Colored Gaussian DAG models

Tobias Boege, Kaie Kubjas, Pratik Misra, Liam Solus

arXiv:2404.04024

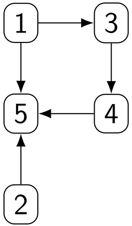
Department of Mathematics
KTH Royal Institute of Technology, Sweden

Applied CATS seminar
KTH Stockholm, 09 April 2024

Gaussian DAG models

- A linear structural equation model defines random variables X recursively via a directed acyclic graph $G = (V, E)$ and Gaussian noise:

$$X_j = \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} X_i + \varepsilon_j, \quad \varepsilon_j \sim \mathcal{N}(0, \omega_j).$$

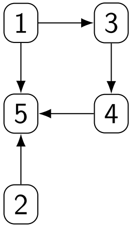


Gaussian DAG models

- A linear structural equation model defines random variables X recursively via a directed acyclic graph $G = (V, E)$ and Gaussian noise:

$$X_j = \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} X_i + \varepsilon_j, \quad \varepsilon_j \sim \mathcal{N}(0, \omega_j).$$

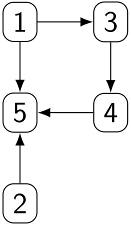
- Parents of node j are regarded as direct causes of j.



Gaussian DAG models

- A linear structural equation model defines random variables X recursively via a directed acyclic graph $G = (V, E)$ and Gaussian noise:

$$X_j = \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} X_i + \varepsilon_j, \quad \varepsilon_j \sim \mathcal{N}(0, \omega_j).$$



- Parents of node j are regarded as direct causes of j.

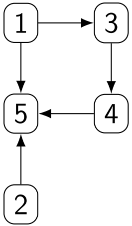
- The vector X is again Gaussian with mean zero. Since G is acyclic, we can solve for the covariance matrix Σ:

$$\Sigma = (I - \Lambda)^{-T} \Omega (I - \Lambda)^{-1}, \quad \text{with } \Lambda \in \mathbb{R}^E \text{ and } \Omega = \mathrm{diag}(\omega).$$

Gaussian DAG models

- A linear structural equation model defines random variables X recursively via a directed acyclic graph $G = (V, E)$ and Gaussian noise:

$$X_j = \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} X_i + \varepsilon_j, \quad \varepsilon_j \sim \mathcal{N}(0, \omega_j).$$



- Parents of node j are regarded as direct causes of j.
- The vector X is again Gaussian with mean zero. Since G is acyclic, we can solve for the covariance matrix Σ:

$$\Sigma = (I - \Lambda)^{-T} \Omega (I - \Lambda)^{-1}, \quad \text{with } \Lambda \in \mathbb{R}^E \text{ and } \Omega = \mathrm{diag}(\omega).$$

- All such matrices form the model $\mathcal{M}(G) \subseteq \mathrm{PD}_V$.

Properties of DAG models

- $\mathcal{M}(G)$ is a smooth submanifold of PD_V and its Zariski closure is irreducible.

Properties of DAG models

- $\mathcal{M}(G)$ is a smooth submanifold of PD_V and its Zariski closure is irreducible.
- The parameters (ω, Λ) are rationally identifiable.

Properties of DAG models

- $\mathcal{M}(G)$ is a smooth submanifold of PD_V and its Zariski closure is irreducible.
- The parameters (ω, Λ) are rationally identifiable.
- The model is equivalently given by the Markov properties of the DAG, e.g.,

$$\mathcal{M}(G) = \{\Sigma \in \mathrm{PD}_V : i \perp\!\!\!\perp j \mid \mathrm{pa}(j) \text{ whenever } ij \notin E\}.$$

Properties of DAG models

- $\mathcal{M}(G)$ is a smooth submanifold of PD_V and its Zariski closure is irreducible.
- The parameters (ω, Λ) are rationally identifiable.
- The model is equivalently given by the Markov properties of the DAG, e.g.,

$$\mathcal{M}(G) = \{\Sigma \in \mathrm{PD}_V : i \perp\!\!\!\perp j \mid \mathrm{pa}(j) \text{ whenever } ij \notin E\}.$$

- Almost all distributions in $\mathcal{M}(G)$ are faithful to G, i.e., do not satisfy more CI statements than the global Markov property.

