

基于 DAG 方法的因果结构识别原理与效率研究^①

聂思玥¹, 刘维奇², 李梦花³

(1. 山西大学 管理与决策研究所, 太原 030006; 2. 山西财经大学, 太原 030006;
3. 山西财经大学 经济学院, 太原 030006)

摘要: 结构分析是宏观经济计量分析的四大任务之一, 变量之间因果结构推断一直是结构分析的重要内容。本文较为系统地阐述了基于有向无环图 (DAG) 方法的因果结构推断的原理, 且对 PC 因果搜索算法进行了详细分析。本文还研究了 DAG 方法的效率, 即对 DAG 因果识别原理中关键的 D-分离的三种情形在线性模型正态误差干扰项的假设下用蒙特卡罗方法模拟分析了 Fisher-z 统计量检验条件独立性的实际检验水平和检验功效, 并对实际中应用 DAG 方法给出了相关建议。

关键词: 有向无环图 (DAG); 条件独立; 检验水平; 检验功效

中图分类号: F224 C812 **文献标识码:** A **文章编号:** (2017) 01-0075-12

0 引言

20 世纪 70 年代之前, 宏观经济计量分析的数据描述、模型预测、结构分析以及政策分析四大任务一般由传统大型联立方程模型承担^[1]。变量之间的结构分析是宏观计量经济分析的四大任务之一, 李子奈指出结构分析旨在揭示经济主体与环境之间的动力学关系, 通俗讲就是揭示变量之间的关系, 是通过对模型结构参数的估计来实现的^[2]。在实证中使用联立方程, 一般依赖于一个先验的经济理论来识别模型中的参数。但经济理论的识别方法被认为带有主观色彩, 且随着联立方程模型对 20 世纪 70 年代的经济衰退和滞胀的预测和政策分析失效, 引来了著名的“卢卡斯批判”, 这样联立方程逐步退出历史舞台。在这种情况下, Sims 提出了新的宏观计量分析架构——向量自回归 (VAR) 模型, 也称简化型 VAR (reduced-form VAR, 对应于后文的结构型 VAR)^[3]。但 Cooley 和 LeRoy 批判简化型 VAR 本身并不能直接观测到系统内变量之间的关系, 因为各个方程中并不包含变量之间当期的关系结构, 这使得结构推断变得困难^[4]。正是基于这些问题的存在, SimsVAR 模型的基础上提出了 SVAR (Structural VAR, 结构型 VAR)^[5]。SVAR 模型旨在通过变量之间当期关系矩阵描述变量之间的因果结构 (causal structure), 但 SVAR 模型在变量当期因果关系的识别上很大程度也依赖于经济理论, 如利用 Cholesky 分解得到的变量之间的 Wold 因果链, 就假定变量之间的因果关系是递归的, 而递归因果结构最重要的是确定变量之间的排序问题。

1 文献综述

有关变量之间因果分析的文献较多, Granger 在这方面的贡献较有代表性。Granger 最初从依据概率理论和数据的统计性质来分析变量之间的因果关系^[6], 之后又分析了时间序列变量之间的因果关系^[7], 认为序列 $\{x_t\}$ 如果有助于改善序列 $\{y_t\}$ 的向前一期预测, 则称 $\{x_t\}$ 是 $\{y_t\}$ 的格兰杰原因。但格兰杰因果被 Hoover 等批判, 认为其无法揭示变量的因果结构, 因为这种因果关系是基于时间顺序的结果, 即

^① 基金项目: 山西省高等学校哲学社会科学项目 (2016202)。

作者简介: 聂思玥 (1982—), 男, 江西吉安人, 经济学博士, 山西大学管理与决策研究所讲师, 研究方向: 金融计量、金融工程与风险管理, Email: niesy119@163.com; 刘维奇 (1963—), 男, 山西忻州人, 管理学博士, 山西大学管理与决策研究所教授, 博士生导师, 研究方向: 金融工程和风险管理, Email: liuwq@sxu.edu.cn; 李梦花 (1982—), 女, 山西运城人, 经济学博士, 山西财经大学经济学院讲师, 研究方向: 货币和宏观经济, Email: lmh0912519@163.com。

使 $\{x_t\}$ 是 $\{y_t\}$ 的格兰杰原因, 也并不意味着存在一个通过控制 $\{x_t\}$ 来达到对 $\{y_t\}$ 进行控制的经济机理; 反过来, 通常这种机理的存在也不意味着变量之间存在格兰杰因果关系^[8]。

近年来, 由加利福尼亚大学洛杉矶分校、卡内基梅隆大学和斯坦福大学等高校学者发展起来的因果图方法在分析因果结构方面不断受到重视, 计量经济学界也正兴起对这方面研究的热潮。这个方法结合了哲学逻辑、统计学等思想, 用图论的方法进行因果结构研究, 代表性成果是有向无环图 (directed acyclic graph, DAG) 分析方法。DAG 图可以对截面数据分析变量之间的当期因果关系, 从而得到变量之间直接控制的内在因果机理。用 DAG 来分析因果关系在人工智能领域的代表性学者有 Spirtes 等^[9] 和 Pearl^[10], 但是在经济学领域的应用还相对较少。在经济学领域, 该方法最早由 Swanson 和 Granger 在 VAR 脉冲响应分析中使用^[11]。随后有大量文献对 DAG 方法在计量经济学领域尤其将 DAG 应用到 SVAR 的因果关系推断方面进行了讨论^[12-16]。Bryant 等分析了对 DAG 方法中变量因果关系进行证伪的一种假设检验方法^[13]。Wang 用 DAG 方法分析了债券市场中变量之间的因果结构^[17]。

国内在 SVAR 模型中使用 DAG 方法进行实证研究的文献正不断出现^[18-20]。这些文献较少涉及 DAG 方法理论的分析, 聂思玥和李梦花^[21]对 DAG 方法中使用的条件独立性检验统计量进行了比较研究, 而本文则对 DAG 方法基本理论进行了较为系统深入地分析, 且结合案例分析了针对该方法目前最常用的 PC 因果搜索算法, 对国内研究 DAG 理论方面形成有益补充。为了分析 DAG 方法的效率, 并考虑到目前主流 DAG 方法都使用 Fisher-Z 统计量进行条件独立性检验, 本文对 DAG 因果识别原理中关键的 D-分离的三种情形在线性模型正态误差干扰项的假设下用蒙特卡罗方法模拟了 Fisher-Z 统计量检验条件独立性的实际检验水平和检验功效, 并对实际应用给出了有用的建议。

2 DAG 方法原理分析

变量的因果结构是由变量之间的直接和间接因果关系组成。DAG 方法运用概率空间的条件独立性检验理论, 对变量两两之间的条件独立性进行判断, 并通过一定的搜索算法, 建立变量间的单向控制模型, 得到最终因果结构图。

