Causal Effect Identification in Alternative Acyclic Directed Mixed Graphs - Supplementary Material

Jose M. Peña JOSE.M.PENA@LIU.SE
Linköping University
Linköping (Sweden)

Proof of Lemma 1 To prove the first statement, note that

$$f(v) = \int [\prod_i f(v_i | pa_G(V_i), u_i)] f(u) \, du = \int [\prod_i f(v_i | pa_G(V_i), u_i)] \prod_j f(u_{S_j}) \, du$$
$$= \prod_j \int \prod_{V_i \in S_j} f(v_i | pa_G(V_i), u_i) f(u_{S_j}) \, du_{S_j} = \prod_j q(s_j)$$

where the second equality follows from Equation 6.

We prove the second statement by induction over the number of variables in V. Clearly, the result holds when V contains a single variable. Assume as induction hypothesis that the result holds for up to n variables. When there are $n+1$ variables, these can be divided into components $S_1, \ldots, S_k, S'$ with factors $q(s_1), \ldots, q(s_k), q(s')$ such that $V_{n+1} \in S'$. As shown above,

$$f(v) = q(s') \prod_j q(s_j)$$

which implies that

$$f(v^{(n)}) = \int f(v) \, dv_{n+1} = [\int q(s') \, dv_{n+1}] \prod_j q(s_j).$$

Note that $f(v^{(n)})$ factorizes according to $G^{V^{(n)}}$ and S_j is a component of $G^{V^{(n)}}$. Therefore,

$$q(s_j) = \prod_{V_i \in S_j} f(v_i | v^{(i-1)})$$

by the induction hypothesis and the fact that $V_1 < \ldots < V_n$ is also a topological order of the nodes in $G^{V^{(n)}}$. Then, $q(s')$ is also identifiable and is given by

$$q(s') = \frac{f(v)}{\prod_j q(s_j)} = \frac{\prod_i f(v_i | v^{(i-1)})}{\prod_j q(s_j)} = \prod_{V_i \in S'} f(v_i | v^{(i-1)}).$$

■

Proof of Lemma 2

$$\begin{aligned}\int q(c)\, d(c \setminus e) &= \int \int [\prod_{V_i \in E} f(v_i | pa_G(V_i), u_i) \prod_{V_i \in C \setminus E} f(v_i | pa_G(V_i), u_i)] f(u)\, du\, d(c \setminus e)\\ &= \int [\prod_{V_i \in E} f(v_i | pa_G(V_i), u_i) \int \prod_{V_i \in C \setminus E} f(v_i | pa_G(V_i), u_i)\, d(c \setminus e)] f(u)\, du\\ &= \int [\prod_{V_i \in E} f(v_i | pa_G(V_i), u_i)] f(u)\, du = q(e)\end{aligned}$$

where the second equality follows from the fact that E is an ancestral set in G^C and, thus, no node in E has a parent in $C \setminus E$. The third equality is due to the fact that the integral over $c \setminus e$ equals 1. This may be easier to appreciate by performing the integral following a topological order of the nodes in $C \setminus E$ with respect to G. ■

Proof of Lemma 3 As mentioned above, $q(c)$ factorizes according to G^C. Therefore, the first statement can be proven in much the same way as the first statement in Lemma 1. The third statement follows from Lemma 2 since $C^{(i)}$ is an ancestral set in G^C.

We prove the second statement by induction over the number of variables in C. Clearly, the result holds when C contains a single variable. Assume as induction hypothesis that the result holds for up to n variables. When there are $n+1$ variables, these can be divided into components $C_1, \ldots, C_k, C'$ with factors $q(c_1), \ldots, q(c_k), q(c')$ such that $V_{n+1} \in C'$. As shown above,

