The Joint Causal Effect in Linear Structural Equation Model and Its Application to Process Analysis

Manabu Kuroki, Masami Miyakawa, Zhihong Cai
Department of Industrial Engineering and Management
Tokyo Institute of Technology
O-okayama, Meguro-ku, Tokyo 152-8552
{mkuroki, miyakawa, sai}@mail.me.titech.ac.jp

Abstract

Consider a case where cause-effect relationships among variables can be described by a causal diagram and the corresponding linear structural equation model. In order to bring a response variable close to a target, this paper proposes a statistical method for inferring a joint causal effect of a conditional plan on the variance of a response variable from non-experimental data. Moreover, based on this method, this paper formulates a conditional plan, which can cancel the influence of covariates on a response variable. The results of this paper could enable us to select an effective plan in linear conditional plans.

Key Words: admissibility; causal effect; graphical modeling; regression model; total effect; optimal plan

1 INTRODUCTION

In quality control, it is important to clarify cause-effect relationships in a process, in order to bring a response variable close to a target by assigning values to treatment variables. The required cause-effect relationships must not only be appropriate to the production process, but also enable us to estimate the effect of a control plan on the response variable reasonably. Statistical causal analysis is one of the effective statistical methods which can answer those kinds of problems.

The main purpose of statistical causal analysis is to evaluate a causal effect through both qualitative causal information and statistical data. Statistical causal analysis started with path analysis (Wright (1923, 1934)), and advanced to structural equation models (Wold (1954), Bollen (1989)). Recently, Pearl (2000) has developed a new framework of causal modeling based on a causal diagram and the corresponding nonparametric structural equation model. In this framework, there are two kinds of control plans: one is an unconditional plan in which the values of treatment variables are held constant, regardless of any values of covariates; the other is a conditional plan in which values of treatment variables are adjustable according to the values of covariates.

Regarding an unconditional plan, Pearl (1995) defined a causal effect of a treatment variable X on a response variable Y. Galles and Pearl (1995) and Pearl (1995) presented graphical criteria for ensuring the identification of a causal effect. Moreover, Pearl and Robins (1995) and Kuroki and Miyakawa (1999c) developed graphical criteria for testing the identification of a causal effect of a set $\boldsymbol{X}$ of treatment variables on Y. This kind of causal effect is called a joint causal effect in this paper. These studies enabled us to evaluate the causal effect of an unconditional plan on Y from nonexperimental data.

When a linear structural equation model is assumed as the data generating process, Pearl (1998) showed that the regression coefficient of X on Y implies a total effect by selecting covariates according to his criterion in the regression model. Kuroki and Miyakawa (1999a) formulated a causal variance effect on Y and presented guidelines for applying these results to process analysis. Moreover, Kuroki and Miyakawa (2002) gave explicit expressions of joint causal mean effects and joint causal variance effects on Y.

Regarding a conditional plan, Kuroki and Miyakawa (1999b) and Pearl (2000) discussed a graphical condition for identifying a causal effect of a conditional plan. In a linear structural equation model, Kuroki and Miyakawa (1999b) formulated a causal variance effect of an effective plan on Y and applied it to adaptive control in quality control.

However, it should be noted that the study of a joint causal effect of a conditional plan is not sufficient. In quality control, it is important to formulate the joint causal mean effect and the joint causal variance effect in order to select an effective plan from nonexperimental studies. Robins (1986, 1987) and Pearl (2000) formulated joint causal effects of a conditional plan through a recursive factorization of a joint distribution, but there has been no formulation of both a joint causal mean effect and a joint causal variance effect of a conditional plan in a linear structural equation model.

In this paper, joint causal mean effects and joint causal variance effects of a control plan will be formulated in a linear structural equation model. Moreover, an effective plan, which can cancel the influence of covariates on a response variable, is suggested based on the formulation. Further, the result will be applied to process analysis. The results of this paper provide us with a selecting criterion for an effective plan from nonexperimental studies.

2 CAUSAL DIAGRAM

A graph is a pair $G = (V, E)$, where V is a finite set of vertices and a set E of arrows is a subset of $V \times V$ of ordered pairs of distinct vertices. For the detailed graph theoretic terminology used in this paper, see Lauritzen (1996) and Pearl (2000).

Suppose that a set of variables $V = \{V_1, \ldots, V_n\}$ and a directed acyclic graph $G = (V, E)$ are given. When each child-parent family in the graph G represents a deterministic function

$$V_i = \sum_{V_j \in pa(V_i)} \alpha_{v_i v_j} V_j + \epsilon_i^*, \quad i = 1, 2, \ldots, n \qquad (1)$$

and a set of equations (1) represents the data generating process, the graph G is called a causal diagram, where $pa(V_i)$ denote parents of V_i in G and stand for sets of variables considered to be direct causes of V_i, and $\epsilon_1^*, \ldots, \epsilon_n^*$ are mutually independent random disturbances. In addition, $\alpha_{v_i v_j}$ is called a path coefficient and $\alpha_{v_i v_j} \neq 0$ for $V_j \in pa(V_i)$.

When cause-effect relationships among variables can be represented by a set of equations (1), the joint distribution of V generated by the process of equations (1) can be factorized recursively according to the graph G as follows (Pearl (1995)):

$$f(v_1, \ldots, v_n) = \prod_{i=1}^{n} f(v_i | pa(v_i)). \qquad (2)$$

If the recursive factorization (2) of a joint distribution is given according to the graph G, conditional independencies implied by the factorization (2) can be obtained from the graph according to the following criterion (Pearl (1988)):

Definition 1 (D-Separation)

Let $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$ be three disjoint subsets of vertices in a causal diagram G. Then $\boldsymbol{Z}$ is said to d-separate $\boldsymbol{X}$ from $\boldsymbol{Y}$ if along every path between a vertex in $\boldsymbol{X}$ and a vertex in $\boldsymbol{Y}$ there exists a vertex w which satisfies one of the following three conditions:

1. w is in $\boldsymbol{Z}$, and one arrow on the path points towards w and the other arrow on the path emerges from w.
2. w is in $\boldsymbol{Z}$, and both arrows on the path emerge from w.
3. Neither w nor any descendant of w is in $\boldsymbol{Z}$, but both arrows on the path point towards w. □

If $\boldsymbol{Z}$ d-separates $\boldsymbol{X}$ from $\boldsymbol{Y}$ in a causal diagram, $\boldsymbol{X}$ is conditionally independent of $\boldsymbol{Y}$ given $\boldsymbol{Z}$ (Geiger et al. (1990)).