2.1 DAG 方法基本原理

为了完整描述这个理论, 首先介绍如下基本定义。

节点: 表示一个变量, 如图 1 (a) 中的 A、B、C、D、E。

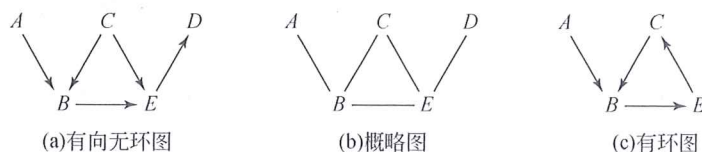


图 1 有向无环图、概略图 (无向图) 和有环图示例

Fig. 1 Illustration of DAG, skeleton diagram and DCG

边: 连接节点的边, 表示被连接的两个变量间存在因果关系, 分为有向边和无向边。有箭头指向的有向边可明确表示节点之间的因果。当全部边都是有向边时, 因果图揭示了图中所有变量之间的因果关系结构, 如图 1 (a) 所示。有向边 $A \rightarrow B$ 表示 A 导致 B, 图 1 (b) 无向边 $D-E$ 表示 D、E 间存在因果关系, 但是指向不明。

父/子节点: 有向边的起源节点称为父节点, 指向节点称为子节点, 在 $A \rightarrow B$ 中, A 为父节点, B 为子节点。

路径: 两个节点之间的因果链, 分为有向路径和无向路径。如图 1 (a) 中 A、E 之间有向路径 $A \rightarrow B \rightarrow E$ 和无向路径 $A \rightarrow B \leftarrow C \rightarrow E$ 两个路径。在有向路径中, 起源节点称为祖先节点, 其他节点称为其子孙节点。

无盾相遇点 (unshielded collider): 三个相邻节点被两条边相连, 一个节点同时是另外两个节点的子节点, 且两个父节点之间无直接相连的边, 则该节点被称为无盾相遇点, 这种结构被称为反叉因果结构。如在图 1 (a) 中的 $A \rightarrow B \leftarrow C$, 此时 A、C 之间没有直接相连的边, 则 B 是一个无盾相遇点。

环：由子孙节点到祖先节点的有向路径，称为环。图 1 (c) 中 $C \rightarrow B \rightarrow E \rightarrow C$ 形成了一个环。在 DAG 中，明确假定研究变量之间不存在环状因果链。

概略图：只有节点和无向边的图，称为概略图，概略图是无向图。图 1 (a) 的概略图见图 1 (b)。

DAG 是一类无环的有向结构图，变量间是单向控制关系，存在下述定理。

定理 1：DAG 所描述的因果结构一定是递归类型的因果结构。

证明：在无环图中，任何节点都只可能是它的子孙节点的“原因”。DAG 是无环图，其所有节点的“原因”都只能是其祖先节点，同时也必然至少存在一个节点作为祖先节点，没有“原因”。而由下三角元素不全为 0 的矩阵所描述的递归类型的因果结构正是这样的一个因果结构，即第 i 个方程的当期解释变量只包含前面第 1, \dots , $i-1$ 个（祖先节点）变量，第 1 个方程无当期解释变量，从而得证。

图 1 (a) 所描述的因果结构，可由矩阵形式表述如下（下面矩阵中 \times 表示非 0 元素）：

$$A_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \times & 1 & \times & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & \times \\ 0 & \times & \times & 0 & 1 \end{bmatrix}, Y = \begin{bmatrix} A \\ B \\ C \\ D \\ E \end{bmatrix}$$

这个矩阵关系看似描述的不是下三角的递归因果结构，但是通过恰当的变量排序，如 A 、 C 、 B 、 E 、 D ，则可以写成下三角矩阵形式的因果结构：

$$A_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \times & \times & 1 & 0 & 0 \\ 0 & \times & \times & 1 & 0 \\ 0 & 0 & 0 & \times & 1 \end{bmatrix}, Y = \begin{bmatrix} A \\ C \\ B \\ E \\ D \end{bmatrix}$$

DAG 图和矩阵 A_0 都描述了变量集 Y 之间的因果结构。两个变量之间的因果关系可从概率空间的条件独立性检验来判定。Suppes^[22] 认为，如果事件 A 使得事件 B 发生的概率增大，即 $P(B|A) > P(B)$ ，则 A 就成为 B 发生的原因。但有时也可能是“共同因素（common cause）” C 造成了 A 与 B 之间的联系；或 A 还通过影响因素 C 来间接影响 B ，可见这些复杂情形会影响两个变量是否存在直接因果关系的判断。条件独立性则可将无关因素作为条件进行屏蔽，得到研究者所关心的两个变量之间的因果关系。Granger 是较早早在因果分析领域引入条件独立性概念的学者，他指出如果式 (1) 成立，则 $\{x_t\}$ 不是 $\{y_t\}$ 的格兰杰原因：

$$y_{t+1} \perp (x_t, x_{t-1}, \dots, x_{t-p}) | (y_t, y_{t-1}, \dots, y_{t-p}; z_t, z_{t-1}, \dots, z_{t-p}) \quad (1)$$

式中， p 为时间滞后阶数， $\{z_t\}$ 为其他解释 $\{y_t\}$ 的变量， $y \perp x | z$ 表示在 z 条件下， y 独立于 x 。可见，变量之间的因果关系可以运用变量之间的条件独立性进行推断。做一般性推广，信息集 M 在信息集 O 的条件下与信息集 N 独立，记作 $M \perp N | O$ ，即式 (2) 成立：

$$M \perp N | O \text{ iff } P(M|N, O) = P(M|O) \quad (2)$$

式中 iff 表示当且仅当， $P(\cdot | \cdot)$ 表示条件概率。式 (2) 的含义是，在知道信息集 O 以后，则对于了解信息集 M ，信息集 N 已不能提供任何多余的信息。若 O 为空集且式 (2) 成立，则表示 M 、 N 独立，无因果关系；但当 O 不为空集，且式 (2) 也不成立时， M 、 N 与条件集 O 之间存在如图 3 (b) 的因果结构；当 O 为空集时式 (2) 不成立，而 O 不为空集时式 (2) 成立，代表 M 、 N 之间无直接因果关系，有间接因果关系，如图 3 [(c), (d), (e)] 所示。

DAG 方法是基于变量集之间的 D -分离（directionally separate）情形，运用条件独立性来推断变量之间的因果结构。和 Wang 及 Kwon 和 Bessler 指出，在 DAG 图中，两个节点 q 、 r 被 D -分离是指因果图中的两节点之间的全部路径均被一个 O 子集所阻断，即 D -分离^[23,24]。因果图中的一条路径被 O 子集 D -分离，当且仅当下面两种情况成立：

- 1) 该路径中包含①因果链： $q \rightarrow s \rightarrow r$ 、 $q \leftarrow s \leftarrow r$ 或②因果叉： $q \leftarrow s \rightarrow r$ 时，子集 O 中包含节点 s ；
- 2) 该路径中包含③反叉因果结构： $q \rightarrow s \leftarrow r$ 时，子集 O 中不包含节点 s 及其子孙节点。

在因果叉②中 $q \leftarrow s \rightarrow r$ 中，无论 q 、 r 之间是否有直接相连边， s 被称为 q 、 r 的公共原因（common cause）。

按上述定义，图 1 (a) 中，节点 A 、 D 被 $\{B、C、E\}$ 集合 D -分离，因 A 、 D 之间的所有路径有两条 $A \rightarrow B \rightarrow E \rightarrow D$ 和 $A \rightarrow B \leftarrow C \rightarrow E \rightarrow D$ ，均被 $\{B、C、E\}$ 集合 D -分离。