Properties of DAG models

- $\mathcal{M}(G)$ is a smooth submanifold of PD_V and its Zariski closure is irreducible.
- The parameters (ω, Λ) are rationally identifiable.
- The model is equivalently given by the Markov properties of the DAG, e.g.,

$$\mathcal{M}(G) = \{\Sigma \in \mathrm{PD}_V : i \perp\!\!\!\perp j \mid \mathrm{pa}(j) \text{ whenever } ij \notin E\}.$$

- Almost all distributions in $\mathcal{M}(G)$ are faithful to G, i.e., do not satisfy more CI statements than the global Markov property.
- Model equivalence $\mathcal{M}(G) = \mathcal{M}(H)$ is combinatorially characterized: if and only if G and H have the same skeleton and v-structures.

Properties of DAG models

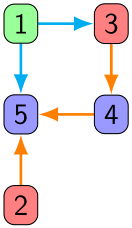
- $\mathcal{M}(G)$ is a smooth submanifold of PD_V and its Zariski closure is irreducible.
- The parameters (ω, Λ) are rationally identifiable.
- The model is equivalently given by the Markov properties of the DAG, e.g.,

$$\mathcal{M}(G) = \{\Sigma \in \mathrm{PD}_V : i \perp\!\!\!\perp j \mid \mathrm{pa}(j) \text{ whenever } ij \notin E\}.$$

- Almost all distributions in $\mathcal{M}(G)$ are faithful to G, i.e., do not satisfy more CI statements than the global Markov property.
- Model equivalence $\mathcal{M}(G) = \mathcal{M}(H)$ is combinatorially characterized: if and only if G and H have the same skeleton and v-structures.
  - Markov equivalence = ambiguity about the direction of causality.

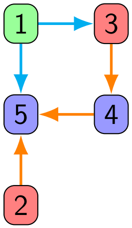
Colored Gaussian DAG models

- In a colored Gaussian DAG model, the vertices and edges of G are partitioned into color classes via a coloring function $c : V \sqcup E \to C_V \sqcup C_E$.



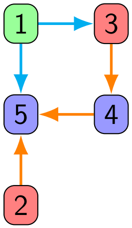
Colored Gaussian DAG models

- In a colored Gaussian DAG model, the vertices and edges of G are partitioned into color classes via a coloring function $c : V \sqcup E \to C_V \sqcup C_E$.
- The parametrization $\Sigma = (I - \Lambda)^{-T}\Omega(I - \Lambda)^{-1}$ stays the same but we reduce the parameter space: $\omega_i = \omega_j$ if $c(i) = c(j)$ and $\lambda_{ij} = \lambda_{kl}$ if $c(ij) = c(kl)$.



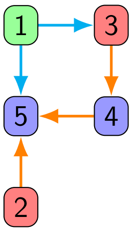
Colored Gaussian DAG models

- In a colored Gaussian DAG model, the vertices and edges of G are partitioned into color classes via a coloring function $c : V \sqcup E \to C_V \sqcup C_E$.
- The parametrization $\Sigma = (I - \Lambda)^{-T}\Omega(I - \Lambda)^{-1}$ stays the same but we reduce the parameter space: $\omega_i = \omega_j$ if $c(i) = c(j)$ and $\lambda_{ij} = \lambda_{kl}$ if $c(ij) = c(kl)$.
- This restricts the parameters to a $\gtrless$ **linear subspace** $\lessgtr$



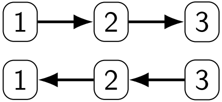
Colored Gaussian DAG models

- In a colored Gaussian DAG model, the vertices and edges of G are partitioned into color classes via a coloring function $c : V \sqcup E \to C_V \sqcup C_E$.
- The parametrization $\Sigma = (I - \Lambda)^{-T}\Omega(I - \Lambda)^{-1}$ stays the same but we reduce the parameter space: $\omega_i = \omega_j$ if $c(i) = c(j)$ and $\lambda_{ij} = \lambda_{kl}$ if $c(ij) = c(kl)$.
- This restricts the parameters to a $\geqslant$**linear subspace**$\leqslant$
- Vertex-only colorings correspond to partial homoscedasticity [WD23].



