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

東京工業大学 社会理工学研究科 黒 木 学

東京工業大学 社会理工学研究科 宮 川 雅 巳

要 旨 本論文では、変数間の因果関係が線形構造方程式モデルと因果ダイアグラムで記述できる場合に、いくつかの処理変量に対して外的操作をおこなったときの反応変量への因果的效果を同時介入効果と呼び、これを推測する問題を考える。Pearl and Robins (1995)によって定義された同時介入効果はノンパラメトリックな分布として与えられている。そこで、本論文では、その平均と分散に着目し、線形構造方程式モデルの下でこれらの特徴量の明示的表現を与える。次に、この同時介入効果の平均と分散を線形回帰モデルを用いて推定するためにはどのような回帰モデルを設定すべきかを考え、同時介入効果の平均と分散がその回帰モデルの母数によってどのように表現されるかを明らかにする。

1. はじめに

線形構造方程式モデルは社会学、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「パス係数の方法」にさかのぼることができる。線形構造方程式モデルはデータ生成過程を表現した統計的因果モデルであり、パス係数は単なる関連性の尺度ではなく「因果の強さ」を表した母数である。

線形構造方程式モデルを図的表現したパスダイアグラムが変数間の定性的因果関係を正しく記述しているとき、観察研究で採取された統計データよりパス係数の値を推定することができる。そのための一つの方法が回帰分析である。このとき、線形回帰モデルの母数である偏回帰係数は観測変数の同時分布の分散共分散行列によって表現される統計的関連性を示す尺度である。したがって、相関関係の程度を表す偏回帰係数が、因果関係の強さを表すパス係数に一致するための十分条件を明らかにする必要がある。ひとつの自明な十分条件は「ある変数を目的変数に取り、その変数の親(その変数への直接の矢線が出る変数)のすべてを説明変数にした回帰モデルでの偏回帰係数は、対応する矢線のパス係数に等しい」である。実は、ある特定のパス係数に関心がある場合には、親のすべてを説明変数にする必要は必ずしもない。また、親以外の変数については、説明変数に含めてもよい場合といけない場合とがある。これについてのより精緻な十分条件がPearl (1998)の与えた単一結合基準である。これはパスダイアグラム上でのグラフ用語によって記述される。このような基準の導入によりパスダイアグラムの有用性が再認識されている(Spirites et al.,

1998).

ところで、パス係数の方法の重要な側面に、直接効果と総合効果の概念および定式化がある (Bollen, 1989). 直接効果は一つの矢線に対応する効果で、その大きさはパス係数で記述される. 総合効果はすべての有向道に対応する効果の総和を意味し、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. 線形回帰モデルにおいて、ある説明変数の偏回帰係数が目的変数への総合効果に一致する十分条件として、バックドア基準が知られている (Pearl, 1995, 1998). この総合効果を一般的な構造方程式モデルへ拡張したものが、Pearl (1995)による介入効果である. この介入効果はある一つの変数に外的操作をおこなったときの因果的效果として与えられたもので、2.3節で定義するようにノンパラメトリックな分布として記述される. したがって、これを推定する立場からは、分布特性に着目することが現実的である. 黒木・宮川(1999a)は、線形構造方程式モデルの下で、介入効果の平均と分散を求めることで、平均に関する係数が従来の総合効果であることを認識するとともに、「分散に対する介入効果」という新たな指標を具現化した. 一方、Pearl and Robins (1995)は、外的操作の可能な変数がいくつかあるときに、これらに同時に外的操作をおこなったときの因果的效果を同時介入効果として定義した. さらに、バックドア基準を拡張した許容性基準という同時介入効果の識別するための十分条件を与えた. Kuroki and Miyakawa (1999)はこれとは別の十分条件を与えた.

さて、同時介入効果は各変数の固定された値に依存するので、線形構造方程式モデルを仮定したときにも、パス係数の積和のような一つの係数として指標化されるものではない. したがって、この場合も分布として定義される同時介入効果の平均と分散に着目し、それらが線形構造方程式モデルの母数によってどのように記述されるかを明示することが有用であると考ええる. さらに、これらを推定する際には、介入効果の場合と同様に、線形回帰モデルを利用できることが望ましい. そのためには、同時介入効果の平均と分散に対する一致推定量を得るために具備すべき線形回帰モデルの要件を明らかにし、かつ線形回帰モデルの母数によって同時介入効果の平均と分散を表現しておく必要がある. これらをおこなうことが本論文の目的である.