Pearl 指出，除了 D -分离以外，DAG 因果图 G 还需要满足两个条件：Markov 因果条件和因果忠实 (faithful) 条件^[10]，才具备推断因果结构的性质。为方便阐述 Markov 因果条件和因果忠实条件，用 $Y_i = f(YI_i, u_i)$ 来代表变量集 Y 中的第 i 个变量的等式方程，其中 YI_i 是 Y_i 的解释变量， u_i 是误差干扰项。

Markov 因果条件：变量集 Y 的因果图定义为 G ， P 是关于 Y 的概率分布函数。 W 是 Y 的一个子集。当且仅当 Y 中的每个子集 W 在其祖先节点变量集的条件下，与不包含 W 的子孙节点变量的其他子集，相互独立；则 G 、 P 满足 Markov 因果条件。Markov 因果条件涵括了两个主要内容：其一，变量之间互为因果的循环结构不存在，即因果图中没有环路；其二， G 中所有变量的公共原因都已包含在因果图 G 中，确保各 u_i 之间相互独立。第二点也称之为因果充分 (causal sufficient) 条件。

因果忠实条件：当方程组 $Y_i = f(A_i, u_i)$ 中的参数发生变化时，因果图 G 和变量集 Y 中的因果关系结构并不随之改变，则称 G 满足因果忠实条件。因果忠实条件的含义是变量之间内在的因果结构是稳定的，可用不标示参数的 DAG 因果图来描述这种因果结构。

D -分离、Markov 因果条件和因果忠实条件构成了 DAG 中变量之间条件独立性分析的基础：首先，被 D -分离的变量 q 、 r ，在分离集 O 的条件下相互独立，即，因果链①中，将 s 视为祖先节点变量集， q 、 r 在条件 s 下相互独立；其次，在反叉因果结构③中，相互独立的变量 q 、 r 在无盾相遇点 s 下变得相关，其原因在于： q 、 r 都是导致 s 的因素，当 q 因素的影响参数变弱时，则由 r 导致 s 的概率会增加，由此可见 q 、 r 在条件 s 下是相关的。因果忠实条件则保证了这种因果关系在参数发生变化时仍然是稳定的。

因此，因果链①和因果叉②中，满足 $q \perp r | s$ ；而反叉因果结构③下，满足 $q \not\perp r | s$ ($\not\perp$ 表示非独立)。如果在因果图中存在环路，如图 1 (c) 所示，虽然 $A \rightarrow B \leftarrow C$ 构成无盾相遇点下的反叉因果结构，但因其不满足 Markov 因果条件， C 节点与 A 节点 (A 非 C 的子孙节点)，在条件 B (B 是 C 的祖先节点) 下，并不相互独立。

正是无盾相遇点 [如图 3 (b) 中反叉因果结构] 在条件独立性分析方面所具备的，与其他两种因果结构 (因果链和因果叉) 不相同的特点，使其成为搜索变量之间因果结构的关键。在图 1 (a) 中，有 1 个无盾相遇点，即在 $A \rightarrow B \leftarrow C$ 中的 B 点；在图 2 (a) 中，有两个无盾相遇点，分别是在 $A \rightarrow B \leftarrow C$ 中的 B 点，和在 $A \rightarrow B \leftarrow E$ 中的 B 点。

上文阐述的是 DAG 方法中条件独立性分析的基本理论。利用上述理论，再结合恰当的搜索算法，即可在一组变量中 (满足 Markov 因果条件和因果忠实条件) 搜索并得到其因果结构。但由于 DAG 下述观测等价定理的存在，搜索得到的最终 DAG 因果结构图可能并不唯一。

定理 2 观测等价定理：任何一个有向无环图 G_1 的概率分布都可由另外一个与 G_1 有相同的概略图和无盾相遇点的有向无环图 G_2 等价地描述，即一组数据可以由有相同的概略图和无盾相遇点的有向无环图中的任何一个等价地描述。

证明：参考 Pearl^[10]，定理 1.2.8。

图 1 (a)、图 2 [(a), (b), (c)] 有相同的概略图——图 1 (b)，但图 1 (a) 只包含一个无盾相遇点，而图 2 [(a), (b), (c)] 都包含两个无盾相遇点。根据观测等价定理，图 2 (a) 与图 2 [(b), (c)] 所描述的数据观测是一致的，而图 1 (a) 则不是。观测等价定理是由 DAG 的假定条件所决定，不随搜索算法的改变而改变。其描述的实质是在 DAG 的既定假定条件下，如图 2 中符合观测等价定理的 3 个 DAG 图有相同的概率似然函数。

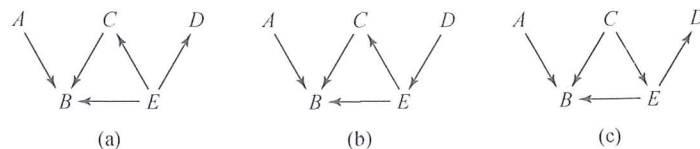


图 2 观测等价定理示例

Fig. 2 Illustration of Observational equivalence theorem

2.1.1 DAG 因果搜索——PC 算法

DAG 图的因果结构是通过变量之间的经验概率分布进行搜索得到。PC 算法按照 D-分离中因果链①、因果叉②和反叉因果结构③的三种情形，使用变量之间方差协方差矩阵或相关矩阵检验条件独立性，得到模型变量之间的因果结构。

1. 算法步骤

假定数据生成过程 DGP_1 、 DGP_2 分别如图 1 (a) 和图 2 (a) 所描述，现以图 1 (a) 和图 2 (a) 为例阐述 PC 搜索算法的基本步骤：

步骤 1，PC 算法首先假定全部变量，其两两之间都存在因果关系，即两两之间都用直接边相连，形成一个完整的无向图。

步骤 2，先通过检验变量两两之间无条件独立性，删除相互独立变量（无任何因果）之间的无向边；再检验变量间的条件独立性，排除两个变量在因果链和因果叉结构中被直接边相连的情况，这样即可得到全部变量的无向因果概略图。通过检验变量两两之间的无条件独立性，如图 1 (a) 中 A、C 间的无向边会被删除；再通过检验变量之间的条件独立性，则图 1 (a) 中 AE、AD、CD、BD 间的无向边将被逐一删除（因 A、E 在条件 B 下相互独立；C、D 在条件 E 下相互独立；B、D 在条件 E 下相互独立；A、D 在条件 B、E 下相互独立）。完成第 2 个步骤就可以得到 DAG 因果图的概略图，如图 1 (b) 所示。而对于图 2 (a)，虽然这个搜索过程的顺序略有不同，但最终得到的概略图也将如图 1 (b) 所示。

步骤 3，运用统计和逻辑分析，对概略图中的所有无向边进行定向。这个步骤需要综合统计检验、逻辑推理、预先假设等方面的知识，是一个复杂过程。一般思路是，先利用无盾相遇点的性质对 DAG 图进行部分定向，再综合其他知识分析图中其他未定向的边，因此可分两小步阐述：