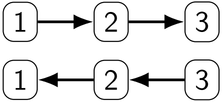
Coloring can disambiguate the causal structure

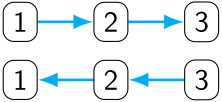
- Coloring reduces Markov-equivalence classes which eases causal discovery.



Coloring can disambiguate the causal structure

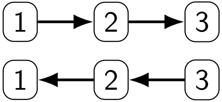
- Coloring reduces Markov-equivalence classes which eases causal discovery.

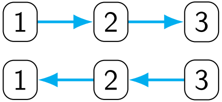




Coloring can disambiguate the causal structure

- Coloring reduces Markov-equivalence classes which eases causal discovery.



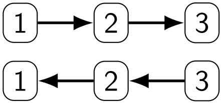


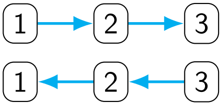
- The vanishing ideal in both cases is

$$I_{13|2} = \langle\ \sigma_{13}\sigma_{22} - \sigma_{12}\sigma_{23}\ \rangle$$

Coloring can disambiguate the causal structure

- Coloring reduces Markov-equivalence classes which eases causal discovery.





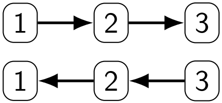
- The vanishing ideal in both cases is

$$I_{13|2} = \langle \, \sigma_{13}\sigma_{22} - \sigma_{12}\sigma_{23} \, \rangle$$

- Generator is invariant under $1 \leftrightarrow 3$.

Coloring can disambiguate the causal structure

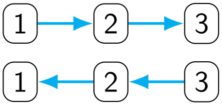
- Coloring reduces Markov-equivalence classes which eases causal discovery.



- The vanishing ideal in both cases is

$$I_{13|2} = \langle\ \sigma_{13}\sigma_{22} - \sigma_{12}\sigma_{23}\ \rangle$$

- Generator is invariant under $1 \leftrightarrow 3$.

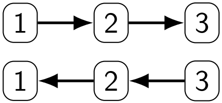


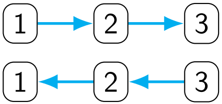
The first vanishing ideal is:

$$I_{13|2} + \langle\ \sigma_{12}\sigma_{22} - \sigma_{11}\sigma_{23}, \sigma_{12}^2 - \sigma_{11}\sigma_{13}\ \rangle$$

Coloring can disambiguate the causal structure

- Coloring reduces Markov-equivalence classes which eases causal discovery.





- The vanishing ideal in both cases is

$$I_{13|2} = \langle\ \sigma_{13}\sigma_{22} - \sigma_{12}\sigma_{23}\ \rangle$$

- Generator is invariant under $1 \leftrightarrow 3$.

The first vanishing ideal is:

$$I_{13|2} + \langle\ \sigma_{12}\sigma_{22} - \sigma_{11}\sigma_{23}, \sigma_{12}^2 - \sigma_{11}\sigma_{13}\ \rangle$$

- Not invariant anymore.

