Biometrika (2008), **95**, 1, *pp.* 17–33
doi: 10.1093/biomet/asm092
Advance Access publication 4 February 2008

Distortion of effects caused by indirect confounding

BY NANNY WERMUTH

Department of Mathematical Statistics, Chalmers/Göteborgs Universitet, Gothenburg, Sweden
wermuth@math.chalmers.se

AND D. R. COX

Nuffield College, Oxford OX1 1NF, U.K.
david.cox@nuffield.ox.ac.uk

SUMMARY

Undetected confounding may severely distort the effect of an explanatory variable on a response variable, as defined by a stepwise data-generating process. The best known type of distortion, which we call direct confounding, arises from an unobserved explanatory variable common to a response and its main explanatory variable of interest. It is relevant mainly for observational studies, since it is avoided by successful randomization. By contrast, indirect confounding, which we identify in this paper, is an issue also for intervention studies. For general stepwise-generating processes, we provide matrix and graphical criteria to decide which types of distortion may be present, when they are absent and how they are avoided. We then turn to linear systems without other types of distortion, but with indirect confounding. For such systems, the magnitude of distortion in a least-squares regression coefficient is derived and shown to be estimable, so that it becomes possible to recover the effect of the generating process from the distorted coefficient.

Some key words: Graphical Markov model; Identification; Independence graph; Linear least-squares regression; Parameter equivalence; Recursive regression graph; Structural equation model; Triangular system.

1. INTRODUCTION

In the study of multivariate dependences as representations of a potential data-generating process, important dependences may appear distorted if common explanatory variables are omitted from the analysis, either inadvertently or because the variables are unobserved. This is an instance of the rather general term confounding.

There are, however, several distinct sources of distortion when a dependence is investigated within a reduced set of variables. The different ways in which such distortions arise need clarification. We do this by first giving examples using small recursive linear systems for which the generating process has a largely self-explanatory graphical representation. Later, these ideas are put in a general setting.

2. SOME INTRODUCTORY EXAMPLES

2·1. *Direct confounding*

The most common case of confounding arises when an omitted background variable is both explanatory to a response of primary interest and also to one of its directly explanatory variables.

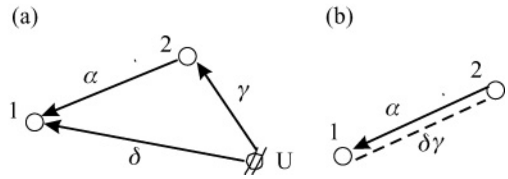

Fig. 1. Simple example of direct confounding. (a) Y_1 dependent on both Y_2 and U; Y_2 dependent on U and U to be omitted. In a linear system for standardized variables, the overall dependence of Y_1 on Y_2 is $\alpha + \delta\gamma$, with confounding effect $\delta\gamma$ due to the unobserved path from Y_1 to Y_2 via U. (b) Graph derived from Fig. 1(a) after omitting U. A dashed line added to $1 \leftarrow 2$ for the induced association. Generating dependence α preserved, but not estimable.

In Fig. 1(a), which shows this simplest instance, the directions of the edges indicate that U is to be regarded as explanatory to both the response Y_1 and to Y_2; Y_2 is, in addition, explanatory to Y_1. We suppose here for simplicity that the random variables have marginally zero means and unit variances.

The generating process is given by three linear equations,

$$Y_1 = \alpha Y_2 + \delta U + \varepsilon_1, \quad Y_2 = \gamma U + \varepsilon_2, \quad U = \varepsilon_3, \tag{1}$$

where each residual, ε_i, has mean zero and is uncorrelated with the explanatory variables on the right-hand side of an equation.

If, as is indicated by the crossed out node in Fig. 1(a), U is marginalized over, the conditional dependence of Y_1 on only Y_2 is obtained, which consists of the generating dependence α and an effect of the indirect dependence of Y_1 on Y_2 via U. This may be seen by direct calculation, assuming that the residuals ε_i have a Gaussian distribution, from

$$E(Y_1 \mid Y_2, U) = \alpha Y_2 + \delta U, \quad E(Y_2 \mid U) = \gamma U, \quad E(U) = 0,$$

leading to

$$E(Y_1 \mid Y_2) = \alpha Y_2 + \delta\, E(U \mid Y_2) = \{\alpha + \delta\gamma \operatorname{var}(U)/\operatorname{var}(Y_2)\}\, Y_2 = (\alpha + \delta\gamma)Y_2. \tag{2}$$

Thus, the generating dependence α is distorted in the conditional dependence of Y_1 on Y_2 alone, unless $\delta = 0$ or $\gamma = 0$. However, this would have been represented by a simpler generating process, in which a missing arrow for $(1, U)$ indicates $\delta = 0$ and a missing arrow for $(2, U)$ shows $\gamma = 0$. Marginal independence of Y_2 and U $(\gamma = 0)$ can be achieved by study design. It is satisfied if Y_2 represents a treatment variable and randomization is used successfully to allocate individuals to treatments. In that case, all direct dependences affecting the treatment variable Y_2 are removed from the generating process, including those of unobserved variables. Effects of the lack of an association between Y_2 and U are explored for more general relationships by Cox & Wermuth (2003) and by Ma et al. (2006). In general, the dependence of Y_1 on U, given Y_2, may vary with the levels y_2 of Y_2.

Conditions under which a generating coefficient α remains unchanged follow also from the recursive relation of linear least-squares regression coefficients (Cochran, 1938), namely

$$\beta_{1\mid 2} = \beta_{1\mid 2\cdot 3} + \beta_{1\mid 3\cdot 2}\beta_{3\mid 2}, \tag{3}$$

where we use a slight modification of Yule's notation for partial regression coefficients. For example, $\beta_{1\mid 2\cdot 3}$ is the coefficient of Y_2 in linear least-squares regression of Y_1 on both Y_2 and Y_3, and we note for Fig. 1 that $\alpha = \beta_{1\,|\,2\cdot 3}$ and $\delta = \beta_{1\,|\,3\cdot 2}$. Cochran's result (3) uses implicitly linear expectations to obtain $\beta_{3\,|\,2}$. As we shall explain later, these linear expectations are well defined for recursive linear least-squares equations, such as (1), which have uncorrelated residuals, but which do not necessarily have Gaussian distributions.

For the later general discussion, we also need a graphical representation of the structure remaining among the observed variables, here of Y_1 and Y_2, as given in Fig. 1(b). The common dependence on U induces an undirected association between the two observed variables, shown by a dashed line. A dashed line represents an association that could have been generated by a single common unobserved explanatory variable. From the generating equations (1), we obtain linear equations with correlated residuals,

$$Y_1 = \alpha Y_2 + \eta_1, \quad Y_2 = \eta_2, \tag{4}$$

where

$$\eta_1 = \delta U + \varepsilon_1, \quad \eta_2 = \gamma U + \varepsilon_2.$$

This shows directly that α cannot be estimated from the three elements of the covariance matrix of (Y_1, Y_2), since the three non-vanishing elements of the residual covariance matrix and α give four parameters for the equations (4). As a consequence, the generating dependence, α, also cannot be recovered from the conditional dependence of Y_1 on Y_2 alone, given by $\beta_{1\,|\,2} = \alpha + \delta\gamma$.

When the dependence represented by generating coefficient α is not estimable as in the induced equations (4), the equations are said to be under-identified. In systems larger than the one in Fig. 1(b), it may be possible to recover the generating dependence from the observed variables, provided there are so-called instrumental variables; for the extensive econometric literature, which builds on early work by Sargan (1958), see Hausmann (1983) or Angrist & Krueger (2001).

As we shall see, direct confounding of a generating dependence of variable pair Y_i, Y_j, say, is absent, in general, if there is no double edge, $i \overset{\leftarrow}{\text{- - -}} j$, induced in the derived graph.

2·2. *Two avoidable types of distortion*

We describe next two further types of distortion that can typically be avoided if the generating process is known and the distortions involve observed variables. One is under-conditioning. It arises by omitting from an analysis those variables that are intermediate between the explanatory variable and an outcome of primary interest. The other is over-conditioning. It arises by using as an explanatory variable to the outcome of primary interest, a variable which is, in fact, itself a response of this outcome.

