A Correctness

Theorem A.1. *Under assumptions 1 to 3, if A is full rank, then $\alpha(c,e)$ is exactly the impact of c on e.*

Proof. Let $I(c,e)$ be the impact of c on e, then when c is not a genuine cause of e, $I(c,e)$ is zero. In the following four steps, we prove for each $c \in X$, where X is the set of potential causes of e:

$$\alpha(c,e) = I(c,e). \tag{A.1}$$

Step 1. Based on assumptions 1 to 3:

$$\begin{aligned}&|T(e|c)| \times E[e|c] = |T(e|c)| \times E[e| \bigwedge_{x \in X} \neg x]+\\&N(e|c) \times I(c,e) + \sum_{x \in X \setminus c} N(e|c,x) \times I(x,e).\end{aligned} \tag{A.2}$$

That is, the total value of e caused by c is equivalent to the total value of e when no cause is present, plus the total impact due to c and other causes that may co-occur with c. Rearranging this, we have:

$$\begin{aligned}&E[e|c] = E[e| \bigwedge_{x \in X} \neg x] + \frac{N(e|c)}{|T(e|c)|} \times I(c,e)+\\&\sum_{x \in X \setminus c} \frac{N(e|c,x)}{|T(e|c)|} \times I(x,e).\end{aligned} \tag{A.3}$$

Step 2. Based on assumptions 1 to 3:

$$|T(e)| \times E[e] = |T(e)| \times E[e| \bigwedge_{x \in X} \neg x] + \sum_{x \in X} N(e|x) \times I(x,e). \tag{A.4}$$

That is, the total value of e, equals the total value of e when no cause is present, plus the total impact of each cause. The above equation can be written as:

$$E[e] = E[e| \bigwedge_{x \in X} \neg x] + \sum_{x \in X} \frac{N(e|x)}{|T(e)|} \times I(x,e). \tag{A.5}$$

Step 3. By replacing $E[e|c]$ and $E[e]$ with the right-hand side of eqs. (A.3) and (A.5), eq. (9) can be written as:

$$(\alpha(c,e) - I(c,e)) + \sum_{x \in X \setminus c} f(e|c,x) \times (\alpha(x,e) - I(x,e)) = 0. \tag{A.6}$$

Step 4. Now, eq. (A.6) for each $c \in X$, where $X = \{c_1, \ldots, c_n\}$, yields the following system of linear equations:

$$A \times \begin{bmatrix} \alpha(c_1,e) - I(c_1,e) \\ \vdots \\ \alpha(c_n,e) - I(c_n,e) \end{bmatrix} = 0, \tag{A.7}$$

where A is a coefficient matrix of X. Since A is full rank, then the above system has a unique solution, which can be written as eq. (A.1). □

Corollary A.1. Under assumptions 1 to 3, if A is full rank, then eqs. (6) and (8) yield these expectations exactly.

Proof. We prove the corollary in three steps.

Step 1. Equation (A.2) can be written as:

$$\begin{aligned}&|T(e|c)| \times E[e|c] = |T(e|c)| \times E[e|c \bigwedge_{x \in X \setminus c} \neg x]+\\&\sum_{x \in X \setminus c} N(e|c,x) \times I(x,e).\end{aligned} \tag{A.8}$$

That is:

$$E[e|c \bigwedge_{x \in X \setminus c} \neg x] = E[e|c] - \sum_{x \in X \setminus c} \frac{N(e|c,x)}{|T(e|c)|} \times I(x,e). \tag{A.9}$$

Step 2. Equation (A.4) can be written as:

$$E[e| \bigwedge_{x \in X} \neg x] = E[e] - \sum_{x \in X} \frac{N(e|x)}{|T(e)|} \times I(x,e). \tag{A.10}$$

Step 3. Equations. (A.9) and (A.10) are exactly $E[e|c \bigwedge_{x \in X \setminus c} \neg x]$ and $E[e| \bigwedge_{x \in X} \neg x]$. Since $\alpha(c,e) = I(c,e)$ (based on theorem A.1), the two equations are the same as eqs. (6) and (8). □

Claim A.1. The coefficient matrix of X_{lis} obtained by algorithm 2 is full rank.

Proof. Without loss of generality, let $c_1, \ldots, c_m$ be the order that causes are added to X_{lis}. Then in each iteration where $c_{max} = c_i$ (with $1 \le i \le m$), the coefficient matrix of $\{c_1, \ldots, c_i\}$ is full rank (otherwise c_i cannot be added to X_{lis}). Thus, A_{lis} is full rank. □

B Time complexity

Claim B.1. With N variables and T timepoints, where $T > N^2$ and each variable is measured (or a value is imputed) for each timepoint, the complexity of calculating $\alpha(c,e)$ is $O(N^2T)$.

Proof. Conceptually, for each e, there are two steps: creating a system of linear equations and solving the system (as shown in algorithm 1). The overall complexity of calculating matrix A and B for all the systems is $O(N^2T)$ (since $f(e|c)$ and $f(e|c,x)$ can be calculated once), and the overall complexity of solving all the systems is $O(N^4)$ (using direct methods such as Gaussian Elimination with complexity of $O(N^3)$). Thus, the overall complexity of calculating $\alpha(c,e)$ for all e is $O(N^2T + N^4)$. As we assume $T > N^2$, the complexity is then $O(N^2T)$. □

Claim B.2. The complexity of algorithm 2 is $O(N^3)$.

Proof. Note that algorithm 2 is used after building a system of linear equations and finding matrix A not full rank. Thus the complexity of steps 3 and 5 are both $O(N)$, since $E[e|c]$, $E[e]$ and A have already been calculated. The complexity of step 6 is $O(N^2)$, since the row echelon form of the coefficient matirx of X_{lis} has already been calculated. Thus, the complexity of the algorithm is $O(N^3)$, as the previous three steps may be repeated at most N times. □

Corollary B.1. The complexity of calculating $\alpha(c,e)$ including algorithm 2 is $O(N^2T)$.

Proof. Based on claims B.1 and B.2, the complexity of calculating $\alpha(c, e)$ including algorithm 2 is $O(N^2T + N^4)$, since the algorithm may be called at most N times. As we assume $T > N^2$, the complexity is still $O(N^2T)$. $\square$