3 PROBLEM DESCRIPTION

Consider a sequential process model depicted in Fig.1. In this model, $\boldsymbol{X} = (X_k, \cdots, X_1)$ stands for a set

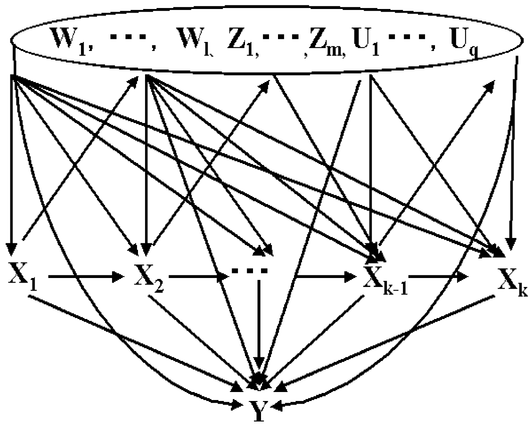

Fig.1: Process Model

of treatment variables, which can be controlled. X_i is controlled before controlling $X_{i+j} (j > 0)$ for any $i (i = 1, \cdots, k)$. Y represents a response variable, such as the final quality characteristic. $\boldsymbol{U} = (U_q, \cdots, U_1)$, $\boldsymbol{Z} = (Z_m, \cdots, Z_1)$ and $\boldsymbol{W} = (W_l, \cdots, W_1)$ represent sets of covariates, such as environment factors, which can be determined before assigning values to the treatment variables. In this process, it is supposed that the assignment of the treatment variables is not randomized but can be conducted according to states of environment factors. Consider a control plan in which $X_i \in \boldsymbol{X}$ is set to be the following linear function of a set $\boldsymbol{W}$ of observed variables:

$$X_i = x_i + \sum_{W_j \in nd(X_i) \setminus X} c_{x_i w_j} W_j, \quad i = 1, 2, \ldots, k, \qquad (3)$$

where $W_j \in \boldsymbol{W}$, both x_i and $c_{x_i w_j}$ are constants, and $nd(X_i)$ denote nondescendants of X_i in a causal diagram, which stand for sets of variables that can be determined before assigning values to the treatment variables. By using matrix expression, equation (3) can be rewritten as follows:

$$\boldsymbol{X} = \boldsymbol{x} + C\boldsymbol{W}, \tag{4}$$

where $\boldsymbol{X}' = (X_k, X_{k-1}, \cdots, X_1)$, $\boldsymbol{W}' = (W_l, W_{l-1}, \cdots, W_1)$, $\boldsymbol{x}' = (x_k, x_{k-1}, \cdots, x_1)$ is a constant vector, and $C = (c_{x_i w_j})$ is a (k, l) constant matrix and $c_{x_i w_j} = 0$, when $W_j \notin nd(X_i) \backslash X$. An external intervention given in equation (3) is called a (linear) control plan, denoted as $set(\boldsymbol{X} = \boldsymbol{x} + C\boldsymbol{W})$, where $\boldsymbol{W}$ is called a set of covariates used for control in this paper.

Here, when V represents a set of variables in the process model, letting $\boldsymbol{S} = V \backslash \boldsymbol{X} \cup \{Y\}$ and

$$\begin{aligned} & g(\boldsymbol{x} + C\boldsymbol{w} | pa(\boldsymbol{X})) \\ = & \prod_{i=1}^{k} f(x_i + \sum_{w_j \in nd(X_i) \backslash X} c_{x_i w_j} w_j | pa(X_i)), \end{aligned}$$

the distribution of Y by the conditional plan is defined based on equation (2) as follows (Spirtes et al.(1993), Pearl and Robins (1995)):

$$\begin{aligned} & f(y | set(\boldsymbol{X} = \boldsymbol{x} + C\boldsymbol{W})) \\ = & \int_{\boldsymbol{s}} \frac{f(y, \boldsymbol{x} + C\boldsymbol{w}, \boldsymbol{s})}{g(\boldsymbol{x} + C\boldsymbol{w} | pa(\boldsymbol{X}))} d\boldsymbol{s}. \end{aligned} \tag{5}$$

Equation (5) is called a joint causal effect of a conditional plan if C is a non-zero matrix; otherwise, it is called a joint causal effect of an unconditional plan. When equation (5) can be determined uniquely from the joint distribution of observed variables, the joint causal effect is said to be identifiable. In general, when a set $\boldsymbol{S}$ is partitioned into three disjoint sets $\boldsymbol{W}$, $\boldsymbol{Z}$, and $\boldsymbol{U}$ ($\boldsymbol{S} = \boldsymbol{W} \cup \boldsymbol{Z} \cup \boldsymbol{U}$), equation (5) can be expressed as depending on X,Y,$\boldsymbol{Z}$ and $\boldsymbol{W}$, but not on $\boldsymbol{U}$. $\boldsymbol{Z}$ is called a set of covariates used for identification in this paper. $\boldsymbol{U}$ is a set of covariates which may not be used to identify the joint causal effect. In this case, we need to observe not only $\boldsymbol{X} \cup \{Y\}$ and $\boldsymbol{W}$ but also $\boldsymbol{Z}$ in order to identify equation (5).

