Rapport de synthèse pour l'algorithme FCI

Jiahua LI
Zeyu TAO

January 2026

1 Introduction

L'inférence causale constitue un outil puissant pour étudier et identifier les relations causales entre les variables aléatoires à partir de données. Il existe différents types d'algorithmes permettant d'automatiser le processus d'identification. Parmi ceux-ci, nous présentons les algorithmes basés sur des contraintes. Cette famille d'algorithmes produit un graphe (e.g., PDAG (Partial Directed Acyclic Graph) ou PAG (Partial Ancestral Graph)) représentant une classe d'équivalence de Markov, où les arêtes capturent des informations causales et le concept de séparation (e.g., d-séparation ou m-séparation) définit les relations d'indépendance conditionnelle entre les variables aléatoires représentées par les sommets du graphe.

2 Algorithme PC (Peter-Clark)

L'algorithme PC, proposé par Spirtes et Glymour en 1993 [1], est l'un des algorithmes basés sur des contraintes permettant d'identifier les relations causales entre les variables aléatoires à partir de données, sous l'hypothèse que les relations causales (orientation des arêtes) peuvent être parfaitement représentées par un DAG (Directed Acyclic Graph), c'est-à-dire, il n'existe pas de variables latentes et de biais de sélection dans les données. Il se compose en deux étapes :

— découvrir le squelette du graphe causal.
— orienter les arêtes.

Deux DAGs appartiennent à la même classe d'équivalence de Markov s'ils représentent les mêmes relations d'indépendance. Par exemple, les DAGs : $X \to Z \to Y$, $X \leftarrow Z \leftarrow Y$ et $X \leftarrow Z \to Y$ représentent la même relation d'indépendance $X \perp\!\!\!\perp Y \mid Z$; le DAGs : $X \to Z \leftarrow Y$ (une v-structure ou un unshielded collider) représente la relation relation d'indépendance $X \perp\!\!\!\perp Y$. En effet, cette proposition est équivalente à : Deux DAGs appartienne à la même classe d'équivalence de Markov s'ils ont les mêmes v-structures.

Fondé sur cette propriété, l'algorithme PC cherche d'abord à identifier et à orienter toutes les v-structures présentées dans le squelette obtenu par la première étape de l'algorithme. Ensuite, il applique des règles de propagation autant que possible pour orienter les autres arêtes de manière à éviter la création de nouvelles v-structures ou de circuits dans le graphe.

À la fin de son exécution, l'algorithme PC retourne un PDAG représentant une classe d'équivalence de Markov. Sous l'hypothèse qu'il n'existe aucune variables latentes, ni aucun biais de sélection, l'algorithme garantit que le vrai graphe causal appartenant à cette classe d'équivalence.

Cependant, en pratique, des variables latentes et des biais de sélection apparaissent naturellement dans les données, ce qui conduit l'algorithme PC, soit à déduire une fausse relation causale, soit à obtenir des relations causales qui ne peuvent être représentées par un DAG, par exemple la présence d'une arête bidirectionnelle $X \leftrightarrow Y$. C'est face à cette limitation que l'algorithme FCI a été proposé. Il s'agit d'une modification de l'algorithme PC permettant de traiter un nombre arbitraire de variables latentes et de biais de sélection.

3 Algorithme FCI (Fast Causal Inference)

La première version de l'algorithme FCI a été initialement proposée par Spirtes et Glymour en 1999 [1], où sa formulation repose sur le concept de Inducing Path Graph. Ensuite, en 2005 [1], Ali, Richardson, Spirtes et Zhang ont réinterprété l'algorithme à l'aide de Maximal Ancestral Graph (MAG), dont la classe d'équivalence de Markov est représentée par un PAG. Enfin, en 2008 [3], Zhang a complété l'algorithme en établissant un ensemble de dix règles de propagation.

Un PAG est un graphe G, où chaque extrémité des arêtes sont représentées par trois types de marques : $\circ$ (CIRCLE), $-$ (TAIL) et $>$ (ARROWHEAD). Cette représentation d'arêtes avec les marques permet de capturer les relations causales d'une classe d'équivalence de Markov avec un nombre arbitraire de variables latentes ou de biais de sélection (ou variable de sélection), plus précisément :

- $X \rightarrow Y$ implique que X est la cause de Y ou d'une variable de sélection, mais Y n'est pas la cause de X ni d'aucune variable de sélection.
- $X \leftrightarrow Y$ implique que X n'est pas la cause de Y ni d'aucune variable de sélection, et Y n'est pas la cause de X ni d'aucune variable de sélection.
- $X - Y$ implique que X est la cause de Y ou de variable de sélection, et Y est la cause de X ou de variable de sélection.