首先，通过条件独立性检验，发现全部无盾相遇点并定向。检验概略图中所有由第 3 个变量连接的 2 个变量（2 个变量间没有直接连接）之间关于第 3 个变量的条件独立性。如图 1 (b) 中 ABC、ABE、BED、CED 检验要求，在 DGP_1 的情形下，图 1 (b) 中 ABC 定向为 $A \rightarrow B \leftarrow C$ ；在 DGP_2 的情形下，图 1 (b) 中 ABC 定向为 $A \rightarrow B \leftarrow C$ ，ABE 定向为 $A \rightarrow B \leftarrow E$ 。

其次，分析其他未定向边。在 DGP_1 的情形下， $A \rightarrow B - E$ 被定向为 $A \rightarrow B \rightarrow E$ （因为若 BE 定向 $B \leftarrow E$ ，将在路径 ABE 中新增了一个无盾相遇点 B）；根据 DAG 无环的假设，需将 CE 定向为 $C \rightarrow E$ ；最后 ED 被定向为 $E \rightarrow D$ ，以保证 E 点不是无盾相遇点。这样 DGP_1 的因果图就成功得到还原。但是并非所有的因果图都能被完整地还原，根据观测等价定理，同一个数据观测可能对应着多个 DAG 因果图。例如，在 DGP_2 的情形下，CE 和 DE 未定向，其定向有四种情况，即 $C \rightarrow E$ ， $C \leftarrow E$ 与 $D \rightarrow E$ ， $D \leftarrow E$ 的四个组合。在无盾相遇点确定的情况下，排除 $C \rightarrow E \leftarrow D$ ，这样就存在三种定向组合，分别是图 2[(a)，(b)，(c)]，符合观测等价定理。

2. 算法分析

PC 算法的核心思想运用条件独立性检验与假设条件对变量间因果关系的形式进行排除分析，最后得到表示变量之间因果结构的有向无环图。图 3 显示了两个变量之间因果关系的全部可能形式，图 3[(b)，(c)，(d)，(e)] 所表示的因果结构正是 D-分离的 3 种情形 [图 3 (d)/(e) 模型同属因果链模型]。PC 算法就是依据条件独立性检验结果对图 3 中变量间的因果关系进行选择。

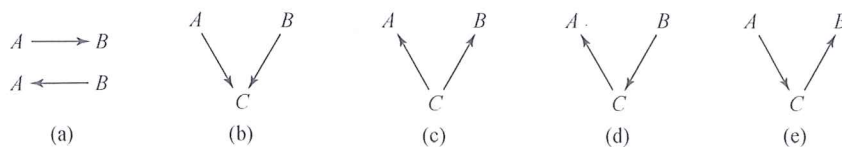


图 3 两变量间的因果关系图

Fig. 3 The causal relationship between two variables

PC 算法的步骤 1 中假定全部变量之间存在直接因果关系，步骤 2、步骤 3 对因果关系的形式进行排除和区分。步骤 2 的无条件独立性检验，就是排除变量间没有直接因果关系的情况；而此时有边连接的变量之间也未必有直接因果关系，如图 3[(c)，(d)，(e)] 中 A、B 都是无条件相关的，但在条件 C 下相互

独立，因此不存在直接因果关系，这些情况在条件独立性检验下将排除，即 A 、 B 之间的边将被删除。剩下的边都表示变量之间存在直接因果关系，得到 DAG 的概略图，且任意变量之间的边一定是图 3 中 5 种情况之一。对这 5 种情况进行分析如下：

第一，所有边都服从图 3 (a)，因为是直接因果，但这个信息无法帮助定向；

第二，其他四种情形中，图 3[(b), (c), (d), (e)] 的结构是类似的，但图 3[(c), (d), (e)] 都是 $A \perp B | C$ ，而图 3 (b) 中 C 是无盾相遇点，即 $A \perp B | C$ ，因此对概略图中这类结构进行条件独立性检验，若 A 、 B 在条件 C 下相关，则一定是图 3 (b) 的情形，可定向为 $A \rightarrow C \leftarrow B$ 。

找出图中全部的无盾相遇点，步骤 3 的第一小步就完成了。在第二小步中，总体原则有两个：一是在后续定向过程中需确保不新增加无盾相遇点；二是保证因果图中不出现环。在这两个原则下，可总结得到以下几个基本定向规则。

(1) 对于图 4 (a)， BC 的定向应是 $C \rightarrow B$ ，否则 C 点成为无盾相遇点；

(2) 对于图 4 (b)， AB 的定向应是 $A \rightarrow B$ ，避免出现环；

(3) 对于图 4 (c)，可确定 CD 定向是 $D \rightarrow C$ (若 $D \leftarrow C$ ，为了避免出现环，须 $A \rightarrow D$ 、 $B \rightarrow D$ ，这样 D 节点是新增的无盾相遇点，不可取)；

(4) 对于图 4 (d)，依据原则 (1) 可确定 BC 的定向应是 $C \rightarrow B$ ，则 DB 定向是 $D \rightarrow B$ (若 $B \rightarrow D$ ，为了避免出现环，须 $A \rightarrow D$ ，这样 D 节点是新增的无盾相遇点，不可取)。

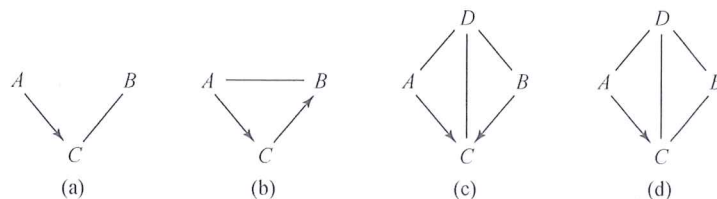


图 4 因果关系搜索的基本定向规则

Fig. 4 Basic orientation of the rules of causality search

3 条件独立性检验

在因果关系搜索的过程中，条件独立性检验非常关键。Scheines 指出在假定 $f(\cdot)$ 是线性模型和 u_i 服从正态分布的条件下，检验条件独立性与检验条件相关性是一致的，可使用 Fisher-Z 检验^[25]：

$$z(r_{XY|O}, T) = \frac{1}{2}(T - N_o - 3)^{\frac{1}{2}} \ln \left(\frac{1 + r_{XY|O}}{1 - r_{XY|O}} \right) \tag{3}$$

式中， T 为变量的样本容量个数； N_o 为条件集 O 中的变量个数； $r_{XY|O}$ 为 A 、 B 在条件集 O 下的相关系数，

$$r_{XY|O} = (r_{XY} - r_{XO}r_{YO}) / (\sqrt{1 - r_{XO}^2} \sqrt{1 - r_{YO}^2}) \tag{4}$$

$z(r_{XY|O}, T)$ 服从正态分布：

$$z(r_{XY|O}, T) - z(\widehat{r_{XY|O}}, T) \sim N(0, 1) \tag{5}$$

对于条件相关的检验，可以用偏相关系数作为统计量进行检验。对式 (6) 用偏回归的方法得到偏相关系数统计量 $r_{u_i u_j | u_k}$ ，再检验 $r_{u_i u_j | u_k} = 0$ ^[11]：

$$u_i = \alpha_j u_j + \alpha_k u_k + \varepsilon_i \tag{6}$$

事实上，在线性模型和正态误差干扰项假设下，Swanson 和 Granger 是通过检验 $\alpha_j = 0$ 是否成立的 t 检验代替检验 $r_{u_i u_j | u_k} = 0$ 。