本論文の構成は以下のとおりである. 2節と3節では既往研究をレビューする. 2節では、介入効果の定義とこれを識別するための十分条件であるバックドア基準を紹介する. 次に、同時介入効果の定義とこれを識別するための十分条件である許容性基準について記述する. 3節では、線形回帰モデルの偏回帰係数が直接効果に一致するための十分条件として単一結合基準を紹介する. 次に、バックドア基準は線形回帰モデルの偏回帰係数が総合効果に一致するための十分条件であることを示す. 4節では、本論文の新しい成果として、線形構造方程式モデルにおいて同時介入効果の平均と分散がパス係数を用いてどのように記述できるかを明らかにする. 5節では、この結果に基づいて、この平均と分散を推定するために設定すべき線形回帰モデルを明らかにし、同時介入効果の平均と分散の明示的表現を与える. 6節では、工業における適用例を示す. 7節では、本論文のまとめをおこなう.

2. 因果ダイアグラムと介入効果

2.1. 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. また, V から α およびその子孫をすべて除いた集合に属す頂点を 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.

2.2. 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

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

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 への矢線をもつ親の集合であり, 誤差変数 $\epsilon_1, \epsilon_2, \dots, \epsilon_m$ は互いに独立であるとする. $\square$

$f(x_i|\text{pa}(x_i))$ 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. ここに, $\text{pa}(X_i)$ が空集合であるときには $f(x_i|\text{pa}(x_i)) = f(x_i)$ とする. 確率変数間の因果的関係が (1) 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. $\square$

【定理 1】 (Geiger et al., 1990)

因果ダイアグラム G において, 頂点集合 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、変数集合 Z を与えたとき X と Y は条件付き独立である。この関係を $X \perp\!\!\!\perp Y | Z$ と表記する。 □

2.3. 介入効果

因果ダイアグラムにおいて、Pearl (1995) は X の値を他の変数の値によらずに一定に固定する外的操作を介入と呼び、このときの Y の分布を介入効果として次のように定義した。

【定義 3】 介入効果

因果ダイアグラム G における頂点集合を $V = \{X, Y, Z_1, Z_2, \dots, Z_p\}$ とする。このとき、

$$f(y|\text{set}(X = x)) = \int_{z_1} \cdots \int_{z_p} \frac{f(x, y, z_1, \dots, z_p)}{f(x|\text{pa}(x))} dz_1 \cdots dz_p$$

を X の Y への介入効果という。ここに、 $\text{set}(X = x)$ は介入によって X の値を x にしたことを意味する。 □

この定義をグラフとして解釈すると、 G から X への矢線を取り除いたグラフを考え、そこで X の値を x に固定したときの Y の分布ということになる。一方、構造方程式モデルの立場から解釈すると、与えられた構造方程式モデルにおいて X を左辺に持つ構造方程式を定数関数 $X = x$ に置き換えた構造方程式モデルを考えたときの、そこでの Y の分布を意味する。

介入効果を推定するためには、一般に X と Y 以外の変数の観測が必要になる。Pearl (1995) は、介入効果が識別可能となる十分条件として以下の定義と定理を与えた。ここに、識別可能とは介入効果が観測変数の同時分布によって記述されることをいう。

【定義 4】 バックドア基準

因果ダイアグラム G において、 X から Y へ有向道があるとする。このとき、次の 2 条件を満たす頂点集合 Z は (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 が X と Y を有向分離する。 □

【定理 2】 (Pearl, 1995)

因果ダイアグラム G において、 (X, Y) についてバックドア基準を満たす頂点集合に対応する変数集合 $Z = \{Z_1, \dots, Z_s\}$ が X, Y とともに観測されていれば、 X の Y への介入効果は識別可能であり

$$f(y|\text{set}(X = x)) = \int_{z_1} \cdots \int_{z_s} f(y|x, z_1, \dots, z_s) f(z_1, \dots, z_s) dz_1 \cdots dz_s$$

で与えられる。 □

2.4. 同時介入効果

Pearl and Robins (1995) は、変数集合 $X = \{X_1, \dots, X_n\}$ の各要素に対して介入をおこなうことを同時介入と呼び、同時介入をおこなったときの変数 Y の分布を同時介入効果として次のように定義した。なお、 $1 \leq i \leq n-1$ なる i に対して、 X_i は X_{i+1} の非子孫であるとする。

【定義 5】同時介入効果

因果ダイアグラム G における頂点集合を $V = \{X_1, \dots, X_n, Y, Z_1, Z_2, \dots, Z_p\}$ とする. $X = \{X_1, \dots, X_n\}$ とするとき,

$$f(y|\text{set}(X = \mathbf{x})) = \int_{z_1} \cdots \int_{z_p} \frac{f(x_1, \dots, x_n, y, z_1, \dots, z_p)}{\prod_{i=1}^n f(x_i|\text{pa}(x_i))} dz_1 \cdots dz_p$$

を X の Y への同時介入効果という. ここに, $\text{set}(X = \mathbf{x})$ は同時介入により X の各要素 X_i を $\mathbf{x}$ の各要素 x_i と固定することを意味する. $\square$

介入効果の場合と同様に, 同時介入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. Pearl and Robins (1995) は, 同時介入効果が識別可能となるための十分条件として以下の定義と定理を与えた.

