Colorful linear structural equation models

Tobias Boege, Kaie Kubjas, Pratik Misra, Liam Solus

arXiv:2404.04024

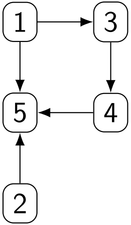
Department of Mathematics and Statistics
UiT The Arctic University of Norway

Algebra seminar UiT
22 November 2024

Linear structural equation models

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

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

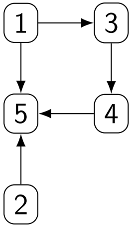


Linear structural equation models

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

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

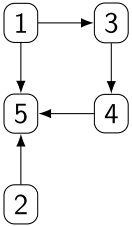
- Parents of node j are regarded as direct causes of j.



Linear structural equation 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.

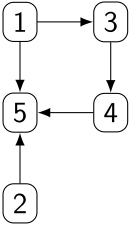
- The vector X is again Gaussian with mean zero. Since G is acyclic, we can solve for the covariance matrix Σ:

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

Linear structural equation 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:

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

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

Reasoning with graphical models

AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE

10.1002/2017WR020412

Key Points:

- We develop a statistical graphical model to characterize the statewide California reservoir system
- We quantify the influence of external physical and economic factors (e.g., statewide PDSI and consumer price index) on the reservoir network
- Further analysis gives a system-wide health diagnosis as a function of PDSI, indicating when heavy management practices may be needed

Supporting Information:

- Supporting Information S1
- Supporting Information S2

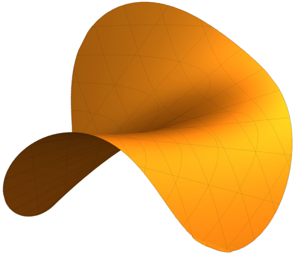
A Statistical Graphical Model of the California Reservoir System

A. Taeb[1], J. T. Reager[2], M. Turmon[2], and V. Chandrasekaran[3]

[1]Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, USA, [2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, [3]Department of Computing and Mathematical Sciences and Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, USA

Abstract The recent California drought has highlighted the potential vulnerability of the state's water management infrastructure to multiyear dry intervals. Due to the high complexity of the network, dynamic storage changes in California reservoirs on a state-wide scale have previously been difficult to model using either traditional statistical or physical approaches. Indeed, although there is a significant line of research on exploring models for single (or a small number of) reservoirs, these approaches are not amenable to a system-wide modeling of the California reservoir network due to the spatial and hydrological heterogeneities of the system. In this work, we develop a state-wide statistical graphical model to characterize the dependencies among a collection of 55 major California reservoirs across the state; this model is defined with respect to a graph in which the nodes index reservoirs and the edges specify the

Statistical models are algebraic varieties*



This set of 3-variate Gaussian distributions in a certain graphical model is defined by a single polynomial equation in its covariance matrix

$$\sigma_{12}(1 + \sigma_{23} - \sigma_{12}^2) = \sigma_{13}.$$

Algebraic statistics studies statistical models and problems using methods of algebraic geometry and computer algebra.

*sometimes

Properties of SEMs

- $\mathcal{M}(G)$ is an irreducible algebraic subvariety and a smooth submanifold of PD_V.

Properties of SEMs

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

Properties of SEMs

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

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

Properties of SEMs

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

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

- 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 SEMs

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

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

- 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 SEMs

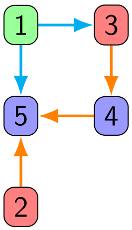
- $\mathcal{M}(G)$ is an irreducible algebraic subvariety and a smooth submanifold of PD_V.
- The parameters (ω, Λ) are rationally identifiable.
- The model is equivalently given by the Markov property of the DAG, e.g.,

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

- 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.

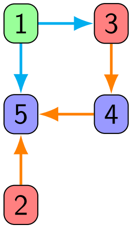
Colored Gaussian DAG models

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



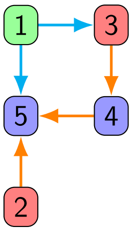
Colored Gaussian DAG models

- 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 $c(ij) = c(kl)$.



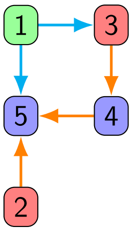
Colored Gaussian DAG models

- 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 linear subspace.



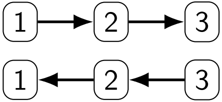
Colored Gaussian DAG models

- 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 linear subspace.
- Vertex-only colorings correspond to partial homoscedasticity [WD23].



Coloring can disambiguate the causal structure

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

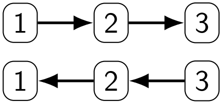


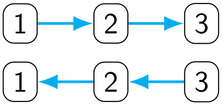
Coloring can disambiguate the causal structure

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

Coloring can disambiguate the causal structure

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



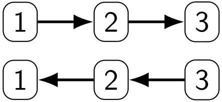


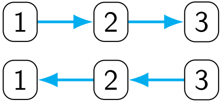
- The vanishing ideal in both cases is

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

Coloring can disambiguate the causal structure

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





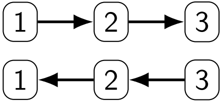
- The vanishing ideal in both cases is

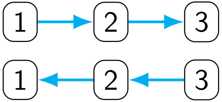
$$I_{13|2} = \langle\ \sigma_{13}\sigma_{22} - \sigma_{12}\sigma_{23}\ \rangle$$

- 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

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

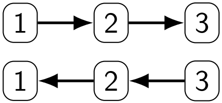
The first vanishing ideal is:

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

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

Coloring can disambiguate the causal structure

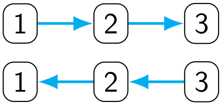
- Coloring reduces Markov-equivalence classes which eases causal discovery.