In Fig.1, although covariates used for control may be selected before assigning values to the treatment variables, the selection of covariates used for identification is dependent on the selection of covariates used for control. When the process model shown in Fig.1 can be described by a causal diagram, we can establish a criterion in order to select both covariates used for control and covariates used for identification. Moreover, based on the selected covariates, the joint causal effect can be inferred from nonexperimental data.

4 JOINT CAUSAL EFFECT OF CONDITIONAL PLAN

4.1 IDENTIFIABILITY

Let $G_{\overline{X}\underline{Z}}$ be the graph obtained by deleting from a graph G all arrows pointing towards vertices in $\boldsymbol{X}$ and all arrows emerging from vertices in $\boldsymbol{Z}$. Pearl and Robins (1995) discussed the identifiability for a joint causal effect of an unconditional plan, and gave the following graphical identifiability criterion:

DEFINITION 2 (ADMISSIBILITY)

Suppose that there exist directed paths from any vertex in $\boldsymbol{X}$ to Y in a causal diagram G, where X_i is a nondescendant of $X_{i+j} (j > 0)$ for any $i (i = 1, \cdots, k)$. If a set $\boldsymbol{T}' = (T_k, T_{k-1}, \cdots, T_1)$ of variables satisfies the following conditions, then $\boldsymbol{T}$ is said to be admissible relative to an ordered sequence of variables $(\boldsymbol{X}, Y)$.

(1) For all $i (i = 1, \cdots, k)$, T_i is a nondescendant of $\{X_i, \cdots, X_k\}$.
(2) In the graph $G_{\underline{X_i}\overline{X_{i+1} \cdots X_k}}$, $\{X_1, \cdots, X_{i-1}, T_1, \cdots, T_i\}$ d-separates X_i from Y. □

The following theorem shows that Definition 2 is useful as a graphical criterion for ensuring the identification of a joint causal effect of a conditional plan:

THEOREM 1

If a set $\boldsymbol{T}$ of observed variables is admissible relative to $(\boldsymbol{X}, Y)$ in a causal diagram G, then the joint causal effect of a control plan $set(\boldsymbol{X} = \boldsymbol{x} + C\boldsymbol{W}) (\boldsymbol{W} \subset \boldsymbol{T})$ on Y is identifiable, and is given by the formula

$$\begin{aligned} & f(y | set(\boldsymbol{X} = \boldsymbol{x} + C\boldsymbol{W})) \\ = & \int_{t_1} \cdots \int_{t_k} f(y | x_1 + \sum_{W_j \in nd(X_1) \backslash X} c_{x_1 w_j} w_j, \cdots, \\ & \quad x_k + \sum_{W_j \in nd(X_k) \backslash X} c_{x_k w_j} w_j, t_1, t_2, \cdots, t_k) \\ & \times \prod_{i=1}^{k} f(t_i | t_1, t_2, \cdots, t_{i-1}, x_1 + \sum_{W_j \in nd(X_1) \backslash X} c_{x_1 w_j} w_j, \cdots, \\ & \quad x_{i-1} + \sum_{W_j \in nd(X_{i-1}) \backslash X} c_{x_k w_j} w_j) dt_1 \cdots dt_k. \end{aligned} \tag{6}$$

□

The proof, which is based on the inference rules (Pearl (1995)), is omitted since it can be given by the similar procedure in section 4.4 of Pearl (2000).

4.2 JOINT CAUSAL EFFECT IN LINEAR STRCTURAL EQUATION MODEL

In practical studies, it is important to evaluate effects of an external intervention on the mean and the variance of a response variable. In this section, consider statistical inference problems of them in a linear structural equation model.

From equation (5),

$$\begin{aligned} & E(Y|set(\boldsymbol{X}=\boldsymbol{x}+C\boldsymbol{W})) \\ & = \int_y yf(y|set(\boldsymbol{X}=\boldsymbol{x}+C\boldsymbol{W}))dy \end{aligned} \quad (7)$$

is called a joint causal mean effect of a control plan $set(\boldsymbol{X}=\boldsymbol{x}+C\boldsymbol{W})$ on Y in this paper. In addition,

$$\begin{aligned} & Var(Y|set(\boldsymbol{X}=\boldsymbol{x}+C\boldsymbol{W})) \\ & = \int_y (y-E(Y|set(\boldsymbol{X}=\boldsymbol{x}+C\boldsymbol{W})))^2 \\ & \quad \times f(y|set(\boldsymbol{X}=\boldsymbol{x}+C\boldsymbol{W}))dy \end{aligned} \quad (8)$$

is called a joint causal variance effect of a control plan $set(\boldsymbol{X}=\boldsymbol{x}+C\boldsymbol{W})$ on Y.

Suppose that $\boldsymbol{X}$, Y and $\boldsymbol{T}$ follow a multivariate normal distribution with mean 0 and variance 1. Then, the integrand in equation (6) indicates a recursive factorization of the joint distribution of $\{Y\}\cup\boldsymbol{T}$, when values of $\boldsymbol{X}$ are set to be linear functions $\boldsymbol{x}+C\boldsymbol{w}$ of $\boldsymbol{w}$ by an external intervention. Consider the following simultaneous regression models according to this distribution.