【定義 6】許容性基準

因果ダイアグラム G において, X の各頂点から Y へ有向道があるとする. 頂点集合 $Z = Z_1 \cup \cdots \cup Z_n$ が次の 2 条件を満たすとき, Z は (X, Y) に対して許容的であるという.

- (1) $1 \leq i \leq n$ なる任意の i に対して, 頂点集合 Z_i は $X_i, \dots, X_n$ の非子孫からなる頂点の集合である.
- (2) $1 \leq i \leq n$ なる任意の i に対して, G より $X_{i+1}, \dots, X_n$ に入る矢線と X_i から出る矢線をすべて取り除いた有向グラフにおいて, $\{X_1, \dots, X_{i-1}\} \cup Z_1 \cup \cdots \cup Z_i$ は X_i と Y を有向分離する.

$\square$

定義 6 において, Z_i は X の要素 X_i に対応して与えられていることに注意する. また, 本論文では Z_i が単一の頂点である場合を中心に議論をおこなうとする.

【定理 3】(Pearl and Robins, 1995)

因果ダイアグラム G において, (X, Y) に対して許容的な頂点集合に対応する変数集合 Z が X, Y とともに観測されていれば, X の Y への同時介入効果は識別可能であり

$$\begin{aligned} f(y|\text{set}(X = \mathbf{x})) &= \int_{z_1} \cdots \int_{z_n} f(y|x_1, \dots, x_n, z_1, \dots, z_n) \\ &\quad \times \prod_{i=1}^n f(z_i|z_1, \dots, z_{i-1}, x_1, \dots, x_{i-1}) dz_1 \cdots dz_n \end{aligned} \quad (2)$$

で与えられる. $\square$

3. 線形構造方程式モデルと直接効果, 総合効果

3.1. 線形構造方程式モデル

V のすべての要素が連続変数の場合に, (1) 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を考える.

$$X_i = \sum_{X_j \in \text{pa}(X_i)} \alpha_{ij} X_j + \epsilon_i \quad (i = 1, \dots, m) \quad (3)$$

ここに, V の各要素は平均 0, 分散 1 に基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、 α_{ij} 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 $\alpha_{ij} \neq 0$ 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。

3.2. 回帰モデルによる直接効果と総合効果の推定

変数間の統計的因果関係を推測することを目的として回帰分析を用いるとき、推定された偏回帰係数は、直接効果や総合効果のように統計的因果推論の立場から解釈の可能なものであることが望ましい。そのためには、どの変数を回帰モデルに取り入れたらよいかを明らかにしておくことが有用である。

X から Y への矢線があり、 X から Y への直接効果の推定に興味があるとき、Pearl (1998) は直接効果と偏回帰係数の関係について以下の定義と定理を与えた。

【定義 7】 単一結合基準

因果ダイアグラム G において X から Y への矢線があるとする。このとき、次の 2 条件を満たす頂点集合 Z は (X, Y) に対して単一結合基準を満たすという。

1. Y から Z の任意の頂点へ有向道がない。
2. G より X から Y への矢線を除いたグラフにおいて、 Z は X と Y を有向分離する。 □

【定理 4】 (Pearl, 1998)

因果ダイアグラム G において、 X から Y への矢線があるとする。 G において頂点集合 Z が (X, Y) に対して単一結合基準を満たすならば、 Y を目的変数、 X と変数集合 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 $\beta_{y|x,z}$ は、 X から Y への直接効果に等しい。 □

一般に、 Z が (X, Y) に対して単一結合基準を満たしていないならば、 $\beta_{y|x,z}$ は X から Y への直接効果を意味しない (Pearl, 1998)。

X から Y への有向道があり、 X から Y への総合効果の推定に興味があるとき、Pearl (1998) は総合効果と偏回帰係数の関係について次の定理を与えた。

【定理 5】 (Pearl, 1998)

因果ダイアグラム G において、 X から Y へ有向道があるとする。 G において頂点集合 Z が (X, Y) に対してバックドア基準を満たすならば、 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 $\beta_{y|x,z}$ は、 X から Y への総合効果に等しい。 □

4. 線形構造方程式モデルにおける同時介入効果の推測

前節では、外的操作の対象となる変数がひとつであるとき、回帰分析を用いて推定された偏回帰係数が直接効果あるいは総合効果を意味するための十分条件を紹介した。しかし、同時介入をおこなったときの Y の平均の推定に回帰分析を用いるとき、そこに現れる偏回帰係数がどのような意味を持つのかは明らかにされていない。本節では、同時介入をおこなったときの Y の平均がパス係数を用いてどのように記述できるかを明らかにする。