Parameter identifiability revisited

- It follows from the recursive factorization and some linear algebra that

$$\omega_j = \mathrm{Var}(X_i \mid X_{\mathrm{pa}(i)}) = \frac{|\Sigma_{j \cup \mathrm{pa}(j)}|}{|\Sigma_{\mathrm{pa}(j)}|}$$

Parameter identifiability revisited

- It follows from the recursive factorization and some linear algebra that

$$\omega_j = \mathrm{Var}(X_i \mid X_{\mathrm{pa}(i)}) = \frac{|\Sigma_{j \cup \mathrm{pa}(j)}|}{|\Sigma_{\mathrm{pa}(j)}|}, \quad \lambda_{ij} = \frac{\mathrm{Cov}(X_i, X_j \mid X_{\mathrm{pa}(j)\setminus i})}{\mathrm{Var}(X_i \mid X_{\mathrm{pa}(j)\setminus i})} = \frac{|\Sigma_{ij|\mathrm{pa}(j)\setminus i}|}{|\Sigma_{\mathrm{pa}(j)}|}.$$

Parameter identifiability revisited

- It follows from the recursive factorization and some linear algebra that

$$\omega_j = \mathrm{Var}(X_i \mid X_{\mathrm{pa}(i)}) = \frac{|\Sigma_{j\cup \mathrm{pa}(j)}|}{|\Sigma_{\mathrm{pa}(j)}|}, \quad \lambda_{ij} = \frac{\mathrm{Cov}(X_i, X_j \mid X_{\mathrm{pa}(j)\setminus i})}{\mathrm{Var}(X_i \mid X_{\mathrm{pa}(j)\setminus i})} = \frac{|\Sigma_{ij|\mathrm{pa}(j)\setminus i}|}{|\Sigma_{\mathrm{pa}(j)}|}.$$

- Study the rational functions

$$\omega_{j|A}(\Sigma) = \frac{|\Sigma_{j\cup A}|}{|\Sigma_A|}, \quad \lambda_{ij|A}(\Sigma) = \frac{|\Sigma_{ij|A\setminus i}|}{|\Sigma_A|}.$$

Parameter identifiability revisited

- It follows from the recursive factorization and some linear algebra that

$$\omega_j = \mathrm{Var}(X_i \mid X_{\mathrm{pa}(i)}) = \frac{|\Sigma_{j \cup \mathrm{pa}(j)}|}{|\Sigma_{\mathrm{pa}(j)}|}, \quad \lambda_{ij} = \frac{\mathrm{Cov}(X_i, X_j \mid X_{\mathrm{pa}(j) \setminus i})}{\mathrm{Var}(X_i \mid X_{\mathrm{pa}(j) \setminus i})} = \frac{|\Sigma_{ij | \mathrm{pa}(j) \setminus i}|}{|\Sigma_{\mathrm{pa}(j)}|}.$$

- Study the rational functions

$$\omega_{j|A}(\Sigma) = \frac{|\Sigma_{j \cup A}|}{|\Sigma_A|}, \quad \lambda_{ij|A}(\Sigma) = \frac{|\Sigma_{ij|A \setminus i}|}{|\Sigma_A|}.$$

- A set A is identifying for a vertex j resp. edge ij if

$$\omega_j = \omega_{j|A}(\Sigma) \text{ resp. } \lambda_{ij} = \lambda_{ij|A}(\Sigma)$$

for all $\Sigma \in \mathcal{M}(G)$.

Parameter identifiability revisited

Theorem

Let $G = (V, E)$ be a DAG. Then:

- $\omega_j = \omega_{j|A}(\Sigma)$ *for every* $\Sigma \in \mathcal{M}(G)$ *if and only if* $\mathrm{pa}(j) \subseteq A \subseteq V \setminus \overline{\mathrm{de}}(j)$. [WD23]

Parameter identifiability revisited

Theorem

Let $G = (V, E)$ be a DAG. Then:

- $\omega_j = \omega_{j|A}(\Sigma)$ *for every* $\Sigma \in \mathcal{M}(G)$ *if and only if* $\mathrm{pa}(j) \subseteq A \subseteq V \setminus \overline{\mathrm{de}}(j)$. *[WD23]*
- *If* $ij \notin E$, *then* $\lambda_{ij} = 0 = \lambda_{ij|A}(\Sigma)$ *for every* $\Sigma \in \mathcal{M}(G)$ *if and only if* $A \setminus i$ *d-separates* i *and* j *in* G. *[Folklore]*

