The intersection property for conditional independence

Alex Fink

Queen Mary University of London

Algebraic Methods in Statistics
Osnabrück, Advent 2017

Conditional independence

Let $X = (X_1, \ldots, X_n)$ be a random var with outcomes $\Omega = \prod_{i=1}^{n} \Omega_i$.
Write $X_A = (X_i)_{i \in A}$, etc.

Let A, B, C be disjoint subsets of the index set $[n]$.

The conditional independence ("CI") statement

$$X_A \perp\!\!\!\perp X_B \mid X_C$$

asserts of X that

$$\mathbf{P}(x_A = a, x_B = b \mid x_C = c) = \mathbf{P}(x_A = a \mid x_C = c) \cdot \mathbf{P}(x_B = b \mid x_C = c)$$

i.e.

$$\mathbf{P}(x_A = a, x_B = b, x_C = c)\mathbf{P}(x_C = c) = \mathbf{P}(x_A = a, x_C = c)\mathbf{P}(x_B = b, x_C = c)$$

for all $a \in \Omega_A$, $b \in \Omega_B$, and $c \in \operatorname{supp} X_c$.

Why CI?

CI is important in understanding observed data:

- identifying irrelevant variables, for dimensionality reduction
- inference of causal relationships

The first attempt to capture all the CI relationships in a dataset was through graphs, each edge being an "atomic" causation.

$X_i \perp\!\!\!\perp X_j \mid X_A$ iff all paths are like:

node $\in A$
node $\notin A$

But this is insufficiently general: not all distributions have a graph.

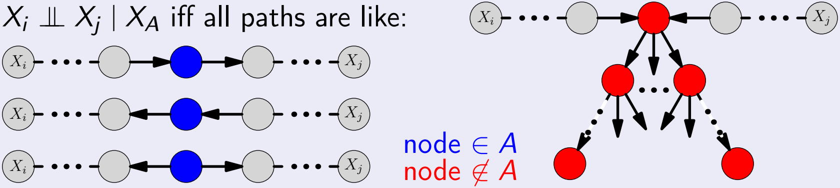
Why CI?

CI is important in understanding observed data:

- identifying irrelevant variables, for dimensionality reduction
- inference of causal relationships

The first attempt to capture all the CI relationships in a dataset was through graphs, each edge being an "atomic" causation.

$X_i \perp\!\!\!\perp X_j \mid X_A$ iff all paths are like:



But this is insufficiently general: not all distributions have a graph.

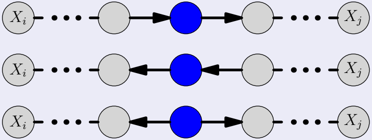
Why CI?

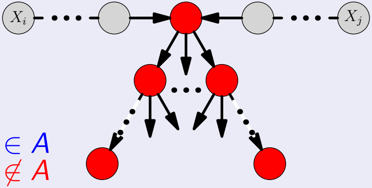
CI is important in understanding observed data:

- identifying irrelevant variables, for dimensionality reduction
- inference of causal relationships

The first attempt to capture all the CI relationships in a dataset was through graphs, each edge being an “atomic” causation.

$X_i \perp\!\!\!\perp X_j \mid X_A$ iff all paths are like:





node $\in A$
node $\notin A$

But this is insufficiently general: not all distributions have a graph.

Discrete distributions

Let X be discrete with outcome probabilities $p_{abcz} = \mathbf{P}(x_A = a, \ldots)$.

The CI statement

$$X_A \perp\!\!\!\perp X_B \mid X_C$$

says that one gets a rank 1 matrix from the tensor (p_{abcd}) by

- flattening in the $A \times B$ direction;
- slicing in the C direction;
- marginalising in the $Z = [n] \setminus (A \cup B \cup C)$ direction.

The ideal of $X_A \perp\!\!\!\perp X_B \mid X_C$ is

$$(p_{a_1 b_1 c+} p_{a_2 b_2 c+} - p_{a_1 b_2 c+} p_{a_2 b_1 c+})$$

where $p_{abc+} = \sum_z p_{abcz}$.

[Pearl–Paz '87] How to capture the combinatorics of the sets of CI statements that hold of some distribution?

Semigraphoids, defined by four conditional independence axioms.