まず、因果ダイアグラム G における頂点集合 V を次の 4 つの部分集合に分割する。

X_1 : 内生変数 (G において入る矢線が存在する変数) で外的操作の対象となる n_1 個の変数からなる集合

X_2 : 外生変数 (G において入る矢線が存在しない変数) で外的操作の対象となる n_2 個の変数からなる集合

X_3 : 内生変数で外的操作の対象ではない n_3 個の変数からなる集合。 Y はこの集合の中に含まれる。

X_4 : 外生変数で外的操作の対象ではない n_4 個の変数からなる集合

ここに、 $V = X_1 \cup X_2 \cup X_3 \cup X_4$, $X_i \cap X_j = \phi (i \neq j)$ である。

ところで、統計的因果分析では、観測変数は処理変量、反応変量、共変量、中間特性に分類できることが多い。観測変数の中で、変動原因を究明する対象となるものを反応変量といい、 Y がこれにあたる。外的操作の対象となる変数を処理変量といい、 X_1 と X_2 が処理変量の集合にあたる。処理変量に対して外的操作をおこなう以前に観測される変量を共変量といい、処理変量から影響を受け、反応変量に影響を与える可能性がある変量を中間特性という。 X_4 は共変量の集合と考えることができる。また、 X_3 の各要素はある処理変量に外的操作をおこなった後に観測されるので、 X_3 は中間特性の集合と考えることができる。

最初に X_i の要素を配し、ついで X_j の要素を配した列ベクトルを $(X_i, X_j)'$ と記し、 A_{ij} を X_j の各要素から X_i への各要素への直接効果を要素に持つ (n_i, n_j) 行列とする。このとき、変数間の直接効果を要素にもつ行列と誤差変数ベクトルの分割表現を用いることにより (3) 式は次式のよう

$$\begin{pmatrix} X_1 \\ X_3 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{pmatrix} \begin{pmatrix} X_1 \\ X_3 \end{pmatrix} + \begin{pmatrix} A_{12} & A_{14} \\ A_{32} & A_{34} \end{pmatrix} \begin{pmatrix} X_2 \\ X_4 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_3 \end{pmatrix}$$

ここで、変数集合 $X_1 \cup X_2$ に対して同時介入をおこなうことを考える。 $X_1 \cup X_2$ に対して同時介入をおこなうことは、この変数集合の各要素を左辺にもつ構造方程式のそれぞれを定数関数に置き換えることを意味する。したがって、最初に X_1 の各要素を定数とした集合 x_1 の要素を配し、次に X_2 の各要素を定数とした集合 x_2 の要素を配し、最後に X_4 の要素を配した列ベクトルを $(x_1, x_2, X_4)'$ と記すとき、同時介入をおこなった後の構造方程式モデルは、整理することによって次のように記述することができる。

$$X_3 = A_{33}X_3 + \begin{pmatrix} A_{31} & A_{32} & A_{34} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ X_4 \end{pmatrix} + \epsilon_3 \quad (4)$$

ここで、 n_3 次単位行列を I_{n_3} と記せば、(4) 式を X_3 についてまとめると、

$$X_3 = (I_{n_3} - A_{33})^{-1} \begin{pmatrix} A_{31} & A_{32} & A_{34} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ X_4 \end{pmatrix} + (I_{n_3} - A_{33})^{-1} \epsilon_3 \quad (5)$$

を得る。

さて、(5) 式において両辺の期待値をとると、

$$\begin{aligned}
 E\left(X_3 \middle| \text{set}\left(\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right)\right) &= (I_{n_3} - A_{33})^{-1} \begin{pmatrix} A_{31} & A_{32} & A_{34} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \\
 &= (I_{n_3} - A_{33})^{-1} \begin{pmatrix} A_{31} & A_{32} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (6)
 \end{aligned}$$

を得る．本論文では，これを平均に対する同時介入効果と呼ぶ．

次に，(5)式に基づいて X_3 の分散共分散行列を求める． X_4 が外生変数の集合であることから， X_4 の各要素と ϵ_3 の各要素は互いに独立である．したがって，

$$\begin{aligned}
 \text{Var}\left(X_3 \middle| \text{set}\left(\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right)\right) \\
 = (I_{n_3} - A_{33})^{-1} A_{34} \Sigma_{x_4, x_4} A'_{34} (I_{n_3} - A'_{33})^{-1} + (I_{n_3} - A_{33})^{-1} \Sigma_{\epsilon_3 \epsilon_3} (I_{n_3} - A'_{33})^{-1} \quad (7)
 \end{aligned}$$

を得る．ここに， Σ_{x_4, x_4} と $\Sigma_{\epsilon_3 \epsilon_3}$ はそれぞれ X_4 と ϵ_3 の分散共分散行列である．本論文では，これを分散に対する同時介入効果と呼ぶ．