Parameter identifiability revisited

Theorem

Let $G = (V, E)$ be a DAG. Then:

- $\omega_j = \omega_{j|A}(\Sigma)$ *for every* $\Sigma \in \mathcal{M}(G)$ *if and only if* $\mathrm{pa}(j) \subseteq A \subseteq V \setminus \overline{\mathrm{de}}(j)$. *[WD23]*
- *If* $ij \notin E$*, then* $\lambda_{ij} = 0 = \lambda_{ij|A}(\Sigma)$ *for every* $\Sigma \in \mathcal{M}(G)$ *if and only if* $A \setminus i$ *d-separates* i *and* j *in* G*. [Folklore]*
- *If* $ij \in E$*, then* $\lambda_{ij} = \lambda_{ij|A}(\Sigma)$ *for every* $\Sigma \in \mathcal{M}(G)$ *if and only if* $i \in A \subseteq V \setminus \overline{\mathrm{de}}(j)$ *and* $A \setminus i$ *d-separates* i *and* j *in the graph* G_{ij} *which arises from* G *by deleting the edge* ij *and the vertices* $\mathrm{de}(j)$.

Parameter identifiability revisited

Theorem

Let $G = (V, E)$ be a DAG. Then:

- $\omega_j = \omega_{j|A}(\Sigma)$ *for every* $\Sigma \in \mathcal{M}(G)$ *if and only if* $\mathrm{pa}(j) \subseteq A \subseteq V \setminus \overline{\mathrm{de}}(j)$. *[WD23]*
- *If* $ij \notin E$*, then* $\lambda_{ij} = 0 = \lambda_{ij|A}(\Sigma)$ *for every* $\Sigma \in \mathcal{M}(G)$ *if and only if* $A \setminus i$ *d-separates* i *and* j *in* G*. [Folklore]*
- *If* $ij \in E$*, then* $\lambda_{ij} = \lambda_{ij|A}(\Sigma)$ *for every* $\Sigma \in \mathcal{M}(G)$ *if and only if* $i \in A \subseteq V \setminus \overline{\mathrm{de}}(j)$ *and* $A \setminus i$ *d-separates* i *and* j *in the graph* G_{ij} *which arises from* G *by deleting the edge* ij *and the vertices* $\mathrm{de}(j)$.

- The polynomials $\mathrm{vcr}(i|A, j|B) = |\Sigma_A||\Sigma_B|(\omega_{i|A} - \omega_{j|B})$ resp. $\mathrm{ecr}(ij|A, kl|B) = |\Sigma_A||\Sigma_B|(\lambda_{ij|A} - \lambda_{kl|B})$ vanish on the model $\mathcal{M}(G, c)$ whenever $c(i) = c(j)$ resp. $c(ij) = c(kl)$ and A and B are identifying.

Model geometry

Theorem

For every colored DAG (G, c) *the model* $\mathcal{M}(G, c)$ *has irreducible Zariski closure and is a smooth submanifold of* PD_V*. It is diffeomorphic to an open ball of dimension* $\mathrm{vc} + \mathrm{ec}$ *(the number of vertex- and edge-color classes).*

Model geometry

Theorem

For every colored DAG (G, c) *the model* $\mathcal{M}(G, c)$ *has irreducible Zariski closure and is a smooth submanifold of* PD_V*. It is diffeomorphic to an open ball of dimension* $\mathrm{vc} + \mathrm{ec}$ *(the number of vertex- and edge-color classes).*

Theorem

The vanishing ideal $P_{G,c}$ *of* $\mathcal{M}(G, c)$ *is* $(I_G + I_c) : S_G$ *where:*