$$\begin{pmatrix} Y \\ \boldsymbol{T} \end{pmatrix} = B_1 \begin{pmatrix} Y \\ \boldsymbol{T} \end{pmatrix} + B_2\boldsymbol{X} + \boldsymbol{\epsilon} \quad (9)$$

$$\boldsymbol{X} = \boldsymbol{x} + D\boldsymbol{T}, \quad (10)$$

where $\boldsymbol{\epsilon}' = (\epsilon_y, \epsilon_{t_k}, \cdots, \epsilon_{t_1})$ is a vector whose components are errors in the corresponding regression models. In addition, $\epsilon_y, \epsilon_{t_k}, \cdots, \epsilon_{t_1}$ are assumed to be mutually independent. $D = (d_{x_it_j})$ is a (k,k) constant matrix and $d_{x_it_j} = 0$, when $T_j \notin nd(X_i)\backslash\boldsymbol{X}$. Moreover, I_p is a p-dimensional identity matrix and $\mathbf{0}_p$ is a p-dimensional zero vector. Let $\beta_{t_ix_j}$ and $\beta_{t_it_j}$ be the regression coefficients of x_j and t_j in the regression model of T_i on $x_1,\cdots,x_{i-1},t_1,\cdots,t_{i-1}$ $(2\le i\le k, 1\le j\le i-1)$, respectively. In addition, letting β_{yx_j} and β_{yt_j} be the regression coefficients of x_j and t_j in the regression model of Y on $x_1,\cdots,x_k,t_1,\cdots,t_k$ $(1\le j\le k)$, respectively, B_1 and B_2 can be given as follows:

$$B_1 = \begin{pmatrix} 0 & \beta_{yt_k} & \beta_{yt_{k-1}} & \cdots & \beta_{yt_2} & \beta_{yt_1} \\ 0 & \ddots & \beta_{t_kt_{k-1}} & \cdots & \beta_{t_kt_2} & \beta_{t_kt_1} \\ \vdots & 0 & \ddots & \beta_{t_it_j} & \vdots & \vdots \\ 0 & \ddots & 0 & \ddots & \beta_{t_3t_2} & \beta_{t_3t_1} \\ 0 & 0 & \ddots & 0 & \ddots & \beta_{t_2t_1} \\ 0 & 0 & \cdots & 0 & \ddots & 0 \end{pmatrix} = \begin{pmatrix} 0 & \boldsymbol{\beta}'_{yt} \\ \mathbf{0}_k & B_{tt} \end{pmatrix}$$

$$B_2 = \begin{pmatrix} \beta_{yx_k} & \beta_{yx_{k-1}} & \cdots & \beta_{yx_2} & \beta_{yx_1} \\ 0 & \beta_{t_kx_{k-1}} & \cdots & \beta_{t_kx_2} & \beta_{t_kx_1} \\ 0 & \ddots & \beta_{t_ix_j} & \vdots & \vdots \\ \vdots & 0 & \ddots & \beta_{t_3t_2} & \beta_{t_3x_1} \\ 0 & 0 & \cdots & 0 & \beta_{t_2x_1} \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta}'_{yx} \\ B_{tx} \end{pmatrix} = \begin{pmatrix} \beta_{yx_k} & \boldsymbol{\beta}'_{yx_{k-1}} \\ \mathbf{0}_k & B_{tx_{k-1}} \end{pmatrix}.$$

$$Var(\boldsymbol{\epsilon}) = \begin{pmatrix} \sigma_{\epsilon_y\epsilon_y} & \mathbf{0}'_k \\ \mathbf{0}_k & \Sigma_{\epsilon_t\epsilon_t} \end{pmatrix}.$$

Then, the following theorem can be obtained:

THEOREM 2

Suppose that $\boldsymbol{T}$ is admissible relative to $(\boldsymbol{X}, Y)$ in a causal diagram G. When $\boldsymbol{X}$, Y and $\boldsymbol{T}$ follow a multivariate normal distribution with mean 0 and variance 1, the joint causal mean effect of a control plan $set(\boldsymbol{X}=\boldsymbol{x}+D\boldsymbol{T})$ on Y is given by the formula

$$\begin{aligned} & E(Y|set(\boldsymbol{X}=\boldsymbol{x}+D\boldsymbol{T})) \\ & = (\beta_{yx_k}, \boldsymbol{\beta}'_{yx_{k-1}} + (\boldsymbol{\beta}'_{yt} + \boldsymbol{\beta}'_{yx}D)F^{-1}B_{tx_{k-1}})\boldsymbol{x}, \end{aligned} \quad (11)$$

and the joint causal variance effect of a control plan $set(\boldsymbol{X}=\boldsymbol{x}+D\boldsymbol{T})$ on Y is given by the formula

$$\begin{aligned} & Var(Y|set(\boldsymbol{X}=\boldsymbol{x}+D\boldsymbol{T})) \\ & = \sigma_{\epsilon_y\epsilon_y} + (\boldsymbol{\beta}'_{yt} + \boldsymbol{\beta}'_{yx}D)F^{-1}\Sigma_{\epsilon_t\epsilon_t}F'^{-1} \\ & \quad \times(\boldsymbol{\beta}_{yt} + D'\boldsymbol{\beta}_{yx}), \end{aligned} \quad (12)$$

where $F = I_k - B_{tt} - B_{tx}D$.

PROOF OF THEOREM 2

Note that the following property is well known regarding the inverse matrix of the partitioned matrix:

$$\begin{pmatrix} M & N \\ P & Q \end{pmatrix}^{-1} = \begin{pmatrix} M^{-1}+M^{-1}NR^{-1}PM^{-1} & -M^{-1}NR^{-1} \\ -R^{-1}PM^{-1} & R^{-1} \end{pmatrix},$$

where $R = Q - PM^{-1}N$. The above property will be used in order to prove Theorem 2.

