞 本論文の執筆にあたって、有益な助言を下さった査読者および貴重な情報を提供していただいた京都大学の蔡志紅博士に深く感謝する。本研究は、文部科学省科学研究費補助金若手研究 (B)、財団法人倉田記念日立科学技術財団および財団法人栢森情報科学振興財団による研究助成を受けて行われた。

参 考 文 献

- Bollen, K.A. (1989): *Structural Equations with Latent Variables*. John Wiley and Sons.
- Bowden, R.J. and Turkington, D.A. (1984): *Instrumental Variables*. Cambridge University Press.
- Bruto, C. and Pearl, J. (2002): Generalized instrumental variables. *Uncertainty in Artificial Intelligence* **18**, 85–93.
- Cochran, W.G. (1938): The omission or addition of an independent variate in multiple linear regression. *Supplement to the Journal of the Royal Statistical Society* **5**, 171–176.
- Geiger, D., Verma, T.S. and Pearl, J. (1990): Identifying independence in bayesian networks. *Networks* **20**, 507–534.
- Grzebyk, M., Wild, P. and Chouaniere, D. (2004): On identification of multi-factor models with correlated residuals. *Biometrika* **91**, 141–151.
- Kuroki, M. and Cai, Z. (2004): Selection of identifiability criteria for total effects by using path diagrams. *Uncertainty in Artificial Intelligence* **20**, 333–340.
- Kuroki, M. and Cai, Z. (2006): On Recovering population's covariance matrix in the presence of selection

- bias, *Biometrika* **93**, 601–611.
- 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.
- 黒木学・宮川雅巳・川田亮平 (2003): 条件付き操作変数法の推定精度と操作変数の選択. 応用統計学 **32**, 89–100.
- 宮川雅巳 (2004): 統計的因果推論—新しい回帰分析の枠組—. 朝倉書店.
- 宮川雅巳・黒木学 (1999): 因果ダイアグラムにおける介入効果推定のための共変量選択. 応用統計学 **28**, 151–162.
- 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.
- Pearl, J. (1988): *Probabilistic Reasoning in Intelligence Systems*. Morgan Kaufmann.
- Pearl, J. (1995): Causal Diagrams for Empirical Research. *Biometrika* **82**, 669–710.
- Pearl, J. (2000): *Causality: Models, Reasoning, and Inference*. Cambridge University Press.
- Rao, C.R. (1973): *Linear Statistical Inference and Its Applications*. John Wiley and Sons.
- Shpitser, I. and Pearl, J. (2006a): Identification of conditional interventional distributions. *Uncertainty in Artificial Intelligence* **22**, 437–444.
- Shpitser, I. and Pearl, J. (2006b): Identification of joint interventional distributions in recursive semi-markovian causal models. *The 21st National Conference on Artificial Intelligence*, 1219–1226.
- Spirtes, P., Glymour, C. and Scheines, R. (2001): *Causation, Prediction, and Search; 2nd Edition*. The MIT Press.
- Stanghellini, E. (1997): Identification of a single-factor models using graphical Gaussian rules. *Biometrika* **84**, 241–244.
- Stanghellini, E. (2004): Instrumental variables in Gaussian directed acyclic graph models with an unobserved confounder. *Environmetrics* **15**, 463–469.
- Stanghellini, E. and Wermuth, N. (2005): On the identification of path analysis models with one hidden variable. *Biometrika* **92**, 337–350.
- Vicard, P. (2000): On the identification of a single factor model with correlated residuals. *Biometrika* **87**, 199–205.
- Wermuth, N. (1989): Moderating effects in multivariate normal distributions. *Methodika*, **3**, 74–93.
- Whittaker, J. (1990): *Graphical Models in Applied Multivariate Statistics*. John and Wiley.
- Wright, S. (1923): The theory of path coefficients: a reply to Nilés' criticism. *Genetics* **8**, 239–255.
- Wright, S. (1934): The method of path coefficients. *Annals of Mathematical Statistics* **5**, 161–215.

(2007 年 3 月 15 日受付 9 月 4 日最終修正 9 月 11 日採択)

著者連絡先: 〒 560-8531 大阪府豊中市待兼山町 1-3 黒木 学

Identifiability Criteria for Total Effects in the Presence of Unmeasured Confounders


Manabu Kuroki

Osaka University

Abstract

In this paper, we assume that cause-effect relationships between variables can be described as a directed acyclic graph and the corresponding linear structural equation model. Then, we consider the identification problem for total effects in the presence of unmeasured confounders. In order to solve the problem, the front door criterion (Pearl, 2000) and the conditional instrumental variable method (Brito and Pearl, 2002) are known as identifiability criteria for total effects. In this paper, we propose new graphical identifiability criteria for total effects based on those for factor models. The results of this paper are useful as the identifiability criteria for total effects and provide a new viewpoint to the identification problem of factor models.

Key words: back door criterion, path diagram

 (Manabu Kuroki)

Received March 15, 2007; Received in final form September 4, 2007; Accepted September 11, 2007.