Les cercles dans un PAG représentent les incertitudes de l'algorithme, c'est-à-dire que les informations d'indépendance conditionnelle contenues dans les données ne permettent pas de conclure des relations causales. Par exemple, $X -\!\circ Y$ implique que X est la cause de Y ou de variable de sélection, mais est ce que Y est la cause de X ou de variable de sélection reste incertain.

Dans un MAG, un chemin discriminant est un chemin $p = (X, \ldots, W, V, Y)$ entre X et Y qui passe par V, tels que :

- p est un chemin d'au moins de taille quatre.
- V et Y sont adjacents, mais V et X ne sont pas adjacents.
- Pour tout sommet W situé entre X et V dans le chemin p, W a une v-structure $(\cdots \rightarrow W \leftarrow \ldots)$ et est un parent de $Y (W \rightarrow Y)$.

Le concept de chemin discriminant est crucial pour l'algorithme FCI, car il permet de caractériser la classe d'équivalence de Markov entre les MAGs. Deux MAGs appartiennent à la même classe d'équivalence s'ils ont les mêmes v-structures, et si un chemin $p = (X, \ldots, W, V, Y)$ est un chemin discriminant dans les deux MAGs, alors V a une v-structure dans un graphe si et seulement si V a une v-structure dans l'autre graphe.

Fondé sur cette propriété, l'algorithme FCI repose globalement sur le même principe que l'algorithme PC :

- Il détermine un squelette initial en appliquant la même procédure que celle de l'algorithme PC. Il identifie ensuite les v-structures, puis applique une autre procédure supplémentaire basée sur le concept de Possible-D-SEP pour obtenir le squelette final. Cette procédure supplémentaire

vise à identifier les relations d'indépendance conditionnelle entre les variables cachées dues à la présence de variables latentes ou de biais de sélection.
— Il identifie à nouveau les v-structures. Ensuite, il applique les dix règles de propagation proposées par Zhang pour obtenir un PAG final représentant une classe d'équivalence de Markov.

En effet, la version de l'algorithme FCI de Zhang est considérée comme la version standard de l'algorithme FCI. À partir de cette version standard, il existe différentes variantes de l'algorithme FCI dans la littérature. En particulier, nous trouvons RFCI (Really Fast Causal Inference) proposé par Colombo, Maathuis, Kalisch et Richardson en 2012 [4], c'est une modification de l'algorithme FCI. Le principe consiste à éviter les tests d'indépendance conditionnelle exhaustifs dans la procédure supplémentaire de l'algorithme FCI, qui prend beaucoup de temps car la taille des Possible-D-SEP peut être assez importante pour des graphes de grandes dimensions. En effet, le graphe retourné par l'algorithme RFCI est moins informatif que celui du FCI, mais les auteurs montrent qu'asymptotiquement les relations causales sont correctes.

4 Conclusion

Ce rapport de synthèse permet de conclure l'étude théorique et la revue bibliographique de l'algorithme FCI, en détaillant sa motivation, ses différentes versions et variantes, ainsi que la procédure de l'algorithme.

Nous avons implémenté la version standard de l'algorithme FCI. Notre code source est disponible dans le répertoire : `https://github.com/JiahuaLiY/project-MADI-2025`.

Références

[1] Chapitre 6 de Causation, Prediction, and Search (2nd ed.). Spirtes, P., Glymour, C., & Scheines, R. (1999). MIT Press.

[2] R.A. Ali, T. Richardson, P. Spirtes, J. Zhang, Towards characterizing Markov equivalence classes for directed acyclic graphs with latent variables, in : Proceedings of the 21th Conference on Uncertainty in Artificial Intelligence, AUAI Press, 2005, pp. 10–17. Full version available at `http://www.stat.washington.edu/www/research/reports/2005/tr476.pdf`.

[3] On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias. Artificial Intelligence, Zhang, J. (2008). 172(16–17), 1873–1896.

[4] Learning high-dimensional directed acyclic graphs with latent and selection variables. Annals of Statistics, Colombo, D., Maathuis, M. H., Kalisch, M., & Richardson, T. S. (2012). 40(1), 294–321.