From equations (9) and (10),

$$\begin{pmatrix} Y \\ \boldsymbol{T} \end{pmatrix} = B_1 \begin{pmatrix} Y \\ \boldsymbol{T} \end{pmatrix} + B_2(\boldsymbol{x}+D\boldsymbol{T}) + \boldsymbol{\epsilon}.$$

Therefore,

$$\begin{aligned} \begin{pmatrix} Y \\ \boldsymbol{T} \end{pmatrix} & = \left(I_{k+1} - B_1 - \begin{pmatrix} 0 & \boldsymbol{\beta}'_{yx}D \\ \mathbf{0}_k & B_{tx}D \end{pmatrix}\right)^{-1} \\ & \quad \times(B_2\boldsymbol{x}+\boldsymbol{\epsilon}) \\ & = \begin{pmatrix} 1 & -\boldsymbol{\beta}'_{yt} - \boldsymbol{\beta}'_{yx}D \\ \mathbf{0}_k & F \end{pmatrix}^{-1} \\ & \quad \times(B_2\boldsymbol{x}+\boldsymbol{\epsilon}) \end{aligned} \quad (13)$$

Here, by taking expectation of both sides of equation (13),

$$\begin{aligned}
&E\left(\left(\begin{array}{c} Y \\ \boldsymbol{T} \end{array}\right)\middle| set(\boldsymbol{X}=\boldsymbol{x}+D\boldsymbol{T})\right) \\
&= \left(\begin{array}{cc} 1 & (\boldsymbol{\beta}'_{yt}+\boldsymbol{\beta}'_{yx}D)F^{-1} \\ \boldsymbol{0}_k & F^{-1} \end{array}\right)\left(\begin{array}{cc} \beta_{yx_k} & \boldsymbol{\beta}'_{yx_{k-1}} \\ \boldsymbol{0}_k & B_{tx_{k-1}} \end{array}\right)\boldsymbol{x} \\
&= \left(\begin{array}{cc} \beta_{yx_k} & \boldsymbol{\beta}'_{yx_{k-1}}+(\boldsymbol{\beta}'_{yt}+\boldsymbol{\beta}'_{yx}D)F^{-1}B_{tx_{k-1}} \\ \boldsymbol{0}_k & F^{-1}B_{tx_{k-1}} \end{array}\right)\boldsymbol{x}.
\end{aligned}$$

From this results, equation (11) can be obtained.

On the other hand, from equation (13),

$$\begin{aligned}
&Var\left(\left(\begin{array}{c} Y \\ \boldsymbol{T} \end{array}\right)\middle| set(\boldsymbol{X}=\boldsymbol{x}+D\boldsymbol{T})\right) \\
&= \left(\begin{array}{cc} 1 & (\boldsymbol{\beta}'_{yt}+\boldsymbol{\beta}'_{yx}D)F^{-1} \\ \boldsymbol{0}_k & F^{-1} \end{array}\right)\left(\begin{array}{cc} \sigma_{\epsilon_y\epsilon_y} & \boldsymbol{0}'_k \\ \boldsymbol{0}_k & \Sigma_{\epsilon_t\epsilon_t} \end{array}\right) \\
&\times\left(\begin{array}{cc} 1 & \boldsymbol{0}'_k \\ F'^{-1}(\boldsymbol{\beta}_{yt}+D'\boldsymbol{\beta}_{yx}) & F'^{-1} \end{array}\right).
\end{aligned}$$

From this result, equation (12) can be obtained. Q.E.D.

By replacing the coefficients of all elements of $\boldsymbol{T}\backslash\boldsymbol{W}$ in equation (10) with zeros, $E(Y|set(\boldsymbol{X}=\boldsymbol{x}+C\boldsymbol{W}))$ and $Var(Y|set(\boldsymbol{X}=\boldsymbol{x}+C\boldsymbol{W}))$ can be obtained.

4.3 JOINT CAUSAL EFFECT OF OPTIMAL PLAN

In practical studies, a control plan which cancels the influence of covariates on a response variable is often conducted in order to bring a response variable close to a target, where such a plan is achieved by choosing values for the coefficient vector which can minimize the variance of Y in the control plans. Therefore, it is important to formulate the control plan. When $\boldsymbol{T}$ is partitioned into disjoint subsets $\boldsymbol{W}$ and $\boldsymbol{Z}$ and set to be $\boldsymbol{T}'=(\boldsymbol{W}',\boldsymbol{Z}')$, by permuting the elements of each parameter matrix in equation (12) according to the sequence of $\boldsymbol{T}$, B_{tt}, B_{tx}, $\Sigma_{\epsilon_t\epsilon_t}$ and $\boldsymbol{\beta}_{yt}$ can be expressed by the following partitioned matrices, respectively:

$$B_{tt}=\left(\begin{array}{cc} B_{zz} & B_{zw} \\ B_{wz} & B_{ww} \end{array}\right),\ B_{tx}=\left(\begin{array}{c} B_{zx} \\ B_{wx} \end{array}\right).$$

$$\Sigma_{\epsilon_t\epsilon_t}=\left(\begin{array}{cc} \Sigma_{\epsilon_z\epsilon_z} & \boldsymbol{0}_{ml} \\ \boldsymbol{0}'_{ml} & \Sigma_{\epsilon_w\epsilon_w} \end{array}\right),\text{and } \boldsymbol{\beta}_{yt}=\left(\begin{array}{c} \boldsymbol{\beta}_{yz} \\ \boldsymbol{\beta}_{yw} \end{array}\right),$$

where $\boldsymbol{0}_{ml}$ is an (m,l) zero matrix. Based on equation (12), the following theorem concerning a joint causal effect of an effective plan can be obtained.

THEOREM 3

Suppose that $\boldsymbol{T}$ is admissible relative to $(\boldsymbol{X},Y)$ in a causal diagram G. When $\boldsymbol{X}$, Y and $\boldsymbol{T}$ follow a multivariate normal distribution with mean 0 and variance 1, , a control plan $set(\boldsymbol{X}=\boldsymbol{x}+C\boldsymbol{W})(\boldsymbol{W}\subset\boldsymbol{T})$, which cancels the influence of $\boldsymbol{W}$ on Y, is given by the formula

$$\begin{aligned}
&C'(B'_{zx}(I_m-B'_{zz})^{-1}\boldsymbol{\beta}_{yz}+\boldsymbol{\beta}_{yx}) \\
&\quad = -B'_{zw}(I_m-B'_{zz})^{-1}\boldsymbol{\beta}_{yz}-\boldsymbol{\beta}_{yw}, \qquad (14)
\end{aligned}$$