4 DAG 方法的效率研究

由第 2 小节的搜索算法分析可知，分析 DAG 方法的效率只需研究路径 D -分离中的三种基本情形的识

别效率，即分析三种情形下的假设检验的实际检验水平和检验功效。对路径 **D**-分离的图 3[(b), (c), (d)/(e)] 识别效率分析时，三种情形可视为三个独立的子模型用蒙特卡罗模拟方法进行分析。

这 3 个模型是对一个问题进行两种不同性质的假设检验，即 $H_0: A, B$ 在条件 C 下相互独立。图 3 [(d)/(e)]、图 3 (c) 在条件 C 下相互独立，数据生成过程为图示中的模型时，可用蒙特卡罗模拟 Fisher-Z 统计量的实际检验水平 (real size)，即真实数据生成过程为 H_0 时，拒绝 H_0 (弃真) 的概率。图 3 (b) 在条件 C 下相关，数据生成过程为图示中的模型时，可用蒙特卡罗模拟 Fisher-Z 统计量的检验功效 (power)，即真实数据生成过程不是 H_0 时，拒绝 H_0 (不取伪) 的概率。冯蕾和聂巧平对实际检验水平和检验功效进行了较为详细的论述^[26]。

1) 图 3 (b) 模型的检验效率分析——检验功效的模拟。图 3 (b) 所示的模型设定如下：

$$\begin{aligned} A &= \{u_i\}, u \sim N(\mathbf{0}, \sigma(A)I), i=1, 2, \dots \\ B &= \{v_i\}, v \sim N(\mathbf{0}, \sigma(B)I), i=1, 2, \dots \\ C &= \beta_0 + \beta_1 A + \beta_2 B + \eta \varepsilon, \varepsilon \sim N(\mathbf{0}, I) \end{aligned} \quad (7)$$

式中， u_i 为 u 中元素； v_i 为 v 中元素； I 为单位阵； u, v 相互独立。模型设定中， A, B 独立，根据回归系数与相关系数的关系 $r_{AC} = \beta_1 \frac{\sigma(A)}{\sigma(C)}$ ， $r_{BC} = \beta_2 \frac{\sigma(B)}{\sigma(C)}$ ，又由式 (6) 可知， $\sigma(C) = \sqrt{\beta_1^2 + \beta_2^2 + \eta^2}$ ，结合 $r_{AB|C}$ 公式，可得到：

$$\begin{aligned} |r_{AB|C}| &= |(r_{AB} - r_{AC}r_{BC}) / (\sqrt{1-r_{AC}^2}\sqrt{1-r_{BC}^2})| \\ &= \left| \left(0 - \beta_1\beta_2 \frac{\sigma(A)\sigma(B)}{\sigma(C)\sigma(C)} \right) / \left(\sqrt{1 - \left(\beta_1 \frac{\sigma(A)}{\sigma(C)} \right)^2} \sqrt{1 - \left(\beta_2 \frac{\sigma(B)}{\sigma(C)} \right)^2} \right) \right| \end{aligned} \quad (8)$$

在模型 (7) 中，可以将 $\beta_1, \beta_2, \sigma(A), \sigma(B)$ 视为因果信号强度控制变量，而 $\sigma(C)$ 可视为噪声强度控制变量。信号强度越大，噪声强度越小时，条件相关系数绝对数 $|r_{AB|C}|$ 越大，越容易检测到真实的因果关系。当 $|r_{AB|C}| \rightarrow 1$ 时，由式 (3)，可知 $|z(r_{AB|C}, T)| \rightarrow \infty$ ；当 $|r_{AB|C}| \rightarrow 0$ 时， $|z(r_{AB|C}, T)| \rightarrow 0$ 。

在模拟中，可控制信号强度不变，调整噪声 $\sigma(C)$ 的强度来达到调整信号噪声比率的目的，得到不同情况的模拟结果，结果见表 1。

表 1 模型 (7) 的因果识别检验功效模拟结果 (1)

Table 1 Test power of model (7) in causality identity simulations (1)

样本容量 噪声强 (η)	50	100	150	200	250
1	0.992	1	1	1	1
1.5	0.792	0.881	0.919	0.953	0.973
2	0.521	0.505	0.516	0.504	0.525
2.5	0.344	0.278	0.224	0.198	0.172
3	0.273	0.17	0.11	0.082	0.059
3.5	0.229	0.12	0.067	0.04	0.029
4	0.205	0.093	0.044	0.026	0.015

注：本表是在设定 $\beta_1=1, \beta_2=-1, \sigma(A)=1, \sigma(B)=1$ 时，分别设定噪声强度 η 为 1、1.5、2、2.5、3、3.5、4，在样本容量为 50、100、150、200、250 下用 stata11 软件各模拟 10000 次得到的结果，设定检验水平为 0.05。

在表 1 中，当 $\eta=1$ 时表明噪声与信号强度相同，DAG 方法在样本量为 50 对该模型的检验功效高达 99.2%。当噪声增强，检验功效迅速下降，但当样本容量增大时，有所改善。而当噪声强度在 1.5 以上时，检验功效很低，增大样本容量也不能改善检验结果。这表明在实际中，当误差干扰项的标准差较大时用 Fisher-Z 检验得到的因果关系并不可靠，犯取伪错误的概率很大，即将实际存在条件相关的因果关系误检验为条件独立。实际上，当误差干扰项的标准差较大时，因果信号相对于噪声很小，已经被噪声所覆盖，样本统计特征以噪声的基本特征为主，可认为此时样本中的变量不存在因果关系，所以当增大样本容量时反而出现检验功效降低的现象。

表2分别展示了 $\beta_1=\beta_2=1.5$ 和 $\sigma_1=\sigma_2=1.5$ 的模拟结果,对比表2,可以发现,信号增强后,对模型识别的检验功效显著增大,具有非常好的稳健性。

表2 模型(7)的因果识别检验功效模拟结果(2)

Table 2 Test power of model (7) in causality identity simulations (2)

噪声强度 \ 样本容量	样本容量					噪声强度 \ 样本容量	样本容量				
	50	100	150	200	250		50	100	150	200	250
η	$\beta_1=1.5, \beta_2=1.5$					η	$\sigma(A)=1.5, \sigma(B)=1.5$				
1	1	1	1	1	1	1	1	1	1	1	
1.5	0.992	1	1	1	1	0.992	1	1	1	1	
2	0.891	0.961	0.987	0.994	0.997	0.884	0.960	0.985	0.993	0.998	
2.5	0.697	0.759	0.805	0.847	0.872	0.693	0.767	0.810	0.844	0.868	
3	0.510	0.520	0.524	0.517	0.516	0.522	0.508	0.520	0.525	0.510	
3.5	0.393	0.328	0.306	0.275	0.247	0.397	0.342	0.305	0.272	0.248	
4	0.322	0.230	0.174	0.149	0.118	0.323	0.237	0.181	0.144	0.115	

注:本表是在分别设定噪声强度 η 为1、1.5、2、2.5、3、3.5、4,在样本容量为50、100、150、200、250下用stata11软件各模拟10000次得到的结果,设定检验水平为0.05。