ところで，(4)式は $X_1 \cup X_2 \cup X_4$ を外生変数の集合とした線形構造方程式モデルにおいて， $X_1 \cup X_2$ の各要素を定数として固定したものと解釈することができる．したがって，(6)式の $(I_{n_3} - A_{33})^{-1} \begin{pmatrix} A_{31} & A_{32} \end{pmatrix}$ の各成分は， $X_1 \cup X_2 \cup X_4$ の各要素を外生変数としたときの， $X_1 \cup X_2$ の各要素から X_3 の各要素への総合効果を意味している．また， X_3 のある要素となる変数への同時介入効果の平均と分散の推測に関心がある場合であっても，その変数を目的変数とした単一の回帰モデルでは，一般にその平均と分散を推定できないことがわかる．さらに，(4)式から， X_1 の要素を左辺に持つ構造方程式はまったく現れていないことがわかる．これは， X_3 の要素に対する同時介入効果の平均と分散の推測を目的として回帰モデルをパス係数の推定に用いるとき， X_1 の要素を目的変数とした線形回帰モデルを考える必要はないことを意味している．

5. 許容性基準に基づく平均および分散に対する同時介入効果の推測

4節では，線形構造方程式モデルの枠組みにおいて，平均および分散に対する同時介入効果のパス係数による記述をおこなった．また，同時介入効果の平均と分散の推測を目的として回帰モデルをパス係数の推定に用いるとき，どのような回帰モデルを用いたらよいかを明らかにした．一般には，(6)式や(7)式に現れるパス係数を常に推定できるとは限らない．そこで，本節では，因果ダイアグラム G において (X, Y) に対して許容的な頂点集合に対応する変数集合 Z が X, Y とともに観測されているとき，回帰モデルを用いて平均と分散に対する同時介入効果を推測する方法論を述べる．

まず，(2)式の積分内の関数は X に同時介入をおこなったときの変数集合 $\{Y\} \cup Z$ の同時分布の逐次的因数分解である．ここで，この逐次的因数分解の各条件付き分布に対応する回帰モデルから構成された，次の連立回帰モデルを考える．

$$\begin{pmatrix} Y \\ Z \end{pmatrix} = B_1 \begin{pmatrix} Y \\ Z \end{pmatrix} + B_2 X + \epsilon^* \quad (8)$$

$$\mathbf{X} = \mathbf{x} \quad (9)$$

ここに, $\mathbf{Z}' = (Z_n, Z_{n-1}, \dots, Z_1)$, $\mathbf{X}' = (X_n, X_{n-1}, \dots, X_1)$ であり, $\mathbf{x}' = (x_n, x_{n-1}, \dots, x_1)$ は定数ベクトルとする. また, $\boldsymbol{\epsilon}^* = (\epsilon_y^*, \epsilon_{z_n}^*, \dots, \epsilon_{z_1}^*)$ を各回帰モデルに対応する誤差変数からなるベクトルとし, 各要素は互いに独立とする.

ここで, $\beta_{z_i x_j}$ と $\beta_{z_i z_j}$ をそれぞれ Z_i を目的変数, $x_1, \dots, x_{i-1}, z_1, \dots, z_{i-1}$ を説明変数とした回帰モデルでの x_j と z_j の偏回帰係数 ($2 \leq i \leq n, 1 \leq j \leq i-1$) とし, $\beta_{y x_j}$ と $\beta_{y z_j}$ をそれぞれ Y を目的変数, $x_1, \dots, x_n, z_1, \dots, z_n$ を説明変数とした回帰モデルでの x_j と z_j の偏回帰係数 ($1 \leq j \leq n$) とするとき, B_1 と B_2 はそれぞれ次のように記述することができる.

$$B_1 = \begin{pmatrix} 0 & \beta_{y z_n} & \beta_{y z_{n-1}} & \cdots & \beta_{y z_2} & \beta_{y z_1} \\ 0 & \ddots & \beta_{z_n z_{n-1}} & \cdots & \beta_{z_n z_2} & \beta_{z_n z_1} \\ \vdots & 0 & \ddots & \beta_{z_i z_j} & \vdots & \vdots \\ 0 & \ddots & 0 & \ddots & \beta_{z_3 z_2} & \beta_{z_3 z_1} \\ 0 & 0 & \ddots & 0 & \ddots & \beta_{z_2 z_1} \\ 0 & 0 & \cdots & 0 & \ddots & 0 \end{pmatrix}$$

$$B_2 = \begin{pmatrix} \beta_{y x_n} & \beta_{y x_{n-1}} & \cdots & \beta_{y x_2} & \beta_{y x_1} \\ 0 & \beta_{z_n x_{n-1}} & \cdots & \beta_{z_n x_2} & \beta_{z_n x_1} \\ 0 & \ddots & \beta_{z_i x_j} & \vdots & \vdots \\ \vdots & 0 & \ddots & \beta_{z_3 x_2} & \beta_{z_3 x_1} \\ 0 & 0 & \cdots & 0 & \beta_{z_2 x_1} \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

このとき, 次の定理が成り立つ.