where m is the number of elements of $\boldsymbol{Z}$. Setting C to C^*, the joint causal variance effect of the conditional plan $set(\boldsymbol{X}=\boldsymbol{x}+C^*\boldsymbol{W})$ on Y is given by the formula

$$\begin{aligned}
&Var(Y|set(\boldsymbol{X}=\boldsymbol{x}+C^*\boldsymbol{W})) \\
&\quad = \sigma_{\epsilon_y\epsilon_y}+\boldsymbol{\beta}'_{yz}(I_m-B_{zz})^{-1}\Sigma_{\epsilon_z\epsilon_z} \\
&\qquad \times(I_m-B'_{zz})^{-1}\boldsymbol{\beta}_{yz}, \qquad (15)
\end{aligned}$$

where $\boldsymbol{Z}=\boldsymbol{T}\backslash\boldsymbol{W}$. □

An external intervention satisfying equation (14) is called an optimal plan in this paper.

PROOF OF THEOREM 3

When we replace the coefficients of all elements of $\boldsymbol{Z}=\boldsymbol{T}\backslash\boldsymbol{W}$ in equation (10) with zeros, $D=(\boldsymbol{0}_{km},C)$ can be obtained. From equation (12), $F=I_k-B_{tt}-B_{tx}D$ and $F'^{-1}(\boldsymbol{\beta}_{yt}+D'\boldsymbol{\beta}_{yx})$ can be expressed by the following partitioned matrices, respectively:

$$\begin{aligned}
&F=I_k-B_{tt}-B_{tx}D \\
&\quad = \left(\begin{array}{cc} I_m-B_{zz} & -B_{zw}-B_{zx}C \\ -B_{wz} & I_l-B_{ww}-B_{wx}C \end{array}\right) \qquad (16)
\end{aligned}$$

and

$$\begin{aligned}
&F'^{-1}(\boldsymbol{\beta}_{yt}+D'\boldsymbol{\beta}_{yx}) \\
&\quad = F'^{-1}\left(\begin{array}{c} \boldsymbol{\beta}_{yz} \\ \boldsymbol{\beta}_{yw}+C'\boldsymbol{\beta}_{yx} \end{array}\right)=\left(\begin{array}{c} \boldsymbol{a}_m \\ \boldsymbol{b}_l \end{array}\right), \qquad (17)
\end{aligned}$$

where $\boldsymbol{a}_m$ and $\boldsymbol{b}_l$ are an m dimensional constant vector and an l dimensional constant vector, respectively. Here, note that the second term of the right hand side of equation (12) is the quadratic form and $\Sigma_{\epsilon_t\epsilon_t}$ is a positive definite diagonal matrix. Since equation (12) can be rewritten as

$$\begin{aligned}
&Var(Y|set(\boldsymbol{X}=\boldsymbol{x}+C\boldsymbol{W})) \\
&\quad = \sigma_{\epsilon_y\epsilon_y}+\boldsymbol{a}'_m\Sigma_{\epsilon_z\epsilon_z}\boldsymbol{a}_m+\boldsymbol{b}'_l\Sigma_{\epsilon_w\epsilon_w}\boldsymbol{b}_l
\end{aligned}$$

through the $(\boldsymbol{a}'_m,\boldsymbol{b}'_l)$, we can provide an external intervention such as $\boldsymbol{b}'_l=\boldsymbol{0}_l$ in order to cancel the influence of $\boldsymbol{W}$ on Y. Therefore, from equation (17),

$$\begin{aligned}
&\left(\begin{array}{c} \boldsymbol{\beta}_{yz} \\ \boldsymbol{\beta}_{yw}+C'\boldsymbol{\beta}_{yx} \end{array}\right) \\
&\quad = \left(\begin{array}{cc} I_m-B'_{zz} & -B'_{wz} \\ -B'_{zw}-C'B'_{zx} & I_l-B'_{ww}-C'B'_{wx} \end{array}\right) \\
&\qquad \times\left(\begin{array}{c} \boldsymbol{a}_m \\ \boldsymbol{0}_l \end{array}\right) \\
&\quad = \left(\begin{array}{c} (I_m-B'_{zz})\boldsymbol{a}_m \\ (-B'_{zw}-C'B'_{zx})\boldsymbol{a}_m \end{array}\right), \qquad (18)
\end{aligned}$$

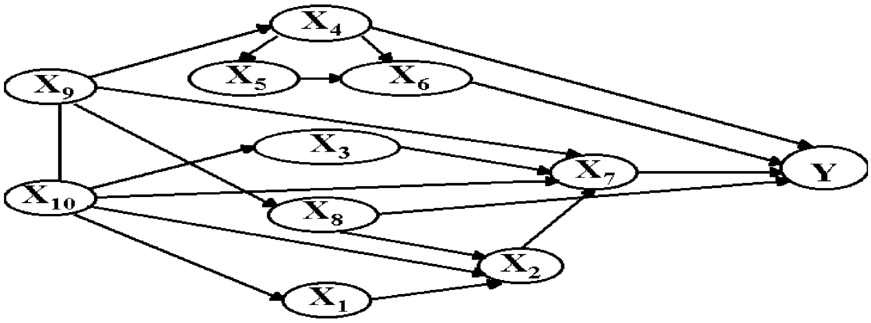

Fig.2 : Causal Diagram (Kuroki and Miyakawa (1999a))

Table 1 : The estimated correlation matrix based on Fig.2(Kuroki and Miyakawa (1999b))

	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	-0.042	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	-0.091	0.326	0.134	-0.614	-0.277	-0.250	-0.044	-0.493	-0.475	0.283	1.000

the following equation can be obtained:

$$\boldsymbol{a}_m = (I_m - B'_{zz})^{-1}\boldsymbol{\beta}_{yz}.$$

Therefore,

$$\begin{aligned} &C'(B'_{zx}(I_m - B'_{zz})^{-1}\boldsymbol{\beta}_{yz} + \boldsymbol{\beta}_{yx}) \\ &= \ -B'_{zw}(I_m - B'_{zz})^{-1}\boldsymbol{\beta}_{yz} - \boldsymbol{\beta}_{yw}, \end{aligned} \qquad (19)$$