$$q(c) = q(c') \prod_j q(c_j)$$

which implies that

$$q(c^{(n)}) = \int q(c)\, dv_{n+1} = [\int q(c')\, dv_{n+1}] \prod_j q(c_j)$$

where the first equality follows from Lemma 2 because $C^{(n)}$ is an ancestral set in G^C. Note that $q(c^{(n)})$ factorizes according to $G^{C^{(n)}}$ and C_j is a component of $G^{C^{(n)}}$. Therefore,

$$q(c_j) = \prod_{V_i \in C_j} \frac{q(c^{(i)})}{q(c^{(i-1)})}$$

by the induction hypothesis and the fact that $V_1 < \ldots < V_n$ is also a topological order of the nodes in $G^{C^{(n)}}$. Then, $q(c')$ is given by

$$q(c') = \frac{q(c)}{\prod_j q(c_j)} = \frac{q(c^{(n+1)})}{\prod_j q(c_j)} = \frac{\prod_{i=1}^{n+1} \frac{q(c^{(i)})}{q(c^{(i-1)})}}{\prod_j q(c_j)} = \prod_{V_i \in C'} \frac{q(c^{(i)})}{q(c^{(i-1)})}.$$

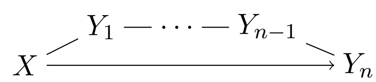
■

Proof of Lemma 4 It suffices to note that

$$q(c|a) = f(c|v \setminus \widehat{\{a,c\}}, a) = \frac{f(a,c|v \setminus \widehat{\{a,c\}})}{f(a|v \setminus \widehat{\{a,c\}})} = \frac{q(a,c)}{\int q(a,c)\,dc}.$$

Moreover, if A is an ancestral set in $G^{A \cup C}$, then $\int q(a,c)\,dc = q(a)$ by Lemma 2. ■

Proof of Theorem 9 For the algorithm to fail, some component C_j cannot be ancestral in line 8. Then, one of the following two cases must occur. Case 1: Assume that C_j is not ancestral in line 8 because it contains a child Y_n of X. Clearly, X is not in C_j by lines 4-5. However, both X and Y_n must be in the component S_i in line 8 for the algorithm to fail, which implies that there is an undirected path between X and Y_n. Case 2: Assume that C_j is not ancestral in line 8 because it contains a child Y_j of Y_i and Y_i is not in C_j. However, both Y_i and Y_j must be in the component S_i in line 8 for the algorithm to fail, then both must be in C_j by lines 4-5. This is a contradiction. Therefore, only the first case can occur, which implies that the algorithm fails only if G has a subgraph of the form



Such a subgraph implies that $f(v \setminus x|\widehat{x})$ is not identifiable from G (Peña and Bendtsen, 2017, Theorem 12). ■

Proof of Lemma 10 Removing edges from an aADMG can only increase the separations represented by the aADMG. Then, if the antecedent of rule 1 is satisfied, so are the antecedents of rules 2 and 3. Then, we can replace the application of rule 1 with the application of rule 2 followed by the application of rule 3, i.e.

$$f(y|\widehat{x}, z, w) = f(y|\widehat{x}, \widehat{z}, w) = f(y|\widehat{x}, w).$$

■

Proof of Lemma 11 We prove the result for Lemma 2. The proof for Lemma 6 is similar. First, note that

$$\int q(c)\,d(c \setminus e) = \int f(c|\widehat{v \setminus c})\,d(c \setminus e) = f(e|\widehat{v \setminus c}).$$