【定理 6】

因果ダイアグラム G において, $(\mathbf{X}, Y)$ に対して許容的な頂点集合に対応する変数集合を $\mathbf{Z}$ とする. $\mathbf{X}, Y, \mathbf{Z}$ が多変量正規分布にしたがうとき, $\mathbf{X}$ による Y の平均に対する同時介入効果は

$$E \left(\left(\begin{array}{c} Y \\ \mathbf{Z} \end{array} \right) \middle| \text{set}(\mathbf{X} = \mathbf{x}) \right) = (I_{n+1} - B_1)^{-1} B_2 \mathbf{x} \text{ の第 1 要素} \quad (10)$$

で与えられ, Y の分散に対する同時介入効果は

$$\text{Var} \left(\left(\begin{array}{c} Y \\ \mathbf{Z} \end{array} \right) \middle| \text{set}(\mathbf{X} = \mathbf{x}) \right) = (I_{n+1} - B_1)^{-1} \Sigma_{\boldsymbol{\epsilon}^* \boldsymbol{\epsilon}^*} (I_{n+1} - B_1')^{-1} \text{ の第 } (1, 1) \text{ 要素} \quad (11)$$

で与えられる. ここに, $\Sigma_{\boldsymbol{\epsilon}^* \boldsymbol{\epsilon}^*}$ は (8) 式における $\boldsymbol{\epsilon}^*$ の分散共分散行列である. □

【証明】

(8) 式, (9) 式より, 次式が直ちに得られる.

$$\begin{pmatrix} Y \\ Z \end{pmatrix} = (I_{n+1} - B_1)^{-1} B_2 x + (I_{n+1} - B_1)^{-1} \epsilon^*.$$

ここで、両辺の期待値を取ることにより、

$$E\left(\begin{pmatrix} Y \\ Z \end{pmatrix} \middle| \text{set}(X = x)\right) = (I_{n+1} - B_1)^{-1} B_2 x.$$

が得られる。同様に、

$$\text{Var}\left(\begin{pmatrix} Y \\ Z \end{pmatrix} \middle| \text{set}(X = x)\right) = (I_{n+1} - B_1)^{-1} \Sigma_{\epsilon^* \epsilon^*} (I_{n+1} - B_1')^{-1}.$$

が直ちに得られる。

Q.E.D.

6. 適用例

6.1. 用いるデータの説明

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。相関行列を表1に示す。

奥野他(1986)の記述より、 $X_1, X_2, X_3, X_4, X_5, X_6$ が制御可能な変量であると推察できる。本論文では、これらの変数のうち、 X_1, X_2, X_4, X_5, X_6 を取り上げて、考察をおこなうことにする。

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1で示された因果モデルをグラフィカルモデリング(例えば Whittaker (1990))によって選択している。ここに、 X_9 と X_{10} の間の因果的關係については、どちらか一方のみの因果的關係を仮定しても本節の議論には影響しないことから、無向の辺のままとした。

このモデルの逸脱度は $\text{dev} = 34.28$, 自由度は $\phi = 36$, p 値は $p = 0.55$ である。

6.2. 解析結果

図1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2に示す。

ここで、表2の相関行列から得られる各処理変量の Y_1 への直接効果と総合効果を表3に与え

表 1. 「ボディ塗装条件の設定」データにおける相関行列(奥野他, 1986)

	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	-0.678	-0.215	0.230	0.040	0.116	0.338	0.002	0.145	-0.496	-0.198
X_2	-0.678	1.000	0.241	-0.442	-0.024	0.005	-0.422	-0.590	-0.509	0.684	0.463
X_3	-0.215	0.241	1.000	-0.201	0.004	-0.067	0.208	-0.007	-0.082	0.307	0.292
X_4	0.230	-0.442	-0.201	1.000	0.191	-0.286	0.287	0.446	0.521	-0.477	-0.614
X_5	0.040	-0.024	0.004	0.191	1.000	0.291	0.117	0.034	-0.048	0.010	-0.151
X_6	0.116	0.005	-0.067	-0.286	0.291	1.000	0.057	-0.123	-0.147	0.178	-0.226
X_7	0.338	-0.422	0.208	0.287	0.117	0.057	1.000	0.251	0.287	-0.122	-0.113
X_8	0.002	-0.590	-0.007	0.446	0.034	-0.123	0.251	1.000	0.761	-0.342	-0.551
X_9	0.145	-0.509	-0.082	0.521	-0.048	-0.147	0.287	0.761	1.000	-0.571	-0.431
X_{10}	-0.496	0.684	0.307	-0.477	0.010	0.178	-0.122	-0.342	-0.571	1.000	0.282
Y_1	-0.198	0.463	0.292	-0.614	-0.151	-0.226	-0.113	-0.551	-0.431	0.282	1.000