Setting C, which satisfies equation (19), to C^*, the joint causal variance effect of the control plan $set(\boldsymbol{X} = \boldsymbol{x} + C^*\boldsymbol{W})$ on Y is given by equation (15). □

5 EXAMPLE

The above results are applicable to analyze the data obtained from a study on setting up painting conditions of car bodies, reported by Okuno et al. (1986). The data was collected with the purpose of setting up the process conditions, in order to increase transfer efficiency. The size of the sample is 38 and the variables of interest are the following:

Painting Condition : Dilution Ratio (X_1), Degree of Viscosity (X_2), Painting Temperature (X_8)

Spraying Condition : Gun Speed (X_3), Spray Distance (X_4), Atomizing Air Pressure (X_5), Pattern Width (X_6), Fluid Output (X_7)

Environment Condition : Temperature (X_9), Degree of Moisture (X_{10})

Response: Transfer Efficiency (Y)

Concerning this process, Kuroki and Miyakawa (1999a) presented the graph shown in Fig.2, whose constructing procedure is as follows: First, according to the idea of Okuno et al.(1986) and the background knowledge, an ordering of sets of variables $V_1 = \{X_9, X_{10}\} \prec V_2 = \{X_8\} \prec V_3 = \{X_1, X_4\} \prec V_4 = \{X_2, X_3, X_5, X_6\} \prec V_5 = \{X_7\} \prec V_6 = \{Y\}$ was provided, where $\prec$ indicates that V_i is precedent to $V_{i+1} (i = 1, \cdots, 5)$. Second, by applying graphical modeling (e.g. Whittaker (1990)) based on this order to the sample correlation matrix in the study of Okuno et al.(1986), the graphical chain model was constructed by considering the simplicity ($dev = 34.28$, $df = 36$, p-value$= 0.55$). In the graphical chain model, the statistical dependencies between X_9 and X_{10} and between X_5 and X_6 were described as edges, and the other dependencies between variables were described as arrows shown in Fig.2. Finally, we substituted the edge between X_5 and X_6 for an arrow, by referring to painting conditions' data. We did not substitute the edge between X_9 and X_{10} for an arrow, since the essential discussion of this paper is not influenced by the direction between X_9 and X_{10}. In this paper, the graph shown in Fig.2 will be asummed to be a causal diagram in the production process.

Kuroki and Miyakawa (1999b) presented the estimated correlation matrix based on the causal diagram as shown in Table 1. $X_1, \cdots, X_6$ are considered to be controllable variables according to Okuno et al.(1986). Therefore, it is supposed that $\{X_1, X_6\}$ and $\{X_4, X_5\}$ are taken from the controllable variables and utilized as treatment variables in order to evaluate their joint causal mean effects and the joint causal variance effects from nonexperimental data, whose results are given in Table 2 and 3. The results in Table 2 and 3 are estimated based on Table 1. Each variable in the causal diagram has zero mean and variance one. In this analysis, any combination of X_8, X_9 and X_{10} is taken as a set of covariates used for control.

Table 2 shows the estimated joint causal mean effect, which is also provided in Kuroki and Miyakawa (2002). In this analysis, a set of treatment variables is the first column of Table 2. A set of variables satisfying the admissibility is the second column of Table 2. The estimated joint causal mean effect is the third column of Table 2. In Table 2, the causal mean effect is also given in the case where X_1,X_4,X_5 and X_6 are taken as one treatment variable from the controllable variables, respectively. It should be noted that the results in Table 2 are not dependent on the plans. Consider a case where a control plan of $\{X_1, X_6\}$ is conducted. The coefficients of x_1 and x_6 in the joint causal effect are consistent with the coefficient of x_1 in the causal effect of X_1 and the coefficient of x_6 in the causal effect of X_6, respectively, since X_1 and X_6 are not related as ancestor-descendant.

Table 2:Causal Mean Effect
(Kuroki and Miyakawa (2002))

Treatment	Identification	Mean
$\{X_1\}$	$\{X_{10}\}$	$0.065x_1$
$\{X_4\}$	$\{X_9\}$	$-0.502x_4$
$\{X_5\}$	$\{X_4\}$	$-0.166x_5$
$\{X_6\}$	$\{X_4\}$	$-0.464x_6$
$\{X_1, X_6\}$	$\{X_4, X_{10}\}$	$0.065x_1 - 0.464x_6$
$\{X_4, X_5\}$	$\{X_9\}$	$-0.471x_4 - 0.166x_5$

On the other hand, in a case where a control plan of $\{X_4, X_5\}$ is conducted, the coefficient of x_4 in the joint causal effect is not consistent with the coefficient of x_4 in the causal effect of X_4, since X_5 is on the directed path from X_4 to Y. However, the coefficient of x_5 in the joint causal effect is consistent with the coefficient of x_5 in the causal effect of X_5, since no other treatment variables are on the directed path from X_5 to Y.

Table 3 shows the estimated joint causal variance effect. A set of covariates used for control is the first column of Table 3. A set of covariates used for identification, whose selection is dependent on a set of covariates used for control, is the second column of Table 3. The estimated joint causal variance effect is the third column of Table 3. Consider a case where an optimal plan of $\{X_1, X_6\}$ is conducted. When a set of covariates used for control is an empty set, the variance of Y exceeds 1. Regarding one treatment variable, such a situation may occur in case where the total effect and the spurious correlation have different signs (Kuroki and Miyakawa (1999a)). It can be also shown that the similar situation may take place, when we choose more than two treatment variables (Kuroki (2002)). On the other hand, when a set of covariates used for control is not empty, the variance of Y is less than 1.

