PAG2ADMG Important Theorems & Proofs

Nishant Subramani
Northwestern University, Evanston IL 60201, USA
nishantsubramani2017@u.northwestern.edu

1 PAG2ADMG Theorems

Theorem 1. *Every ADMG $\mathscr{A}$ in the Markov equivalence class M has the same skeleton of the PAG $\mathscr{P}$ that describes M.*

Proof. It is sufficient to show that for any pair of vertices in a PAG that lacks any type of edge between them, there exists no edge between the same pair of vertices in each and every ADMG that describes the same independencies and dependencies encoded by this PAG.

By definition, two vertices X and Y in a PAG lack an edge if and only if there is some possibly empty set Z that does not contain X or Y that m-separates these two vertices.

First a few quantities are defined:

1. Let **V** represent the set of all vertices in the PAG $\mathscr{P}$.
2. Let the same set **V** also represent the set of all vertices in every ADMG that describes the same independencies and dependencies encoded by P over V.
3. Let **V'** = **V** $\setminus$ X, Y.
4. Let **S** represent the set of ADMGs that describe the same independencies and dependencies encoded by $\mathscr{P}$.

The proof the above theorem is completed below by contradiction.

Assume that there exist two vertices X and Y in $\mathscr{P}$ that lack an edge that are contained in a set **V**. Assume there exists a directed edge between X and Y ($X \rightarrow Y \oplus Y \rightarrow X$) in an ADMG $\mathscr{A}$ which is a member of **S**. Now there exists no set of vertices in **V'** that m-separates X and Y in $\mathscr{A}$, and thus introduces new independence and dependence relationships. So $\mathscr{A}$ cannot be a member of **S**. Similarly if there exists a bi-directed edge between X and Y in $\mathscr{A}$ which is a member of **S**, there exists no set of vertices in **V'** that m-separates X and Y

in $\mathscr{A}$, and thus introduces new independence and dependence relationships. So $\mathscr{A}$ cannot be a member of **S**.

Thus the only remaining edge type is the lack of an edge. So for every pair of vertices X and Y in $\mathscr{P}$ contained in the set **V** that lacks an edge between them, each and every ADMG $\mathscr{A}$ in the set **S** cannot have an edge between X and Y.

This proves that the skeleton of every ADMG $\mathscr{A}$ which is a member of the set **S** will at most contain all of the edges represented by the skeleton of $\mathscr{P}$, and thus the skeleton of every ADMG $\mathscr{A}$ will be a subset of the skeleton of $\mathscr{P}$.

□

Theorem 2. *If there exists a directed edge between a pair of nodes X and Y in a PAG $\mathscr{P}$, then every ADMG $\mathscr{A}$ in M must contain the same directed edge in the same orientation.*

Proof. First a few quantities are defined:

1. Let **S** represent the set of ADMGs that describe the same independencies and dependencies encoded by $\mathscr{P}$
2. Let $\mathscr{A}$ represent an ADMG that is contained in the set **S**.

Some formalities and conversions based on the definitions of these graphs must be outlined:

1. By definition any ADMG $\mathscr{A}$' can be converted into a DAG $\mathscr{D}$ by replacing all bi-directed edges with a latent variable confounder.
2. Then $\mathscr{D}$ can be converted into a PAG $\mathscr{P}$'.
3. $\mathscr{P}$' must equal $\mathscr{P}$ if $\mathscr{A}$' is in the set **S** because the PAG is a unique representation of a Markov Equivalence Class of MAGs.
4. Assume there exists a directed edge from X to Y in $\mathscr{P}$ Assume that an ADMG $\mathscr{A}$ in **S** is converted into a DAG $\mathscr{D}$ and then into a PAG $\mathscr{P}$'. By definition the only way a directed edge can exist between X and Y in $\mathscr{P}$' is if every MAG $\mathscr{M}$' in the Markov Equivalence Class of MAGs encoded by $\mathscr{P}$' contains this directed edge from X to Y.
5. The only way every MAG $\mathscr{M}$' can contain this edge is if X in an ancestor of Y in $\mathscr{D}$ and there exists an inducing path from X to Y relative to the set of hidden variables in $\mathscr{D}$.

The proof the above theorem is completed below by contradiction.