We give two simple examples in Fig. 2, again for standardized variables that are linearly related. The boxed-in node, $\boxed{\bigcirc}$, indicates conditioning on given levels of a variable, and a crossed out node, $\not\bigcirc$, means, as before, marginalizing.

In Fig. 2(a), with Y_3 representing a treatment variable, interest could often be in what is called the total effect of Y_3 on Y_1. Then, marginalizing over the intermediate variable Y_2 is appropriate and $\beta_{1\,|\,3} = \delta + \alpha\gamma$ is estimated. Suppose, however, that the generating dependence of response Y_1 on Y_3 , given Y_2 is of main interest; then the direct effect δ in the data-generating process is to be estimated, and a decomposition of the total effect $\beta_{1\,|\,3}$ becomes essential, into the direct effect, δ, and the indirect effect, $\alpha\gamma$, via the intermediate variable Y_2. In this case, omission of Y_2 would be an instance of under-conditioning, leading to a mistaken interpretation; for example, see Wermuth & Cox (1998b).

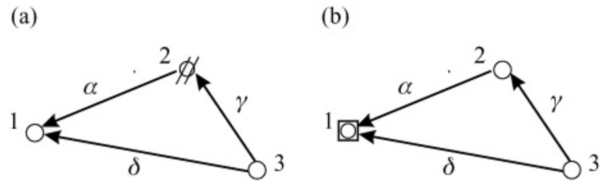

Fig. 2. Distortions due to under- and over-conditioning. (a) Generating dependence $\beta_{1\,|\,3\cdot 2} = \delta$ distorted with $\beta_{1\,|\,3}$, i.e. after removing Y_2 from conditioning set of Y_1 ; (b) Generating dependence $\beta_{2\,|\,3} = \gamma$ distorted with $\beta_{2\,|\,3\cdot 1}$, i.e. after including Y_1 into the conditioning set of Y_2.

For Fig. 2(b), the following form of the recursive relation of least-squares regression coefficients $\beta_{2\,|\,3\cdot 1} = \beta_{2\,|\,3} - \beta_{2\,|\,1\cdot 3}\beta_{1\,|\,3}$ gives, together with $\beta_{2\,|\,1\cdot 3} = \beta_{1\,|\,2\cdot 3}\sigma_{22\,|\,3}/\sigma_{11\,|\,3}$,

$$\beta_{2\,|\,3\cdot 1} = \gamma - \{(1-\gamma^2)/(1-\rho_{13}^2)\}\alpha\rho_{13}, \quad \text{with } \rho_{13} = \delta + \alpha\gamma.$$

The generating dependence could not be recovered if no information were available for Y_2 in Fig. 2(a) or for Y_1 in Fig. 2(b).

More complex forms of over-conditioning result by both marginalizing and conditioning. The simplest more general form is the presence of the following path:

$$i \longrightarrow \boxed{\circ} \longleftarrow \not\circ \longrightarrow \boxed{\circ} \longleftarrow j.$$

With any type of over-conditioning, the roles given by the generating process are interchanged for some variables, since a response to an outcome variable becomes included in the conditioning set of this outcome. Presence of strong distortions due to over-conditioning typically leads directly to a mistaken interpretation.

As we have seen, consequences of under- and over-conditioning can be quite different. However, after a set of variables is omitted from the generating process, both over- and under-conditioning for a response are avoided by the same strategy: by considering the conditional dependence on all and only those of the observed variables that are explanatory for response, either directly or indirectly via intermediate variables.

In the following two examples of indirect confounding, there is no direct confounding and there is no distortion due to over- or to under-conditioning.

2·3. *Indirect confounding in an intervention study*

A simple system without direct confounding, but with distortions of the generating dependence of Y on $T_{\rm p}$, is shown in Fig. 3.

It concerns an intervention study and is adapted from Robins & Wasserman (1997), who showed that the generating dependence of the main outcome variable, Y, on past treatment, $T_{\rm p}$, given both the more recent treatment, $T_{\rm r}$, and the unobserved health status, U, of a patient, cannot be estimated consistently by any least-squares regression coefficient in the observed variables, in spite of the use of randomization when administering the two treatments sequentially.

A past treatment $T_{\rm p}$ is decoupled from U due to full randomized allocation of treatments to individuals, and there is an intermediate binary outcome, A. The recent treatment, $T_{\rm r}$, is decoupled from both $T_{\rm p}$ and U, but not from A, since randomized allocation of treatments to individuals is at this stage, assumed to be conditional on the level of the intermediate outcome variable A.

For some detailed discussion of the structure represented by the graph in Fig. 3, we turn now to a linear system of standardized variables in which observed variables (1, 2, 3, 4) correspond to

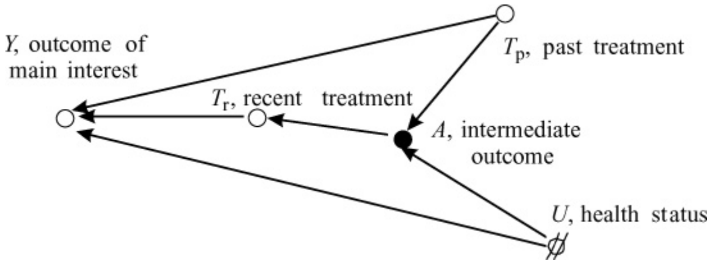

Fig. 3. Generating process in five variables, missing edge for (T_p, U) due to full randomized allocation of individuals to treatments, and missing edges for (T_r, U) and (T_r, T_p) due to randomization conditionally, given A. With U unobserved, no direct confounding results, but the generating dependence of Y on T_p (but not of Y on T_r) becomes indirectly confounded.

(Y, T_r, A, T_p) and obtain the equations with uncorrelated residuals, defined implicitly by Fig. 4(a), as

$$Y_1 = \lambda Y_2 + \alpha Y_4 + \delta U + \varepsilon_1, \quad Y_2 = \nu Y_3 + \varepsilon_2, \quad Y_3 = \theta Y_4 + \gamma U + \varepsilon_3, \quad Y_4 = \varepsilon_4, \quad U = \varepsilon_U. \tag{5}$$

From Fig. 4(a), the graph for the remaining four observed variables in Fig. 4(b) is derived by replacing the path $1 \prec\!\!-\!\!- U \;-\!\!-\!\!\succ 3$ by a dashed line for (1,3).

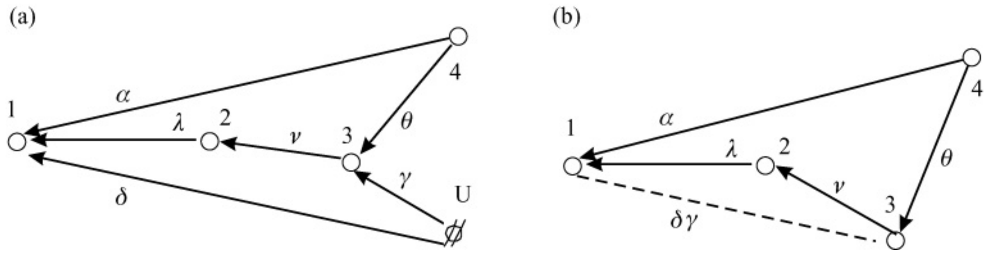

Fig. 4. (a) The graph of Fig. 3 for a linear system in standardized variables and (b) the derived graph with an induced association for (1,3), shown as a dashed line; no direct confounding of α, the generating dependence of 1 on 4, but the confounding path, $1 - - - 3 \prec\!\!-\!\!- 4$, turns $\beta_{1|4\cdot23}$ into a distorted measure of $\alpha = \beta_{1|4\cdot3U}$.