Table 3:Joint Causal Variance Effect

(a):$\{X_1, X_6\}$

Control	Identification	Variance
	$\{X_4, X_{10}\}$	1.002
$\{X_8\}$	$\{X_4, X_{10}\}$	0.750
$\{X_9\}$	$\{X_4, X_{10}\}$	0.685
$\{X_{10}\}$	$\{X_4\}$	0.876
$\{X_8, X_9\}$	$\{X_4, X_{10}\}$	0.644
$\{X_8, X_{10}\}$	$\{X_4\}$	0.673
$\{X_9, X_{10}\}$	$\{X_4\}$	0.683
$\{X_8, X_9, X_{10}\}$	$\{X_4\}$	0.642

(b):$\{X_4, X_5\}$

Control	Identification	Variance
	$\{X_9\}$	0.609
$\{X_8\}$	$\{X_9\}$	0.524
$\{X_9\}$		0.563
$\{X_{10}\}$	$\{X_9\}$	0.592
$\{X_8, X_9\}$		0.523
$\{X_8, X_{10}\}$	$\{X_9\}$	0.523
$\{X_9, X_{10}\}$		0.563
$\{X_8, X_9, X_{10}\}$		0.522

When a set of covariates used for control includes both X_8 and X_9, there is no noticeable difference between the variances of Y in optimal plans. In addition, when $\{X_9\}$ or $\{X_9, X_{10}\}$ is used as a set of covariates used for control, there is no noticeable difference between the variances of Y in optimal plans. In other cases, the noticeable difference between the variances of Y can be observed.

Finally, consider a case where an optimal plan of $\{X_4, X_5\}$ is conducted. The variances of Y are less than 1 regardless of the selection of covariates used for control. Regarding one treatment variable, in case where the covariates satisfying the back door criterion are utilized as covariates used for control, the causal variance effect is known to be less than 1 (Kuroki and Miyakawa (1999b)). It can be also shown that the similar situation takes place, when we choose more than two treatment variables (Kuroki (2002)).

When the set of covariates used for control includes X_8, the variances of Y are close to 0.523, since the regression coefficients of the covariates used for identification in equation (15) take small values. Similarly, when we use X_9 or $\{X_9, X_{10}\}$ as a set of covariates used for control, the variances of Y are close to each other.

6 Conclusion

This paper showed that the admissibility condition is useful as an identifiability criterion for the joint causal effect of a conditional plan. In addition, it provided a method for determining a set of control plans that cancel the influence of covariates used for control in linear structural equation models. Furthermore, the results were applied to quality control. In many situations, it is necessary to adjust the value of a response variable to a specific value by using a control plan. This paper is helpful for evaluating the effect of the control plan on the response variables, when the graph structure of cause-effect relationships among variables is known.

In this paper, the described model does not involve issues of dynamics, so the time to reach a steady state is not modeled. As a practical matter, such dynamics may be also important to model in real application. Further investigation is necessary to deal with such problems, including relaxing the assumption of linear structural equation models.

ACKNOWLEGDEMENT

The comments of the reviewers on preliminary versions of this paper are acknowledged. This research was supported by Grant-in-aid for Scientific Research 14780163, Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

Bollen, K.A.(1989). *Structural Equations with Latent Variables*, John Wiley & Sons, New York.

Galles, D. and Pearl, J. (1995). Testing Identifiability of Causal Effects. In P. Besnard and S. Hanks, editors, *Uncertainty in Artificial Intelligence 11*, 185-195. Morgan Kaufmann, San Francisco.

Geiger, D., Verma, T. and Pearl, J. (1990). Identifying Independence in Bayesian Networks. *Networks*, **20**, 507-534.

Kuroki, M and Miyakawa, M. (1999a). Estimation of Causal Effects in Causal Diagrams and its Application to Process Analysis (In Japanese). *Journal of the Japanese Society for Quality Control*, **29**, 237-247.

Kuroki, M and Miyakawa, M. (1999b). Estimation of Conditional Intervention Effects in Adaptive Control (In Japanese). *Journal of the Japanese Society for Quality Control*, **29**, 476-486.

Kuroki, M. and Miyakawa, M. (1999c). Identifiability Criteria for Causal Effects of Joint Interventions, *Journal of the Japanese Statistical Society*, **29**, 105-117.

Kuroki, M and Miyakawa, M. (2002). Joint Intervention Effects in Linear Structural Equation Model Through Linear Regression Parameters (In Japanese). *Japanese Journal of Applied Statistics*, **31**, to appear.

Kuroki, M. (2002). Some Properties of the Causal Variance Effect of Control Plan in Causal Path Diagrams. (In preparation).

Lauritzen, S. L. (1996). *Graphical Models.* Clarendon Press, Oxford.

Okuno, T., Katayama, Z., Kamigori, N., Itoh, T., Irikura, N. and Fujiwara, N. (1986). *Kougyou ni okeru Tahenryou Data no Kaiseki* (In Japanese), Nikkagiren, Tokyo.

Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems. Morgan Kaufmann, San Francisco.

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, New York.

Pearl, J. and Robins, J. M. (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, San Francisco.

Robins. J. M. (1986). A New Approach to Causal Inference in Mortality Studies with a Sustained Exposure Period: Applications to Control of the Healthy Workers Survivor Effect. *Mathematical Modeling*, **7**, 1393-1512.

Robins, J. M. (1987). Addendum to "A New Approach to Causal Inference in Mortality Studies with a Sustained Exposure Period: Applications to Control of the Healthy Workers Survivor Effect". *Computers and Mathematics with Applications*, **14**, 923-945.

Spirtes, P., Glymour, C., and Schienes, R. (1993). *Causation, Prediction, and Search.* Springer-Verlag, New York.

Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics*, Wiley, Chichester.

Wold, H.O.(1954). Causality and Econometrics, *Econometrika*, 22, 162-177.

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, *The Annals of Mathematical Statistics.*, 5, 161-215.