Assume $\mathscr{P}$ contains a directed edge from X to Y. Assume $\mathscr{A}$ is in **S** and contains a directed edge from Y to X. $\mathscr{D}$ then also contains this directed edge Y to X. D is converted to a PAG $\mathscr{P}$'. $\mathscr{M}$' cannot contain a directed edge from

X to Y because X is no longer an ancestor of Y in $\mathscr{D}$. Since at least one MAG $\mathscr{M}$' lacks a directed edge from X to Y, $\mathscr{P}$' lacks this edge, and thus $\mathscr{P}$' and $\mathscr{P}$ encode a different set of Markov Equivalent MAGs. Thus, if there exists a directed edge in $\mathscr{P}$ from X to Y, there cannot exist a directed edge from Y to X in $\mathscr{A}$ if it is in the set $\mathbf{S}$.

Assume $\mathscr{P}$ contains a directed edge from X to Y. Assume $\mathscr{A}$ is in $\mathbf{S}$ and contains a bi-directed edge between X and Y. $\mathscr{D}$ now contains a latent variable confounder H and has a directed edge from H to X, and from H to Y. $\mathscr{D}$ is converted into a PAG $\mathscr{P}$'. $\mathscr{M}$ cannot contain a directed edge from X to Y because there cannot exist a directed path in $\mathscr{D}$ from X to Y such that all variables along that directed path (except for the endpoints) are colliders or all variables along that directed path are in the latent variable set. By definition, a directed path between X and Y with at least one other variable cannot be directed if that other variable is a collider. If that inducing path definition is to hold, a directed edge must be present between X and Y in $\mathscr{D}$ which is not the case, or the directed path between X and Y must contain latent variables $\{H_1, \ldots, H_n\}$ where n is at least 1 such that every vertex H_i has a directed edge from H_i to $H_i + 1$ and the directed path must have directed edges from X to H_1 and from H_n to Y. This path cannot exist because all variables in the latent set are common confounders. Thus $\mathscr{P}$' does not equal $\mathscr{P}$ and thus encodes a different set of Markov Equivalent MAGs. Thus if there exists a directed edge in $\mathscr{P}$ from X to Y, there cannot exist a bi-directed edge between X and Y in $\mathscr{A}$ if it is in the set $\mathbf{S}$.

Assume $\mathscr{P}$ contains a directed edge from X to Y. Assume $\mathscr{A}$ is in $\mathbf{S}$ and contains no edge between X and Y. $\mathscr{D}$ now contains no edge from X to Y. $\mathscr{D}$ is converted into a PAG $\mathscr{P}$'. $\mathscr{M}$ cannot contain a directed edge from X to Y because there cannot exist a directed path in $\mathscr{D}$ from X to Y such that all variables along the directed path except for the endpoints are colliders or all variables along that directed path are in the latent variable set. By definition, a directed path between X and Y with at least one other variable cannot be directed if that other variable is a collider. If that inducing path definition is to hold, a directed edge must be present between X and Y in $\mathscr{D}$ which is not the case, or the directed path between X and Y must contain latent variables $\{H_1, \ldots, H_n\}$ where n is at least 1 such that every vertex H_i has a directed edge from H_i to H_i+1 and the directed path must have directed edges from X to H_1 and from H_n to Y. This path cannot exist because the only variables that are in the latent set are common confounders where a bi-directed edge is present in $\mathscr{A}$ and are not connected to any other observed vertices. Thus if there exists a directed edge in $\mathscr{P}$ from X to Y, there cannot exist no edge between X and Y in $\mathscr{A}$ if it is in the set $\mathbf{S}$.

Thus the only possibility is that a directed edge exists from X to Y in $\mathscr{A}$ if it is to be in the set $\mathbf{S}$. □

Theorem 3. *If there exists a bi-directed edge between a pair of nodes X and Y in a PAG $\mathscr{P}$, then every ADMG $\mathscr{A}$ in M cannot contain a directed edge between X and Y in any orientation.*

Proof. The following proof of this theorem is completed by contradiction. The same definitions and formalities stated in the above proofs hold.

Assume $\mathscr{P}$ contains a bi-directed edge between X and Y. Assume $\mathscr{A}$ is in **S** and contains a directed edge from X to Y. $\mathscr{D}$ then also contains this directed edge X to Y. $\mathscr{D}$ is converted to a PAG $\mathscr{P}$'. $\mathscr{M}$' cannot contain a bi-directed edge between X and Y because X is an ancestor of Y in $\mathscr{D}$. Since at least one MAG $\mathscr{M}$' lacks a bi-directed edge between X and Y, $\mathscr{P}$' lacks this edge, and thus $\mathscr{P}$' and $\mathscr{P}$ encode a different set of Markov Equivalent MAGs. Thus, if there exists a bi-directed edge in $\mathscr{P}$ between X and Y, there cannot exist a directed edge from X to Y in $\mathscr{A}$ if it is in the set **S**. By symmetry there cannot exist a directed edge from Y to X in $\mathscr{A}$ if it is in the set **S**. □

Theorem 4. *If there exists a circle-arrow edge between a pair of nodes X and Y such that the circle mark is at X and the arrow mark is at Y, then every ADMG $\mathscr{A}$ in M cannot contain a directed edge from Y to X.*

Proof. The following proof of this theorem is completed by contradiction. The same definitions and formalities stated in the above proofs hold.