The correlation matrix of the observed variables is, by direct computation or by tracing paths, i.e. by repeated use of recursive relations for least-squares coefficients,

$$\mathrm{corr}(Y_1, Y_2, Y_3, Y_4) = \begin{pmatrix} 1 & \lambda + \alpha\theta\nu + \delta\gamma\nu & \lambda\nu + \alpha\theta + \delta\gamma & \alpha + \lambda\nu\theta \\ \cdot & 1 & \nu & \nu\theta \\ \cdot & \cdot & 1 & \theta \\ \cdot & \cdot & \cdot & 1 \end{pmatrix},$$

where the dots indicate entries in a symmetric matrix, left out as redundant.

Furthermore, each of the four possible least-squares coefficients of Y_1, regressed on Y_4 and on some or none of Y_2 and Y_3, is a distorted measure of the generating dependence α, since

$$\beta_{1|4} = \alpha + \lambda\nu\theta, \qquad \beta_{1|4\cdot2} = \alpha - \nu^2\theta\delta\gamma/(1-\nu^2\theta^2), \quad \beta_{1|4\cdot23} = \beta_{1|4\cdot3} = \alpha - \theta\delta\gamma/(1-\theta^2)\cdot$$

This verifies the result of Robins & Wasserman (1997) for a purely linear system and explains how effect reversal can occur, depending on the signs and relative magnitudes of the two terms in these last formulae. The distortion of α in $\beta_{1|4\cdot23}$, which is due to what we call indirect confounding, results by the combination of conditioning on Y_3, which is indirectly explanatory for the response Y_1, and of marginalizing over U, which is a common explanatory variable for both Y_1 and Y_3.

The explicit expressions for the least-squares coefficients show, in addition, that the generating coefficient α may, in this case, be recovered from the observed variables, for instance, with $\beta_{2\,|\,4} = \nu\theta$ and $\beta_{1\,|\,2\cdot 34} = \lambda$; the same holds for least-squares estimates, so that the generating coefficient α is identifiable for the given process, see § 6·2.

2·4. *Indirect confounding in an observational study*

The data-generating process in Fig. 5 is for determinants of quality of life after removal of the bladder because of a tumour. There are five quantitative variables and one binary variable A, which captures whether the bladder substitute leads to continent or incontinent urine diversion. When both U and V are unobserved, there is no direct confounding, but indirect confounding for the generating dependence of physical quality of life after surgery, Y, on the type of diversion, A.

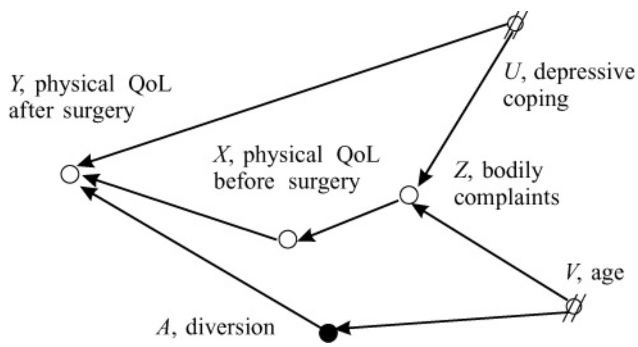

Fig. 5. A potential generating process for physical quality of life of male patients with a bladder tumour, after surgical removal of the bladder; data from Hardt et al. (2004).

The confounding path in the derived graph is different from the one in the intervention study: it is the path $Y - - - Z - - - A$ with implicit conditioning on Z, since Z is indirectly explanatory for Y via the intermediate variable X.

2·5. *Objective of paper*

From these examples, several questions arise for general generating processes, including those that contain both continuous and discrete variables as responses. Is there a class of structures in which indirect confounding of a generating dependence can occur when there is no other distortion, Figs 3 to 5 being just examples? Can the distortions be then so severe that qualitatively different conclusions on the direction and strength of dependencies arise? Are there general conditions under which we can always quantify the amount of indirect confounding in linear least-squares coefficients, so that the generating coefficients can be recovered at least in linear systems? The main objective of the present paper is to give affirmative answers to these questions.

3. GRAPHICAL AND MATRIX REPRESENTATIONS

3·1. *Parent graphs and triangular systems of densities*

A graphical representation of a stepwise generating process consists of nodes, drawn as circles for continuous and as dots for discrete variables, and of directed edges, drawn as arrows. It has an ordered node set $V = (1, 2, \ldots, d)$, such that a component variable Y_i of a vector variable Y_V corresponds to node i and, for $i < j$, the relationship between variables Y_i and Y_j is interpreted with Y_j being potentially explanatory to response Y_i.

For each node i, there is a subset $\mathrm{par}(i)$ of $r(i) = (i + 1, \ldots, d)$, called the parent set of i, with the corresponding variables said to be directly explanatory for Y_i. An ij-arrow starts at node j and points to node i, if and only if node j is a parent of node i; the graph, denoted by G^V_{par}, is named the parent graph.

A joint density f_V, written compactly in terms of nodes and of the form

$$f_V = \prod_{i=1}^{d} f_{i\,|\,\mathrm{par}\,(i)}, \tag{6}$$

is then generated over the given parent graph by starting with the last background variable, Y_d, continuing with Y_{d-1}, up to Y_1, the response of primary interest. In that way, the independence structure is fully described by the parent graph: if the ij-arrow, i.e. the edge for node pair (i, j), is missing, then Y_i is independent of Y_j, given $Y_{\mathrm{par}\,(i)}$, written in terms of nodes as $i \perp\!\!\!\perp j \mid \mathrm{par}\,(i)$. If the ij-arrow is present, then Y_i is dependent on Y_j, given $Y_{\mathrm{par}\,(i)\setminus j}$.

For the later results, some further definitions for graphs are useful. An ij-path is a sequence of edges which join the path endpoint nodes i and j via distinct nodes. Nodes along a path, called its inner nodes, exclude the path endpoints. An edge is regarded as a path without inner nodes. For an ij-path which starts with an arrow-end at node j, meets an arrow-end at each inner node and ends with an arrow-head at i, node j is called an ancestor of i and the set of ancestors is denoted by $\mathrm{anc}(i)$. Variables attached to such inner nodes are intermediate between Y_i and Y_j. A node j, which is an ancestor but not a parent of i, indicates that Y_j is only indirectly explanatory for Y_i.

3·2. *Linear triangular systems*

Instead of a joint density, a linear triangular system of equations may be generated over a given parent graph. Then, for mean-centred variables, the linear conditional expectation of Y_i on $Y_{r(i)}$, where as before $r(i) = (i + 1, \ldots d)$, is

$$E_{\mathrm{lin}}(Y_i \mid Y_{r(i)}) = \Pi_{i\,|\,\mathrm{par}(i)} Y_{\mathrm{par}(i)}, \tag{7}$$

if the residuals, denoted by ε_i, are uncorrelated with Y_j for all j in $r(i)$ (Cramér, 1946, p. 302) and if there is a direct contribution to linear prediction of Y_i only for j, a parent node of i. Thus, E_{lin} is to be interpreted as forming a linear least-squares regression and $\Pi_{i\,|\,\mathrm{par}(i)}$ denotes a row vector of nonzero linear least-squares regression coefficients.

A missing ij-arrow means, in this case, that Y_i is linearly independent of Y_j , given $Y_{\mathrm{par}(i)}$, and this is reflected in $\beta_{i\,|\,j\cdot\mathrm{par}(i)} = 0$. The linear equations, corresponding to (7), are in matrix form

$$AY = \varepsilon, \tag{8}$$

where A is an upper-triangular matrix with unit diagonal elements, and ε is a vector of zero-mean, uncorrelated random variables, called residuals. The diagonal form of the positive definite residual covariance matrix, $\mathrm{cov}(\varepsilon) = \Delta$, defines linear least-squares regression equations, such that the nonzero off-diagonal elements A_{ij} of A are

$$-A_{ij} = \beta_{i\,|\,j\cdot r(i)\setminus j} = \beta_{i\,|\,j\cdot \mathrm{par(i)}\setminus j}. \tag{9}$$

The concentration matrix implied by (8) is $\Sigma^{-1} = A^{\mathrm{T}}\Delta^{-1}A$. The matrix pair (A, Δ^{-1}) is also called a triangular decomposition of Σ^{-1}. It is unique for the fixed order given by V. For a given Σ^{-1} of dimension d, there are $d!$ possible triangular decompositions, so that linear least-squares coefficients $\beta_{i\mid j\cdot C}$ are defined for any subset C of V without nodes i and j. Thus, for the examples set out in § 2, Gaussian distributions of the residuals are not needed; the same results are achieved for linear triangular systems (8) which have uncorrelated residuals.