- $I_G = \langle |\Sigma_{ij|\mathrm{pa}(j)}| : ij \notin E\rangle$ *is the conditional independence ideal of* G*,*
- $I_c = \langle \mathrm{vcr}(i|\mathrm{pa}(i), j|\mathrm{pa}(j)) : c(i) = c(j)\rangle + \langle \mathrm{ecr}(ij|\mathrm{pa}(j), kl|\mathrm{pa}(l)) : c(ij) = c(kl)\rangle$ *is the coloring ideal of* G*,*
- $S_G = \{\prod_{j\in V} |\Sigma_{\mathrm{pa}(j)}|^{k_j} : k_j \in \mathbb{N}\}$ *is the monoid of parental principal minors.*

Model geometry

Theorem

For every colored DAG (G, c) the model $\mathcal{M}(G, c)$ has irreducible Zariski closure and is a smooth submanifold of PD_V. It is diffeomorphic to an open ball of dimension $\mathrm{vc} + \mathrm{ec}$ (the number of vertex- and edge-color classes).

Theorem

The vanishing ideal $P_{G,c}$ of $\mathcal{M}(G, c)$ is $(I_G + I_c) : S_G$ where:

- $I_G = \langle |\Sigma_{ij|\mathrm{pa}(j)}| : ij \notin E \rangle$ *is the conditional independence ideal of G,*
- $I_c = \langle \mathrm{vcr}(i|\mathrm{pa}(i), j|\mathrm{pa}(j)) : c(i) = c(j) \rangle + \langle \mathrm{ecr}(ij|\mathrm{pa}(j), kl|\mathrm{pa}(l)) : c(ij) = c(kl) \rangle$ *is the coloring ideal of G,*
- $S_G = \{\prod_{j \in V} |\Sigma_{\mathrm{pa}(j)}|^{k_j} : k_j \in \mathbb{N}\}$ *is the monoid of parental principal minors.*

- Resolves the colored generalization of a conjecture of Sullivant; see also [RP14].

Implicitization up to saturation

Lemma

Let R, R' be rings, $S \subseteq R$ multiplicatively closed, and:

- *maps $\phi : R \to R'$ and $\psi : R' \to S^{-1}R$ with $\psi \circ \phi = \mathrm{id}_R$,*
- *for a prime ideal $I' = \langle f_1, \dots, f_k \rangle$, write $\psi(f_i) = {}^{g_i}/_{h_i}$ and set $J = \langle g_i \rangle$.*

If $I := \phi^{-1}(I') \in \mathrm{Spec}(S^{-1}R/J)$, then $I = J : S$.

Implicitization up to saturation

Lemma

Let R, R' be rings, $S \subseteq R$ multiplicatively closed, and:

- *maps $\phi : R \to R'$ and $\psi : R' \to S^{-1}R$ with $\psi \circ \phi = \mathrm{id}_R$,*
- *for a prime ideal $I' = \langle f_1, \dots, f_k \rangle$, write $\psi(f_i) = {g_i}/{h_i}$ and set $J = \langle g_i \rangle$.*

If $I := \phi^{-1}(I') \in \mathrm{Spec}(S^{-1}R/J)$, then $I = J : S$.

- For example, $\phi =$ parametrization of $\mathcal{M}(K_n)$, $\psi =$ parameter identification map and $I' =$ linear equations on parameters from missing edges and color classes.

Implicitization up to saturation

Lemma

Let R, R' be rings, $S \subseteq R$ multiplicatively closed, and:

- *maps $\phi : R \to R'$ and $\psi : R' \to S^{-1}R$ with $\psi \circ \phi = \mathrm{id}_R$,*
- *for a prime ideal $I' = \langle f_1, \ldots, f_k \rangle$, write $\psi(f_i) = {}^{g_i}/_{h_i}$ and set $J = \langle g_i \rangle$.*

If $I := \phi^{-1}(I') \in \mathrm{Spec}(S^{-1}R/J)$, then $I = J : S$.

- For example, $\phi =$ parametrization of $\mathcal{M}(K_n)$, $\psi =$ parameter identification map and $I' =$ linear equations on parameters from missing edges and color classes.
- The lemma computes the vanishing ideal up to a saturation of rationally identifiable models with additional equation constraints.