此外, $\beta_1, \beta_2, \sigma(A), \sigma(B)$ 以及 $\sigma(C)$ 所描述的信号噪声关系,上述控制变量也决定了系数 β_1, β_2 的 t 统计量。 $\beta_1, \beta_2, \sigma(A), \sigma(B)$ 越大, $\sigma(C)$ 越小, t 统计量绝对值越大,表明因果关系的信号强,噪声弱。用模型(7)来解释这点也很容易, t 统计量绝对值越大,模型中系数 β_1, β_2 的显著性越强,从而模型(7)所表示的因果结构存在的概率越大。

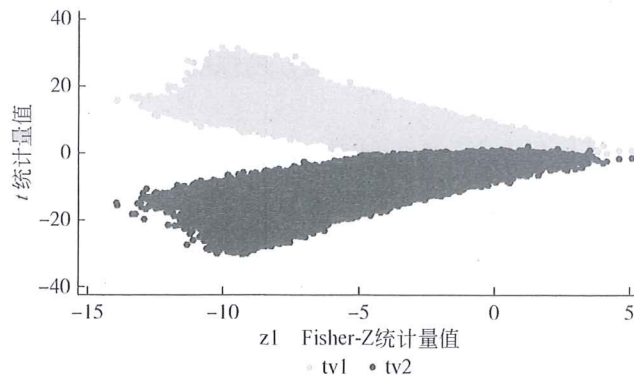


图5 模型 t 统计量与Fisher-Z统计量散点图

Fig. 5 Scatter of the t statistics and the Fisher-Z statistics

图5中, tv1、tv2分别是 β_1, β_2 在表1模拟中的 t 统计量值(图中纵轴), z1(横轴)是表1模拟中的Fisher-Z统计量值,总计由35万组数据绘制而成。图5显示,模拟结果很好地吻合了 t 统计量绝对值与Fisher-Z统计量存在正相关关系。

2) 图3[(c)、(d)/(e)]模型的检验效率分析——实际检验水平的模拟。图3(d)/(e)模型设定如下:

$$\begin{aligned}
 A &= \{u_i\}, u \sim N(\mathbf{0}, \sigma(A)I), i=1, 2, \dots \\
 C &= \beta_{01} + \beta_1 A + \eta_1 v, v \sim N(\mathbf{0}, I) \\
 B &= \beta_{02} + \beta_2 C + \eta_2 w, w \sim N(\mathbf{0}, I)
 \end{aligned}
 \tag{9}$$

图3(c)模型设定如下:

$$\begin{aligned}
 C &= \{u_i\}, u \sim N(\mathbf{0}, \sigma(C)\mathbf{I}), i=1, 2, \dots \\
 A &= \beta_{01} + \beta_1 C + \eta_1 v, v \sim N(\mathbf{0}, \mathbf{I}) \\
 B &= \beta_{02} + \beta_2 C + \eta_2 w, w \sim N(\mathbf{0}, \mathbf{I})
 \end{aligned}
 \tag{10}$$

定理 3 在因果链、因果叉模型中，条件相关系数为 0。

证明：对于因果链 (9) 模型：

$$\begin{aligned}
 r_{AB} &= \frac{\text{cov}(A, B)}{\sqrt{\text{var}(A)\text{var}(B)}} = \frac{\text{var}(A+B) - \text{var}(A) - \text{var}(B)}{2\sqrt{\text{var}(A)\text{var}(B)}} \\
 &= \frac{[(\beta_1\beta_2 + 1)\sigma(A)]^2 + \beta_2^2\eta_1^2 + \eta_2^2 - \sigma(A)^2 - \beta_2^2[(\beta_1\sigma(A))^2 + \eta_1^2] - \eta_2^2}{2\sqrt{\sigma(A)^2(\beta_2^2\{[\beta_1\sigma(A)]^2 + \eta_1^2\} + \eta_2^2)}} \\
 &= \frac{\beta_1\beta_2\sigma^2(A)}{\sqrt{\sigma(A)^2(\beta_2^2\{[\beta_1\sigma(A)]^2 + \eta_1^2\} + \eta_2^2)}} = \frac{1}{\sqrt{\left[1 + \frac{\eta_1^2}{\beta_1^2\sigma^2(A)} + \frac{\eta_2^2}{\beta_1\beta_2\sigma^2(A)}\right]}},
 \end{aligned}$$

而在因果链模型中 r_{AC} 、 r_{BC} 分别是，

$$r_{AC} = \beta_1 \frac{\sigma(A)}{\sqrt{\beta_1^2\sigma^2(A) + \eta_1^2}} = \frac{1}{\sqrt{1 + \frac{\eta_1^2}{\beta_1^2\sigma^2(A)}}} \text{ 和 } r_{BC} = \beta_2 \frac{\sigma(C)}{\sqrt{\beta_2^2\sigma^2(C) + \eta_2^2}} = \frac{1}{\sqrt{1 + \left(\frac{\eta_2}{\beta_2\sigma(C)}\right)^2}}$$

从而得到：

$$\begin{aligned}
 r_{AC}r_{BC} &= \frac{1}{\sqrt{1 + \frac{\eta_1^2}{\beta_1^2\sigma^2(A)}}} \frac{1}{\sqrt{1 + \left(\frac{\eta_2}{\beta_2\sigma(C)}\right)^2}} = \frac{1}{\sqrt{1 + \frac{\eta_1^2}{\beta_1^2\sigma^2(A)} + \left(\frac{\eta_2}{\beta_2\sigma(C)}\right)^2} \frac{\sigma^2(C)}{\beta_1^2\sigma^2(A)}} \\
 &= r_{AB}
 \end{aligned}
 \tag{11}$$

将式 (11) 代入式 (4)：得到 $r_{AB|C} = (r_{AB} - r_{AC}r_{BC}) / (\sqrt{1 - r_{AC}^2}\sqrt{1 - r_{BC}^2}) = 0$ 。

对于因果叉模型 (10)：

$$\begin{aligned}
 r_{AB} &= \frac{\text{cov}(A, B)}{\sqrt{\text{var}(A)\text{var}(B)}} = \frac{\text{var}(A+B) - \text{var}(A) - \text{var}(B)}{2\sqrt{\text{var}(A)\text{var}(B)}} \\
 &= \frac{\beta_1\beta_2\sigma^2(C)}{\sqrt{(\beta_1^2\sigma^2(C) + \eta_1^2)(\beta_2^2\sigma^2(C) + \eta_2^2)}} = \frac{1}{\sqrt{\left(1 + \left(\frac{\eta_1}{\beta_1\sigma(C)}\right)^2\right)\left(1 + \left(\frac{\eta_2}{\beta_2\sigma(C)}\right)^2\right)}}
 \end{aligned}$$

而在因果链模型中 r_{AC} 、 r_{BC} 分别是

$$r_{AC} = \beta_1 \frac{\sigma(C)}{\sigma(A)} = \frac{1}{\sqrt{1 + \left(\frac{\eta_1}{\beta_1\sigma(C)}\right)^2}}, r_{BC} = \beta_2 \frac{\sigma(C)}{\sigma(B)} = \frac{1}{\sqrt{1 + \left(\frac{\eta_2}{\beta_2\sigma(C)}\right)^2}}$$