3·3. *Edge matrices and structural zeros*

The binary matrix representation of the parent graph of a linear triangular system is $\mathcal{A} = \mathrm{In}[A]$, where the indicator operator In replaces every nonzero element in a matrix by a one. It is called an edge matrix, since off-diagonal ones represent edges present in the graph. The edge matrix of G^V_{par} of a joint density (6) is the $d \times d$ upper triangular binary matrix with elements $\mathcal{A}_{ij}$ defined by

$$\mathcal{A}_{ij} = \begin{cases} 1 & \text{if and only if } i \prec\!\!-\, j \text{ in } G^V_{\text{par}} \text{ or } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

It is the transpose of the usual adjacency matrix representation of the graph with additional ones along the diagonal to simplify transformations, such as marginalizing.

New edge matrices are induced after changing conditioning sets of dependencies, as given by the parent graph. For each ij-one in an induced-edge matrix, at least one ij-path can be identified in the given parent graph that leads in the family of linear systems generated over the parent graph to a nonvanishing parameter for pair (Y_i, Y_j); see, for example, § 3 of Wermuth & Cox (1998a). Whenever no such path exists, an ij-zero is retained in the induced-edge matrix. It indicates for the linear system that the corresponding parameter is structurally zero, i.e. that it remains zero as a consequence of the generating process. Therefore, we denote edge matrices by calligraphic letters that correspond to the latin letters of the associated parameter matrices in linear systems.

4. SOME PRELIMINARY RESULTS

4·1. *Some early results on triangular systems*

Linear triangular systems (7) have been introduced as path analyses in genetics (Wright, 1923, 1934) and as linear recursive equations with uncorrelated residuals in econometrics (Wold, 1954). Early studies of their properties include Tukey (1954), Wermuth (1980) and Kiiveri et al. (1984). They form a subclass of linear structural equations (Goldberger, 1991, Ch. 33, p. 362)

Linear triangular systems (7) and triangular systems of densities (6) are both well suited to describe development without and with interventions. Both form a subclass of graphical Markov models (Cox & Wermuth, 1993, 1996; Wermuth, 2005). For graphical models based on a special type of distributional assumption, namely the conditional Gaussian distribution, see Edwards (2000), Lauritzen (1996) and Lauritzen & Wermuth (1989).

4·2. *Omitting variables from a triangular system*

For a split of Y_V into any two component variables Y_M and Y_N with $N = V \setminus M$, the density f_V in (6) can be factorized in the form

$$f_V = f_{M\mid N} f_N.$$

One may integrate over Y_M to obtain the structure in the joint marginal density f_N of Y_N, implied by the generating process.

Essential aspects of this structure are captured by the changes resulting for the parent graph. We denote by $\tilde{V} = (M, N)$ the correspondingly ordered node set and by G^N_{rec}, the graph of recursive dependencies derived for the reduced node set N, where the order of nodes within N is preserved from V. The objective is to deduce the independence structure of G^N_{rec} from that of the parent graph and to define different types of association introduced by marginalizing over Y_M.

4·3. *Two matrix operators*

To set out changes in edge matrices, we start with a general type of edge matrix $\mathcal{F}$, such as for A, Σ^{-1} and Σ, denote the associated linear parameter matrix by F, and apply two matrix operators, called partial closure and partial inversion.

Let F be a square matrix of dimension d with principal submatrices that are all invertible and let a be any subset of V. Let, further, $b = V \setminus a$ and let an associated binary edge matrix $\mathcal{F}$ be partitioned according to (a, b). We also partition F and B accordingly and denote by B_{ab} the submatrix of B with rows pertaining to node components a and columns to components b. Then the operator, called partial inversion, transforms F into $\mathrm{inv}_a F$ and the operator, called partial closure, transforms $\mathcal{F}$ into the associated edge matrix $\mathrm{zer}_a \mathcal{F}$. They are defined as follows.

DEFINITION 1 (Partial inversion and partial closure; Wermuth et al., 2006b). *The operators of partial inversion and partial closure are*

$$\mathrm{inv}_a F = \begin{pmatrix} F_{aa}^{-1} & -F_{aa}^{-1}F_{ab} \\ F_{ba}F_{aa}^{-1} & F_{bb} - F_{ba}F_{aa}^{-1}F_{ab} \end{pmatrix}, \qquad \mathrm{zer}_a \mathcal{F} = \mathrm{In}\left[\begin{pmatrix} \mathcal{F}_{aa}^{-} & \mathcal{F}_{aa}^{-}\mathcal{F}_{ab} \\ \mathcal{F}_{ba}\mathcal{F}_{aa}^{-} & \mathcal{F}_{bb} + \mathcal{F}_{ba}\mathcal{F}_{aa}^{-}\mathcal{F}_{ab}, \end{pmatrix}\right],$$

where

$$\mathcal{F}_{aa}^{-} = \mathrm{In}[(k\mathcal{I}_{aa} - \mathcal{F}_{aa})^{-1}], \tag{11}$$

with $k - 1$ denoting the dimension of $\mathcal{F}_{aa}$ and $\mathcal{I}$ an identity matrix.

Adding a sufficiently large constant along the diagonal in (11) ensures that an invertible matrix is obtained, that the inverted matrix has nonnegative elements and that this inverse has a zero entry, if and only if there is a structural zero in F^{-1}. If $\mathcal{F}_{aa} = \mathcal{A}_{aa}$ is upper-triangular, then an ij-one is generated in $\mathcal{A}_{aa}^{-}$, if and only if j is an ancestor of i in the graph with edge matrix $\mathcal{A}_{aa}$. If instead $\mathcal{F}_{aa}$ is symmetric, an ij-one is generated in $\mathcal{F}_{aa}^{-}$ if and only if there is an ij-path in the graph with edge matrix $\mathcal{F}_{aa}$.

Both operators can be applied to a sequence of distinct subsets of a in any order to give $\mathrm{inv}_a F$ and $\mathrm{zer}_a \mathcal{F}$. Closing paths repeatedly has no effect, but partial inversion is undone by applying it repeatedly, i.e. $\mathrm{inv}_a(\mathrm{inv}_a F) = F$. Another important property of partial inversion is that

$$F\begin{pmatrix} y_a \\ y_b \end{pmatrix} = \begin{pmatrix} z_a \\ z_b \end{pmatrix} \quad \text{implies} \quad \mathrm{inv}_a F \begin{pmatrix} z_a \\ y_b \end{pmatrix} = \begin{pmatrix} y_a \\ z_b \end{pmatrix}. \tag{12}$$

Repeated application of these two operators will be used here to identify relevant properties of triangular systems. In essence, partial inversion isolates the random variable of interest after marginalization and partial closure specifies the implied graphical structure.

4·4. *The induced recursive regression graph*

For a linear triangular system, $AY = \varepsilon$ in (7), with diagonal covariance matrix Δ of the residuals ε and a parent graph with edge matrix $\mathcal{A}$, let M denote any subset of V to be marginalized over, let $N = V \setminus M$ and let $\tilde{A}, \tilde{\mathcal{A}}$ be the matrices $A, \mathcal{A}$ arranged and partitioned according to $\tilde{V} = (M, N)$ and preserving within subsets the same order as in V. Then we define

$$B = \mathrm{inv}_M \tilde{A}, \qquad \mathcal{B} = \mathrm{zer}_M \tilde{\mathcal{A}}, \tag{13}$$

to obtain the induced equations in terms of Δ and B, and the graph in terms of $\mathcal{B}$. After applying property (12) to $\tilde{A}\tilde{Y} = \tilde{\varepsilon}$, we have that

$$\mathrm{inv}_M \tilde{A}\begin{pmatrix} \varepsilon_M \\ Y_N \end{pmatrix} = \begin{pmatrix} Y_M \\ \varepsilon_N \end{pmatrix}.$$

The bottom row of this equation gives the observed variables Y_N that remain after marginalizing over Y_M as a function of components of ε and B. We summarize this in Lemma 1.