Implicitization up to saturation

Lemma

Let R, R' be rings, $S \subseteq R$ multiplicatively closed, and:

- *maps $\phi : R \to R'$ and $\psi : R' \to S^{-1}R$ with $\psi \circ \phi = \mathrm{id}_R$,*
- *for a prime ideal $I' = \langle f_1, \ldots, f_k \rangle$, write $\psi(f_i) = {}^{g_i}/_{h_i}$ and set $J = \langle g_i \rangle$.*

If $I := \phi^{-1}(I') \in \mathrm{Spec}(S^{-1}R/J)$, then $I = J : S$.

- For example, $\phi =$ parametrization of $\mathcal{M}(K_n)$, $\psi =$ parameter identification map and $I' =$ linear equations on parameters from missing edges and color classes.
- The lemma computes the vanishing ideal up to a saturation of rationally identifiable models with additional equation constraints.
- Knowing a parametrization and generators for the vanishing ideal up to saturation is sufficient in practice for model distinguishability.

Implicitization up to saturation

Lemma

Let R, R' be rings, $S \subseteq R$ multiplicatively closed, and:

- *maps $\phi : R \to R'$ and $\psi : R' \to S^{-1}R$ with $\psi \circ \phi = \mathrm{id}_R$,*
- *for a prime ideal $I' = \langle f_1, \ldots, f_k \rangle$, write $\psi(f_i) = {}^{g_i}/_{h_i}$ and set $J = \langle g_i \rangle$.*

If $I := \phi^{-1}(I') \in \mathrm{Spec}(S^{-1}R/J)$, then $I = J : S$.

- For example, $\phi =$ parametrization of $\mathcal{M}(K_n)$, $\psi =$ parameter identification map and $I' =$ linear equations on parameters from missing edges and color classes.
- The lemma computes the vanishing ideal up to a saturation of rationally identifiable models with additional equation constraints.
- Knowing a parametrization and generators for the vanishing ideal up to saturation is sufficient in practice for model distinguishability.
- Conceivable to extend to non-linear equations and inequalities.

Faithfulness

Fix a colored DAG (G, c) and $\Sigma \in \mathcal{M}(G, c)$.

- Σ is faithful to G if it satisfies no more CI statements than the d-separations in G.

Faithfulness

Fix a colored DAG (G, c) and $\Sigma \in \mathcal{M}(G, c)$.

- Σ is faithful to G if it satisfies no more CI statements than the d-separations in G.
- Σ is faithful to c if it satisfies no more vcr or ecr relations than those from c.

Faithfulness

Fix a colored DAG (G, c) and $\Sigma \in \mathcal{M}(G, c)$.

- Σ is faithful to G if it satisfies no more CI statements than the d-separations in G.
- Σ is faithful to c if it satisfies no more vcr or ecr relations than those from c.

Theorem ([WD23; STD10])

- *Generic $\Sigma \in \mathcal{M}(G, c)$ is faithful to c.*

Faithfulness

Fix a colored DAG (G, c) and $\Sigma \in \mathcal{M}(G, c)$.

- Σ is faithful to G if it satisfies no more CI statements than the d-separations in G.
- Σ is faithful to c if it satisfies no more vcr or ecr relations than those from c.

Theorem ([WD23; STD10])

- *Generic $\Sigma \in \mathcal{M}(G, c)$ is faithful to c.*
- *Generic $\Sigma \in \mathcal{M}(G, c)$ is faithful to G if c is a vertex-coloring or an edge-coloring.*

Faithfulness

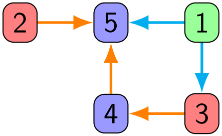
Fix a colored DAG (G, c) and $\Sigma \in \mathcal{M}(G, c)$.

- Σ is faithful to G if it satisfies no more CI statements than the d-separations in G.
- Σ is faithful to c if it satisfies no more vcr or ecr relations than those from c.