図 1. 「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)

表 2. 図 1 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

る。表 3 では、第一列目に外的操作の対象とする処理変量、第二列目に単一結合基準を満たす共変量集合、第三列目には直接効果の推定値、第四列目にはバックドア基準を満たす共変量集合、第五列目には総合効果の推定値を配している。これらの結果は定理 4 および定理 5 に基づいて得られるものである。また、最後の列には各処理変量に対して外的操作をおこなったときの Y_1 の分散

の推定値を配している。この結果は、定理 6 を処理変量が単一の場合へ適用することによって得られるものである。

表 3. 直接効果と総合効果

処理変量	単一結合基準	直接効果	バックドア基準	総合効果	分散
X_1	$\{X_2, X_8, X_{10}\}$	0.000	$\{X_{10}\}$	0.065	1.016
X_2	$\{X_7, X_8, X_9, X_{10}\}$	0.000	$\{X_8, X_{10}\}$	-0.124	1.096
X_4	$\{X_6, X_9\}$	-0.635	$\{X_9\}$	-0.502	0.636
X_5	$\{X_4, X_6\}$	0.000	$\{X_4\}$	-0.166	0.935
X_6	$\{X_4\}$	-0.464	$\{X_4\}$	-0.464	0.983

図 1 と表 3 から、 X_4 のように Y_1 への矢線以外の有向道がある場合には、総合効果と直接効果に差異が生じていることがわかる。特に、 X_1, X_2, X_5 のように、 Y_1 への矢線がない場合には、直接効果は 0 となる。また、 X_6 のように有向道が矢線のみである場合には、直接効果と総合効果は一致することが確認できる。一方、外的操作をおこなう前には Y_1 の分散は 1 であるにもかかわらず、外的操作をおこなった後の分散が 1 を超えているものがある。このような状況は、総合効果と擬似相関が異符号である場合に起こることがある(黒木・宮川(1999a))。

ここで、 X_1, X_2, X_4, X_5, X_6 のうち任意の二つの処理変量を取り上げて同時介入をおこなうことを考える。このとき、これらの変量集合に対して同時介入をおこなったときの Y_1 の平均と分散を、処理変量集合ごとに推定した結果が表 4 である。ただし $\text{Var}(Y_1) = 1$ に基準化している。表 4 では、第一列目に外的操作の対象とする処理変量集合、第二列目に許容的な共変量集合、第三列目には平均に対する同時介入効果の推定式、第四列目には分散に対する同時介入効果の推定値を配している。これらの結果は (10) 式および (11) 式の右辺に現れるパラメータを表 2 の相関行列を用いて表現することにより得られたものであり、 Y_1 を目的変数とした回帰モデルのみでは推定できないものである。

表 4. 同時介入効果

処理変量集合	許容性基準	平均	分散
$\{X_1, X_2\}$	$\{X_8, X_{10}\}$	$0x_1 - 0.124x_2$	1.096
$\{X_1, X_4\}$	$\{X_9, X_{10}\}$	$0.065x_1 - 0.502x_4$	0.642
$\{X_1, X_5\}$	$\{X_4, X_{10}\}$	$0.065x_1 - 0.166x_5$	0.951
$\{X_1, X_6\}$	$\{X_4, X_{10}\}$	$0.065x_1 - 0.464x_6$	1.002
$\{X_2, X_4\}$	$\{X_8, X_9, X_{10}\}$	$-0.124x_2 - 0.502x_4$	0.691
$\{X_2, X_5\}$	$\{X_4, X_8, X_{10}\}$	$-0.124x_2 - 0.166x_5$	1.029
$\{X_2, X_6\}$	$\{X_4, X_8, X_{10}\}$	$-0.124x_2 - 0.464x_6$	1.090
$\{X_4, X_5\}$	$\{X_9\}$	$-0.471x_4 - 0.166x_5$	0.609
$\{X_4, X_6\}$	$\{X_9\}$	$-0.635x_4 - 0.464x_6$	0.438
$\{X_5, X_6\}$	$\{X_4\}$	$0x_5 - 0.464x_6$	0.983

図 1、表 3 および表 4 から、平均に対する同時介入効果について次のようなことが考察できる。 $\{X_1, X_2\}, \{X_4, X_5\}$ 等のように外的操作の対象となる処理変量間に子孫-先祖の関係があるときには、子孫に対応する変数にかかる係数は総合効果と一致しているが、先祖に対応する変数にかかる係