LEMMA 1 (The induced recursive regression graph. Wermuth & Cox, 2004; Corollary 1 and Theorem 3). *The recursive equations in Y_N obtained from a linear triangular system* (7)*, which are orthogonal to the equations in Y_M corrected for linear least-squares regression on Y_N, have equation parameters defined by B_{NN} and residual covariances defined by $K_{NN} = \mathrm{cov}(\eta_N)$. The induced recursive regressions equations are*

$$B_{NN}Y_N = \eta_N, \qquad \eta_N = \varepsilon_N - B_{NM}\varepsilon_M, \qquad K_{NN} = \Delta_{NN} + B_{NM}\Delta_{MM}B_{NM}^{\mathrm{T}}. \tag{14}$$

The edge matrix components of the recursive regression graph G^N_{rec}, induced by triangular systems (6) *or* (7) *after marginalizing over Y_M, are*

$$\mathcal{K}_{NN}, \quad \mathcal{K}_{NN} = \mathrm{In}[\mathcal{I}_{NN} + \mathcal{B}_{NM}\mathcal{B}_{NM}^{\mathrm{T}}]. \tag{15}$$

The key issue here is that the edge matrix components of G^N_{rec} derive exclusively from special types of path in G^N_{par}, represented by $\mathcal{A}$, and, since probabilistic independence statements defined by a parent graph combine in the same way as linear independencies specified by the same graph, the edge matrix induced by the linear system holds for all densities generated over the same G^V_{par}; see Marchetti and Wermuth (2008).

In general, edge matrices indicate both edges present in a graph, by ij-ones, and structurally zero parameters in a linear system, by ij-zeros. The types of induced edge are specified by using the following convention.

DEFINITION 2 (Types of edge represented by a matrix). *An ij-one in $\mathcal{F}$, an edge matrix derived from $\mathcal{A}$ for an induced linear parameter matrix F, represents*

(i) *an arrow, $i \prec\!\!-\!\!- j$, if F is an equation parameter matrix,*

(ii) *an ij-dashed line, $i - - - j$, if F is a residual covariance matrix.*

Thus, for instance, ij-arrows result with $\mathcal{B}_{NN}$ and ij-dashed lines with $\mathcal{K}_{NN}$. The type of induced edge relates more generally to the defining matrix products.

DEFINITION 3 (Types of edge resulting by edge matrix products). *Let an edge matrix product define an association-inducing path for a family of linear systems. Then the generated edge inherits the edge ends of the left-hand and of the right-hand matrix in the product.*

Thus, $\mathcal{A}_{NM}\mathcal{A}^-_{MM}\mathcal{A}_{MN}$ results in arrows and $\mathcal{B}_{NM}\mathcal{B}^{\mathrm{T}}_{NM}$ leads to dashed lines as a condensed notation for generating paths which have arrow-heads at both path endpoints.

4·5. *Linear parameter matrices and induced-edge matrices*

For the change, for instance, from parameter matrices B_{NN}, K_{NN} in the linear systems (14) to induced-edge matrix components $\mathcal{B}_{NN}, \mathcal{K}_{NN}$ in (15), one wants to ensure that every matrix product and every sum of matrix products has nonnegative elements, so that no additional zero is created and possibly, all zeros are retained. This is summarized as follows.

LEMMA 2 (Transforming parameter matrices in linear systems into edge matrices). *Let induced parameter matrices be defined by parameter components of a linear system $FY = \zeta$, with correlated residuals, such that the matrix products hide no self-cancellation of an operation, such as a matrix multiplied by its inverse. Let, further, the structural zeros of F be given by $\mathcal{F}$. Then the induced-edge matrix components are obtained by replacing, in the defining equations,*

(i) *every inverse matrix, F^{-1}_{aa} say, by the binary matrix of its structural zeros $\mathcal{F}^-_{aa}$,*

(ii) *every diagonal matrix by an identity matrix of the same dimension,*

(iii) *every other submatrix, $-F_{ab}$ or F_{ab} say, by the binary submatrix of structural zeros, $\mathcal{F}_{ab}$,*
(iv) *and then applying the indicator function.*

Thus, for instance, $\mathcal{B}_{NN} = \text{In}[\mathcal{A}_{NN} + \tilde{\mathcal{A}}_{NM}\mathcal{A}^{-}_{MM}\tilde{\mathcal{A}}_{MN}]$ may be obtained in this way from $B_{NN} = A_{NN} - \tilde{A}_{NM}A^{-1}_{MM}\tilde{A}_{MN}$.

The more detailed results that follow are obtained by starting with equations (14) for the observed vector variable Y_N, applying the two matrix operators, the above stated types of transformation and by orthogonalizing correlated residuals. The last task is achieved for an arbitrary subset a of V and $b = V \setminus a$, after transforming η_b into residuals corrected for marginalization over Y_a, i.e. by taking $\eta_{b-a} = \eta_b - C_{ba}\eta_a$, and then by conditioning η_a on η_{b-a}, i.e. by obtaining $\eta_{a\,|\,b-a} = \eta_a - \text{cov}(\eta_a, \eta_{b-a})\{\text{cov}(\eta_{b-a})\}^{-1}\eta_{b-a}$. For this, one uses the appropriate residual covariance matrix partially inverted with respect to b.

4·6. *Consequences of the induced recursive regression graph*

In Lemma 1, we have specified the structure for Y_N, obtained after marginalizing over Y_M. To study Y_N in more detail, let a be any subset of N and $b = N \setminus a$. Let, further, $\Pi_{a\,|\,b}$ be the matrix of regression coefficients obtained by linear least-squares regression of Y_a on Y_b. An element $\Pi_{a\,|\,b}$ for i of a and j of b is the least-squares regression coefficient $\beta_{i\,|\,j\cdot b\setminus j}$. The edge matrix corresponding to $\Pi_{a\,|\,b}$ is denoted by $\mathcal{P}_{a\,|\,b}$, with element $\mathcal{P}_{i\,|\,j\cdot b\setminus j}$.

Suppose now that the parameter matrices of linear recursive equations (14) and corresponding edge matrices (15) are arranged and partitioned according to $\tilde{N} = (a, b)$ and that, within subsets, the order of nodes remains as in N. Then we define

$$C_{NN} = \text{inv}_a\, \tilde{B}_{NN}, \qquad \mathcal{C}_{NN} = \text{zer}_a\, \tilde{\mathcal{B}}_{NN}, \tag{16}$$

and, with $W_{NN} = \text{cov}(\eta_a, \eta_b - C_{ba}\eta_a)$ and $\mathcal{W}_{NN}$ the corresponding induced-edge matrix,

$$Q_{NN} = \text{inv}_b\, \tilde{W}_{NN}, \qquad \mathcal{Q}_{NN} = \text{zer}_b\, \tilde{\mathcal{W}}_{NN}, \tag{17}$$

to obtain with $\mathcal{P}_{a\,|\,b}$ the independence statements induced by G^N_{rec} for Y_a , given Y_b.