Moreover,

$$q(e) = f(e|\widehat{v \setminus e}) = f(e|\widehat{v \setminus c})$$

where the second equality follows from rule 3 since $E \perp_{G_{\underrightarrow{\overrightarrow{V \setminus C}}\,\overline{\overrightarrow{C \setminus E}}}} C \setminus E|\emptyset$. To see that this separation holds, assume that there is a route ρ in $G_{\underrightarrow{\overrightarrow{V \setminus C}}\,\overline{\overrightarrow{C \setminus E}}}$ between a node in E and a node in $C \setminus E$. Note that ρ cannot only contain nodes in C, because the nodes in $C \setminus E$ only have outgoing directed edges in $G_{\underrightarrow{\overrightarrow{V \setminus C}}\,\overline{\overrightarrow{C \setminus E}}}$, which implies that E is not ancestral set in G^C, which contradicts the assumptions in Lemma 2. So, ρ must contain some node in $V \setminus C$. Note however that some node in $V \setminus C$ must be a collider in ρ because, in $G_{\underrightarrow{\overrightarrow{V \setminus C}}\overline{\overrightarrow{C \setminus E}}}$, the nodes in $V \setminus C$ only have undirected edges whereas the nodes in $C \setminus E$ only have outgoing directed edges. Therefore, ρ is not connecting given $\emptyset$. ■

Proof of Lemma 12 We prove the result for Lemma 1. The proofs for Lemmas 3, 5 and 7 are similar. Moreover, we only prove the first statement in Lemma 1, because the proof of the second statement provided in Lemma 1 only involves standard probability manipulations. Likewise, we do not need to prove the third statement of Lemmas 3 and 7 because, as shown in the proof of those lemmas, it follows from Lemma 2, which follows from rule 3 as shown in Lemma 11.

Let V be partitioned into components $S_1, \ldots, S_k$ for the aADMG G. Moreover, assume without loss of generality that if the edge $A \to B$ is in G, then $A \in S_i$ and $B \in S_j$ with $i \leq j$. Let $S_{<j} = \bigcup_{i<j} S_i$ and $S_{\leq j} = \bigcup_{i \leq j} S_i$. Note that

$$f(v) = \prod_j f(s_j | s_{<j}).$$

Moreover,

$$f(s_j | s_{<j}) = f(s_j | \widehat{v \setminus s_{\leq j}}, s_{<j})$$

by rule 3 since $S_j \perp_{G_{\overline{\overrightarrow{V \setminus S_{\leq j}}}}} V \setminus S_{\leq j} | S_{<j}$. To see that this separation holds, assume that there is a route ρ in $G_{\overline{\overrightarrow{V \setminus S_{\leq j}}}}$ between a node in S_j and a node in $V \setminus S_{\leq j}$. Note that the nodes in $V \setminus S_{\leq j}$ only have outgoing directed edges in $G_{\overline{\overrightarrow{V \setminus S_{\leq j}}}}$. Therefore, ρ implies that some node in $V \setminus S_{\leq j}$ is an ancestor in G of some node in $S_{\leq j}$, which contradicts our assumption above.

Finally, note that

$$f(s_j | \widehat{v \setminus s_{\leq j}}, s_{<j}) = f(s_j | \widehat{v \setminus s_j}) = q(s_j)$$

where the first equality follows from rule 2 because $S_j \perp_{G_{\underrightarrow{\overrightarrow{V \setminus S_{\leq j}}}\underrightarrow{S_{<j}}}} S_{<j} | \emptyset$. To see that this separation holds, assume that there is a route ρ in $G_{\underrightarrow{\overrightarrow{V \setminus S_{\leq j}}}\underrightarrow{S_{<j}}}$ between a node in S_j and a node in $S_{<j}$. Then, there exist two nodes $A \in S_j$ and $B \in S_{<j}$ that are adjacent in ρ or there exist two nodes $A' \in S_j$ and $B' \in V \setminus S_{\leq j}$ that are adjacent in ρ. However, either case implies a contradiction:

- $A - B$ contradicts that S_j is a component.
- $A \to B$ contradicts our assumption above.
- $A \leftarrow B$ contradicts that B has no outgoing directed edge in $G_{\underrightarrow{\overrightarrow{V \setminus S_{\leq j}}}\underrightarrow{S_{<j}}}$.
- $A' - B'$ contradicts that S_j is a component
- $A' \to B'$ and $A' \leftarrow B'$ contradict that B' only has undirected edges in $G_{\underrightarrow{\overrightarrow{V \setminus S_{\leq j}}}\underrightarrow{S_{<j}}}$.

■