- 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:

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

- Not invariant anymore.

Parameter identifiability revisited

Consider the functions $\omega_{j|A}(\Sigma) = \mathrm{Var}(X_i \mid X_A)$ and $\lambda_{ij|A}(\Sigma) = \frac{\mathrm{Cov}(X_i, X_j | X_{A \setminus i})}{\mathrm{Var}(X_i | X_{A \setminus i})}$.

Theorem

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

Parameter identifiability revisited

Consider the functions $\omega_{j|A}(\Sigma) = \mathrm{Var}(X_i \mid X_A)$ and $\lambda_{ij|A}(\Sigma) = \frac{\mathrm{Cov}(X_i, X_j | X_{A\setminus i})}{\mathrm{Var}(X_i | X_{A\setminus i})}$.

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

Consider the functions $\omega_{j|A}(\Sigma) = \mathrm{Var}(X_i \mid X_A)$ and $\lambda_{ij|A}(\Sigma) = \frac{\mathrm{Cov}(X_i, X_j | X_{A\setminus i})}{\mathrm{Var}(X_i | X_{A\setminus i})}$.

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

Consider the functions $\omega_{j|A}(\Sigma) = \mathrm{Var}(X_i \mid X_A)$ and $\lambda_{ij|A}(\Sigma) = \frac{\mathrm{Cov}(X_i, X_j | X_{A\setminus i})}{\mathrm{Var}(X_i | X_{A\setminus i})}$.

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

Consider the functions $\omega_{j|A}(\Sigma) = \mathrm{Var}(X_i \mid X_A)$ and $\lambda_{ij|A}(\Sigma) = \frac{\mathrm{Cov}(X_i, X_j | X_{A\setminus i})}{\mathrm{Var}(X_i | X_{A\setminus i})}$.

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* $\mathrm{de}(j)$.

This expresses the causal and coloring conditions $\lambda_{ij} = 0$, $\omega_i = \omega_j$ and $\lambda_{ij} = \lambda_{kl}$ as polynomial conditions cir, vcr and ecr in Σ.

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:

- $I_G = \langle \mathrm{cir}(ij|\mathrm{pa}(j)) : ij \notin E \rangle$ *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:

- $I_G = \langle \mathrm{cir}(ij|\mathrm{pa}(j)) : ij \notin E\rangle$ *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].

Application to implicitization

```
needsPackage "GraphicalModels";

V = {0,1,2,3,4,5};
G = digraph(V,{{0,2},{0,3},{1,2},{1,3},{2,3},{3,4},{0,5},{1,5},{2,5},{3,5},{4,5}});
R = gaussianRing G; S = covarianceMatrix R;
allE = set(flatten for i in 0..#V-1 list for j in i+1..#V-1 list (V#i,V#j));

-- Vanishing ideal via built-in elimination method: not finished after 20 minutes
time I1 = gaussianVanishingIdeal R;

-- Vanishing ideal via saturation: 0.186855 seconds
time (
  prs = for i in V list (
    P := toList parents(G, i);
    if #P == 0 then 1 else det submatrix(S, P, P)
  );
  J = ideal for ij in toList(allE-set(edges G)) list (
    P := toList parents(G, ij#1);
    det submatrix(S, {ij#0}|P, {ij#1}|P)
  );
  I2 = fold(saturate, J, prs);
);
```

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

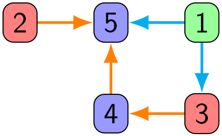
Fix a colored DAG (G, c) and $\Sigma \in \mathcal{M}(G, c)$.

- 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 $X_1 \perp\!\!\!\perp X_4 \mid X_5$.
  Not faithful to G!



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 each node.

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 $|c^{-1}(j)| \geq 2$.

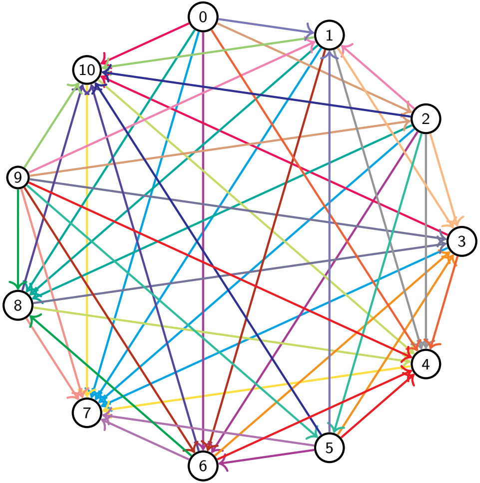
An edge-colored DAG (G, c) is BPEC if:

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

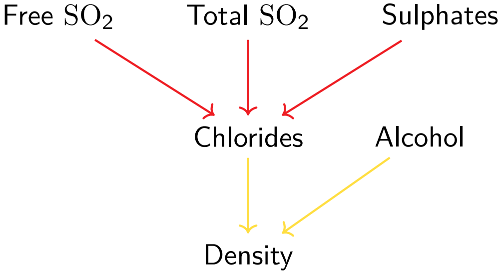
Theorem

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.

Wine tasting



A sensible subgraph:



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).