LEMMA 3 (The induced-edge matrix of conditional dependence of Y_a on Y_b; Wermuth & Cox, 2004, Theorem 1). *The linear least-squares regression coefficient matrix $\Pi_{a\,|\,b}$ induced by a system of linear recursive regression (14) in G^N_{rec} is*

$$\Pi_{a\,|\,b} = C_{ab} + C_{aa}Q_{ab}C_{bb}, \tag{18}$$

and the edge matrix $\mathcal{P}_{a\,|\,b}$ induced by a recursive regression graph G^N_{rec} is

$$\mathcal{P}_{a\,|\,b} = \text{In}[\mathcal{C}_{ab} + \mathcal{C}_{aa}\mathcal{Q}_{ab}\mathcal{C}_{bb}]. \tag{19}$$

The independence interpretation of recursive regression graphs results from Lemma 3, if a is split further into any two nonempty subsets α and d and b into two nonempty subsets β and c. For the dependence of Y_α on Y_β, given Y_c, i.e. for a conditional dependence in the marginal density $f_{\alpha\beta C}$, one obtains the induced-edge matrix $\mathcal{P}_{\alpha\,|\,\beta\cdot c}$ as a submatrix of $\mathcal{P}_{a\,|\,b}$:

$$\mathcal{P}_{a\,|\,b} = \begin{pmatrix} \mathcal{P}_{\alpha\,|\,\beta\cdot c} & \mathcal{P}_{\alpha\,|\,c\cdot\beta} \\ \mathcal{P}_{d\,|\,\beta\cdot c} & \mathcal{P}_{d\,|\,c\cdot\beta} \end{pmatrix}.$$

COROLLARY 1 (Independence induced by a recursive regression graph). *The following statements are equivalent consequences of the induced recursive regression graph G^N_{rec} :*
(i) $\mathcal{P}_{\alpha\,|\,\beta\cdot c} = 0$;
(ii) $\alpha \perp\!\!\!\perp \beta \mid c$ *is implied for all triangular systems of densities generated over a parent graph*;
(iii) $\text{In}[\mathcal{C}_{\alpha\beta} + \mathcal{C}_{\alpha a}\mathcal{Q}_{ab}\mathcal{C}_{b\beta}] = 0$.

It follows further from (19) that the conditional dependence of Y_i on Y_j, given $Y_{b\setminus j}$ coincides with the generating dependence corresponding to $\mathcal{A}_{ij} = 1$, if $\mathcal{C}_{ia}\mathcal{Q}_{ab}\mathcal{C}_{bj} = 0$ and $\mathcal{A}_{im}\mathcal{A}^{-}_{mm}\mathcal{A}_{mj} = 0$, where $m = M \cup a$. The first condition means absence of an ij-path in G^N_{rec} via induced associations captured by $\mathcal{Q}_{ab}$. The second condition means that, in G^V_{par}, node j is not an ancestor of i, such that all inner nodes of the path are in m. With appropriate choices of i and b, the distortions described in the examples of § 2 are obtainable by using (18).

However, to correct for the distortions, one needs to know when the parameters induced for G^N_{rec} are estimable. From the discussion in § 2·2, this is not possible, in general. We therefore turn next to systems without over- and under-conditioning concerning components of Y_N.

5. DISTORTION IN THE ABSENCE OF OVER- AND UNDER-CONDITIONING

For any conditional dependence of Y_i on Y_b, we let b coincide with the observed ancestors of node i, i.e. we take $b = \text{anc}(i) \cap N$ and node i in $a = N \setminus b$. The corresponding modification of Lemma 3 results after observing that, in this case, there is no path from $d = a \setminus i$ to node i, so that $(\text{inv}_d B_{NN})_{S,S} = B_{SS}$ with $S = N \setminus d$ and that, in addition, $C_{ba} = B_{ba} = 0$.

PROPOSITION 1 (The conditional dependence of Y_i on Y_b in the absence of over- and under-conditioning). *The graph G^N_{rec}, obtained after marginalizing over variables Y_M induces the following edge vector for the conditional dependence of Y_i on Y_b*:

$$\mathcal{P}_{i\,|\,b} = \text{In}[\mathcal{B}_{ib} + (\mathcal{K}_{ib}\mathcal{K}^{-}_{bb})\mathcal{B}_{bb}]. \tag{20}$$

In addition, the linear system to G^N_{rec} induces the following vector of least-squares regression coefficients:

$$\Pi_{i\,|\,b} = B_b + (K_{ib}K^{-1}_{bb})B_{bb}. \tag{21}$$

The conditional dependence of Y_i on Y_j, given $Y_{b\setminus j}$, measures the generating dependence corresponding to $\mathcal{A}_{ij} = 1$ without distortions

(i) *due to unobserved intermediate variables if $\mathcal{A}_{iM}\mathcal{A}^{-}_{MM}\mathcal{A}_{Mj} = 0$, and*

(ii) *due to direct confounding if $\mathcal{K}_{ij} = 0$, and*

(iii) *due to indirect confounding if $(\mathcal{K}_{ib}\mathcal{K}^{-}_{bb})\mathcal{B}_{bj} = 0$.*

Distortions of type (i) are avoided if the observed node set N consists of the first d_N nodes of $V = (1, \ldots, d)$. Then no path can lead from a node in N to an omitted node in M, so that $\mathcal{A}_{NM} = 0$ and hence $B_{NN} = A_{NN}$, $\mathcal{B}_{NN} = \mathcal{A}_{NN}$.

COROLLARY 1 (Paths of indirect confounding in the absence of over- and under-conditioning). *In a graph G^N_{rec}, only the following two types of path may introduce distortions due to indirect confounding*:

(i) $(\mathcal{K}_{ab}\mathcal{K}^{-}_{bb})_{ij} \neq 0$,

(ii) $(\mathcal{K}_{ib}\mathcal{K}^{-}_{bb})\mathcal{B}_{bj} \neq 0$.

When three dots indicate that there may be more edges of the same type, coupling more distinct nodes, then typical paths of type (i) and (ii) are, respectively,

$$i\;{-}{-}{-}\;\boxed{\circ}\;\ldots\;\boxed{\circ}\;{-}{-}{-}\;\boxed{\circ}\;{-}{-}{-}\;j, \qquad i\;{-}{-}{-}\;\boxed{\circ}\;\ldots\;\boxed{\circ}\;{-}{-}{-}\;\boxed{\circ} \longleftarrow j,$$

where each node $\boxed{\circ}$ along the path is conditioned on and represents a node which is a forefather of node i, i.e. an ancestor but not a parent of i. In Fig. 4(b), the confounding path $1\;{-}{-}{-}\;3 \longleftarrow 4$ is of type (ii). In Fig. 6(b) below, the confounding path is of type (i).

6. Indirect confounding in linear generating processes

6·1. *Distortions and constraints*

Confounding ij-paths in G^N_{rec}, as specified in Corollary 2 for $\mathcal{A}_{ij} = 1$, have as inner nodes exclusively forefather nodes of i and induce in families of linear generating processes associations for pair Y_i, Y_j, in addition to the generating dependence. However, as we shall see in § 6·2, a generating coefficient can be recovered from a least-squares regression coefficient in the observed variables, provided there is no other source of distortion.

Corollary 3 (Indirect confounding in a linear least-squares regression coefficient when other sources of distortion are absent). *Suppose, first, that the recursive regression graph G^N_{rec} is without a double edge, $\overset{\prec}{- - -}$, secondly, that only background variables Y_M with $M = (d_N + 1, \ldots, d)$ are omitted from* (8)*, and, thirdly, that conditioning of Y_i is on $Y_{\text{anc}(i)\cap N}$. Then, if there is a confounding ij-path in G^N_{rec}, a nonzero element in $P_{i\,|\,b}$ contains*

(i) *a distortion due to indirect confounding for $\mathcal{A}_{ij} = 1$;*

(ii) *a merely induced dependence for $\mathcal{A}_{ij} = 0$;*

(iii) *the generating coefficient is recovered from $\beta_{i\,|\,j\cdot b\setminus j}$ with*

$$-A_{ij} = \beta_{i\,|\,j\cdot b\setminus j} - \left(K_{ib}K^{-1}_{bb}\right)A_{bj}. \tag{22}$$

A different way of expressing case (ii) is to say that the induced conditional dependence corresponds to a constrained least-squares regression coefficient.

6·2. *Parameter equivalent equations*

To show that the corrections in (22) are estimable, we turn to the slightly more general situation in which the only absent distortion is direct confounding, i.e. G^N_{rec} is without a double edge, and obtain parameter equivalence between two types of linear equation, since the parameters of the first set can be obtained in terms of those in the second set and vice versa.