Symmetry $X_A \perp\!\!\!\perp X_B \mid X_C \implies X_B \perp\!\!\!\perp X_A \mid X_C$

Decomposition $X_A \perp\!\!\!\perp X_{B\cup C} \mid X_D \implies X_A \perp\!\!\!\perp X_B \mid X_D$

Weak union $X_A \perp\!\!\!\perp X_{B\cup C} \mid X_D \implies X_A \perp\!\!\!\perp X_B \mid X_{C\cup D}$

Contraction $(X_A \perp\!\!\!\perp X_B \mid X_{C\cup D} \text{ and } X_A \perp\!\!\!\perp X_C \mid X_D) \implies X_A \perp\!\!\!\perp X_{B\cup C} \mid X_D$

(These don't completely characterise distributions; no finite list of axioms can. But they are the complete list with ≤ 2 conjuncts. [Studený '92, '97])

The intersection axiom

[Pearl–Paz '87] How to capture the combinatorics of the sets of CI statements that hold of some distribution? (semigraphoids, graphoids)

The intersection axiom *almost* holds:

$$X_A \perp\!\!\!\perp X_B \mid X_{C\cup D},\ X_A \perp\!\!\!\perp X_C \mid X_{B\cup D} \overset{?}{\Longrightarrow} X_A \perp\!\!\!\perp X_{B\cup C} \mid X_D$$

Let's analyse it in the discrete case.

$$\mathcal{I} := (p_{i_1j_1k}p_{i_2j_2k} - p_{i_2j_1k}p_{i_1j_2k},\ p_{i_1jk_1}p_{i_2jk_2} - p_{i_2jk_1}p_{i_1jk_2})$$
$$\overset{?}{\supseteq} (p_{i_1j_1k_1}p_{i_2j_2k_2} - p_{i_2j_1k_1}p_{i_1j_2k_2})$$

If the probability density is positive everywhere, then the intersection axiom holds. ([DSS '08] discrete; [Pearl '09] continuous)

Question

What weaker conditions on positivity suffice?

The intersection axiom

[Pearl–Paz '87] How to capture the combinatorics of the sets of CI statements that hold of some distribution? (semigraphoids, graphoids)

The intersection axiom *almost* holds:

$$X_1 \perp\!\!\!\perp X_2 \mid X_3,\ X_1 \perp\!\!\!\perp X_3 \mid X_2 \overset{?}{\Longrightarrow} X_1 \perp\!\!\!\perp (X_2, X_3)$$

i.e.

$$\mathcal{I} := (p_{i_1 j_1 k} p_{i_2 j_2 k} - p_{i_2 j_1 k} p_{i_1 j_2 k},\ p_{i_1 j k_1} p_{i_2 j k_2} - p_{i_2 j k_1} p_{i_1 j k_2}) \overset{?}{\supseteq} (p_{i_1 j_1 k_1} p_{i_2 j_2 k_2} - p_{i_2 j_1 k_1} p_{i_1 j_2 k_2})$$

If the probability density is positive everywhere, then the intersection axiom holds. ([DSS '08] discrete; [Pearl '09] continuous)

Question

What weaker conditions on positivity suffice?

The intersection axiom

[Pearl–Paz '87] How to capture the combinatorics of the sets of CI statements that hold of some distribution? (semigraphoids, graphoids)

The intersection axiom *almost* holds:

$$X_1 \perp\!\!\!\perp X_2 \mid X_3,\ X_1 \perp\!\!\!\perp X_3 \mid X_2 \overset{?}{\Longrightarrow} X_1 \perp\!\!\!\perp (X_2, X_3)$$

In fact

$$\sqrt{\mathcal{I}} = \sqrt{(p_{i_1j_1k}p_{i_2j_2k} - p_{i_2j_1k}p_{i_1j_2k},\ p_{i_1jk_1}p_{i_2jk_2} - p_{i_2jk_1}p_{i_1jk_2})}$$
$$\not\supseteq (p_{i_1j_1k_1}p_{i_2j_2k_2} - p_{i_2j_1k_1}p_{i_1j_2k_2})$$

If the probability density is positive everywhere, then the intersection axiom holds. ([DSS '08] discrete; [Pearl '09] continuous)

Question

What weaker conditions on positivity suffice?