Theorem ([WD23; STD10])

- *Generic $\Sigma \in \mathcal{M}(G, c)$ is faithful to c.*
- *Generic $\Sigma \in \mathcal{M}(G, c)$ is faithful to G if c is a vertex-coloring or an edge-coloring.*

- The example on the right colors vertices and edges.
  The generic matrix in the model satisfies $1 \perp\!\!\!\perp 4 \mid 5$.
  No faithful distribution!



Structure identifiability

Theorem ([WD23])

If (G, c) and (H, c) are vertex-colored DAGs, then $\mathcal{M}(G, c) = \mathcal{M}(H, c)$ if and only if G and H are Markov-equivalent and $\mathrm{pa}_G(j) = \mathrm{pa}_H(j)$ for all $j \in V$ with $|c^{-1}(j)| \geq 2$.

Structure identifiability

Theorem ([WD23])

If (G, c) and (H, c) are vertex-colored DAGs, then $\mathcal{M}(G, c) = \mathcal{M}(H, c)$ if and only if G and H are Markov-equivalent and $\mathrm{pa}_G(j) = \mathrm{pa}_H(j)$ for all $j \in V$ with $|c^{-1}(j)| \geq 2$.

An edge-colored DAG (G, c) is BPEC if:

Structure identifiability

Theorem ([WD23])

If (G, c) and (H, c) are vertex-colored DAGs, then $\mathcal{M}(G, c) = \mathcal{M}(H, c)$ if and only if G and H are Markov-equivalent and $\mathrm{pa}_G(j) = \mathrm{pa}_H(j)$ for all $j \in V$ with $|c^{-1}(j)| \geq 2$.

An edge-colored DAG (G, c) is BPEC if:

- proper: all edge color classes have size at least two,

Structure identifiability

Theorem ([WD23])

If (G, c) and (H, c) are vertex-colored DAGs, then $\mathcal{M}(G, c) = \mathcal{M}(H, c)$ if and only if G and H are Markov-equivalent and $\mathrm{pa}_G(j) = \mathrm{pa}_H(j)$ for all $j \in V$ with $|c^{-1}(j)| \geq 2$.

An edge-colored DAG (G, c) is BPEC if:

- proper: all edge color classes have size at least two,
- blocked: color classes partition parent sets of nodes.

Structure identifiability

Theorem ([WD23])

If (G, c) and (H, c) are vertex-colored DAGs, then $\mathcal{M}(G, c) = \mathcal{M}(H, c)$ if and only if G and H are Markov-equivalent and $\mathrm{pa}_G(j) = \mathrm{pa}_H(j)$ for all $j \in V$ with $|c^{-1}(j)| \geq 2$.

An edge-colored DAG (G, c) is BPEC if:

- proper: all edge color classes have size at least two,
- blocked: color classes partition parent sets of nodes.

Theorem

If (G, c) and (H, d) are two BPEC-DAGs, then $\mathcal{M}(G, c) = \mathcal{M}(H, d)$ implies $(G, c) = (H, d)$. In particular, the Markov-equivalence classes of BPEC-DAGs are singletons and the causal structure is identifiable.

References

[BKMS24] Tobias Boege, Kaie Kubjas, Pratik Misra, and Liam Solus. Colored Gaussian DAG models. arXiv:2404.04024 [math.ST]. 2024.

[RP14] Hajir Roozbehani and Yury Polyanskiy. Algebraic Methods of Classifying Directed Graphical Models. arXiv:1401.5551 [cs.IT]. 2014.

[STD10] Seth Sullivant, Kelli Talaska, and Jan Draisma. "Trek separation for Gaussian graphical models". In: Ann. Stat. 38.3 (2010), pp. 1665–1685. ISSN: 0090-5364. DOI: 10.1214/09-AOS760.

[WD23] Jun Wu and Mathias Drton. "Partial Homoscedasticity in Causal Discovery with Linear Models". In: IEEE Journal on Selected Areas in Information Theory (2023).