Equations (14) give, for Y_1 and $b = \text{anc}(1) \cap N$,

$$Y_1 = B_{1b}Y_b + \eta_1, \qquad \text{cov}(\eta_1, \eta_b) = K_{1b}, \qquad \text{var}(\eta_1) = K_{11}. \tag{23}$$

They imply, with

$$\Sigma^{-1} = B^{\mathrm{T}}K^{-1}B, \qquad K^{-1} = G^{\mathrm{T}}D^{-1}G, \qquad D \text{ diagonal}, \qquad G \text{ upper triangular},$$

that (BG, D^{-1}) is the triangular decomposition of Σ^{-1}, which gives, as least-squares equation for Y_1,

$$Y_1 = \Pi_{1\,|\,b}Y_b + \epsilon_1, \qquad \text{cov}(\epsilon_1, \epsilon_b) = 0, \qquad \text{var}(\epsilon_1) = D_{11}. \tag{24}$$

For G^N_{rec} without a double edge, if Σ_{bb} and the parameters of equation (23) are given, then so is (BG, D^{-1}) and hence also the possibly constrained regression equation (24). Conversely, if $\Pi_{1\,|\,b}$ and Σ_{bb} are given, we define $L_{bb} = K^{-1}_{bb}B_{bb}$, call c the observed parents of node 1 and partition $\Pi_{1\,|\,b}$ either with c and d, where each element of K_{1d} is nonzero, or with c, d and e, if there is a vector with $K_{1e} = 0$. Next, we observe that both equations

$$(\Pi_{1\,|\,c\cdot d}\ \ \Pi_{1\,|\,d\cdot c}) = -(H_{1c}\ \ 0) + (0\ \ K_{1d})L_{bb},$$

$$(\Pi_{1\,|\,c\cdot de}\ \ \Pi_{i\,|\,d\cdot ce}\ \ \Pi_{1\,|\,e\cdot cd}) = -(H_{1c}\ \ 0\ \ 0) + (0\ \ K_{1d}\ \ 0)L_{bb},$$

can be solved for H_{1c} and K_{1d} and that K_{11} results from $K = B\Sigma B^{\mathrm{T}}$. This one-to-one correspondence is extended by starting with the equation for Y_{d_N-1} and successively proceeding to the equation for Y_1.

PROPOSITION 2. *For a recursive regression graph G^N_{rec} without a double edge, the ith equation is parameter equivalent to the possibly constrained least-squares regression equation obtained from the triangular decomposition of Σ^{-1}_{SS} with $S = (i, \ldots, N)$.*

Thus, given independent observations on a linear system for G^N_{rec} without a double edge, all parameters can be estimated. One may apply general software for structural equations, or estimation may be carried out within the statistical environment R (Marchetti, 2006), which uses the EM algorithm as adapted by Kiiveri (1987).

The result about parameter equivalence strengthens identification criteria, since it gives the precise relationships between two sets of parameters. Previously, different graphical criteria for identification in linear triangular systems with some unobserved variables have been derived by Brito & Pearl (2002) and by Stanghellini & Wermuth (2005).

Propositions 1 and 2 imply, in particular, that, in the absence of direct confounding and of over- and of under-conditioning, a generating dependence α_{ij} of a linear system (8) may actually be recovered from special types of distorted least-squares regression coefficients computed from the reduced set Y_N of observed variables. However, this can be done only if the presence of indirect confounding has been detected and both Y_i and Y_j are observed.

7. THE INTRODUCTORY EXAMPLES CONTINUED

7·1. *Indirect confounding in an intervention study*

We now continue, first, the example of § 2·3 that illustrates indirect confounding in an intervention study. For node 1 in G^N_{rec}, shown in Fig. 4(b), the conditioning set $b = (2, 3, 4)$ avoids over- and under-conditioning. The omitted variable in G^N_{rec} shown in Fig. 4(a) is the last background variable, and it induces after marginalizing the confounding path $1 \text{---} 3 \prec\!\!\!- 4$, of type (ii) in Corollary 1, but no double edge in G^N_{rec}.

Thus, equation (22) applies and $-A_{14} = \beta_{1|4\cdot23} - (K_{13}/K_{33})A_{34}$ gives, with $\alpha = \beta_{1|4\cdot23} + \delta\gamma\theta/(1-\theta^2)$, the correction needed to recover α from $\beta_{1|4\cdot23}$. Since Fig. 4(b) does not contain a confounding path for $1 \prec\!\!\!- 2$, the coefficient $\beta_{1|2\cdot34}$ is an unconfounded measure of λ.

7·2. *Indirect confounding in an observational study*

For the example of indirect confounding in Fig. 5 in § 2.3, the graph in Fig. 6(a) gives the same type of parent graph as the one in Fig. 5, but for standardized variables related linearly, and Fig. 6(b) shows the induced graph G^N_{rec}.

The linear equations for the parent graph in Fig. 6(a) contain four observed variables and two uncorrelated unobserved variables U and V:

$$Y_1 = \lambda Y_2 + \alpha Y_3 + \omega U + \varepsilon_1, \quad Y_2 = \nu Y_4 + \varepsilon_2, \quad Y_3 = \delta V + \varepsilon_3, \quad Y_4 = \gamma U + \theta V + \varepsilon_4. \tag{25}$$

The equation parameter α is a linear least-squares regression coefficient, $\alpha = \beta_{1|3\cdot 2U} = \beta_{1|3\cdot 24U}$, since Y_2, Y_3 and U are the directly explanatory variables of the response Y_1 and there is no direct contribution of variable Y_4. The induced equations implicitly defined by the graph of Fig. 6(b) are obtained from the generating equations (25), if we use

$$\eta_1 = \omega U + \varepsilon_1, \quad \eta_2 = \varepsilon_2, \quad \eta_3 = \delta V + \varepsilon_3, \quad \eta_4 = \gamma U + \theta V + \varepsilon_4,$$

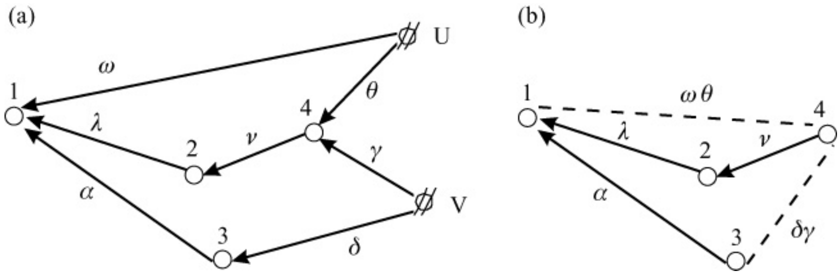

Fig. 6. (a) The parent graph of Fig. 5 with variables relabelled and linear generating coefficients attached; (b) the graph G^N_{rec} induced by (a) without direct confounding, but with indirect confounding of the generating dependence α of 1 on 3 via the confounding path 1 - - - 4 - - - 3.

to give, as equations in the four remaining observed variables,

$$Y_1 = \lambda Y_2 + \alpha Y_3 + \eta_1, \qquad Y_2 = \nu Y_4 + \eta_2, \qquad Y_3 = \eta_3, \qquad Y_4 = \eta_4. \tag{26}$$

The two nonzero residual covariances K_{14} and K_{34} generate the following two nonzero elements in $K_{1b}K_{bb}^{-1}$; for explicit results with longer covariance chains, see Wermuth et al. (2006a). In fact,

$$K_{1b}K_{bb}^{-1} = \left[0,\ -K_{14}K_{34}/\left(K_{33}K_{44} - K_{34}^2\right),\ K_{14}K_{33}/\left(K_{33}K_{44} - K_{34}^2\right)\right].$$

From (22), we obtain the required correction of $\beta_{1\,|\,3\cdot 24}$ to recover $\alpha = -A_{13}$ as

$$\alpha = \beta_{1\,|\,3\cdot 24} + K_{14}K_{34}/\left(K_{33}K_{44} - K_{34}^2\right)\cdot$$

Since there is no confounding path for $1 \prec\!\!-\!\!- 2$, the coefficient $\beta_{1\,|\,2\cdot 34}$ is an unconfounded measure of $\lambda = \beta_{1\,|\,2\cdot 3UV}$.