数は総合効果とは異なっている。例えば、 $\{X_4, X_5\}$ に対して同時介入をおこなったとき、 x_5 にかかる係数は総合効果と一致するが、 x_4 にかかる係数は総合効果や直接効果とは明らかに異なっていることが確認できる。また、 $\{X_4, X_6\}$ に対して同時介入をおこなったとき、 x_4, x_6 にかかる係数はともに直接効果と一致している。一方、 $\{X_1, X_2\}$ に対して同時介入をおこなったときには、 X_1 から Y_1 へのすべての有向道上に X_2 があるので、同時介入効果の平均に対する推定式において x_1 にかかる係数は 0 となっている。同様に、 X_5 から Y_1 へのすべての有向道上に X_6 があるので、 $\{X_5, X_6\}$ に対して同時介入をおこなったときの Y_1 の平均に対する推定式においても x_5 にかかる係数は 0 となっている。なお、処理変量間に子孫-先祖の関係がない場合には、両方の係数とも総合効果と一致している。

分散に対する同時介入効果については次のようなことが考察できる。 $\{X_1, X_2\}, \{X_5, X_6\}$ のように外的操作の対象となる処理変量間に子孫-先祖の関係があり、かつ先祖に対応する変数から Y_1 へのすべての有向道上に子孫に対応する変数がある場合には、同時介入による Y_1 の分散は子孫に対応する変数にのみ外的操作をおこなったときの Y_1 の分散と一致している。例えば、 X_1 から Y_1 へのすべての有向道上に X_2 がある。このとき、 $\{X_1, X_2\}$ に対して同時介入をおこなったときの Y_1 の分散は、 X_2 に外的操作をおこなったときの Y_1 の分散と一致している。同様に、 $\{X_5, X_6\}$ に対して同時介入をおこなったときの Y_1 の分散は、 X_6 に外的操作をおこなったときの Y_1 の分散と一致している。

7. 結 論

本論文では、線形構造方程式モデルの枠組みにおいて、同時介入効果の平均と分散がパス係数によりどのように記述できるかを明らかにした。また、平均および分散に対する同時介入効果を推定するために設定すべき線形回帰モデルを明らかにし、同時介入効果の平均と分散の明示的表現を与えた。この結果は、同時介入効果の平均と分散を偏回帰係数を用いて記述するとき、統計的因果推論の立場からその記述に解釈をすることができるという意味で有用であると考えられる。

謝 辞 有益な助言をくださった査読者の方々に感謝の意を表します。

参 考 文 献

- Bollen, K.A. (1989): *Structural Equations with Latent Variables*. John Wiley & Sons, New York.
- 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.
- Kuroki, M. and Miyakawa, M. (1999): Identifiability criteria for causal effects of joint interventions, *Journal of the Japanese Statistical Society* **29**, 105-117.
- 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. and Robins, J. (1995): Probabilistic evaluation of sequential plans from causal models with hidden variables. In Besnard, P. and Hanks, S. editors, *Uncertainty in Artificial Intelligence* **11**, 669-709. Morgan Kaufmann.
- Spirites, P., Richardson, T., Meek, C., Scheines, R., and Glymour, C. (1998): Using path diagrams as a structural equation modelling tool. *Sociological Methods and Research* **27**, 182-225.

- Whittaker, J. (1990): *Graphical Models in Applied Multivariate Statistics*. John Wiley & Sons, Chichester.
- Wright, S. (1923): The theory of path coefficients: a reply to Niles' criticism, *Genetics* **8**, 239–255.
- Wright, S. (1934): The method of path coefficients. *Ann. Math. Statist.* **5**, 161–215.

(2001 年 10 月 10 日受付 2002 年 4 月 16 日最終修正 4 月 20 日採択)

著者連絡先: 〒 152-8552 目黒区大岡山 2-12-1
東京工業大学大学院 社会理工学研究科 経営工学専攻 黒木 学

Joint Intervention Effects in Linear Structural Equation Model through Linear Regression Parameters

Manabu Kuroki^{1,*} and Masami Miyakawa¹

¹ Department of Industrial Engineering and Management, Tokyo Institute of Technology,
2–12–1, Oh-okayama, Meguro-ku, Tokyo, 152–8552, Japan

Abstract

In case where causal relationships among variables can be described by a causal diagram and the corresponding linear structural equation model, this paper deals with statistical inference problems of joint intervention effects, which are the effects of external interventions in some treatment variables. The joint intervention effects defined by Pearl and Robins (1995) are formulated as a nonparametric distribution. This paper formulates both the mean and the variance of the joint intervention effects through the path coefficients in a linear structural equation model. In addition, it clarifies how linear regression models should be used in order to estimate both the mean and the variance of the joint intervention effects. Furthermore, this paper gives explicit expression of the mean and the variance of the joint intervention effects through linear regression parameters.

Key words: admissibility, confounding factor, D-separation, observational study.

* Corresponding author.

(Manabu Kuroki).

Received October 10, 2001; Received in final form April 16, 2002; Accepted April 20, 2002.