从而得到：

$$r_{AB} = r_{AC}r_{BC}, \text{ 从而 } r_{AB|C} = 0. \text{ 证毕。} \tag{12}$$

如前文所述，对这两个模型进行模拟检验的是 Fisher-Z 统计量的实际检验水平，表 3 是这两个模型在 $\alpha=0.05$ 时实际检验水平的模拟结果。

表 3 模型 (9) 和模型 (10) 的因果识别检验功效模拟结果 (1)

Table 3 Test power of model (9) and (10) in causality identity simulations (1)

噪声强度 \ 样本容量	样本容量					噪声强度 \ 样本容量	样本容量				
	50	100	150	200	250		50	100	150	200	250
$\eta_1 = \eta_2$	模型 (9)					$\eta_1 = \eta_2$	模型 (10)				
1	0.172	0.048	0.017	0.004	0.002	1	0.180	0.048	0.016	0.005	0.002
1.5	0.170	0.052	0.017	0.006	0.002	1.5	0.175	0.048	0.015	0.006	0.002

续表

噪声强度 \ 样本容量	样本容量					噪声强度 \ 样本容量	样本容量				
	50	100	150	200	250		50	100	150	200	250
2	0.178	0.050	0.016	0.006	0.001	2	0.179	0.051	0.014	0.006	0.002
2.5	0.187	0.050	0.015	0.004	0.002	2.5	0.172	0.053	0.015	0.005	0.002
3	0.174	0.049	0.018	0.005	0.001	3	0.174	0.051	0.017	0.006	0.002
3.5	0.175	0.052	0.017	0.005	0.001	3.5	0.183	0.049	0.018	0.005	0.002
4	0.180	0.052	0.017	0.005	0.001	4	0.172	0.047	0.014	0.004	0.002

注：本表是在设定模型 (9) 和模型 (10) 中 $\beta_{01}=0.7$ 、 $\beta_{02}=1.3$ 、 $\beta_1=1.7$ 、 $\beta_2=1.1$ 、 $\sigma(A)=1$ (模型 9)、 $\sigma(C)=1$ (模型 10) 时，分别设定噪声强度 $\eta_1=\eta_2=1、1.5、2、2.5、3、3.5、4$ 时，在样本容量为 50、100、150、200、250 下用 stata11 软件各模拟 10000 次得到的结果。

对比表 3 中两个模型的结果，Fisher-Z 统计量对两个模型的实际检验水平基本一致。在样本容量大于 150 的情况下，Fisher-Z 统计量对这两个模型具有非常好的实际检验水平；模拟结果还显示，在样本容量为 100 时，实际检验水平与设定的检验水平 $\alpha=0.05$ 基本一致。

表 4 模型 (9) 和模型 (10) 的因果识别检验功效模拟结果 (2)

Table 4 Test power of model (9) and (10) in causality identity simulations (2)

噪声强度 \ 样本容量	样本容量					噪声强度 \ 样本容量	样本容量				
	50	100	150	200	250		50	100	150	200	250
η_1	$\eta_2=1$					η_1	$\eta_2=2$				
1	0.177	0.049	0.015	0.006	0.002	1	0.177	0.051	0.016	0.005	0.002
1.5	0.168	0.047	0.016	0.006	0.002	1.5	0.172	0.050	0.016	0.005	0.001
2	0.177	0.050	0.015	0.006	0.002	2	0.175	0.052	0.016	0.005	0.002
2.5	0.181	0.050	0.016	0.005	0.002	2.5	0.176	0.051	0.015	0.004	0.002
3	0.172	0.052	0.017	0.005	0.002	3	0.176	0.049	0.017	0.005	0.002
3.5	0.182	0.049	0.015	0.005	0.002	3.5	0.179	0.051	0.015	0.004	0.002
4	0.178	0.050	0.015	0.005	0.002	4	0.171	0.052	0.017	0.005	0.002
η_1	$\eta_2=3$					η_1	$\eta_2=4$				
1	0.177	0.049	0.015	0.006	0.002	1	0.167	0.055	0.014	0.005	0.002
1.5	0.168	0.047	0.016	0.006	0.002	1.5	0.171	0.052	0.015	0.004	0.001
2	0.177	0.050	0.015	0.006	0.002	2	0.181	0.048	0.016	0.005	0.002
2.5	0.181	0.050	0.016	0.005	0.002	2.5	0.174	0.050	0.019	0.005	0.002
3	0.172	0.052	0.017	0.005	0.002	3	0.171	0.052	0.015	0.005	0.002
3.5	0.182	0.049	0.015	0.005	0.002	3.5	0.181	0.053	0.013	0.005	0.002
4	0.178	0.050	0.015	0.005	0.002	4	0.165	0.048	0.015	0.006	0.002

注：鉴于模型 (9) 和模型 (10) 有基本一致的实际检验水平，本次只对模型 (9) 的这种组合变动情况进行模拟。

表 4 是在其他设定不变时， η_1 和 η_2 变动组合的模拟结果。对比表 3、表 4 可发现，模型 (9) 和模型 (10) 在既定样本容量下的实际检验水平比较稳定，噪声强度对实际检验水平几乎无影响。从定理 3 的证明中可知，控制噪声的 η_1 和 η_2 在最终的条件相关表达式中并不存在，因而只要保持模型设定形式不变，噪声的大小对上述两个模型的检验并不构成影响。

5 结论与建议

在经济学领域，DAG 方法目前主要是和 SVAR 模型结合使用，用以分析变量当期结构关系矩阵。本文定理 1 表明用 DAG 方法分析 SVAR 模型中变量当期结构关系矩阵，本质上是对 wold 因果链的更进一步

约束。而从图 1 (a) 的矩阵与因果图关系可以看到, DAG 方法可较好地解决 wold 因果链排序问题。DAG 方法要求满足变量间因果关系满足 Markov 因果条件和因果忠实条件, 其理论的核心原理是 D-分离的三种情形下的条件独立性分析, 这是变量之间因果关系分析的基础, 而变量之间反叉因果结构中无盾相遇点则是搜索因果关系并得到最终因果结构图的关键。本文定理 2 说明 DAG 方法由于其基本假设所导致的因果结构图不唯一 (观测等价), 从实用分析角度来说可将这个视为 DAG 方法的一个缺陷。

此外, 本文研究了 DAG 方法对图 3 中变量之间因果结构的识别效率。模拟发现, 在小样本下, 检验条件独立性的 Fisher-Z 统计量对识别图 3 (b) 中的反叉因果结构的效率非常高。只要保持信号强度高于噪声强度, 如表 2 第一行所示, 即便在样本容量为 50 的情况下对该模型识别的检验功效都接近于 1。而在识别图 3 [(c)、(d)/(e)] 的因果链和因果叉结构时, 检验水平的模拟分析结果表明, 在模型设定既定的情况下, 噪声对这类因果结构的识别无影响, 在样本容量为 100 时, 实际检验水平与设定的检验水平 $\alpha = 0.05$ 基本一致。而当样本容量小于 100 时, 实际检验水平降低较大。为此, 特对实践中使用 DAG 方法给以下三点建议:

1) 由于实际中一般假定变量因果满足 Markov 因果条件和因果忠实条件, 而这两个条件要求变量间的因果结构是完整、充分的。因此在分析实际问题时, 建议将与研究变量有关联的变量尽可能考虑进来, 由 DAG 方法识别排除不在因果结构中的变量。

2) 在实际研究中, 研究人员可参考本文的模拟结果对估计结果进行分析, 但建议研究人员用 Bootstrap 方法对估计出来的模型进行自抽样效率分析。

3) 考虑到在检验水平的模拟中, 样本容量不低于 100 才能保证实际检验水平达到 0.05, 因而建议使用样本容量大于 100 的数据进行分析。当样本容量小于 100 时, 可以考虑将检验水平设定为 10%, 以提高实际检验功效。

总之, DAG 方法是当期因果结构分析的有力工具, 也为分析 SVAR 模型提供了新的途径。但在实践中, 只有清晰地了解其原理, 注意其局限性和适用性, 才能正确使用好该方法, 将该方法的最大优势发挥出来。

参考文献:

- [1] Stock J H, Watson M W. Vector autoregressions [J]. *Journal of Economic Perspectives*, 2001, 15 (4): 101-115.
- [2] 李子奈. 计量经济学模型方法论的若干问题 [J]. *经济学动态*, 2007, (10): 22-30.
Li Z N. Some Problems listed in econometrics model and its methods [J]. *Economics Information*, 2007, (10): 22-30. (in Chinese)
- [3] Sims C A. Macroeconomics and reality [J]. *Econometrica*, 1980, 48 (1): 1-48.
- [4] Cooley T F, LeRoy S F, A theoretical macro econometrics: A critique [J]. *Journal of Monetary Economics*, 1985, 16 (3): 283-308.
- [5] Sims C A. Are forecasting models usable for policy analysis [J]. *Quarterly Review*, (Win), 1986, 2-16.
- [6] Granger C W. Investigating causal relations by econometric models and cross-spectral methods [J]. *Econometrica: Journal of the Econometric Society*, 1969, 424-438.
- [7] Granger C W. Testing for causality: A personal viewpoint [J]. *Journal of Economic Dynamics and Control*, 1980, 2: 329-352.
- [8] Hoover K D. Causality in Economics and Econometrics [M]. Basingstoke: Palgrave Macmillan, 2008.
- [9] Spirtes P, Glymour C N, Scheines R. Causation Prediction & Search 2e [M]. Cambridge: MIT press, 2000.
- [10] Pearl J. Causality: Models, reasoning and inference [M]. Cambridge: Cambridge Univ. Press, 2000.
- [11] Swanson N R, Granger C W. Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions [J]. *Journal of the American Statistical Association*, 1997, 92 (437): 357-367.
- [12] Hoover K D. Automatic inference of the contemporaneous causal order of a system of equations [J]. *Econometric Theory*, 2005, 21 (01): 69-77.
- [13] Bryant H L, Bessler D A, Haigh M S. Disproving causal relationships using observational data [J]. *Oxford Bulletin of Economics and Statistics*, 2009, 71 (3): 357-374.
- [14] Demiralp S, Hoover K D, Searching for the causal structure of a vector autoregression [J]. *Oxford Bulletin of Economics and Statistics*, 2003, 65: 745-767.

- [15] Haigh M S, Bessler D A. Causality and price discovery: An application of directed acyclic graphs [J]. the Journal of Business, 2004, 77 (4): 1099-1121.
- [16] Moneta A, Spirtes P. Graphical models for the identification of causal structures in multivariate time series models proc [C]. Joint Conference on Information Sciences, 2006.
- [17] Wang Z. The causal structure of bond yields [J]. The Quarterly Review of Economics and Finance, 2012, 52 (1): 93-102.
- [18] 胡援成, 张朝洋. 美元贬值对中国通货膨胀的影响: 传导途径及其效应 [J]. 经济研究 2012 (4): 101-112.
Hu Y C, Zhang C Y. Impacts of dollar depreciation on China's inflation transmission channels and Effects [J]. Economic Research Journal, 2012, (4): 101-112. (in Chinese)
- [19] 邓卫广. 中国输出通货紧缩了吗? ——DAG 下通货膨胀传递的实证分析 [J]. 经济科学 2007, (06): 43-56.
Deng W G. Has China delivered deflation? —A DAG based empirical analysis of inflation transmission [J]. Economic Science, 2007, (6): 43-56. (in Chinese)
- [20] 杨子晖. 中国输出了“通货紧缩”或“通货膨胀” [J]. 数量经济技术经济研究. 2009 (09): 3-17.
Yang H. Is China “Exporting Deflation” Or “Exporting Inflation” [J]. The Journal of Quantitative & Technical Economics, 2009, (09): 3-17. (in Chinese)
- [21] 聂思玥, 李梦花. 几个条件独立性检验统计量的构造与比较分析 [J]. 数量经济技术经济研究. 2014, (02): 137-147.
Nie S Y, Li M H. Construction and comparative analysis of several conditional independence test statistics [J]. The Journal of Quantitative & Technical Economics, 2014, (02): 137-147. (in Chinese)
- [22] Suppes P. A Probabilistic Theory of Causality [M]. Amsterdam: North-Holland. 1970.
- [23] Bessler D A, Wang Z. D-separation, forecasting, and economic science: A conjecture [J]. Theory and Decision, 2012, 73 (2): 295-314.
- [24] Kwon D, Bessler D A. Graphical methods, inductive causal inference, and econometrics: A literature review [J]. Computational Economics, 2011, 38 (1): 85-106.
- [25] Scheines R, Spirtes P, Glymour C, et al. The TETRAD project: Constraint based aids to causal model specification [J]. Multivariate Behavioral Research, 1998, 33 (1): 65-117.
- [26] 冯蕾, 聂巧平. 结构突变对 KPSS 检验水平与检验功效的影响——基于有限样本情形的模拟及实证研究 [J]. 统计研究, 2009, (9): 96-100.
Feng L, Nie Q P. Structural change on KPSS test for stationarity of finite sample by monte carlo simulations and empirical study [J]. Statistical Research, 2009, (9): 96-100. (in Chinese)

Theory of Causality Structure Identifying Based on DAG Methods and Efficiency Research

Nie Siyue¹, Liu Weiqi², Li Menghua³

1. Institute of management and Decision, Shanxi University, Taiyuan 030006, China;
2. Faculty of Finance & Banking, Shanxi University of Finance and Economics, Taiyuan 030006, China;
3. Faculty of Finance & Banking, Shanxi University of Finance and Economics, Taiyuan 030006, China.

Abstract: Structural analysis is one of the four main tasks in macro-economic analysis, and the causal structure inferring has always been important content of structural analysis. This article systematically illustrated the principles of the directed acyclic graph (DAG) method based on causal structure analysis, and had a detailed analysis of the PC cause search algorithm. This article also studied the efficiency of the DAG method, that is, under linear model with Normal Errors assumptions in the three situations of D-separation, we simulated the real testing size and testing power of the conditional independence hypothesis test using Fisher-z Statistics with a Monte Carlo method. At last, some suggestions are given on the applications of the DAG method.

Key words: Directed Acyclic Graph (DAG); Conditional Independence; Testing Size; Testing Power