The following numerical example of the generating process in Fig. 6(a) shows a case of strong effect reversal. The negative values of the linear least-squares coefficients in the generating system are elements of A. The matrix pair (A, Δ^{-1}) is the triangular decomposition of Σ^{-1}, so that $\Sigma^{-1} = A^{\mathrm{T}}\Delta^{-1}A$. The nonzero off-diagonal elements of A and the diagonal elements of Δ are

$$A_{12} = -0{\cdot}30, \quad A_{13} = -0{\cdot}36, \quad A_{15} = -0{\cdot}90, \quad A_{24} = -0{\cdot}60,$$
$$A_{36} = -0{\cdot}90, \quad A_{45} = 0{\cdot}65, \quad A_{46} = 0{\cdot}75,$$
$$\text{diag}(\Delta) = (0{\cdot}2685,\ 0{\cdot}6400,\ 0{\cdot}1900,\ 0{\cdot}0150,\ 1,\ 1).$$

The observed variables correspond to rows and columns 1 to 4 of A, variable U to column 5 and variable V to column 6.

The correlation matrix Σ_{NN} of the four observed variables and the residual covariance matrix, $K_{NN} = A_{NN}\Sigma_{NN}A_{NN}^{\mathrm{T}}$, are

$$\Sigma_{NN} = \begin{pmatrix} 1 & -0{\cdot}1968 & 0{\cdot}2385 & -0{\cdot}6480 \\ \cdot & 1 & -0{\cdot}4050 & 0{\cdot}6000 \\ \cdot & \cdot & 1 & -0{\cdot}6750 \\ \cdot & \cdot & \cdot & 1 \end{pmatrix}, \quad K_{NN} = \begin{pmatrix} 1{\cdot}0785 & 0 & 0 & -0{\cdot}5850 \\ \cdot & 0{\cdot}6400 & 0 & 0 \\ \cdot & \cdot & 1 & -0{\cdot}6750 \\ \cdot & \cdot & \cdot & 1 \end{pmatrix}.$$

Nothing peculiar can be detected in the correlation matrix of the observed variables: there is no very high individual correlation and there is no strong multi-collinearity. The two nonzero elements K_{14} and K_{34} correspond to the two dashed lines in Fig. 6(b).

The generating coefficient of dependence of Y_1 on Y_3, given Y_2 and U, is $-A_{13} = \beta_{1\,|\,3\cdot 2U} = 0{\cdot}36$. The least-squares regression coefficient of Y_3, when Y_1 is regressed on only the observed variables, is $\beta_{1\,|\,3\cdot 24} = -0{\cdot}3654$. This coefficient is of similar strength to that of the generating dependence $\beta_{1\,|\,3\cdot 2U}$, but reversed in sign. This illustrates how severe the effect of indirect confounding can be if it remains undetected: one may come to a qualitatively wrong conclusion.

ACKNOWLEDGEMENT

We thank Giovanni Marchetti, Ayesha Ali, Professor D. M. Titterington and referees for insightful and constructive comments. We are grateful for the support of our cooperation by the Swedish Research Society, directly and via the Chalmers Stochastic Center, and by the Swedish Strategic Fund via the Gothenburg Mathematical Modelling Center.

REFERENCES

ANGRIST, J. D. & KRUEGER, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *J. Econ. Perspect.* **15**, 65–89.

BRITO, C. & PEARL, J. (2002). A new identification condition for recursive models with correlated errors. *Struct. Equ. Model.* **9**, 459–74.

COCHRAN, W. G. (1938). The omission or addition of an independent variate in multiple linear regression. *J. R. Statist. Soc. Suppl.* **5**, 171–6.

COX, D. R. & WERMUTH, N. (1993). Linear dependencies represented by chain graphs (with Discussion). *Statist. Sci.* **8**, 204–18; 247–77.

COX, D. R. & WERMUTH, N. (1996). *Multivariate Dependencies: Models, Analysis, and Interpretation.* London: Chapman and Hall.

COX, D. R. & WERMUTH, N. (2003). A general condition for avoiding effect reversal after marginalization. *J. R. Statist. Soc.* B **56**, 934–40.

CRAMÉR, H. (1946). *Mathematical Methods of Statistics.* Princeton, NJ: Princeton University Press.

EDWARDS, D. (2000). *Introduction to Graphical Modeling*, 2nd ed. New York: Springer.

GOLDBERGER, A. S. (1991). *A Course in Econometrics.* Cambridge, MA: Harvard University Press.

HARDT, J., PETRAK, F., FILIPAS, D. & EGLE, U. T. (2004). Adaption to life after surgical removal of the bladder – An application of graphical Markov models for analysing longitudinal data. *Statist. Med.* **23**, 649–66.

HAUSMAN, J. A. (1983). Instrumental variable estimation. In *Encyclopedia of Statistical Sciences* 4, Ed. S. Kotz, N. L. Johnson and C. B. Read, pp. 150–3. New York: Wiley.

KIIVERI, H. T. (1987). An incomplete data approach to the analysis of covariance structures. *Psychometrika* **52**, 539–54.

KIIVERI, H. T., SPEED, T. P. & CARLIN, J. B. (1984). Recursive causal models. *J. Aust. Math. Soc.* A **36**, 30–52.

LAURITZEN, S. L. (1996). *Graphical Models*. Oxford: Oxford University Press.

LAURITZEN, S. L. & WERMUTH, N. (1989). Graphical models for associations between variables, some of which are qualitative and some quantitative. *Ann. Statist.* **17,** 31–54.

MA, Z., XIE, X. & GENG, Z. (2006). Collapsibility of distribution dependence. *J. R. Statist. Soc.* B **68**, 127–33.

MARCHETTI, G. M. (2006). Independencies induced from a graphical Markov model after marginalization and conditioning: The R package ggm. *J. Statist. Software* **15**, issue 6.

MARCHETTI, G. M. & WERMUTH N. (2008). Matrix representations and independencies in directed acyclic graphs. *Annals of Statistics* **36**, to appear.

ROBINS, J. & WASSERMAN, L. (1997). Estimation of effects of sequential treatments by reparametrizing directed acyclic graphs. In *Proc. 13th Annual Conf. Uncertainty Artificial Intelligence,* Ed. D. Geiger and O. Shenoy, pp. 409–20. San Francisco, CA: Morgan and Kaufmann.

SARGAN, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica* **26**, 393–415.

STANGHELLINI, E. & WERMUTH, N. (2005). On the identification of path analysis models with one hidden variable. *Biometrika* **92**, 337–50.

TUKEY, J. W. (1954). Causation, regression, and path analysis. In *Statistics and Mathematics in Biology*, Ed. O. Kempthorne, T. A. Bancroft, J. W. Gowen and J. L. Lush, pp. 35–66. Ames: The Iowa State College Press.

WERMUTH, N. (1980). Linear recursive equations, covariance selection, and path analysis. *J. Am. Statist. Assoc.* **75**, 963–97.

WERMUTH, N. (2005). Graphical chain models. In *Encyclopedia of Behavioral Statistics, II*, Ed. B. Everitt and David C. Howell, pp. 755–7. Chichester: Wiley.

WERMUTH, N. & COX, D. R. (1998a). On association models defined over independence graphs. *Bernoulli* **4**, 477–95.

WERMUTH, N. & COX, D. R. (1998b). Statistical dependence and independence. In *Encyclopedia of Biostatistics,* Ed. P. Armitage and T. Colton, pp. 4260–7. New York: Wiley.

WERMUTH, N. & COX, D. R. (2004). Joint response graphs and separation induced by triangular systems. *J. R. Statist. Soc.* B **66**, 687–717.

WERMUTH, N., COX, D. R. & MARCHETTI, G. (2006a). Covariance chains. *Bernoulli* **12**, 841–62.

WERMUTH, N., WIEDENBECK, M. & COX, D. R. (2006b). Partial inversion for linear systems and partial closure of independence graphs. *BIT, Numer. Math.* **46**, 883–901.

WOLD, H. O. (1954). Causality and econometrics. *Econometrica* **22**, 162–77.

WRIGHT, S. (1923). The theory of path coefficients: A reply to Niles' criticism. *Genetics* **8**, 239–55.

WRIGHT, S. (1934). The method of path coefficients. *Ann. Math. Statist.* **5**, 161–215.

[*Received April* 2006. *Revised June* 2007]