Assume $\mathscr{P}$ contains an edge between X and Y such that there is a circle mark at X and an arrowhead mark at Y. Assume $\mathscr{A}$ is in **S** and contains a directed edge from Y to X. $\mathscr{D}$ then also contains this directed edge Y to X. $\mathscr{D}$ is converted to a PAG $\mathscr{P}$'. $\mathscr{M}$' contains this directed edge from Y to X because there is an inducing path from Y to X and Y is an ancestor of X in $\mathscr{D}$. Since at least one member in the Markov Equivalence Class encoded by $\mathscr{P}$' ($\mathscr{M}$') contains a directed edge from Y to X, $\mathscr{P}$' cannot have an edge between X and Y that has an arrowhead mark at Y because there exists at least one MAG ($\mathscr{M}$') that has a tail mark at Y. Thus by definition, $\mathscr{P}$' cannot contain an edge between X and Y that has an arrowhead mark at Y and a circle mark at X. By definition PAGs are unique encodings of a Markov Equivalence Class of MAGs and thus $\mathscr{P}$ and $\mathscr{P}$' describe different Markov Equivalence Class of MAGs and $\mathscr{A}$ is not in the set **S**. Thus if there exists an edge between X and Y in $\mathscr{P}$ such that there is a circle mark at X and an arrowhead mark at Y, then $\mathscr{A}$ cannot contain a directed edge from Y to X if it is in the set **S**.

□

Theorem 5. *If there exists two ADMGs $\mathscr{A}_1$ and $\mathscr{A}_2$ in M such that $\mathscr{A}_1$ contains a directed edge between a pair of nodes X and Y and $\mathscr{A}_2$ contains a bi-directed edge between X and Y, then this is the only time that there must be an ADMG $\mathscr{A}_3$ that contains both a bi-directed edge and a directed edge in the same orientation between X and Y.*

Proof. The proof of this theorem has the following two directions:

1. If there exists two ADMGs $\mathscr{A}_1$ and $\mathscr{A}_2$ in M such that $\mathscr{A}_1$ contains a directed edge between a pair of nodes X and Y and $\mathscr{A}_2$ contains a bi-directed edge between X and Y, then there must be an ADMG $\mathscr{A}_3$ that contains both a bi-directed edge and a directed edge in the same orientation between X and Y.

2. If there exists an ADMG $\mathscr{A}_3$ that contains both a bi-directed edge and a directed edge in the some orientation between X and Y, then there must exist two ADMGs $\mathscr{A}_1$ and $\mathscr{A}_2$ in M such that $\mathscr{A}_1$ contains a directed edge between a pair of nodes X and Y in the same orientation and $\mathscr{A}_2$ contains a bi-directed edge between X and Y.

First we prove item 1 using induction.

Base Case: Number of nodes = 2; Number of distinct pairs of nodes = 1. The three graphs trivially belong to the same Markov Equivalence class M.

Inductive Hypothesis: For number of distinct pairs of nodes = k. Item 1 holds.

Prove that for number of distinct pair of nodes = k+1, item 1 holds.

$\mathscr{A}_1$ and $\mathscr{A}_2$ are both in M. They are the same graph except for one pair of nodes in which they have a directed edge and bi-directed edge respectively. Since the existence of either edge does not add any new independence relations and change M as evidenced by both $\mathscr{A}_1$ and $\mathscr{A}_2$ being in M, having $\mathscr{A}_3$ with both edges in the same location also belongs in M.

Next we prove item 2 with a similar inductive argument.

Base Case: Number of nodes = 2; Number of distinct pairs of nodes = 1. The three graphs trivially belong to the same Markov Equivalence class M.

Inductive Hypothesis: For number of distinct pairs of nodes = k. Item 2 holds.

Prove that for number of distinct pair of nodes = k+1, item 2 holds.

Since $\mathscr{A}_3$ contains both a directed edge and bi-directed at some location between a distinct pair of nodes k+1. This can only exist if the existence of one edge by itself is non-essential. If the existence of one edge by itself was essential, there would be no need for this double-edge because the second edge is redundant. Thus $\mathscr{A}_1$ and $\mathscr{A}_2$ must also exist in M as neither edge explicitly is essential for M. $\square$