Analysing the discrete case

Question

[Drton–Sturmfels–Sullivant '08] What are the primary components of

$$\mathcal{I} = (p_{i_1 j_1 k} p_{i_2 j_2 k} - p_{i_2 j_1 k} p_{i_1 j_2 k},\ p_{i_1 j k_1} p_{i_2 j k_2} - p_{i_2 j k_1} p_{i_1 j k_2})?$$

One of them is the ideal of $X_1 \perp\!\!\!\perp X_2 \mid X_3$:

$$\mathcal{I} : (p_{111} \cdots p_{|\Omega_1|,|\Omega_2|,|\Omega_3|})^\infty = (p_{i_1 j_1 k_1} p_{i_2 j_2 k_2} - p_{i_2 j_1 k_1} p_{i_1 j_2 k_2}).$$

The other components will be binomial ideals as well [Eisenbud–Sturmfels '96].

Moral theorem

If $X_A \perp\!\!\!\perp X_B \mid X_{C \cup D}$ and $X_A \perp\!\!\!\perp X_C \mid X_{B \cup D}$, then $X_A \perp\!\!\!\perp X_{B \cup C} \mid (X_D, \mathscr{C})$, where $\mathscr{C}$ is the "connected component" of $\mathrm{supp}(X_{B \cup C})$ containing $x_{B \cup C}$.

Question

[Drton–Sturmfels–Sullivant '08] What are the primary components of

$$\mathcal{I} = (p_{i_1 j_1 k} p_{i_2 j_2 k} - p_{i_2 j_1 k} p_{i_1 j_2 k},\ p_{i_1 j k_1} p_{i_2 j k_2} - p_{i_2 j k_1} p_{i_1 j k_2})?$$

One of them is the ideal of $X_1 \perp\!\!\!\perp X_2 \mid X_3$:

$$\mathcal{I} : (p_{111} \cdots p_{|\Omega_1|,|\Omega_2|,|\Omega_3|})^\infty = (p_{i_1 j_1 k_1} p_{i_2 j_2 k_2} - p_{i_2 j_1 k_1} p_{i_1 j_2 k_2}).$$

The other components will be binomial ideals as well [Eisenbud–Sturmfels '96].

Moral theorem

If $X_A \perp\!\!\!\perp X_B \mid X_{C \cup D}$ and $X_A \perp\!\!\!\perp X_C \mid X_{B \cup D}$, then $X_A \perp\!\!\!\perp X_{B \cup C} \mid (X_D, \mathscr{C})$, where $\mathscr{C}$ is the "connected component" of $\operatorname{supp}(X_{B \cup C})$ containing $x_{B \cup C}$.

Question

[Drton–Sturmfels–Sullivant '08] What are the primary components of

$$\mathcal{I} = (p_{i_1 j_1 k} p_{i_2 j_2 k} - p_{i_2 j_1 k} p_{i_1 j_2 k},\ p_{i_1 j k_1} p_{i_2 j k_2} - p_{i_2 j k_1} p_{i_1 j k_2})?$$

One of them is the ideal of $X_1 \perp\!\!\!\perp X_2 \mid X_3$:

$$\mathcal{I} : (p_{111} \cdots p_{|\Omega_1|,|\Omega_2|,|\Omega_3|})^\infty = (p_{i_1 j_1 k_1} p_{i_2 j_2 k_2} - p_{i_2 j_1 k_1} p_{i_1 j_2 k_2}).$$

The other components will be binomial ideals as well [Eisenbud–Sturmfels '96].

Moral theorem

If $X_A \perp\!\!\!\perp X_B \mid X_{C \cup D}$ and $X_A \perp\!\!\!\perp X_C \mid X_{B \cup D}$, then $X_A \perp\!\!\!\perp X_{B \cup C} \mid (X_D, \mathscr{C})$, where $\mathscr{C}$ is the "connected component" of $\mathrm{supp}(X_{B \cup C})$ containing $x_{B \cup C}$.

The primary decomposition of $\mathcal{I}$

Theorem (Fink '11); conjecture (Cartwright, Engström)

$\mathcal{I}$ has the primary decomposition $\mathcal{I} = \bigcap_G P_G$ *running over* *admissible graphs G.*

Each P_G is prime, so $\mathcal{I}$ is radical.

A bipartite graph on $\Omega_2 \amalg \Omega_3$ is *admissible* if adding any edge unites two connected components.

$$P_G = (p_{i_1j_1k_1}p_{i_2j_2k_2} - p_{i_2j_1k_1}p_{i_1j_2k_2} : (j_1, k_1) \text{ and } (j_2, k_2) \in G \text{ connected}) + (p_{ijk} : (j, k) \notin G)$$

Right: the tensor (p_{ijk}) viewed along the i direction.

	1	2	3	4
1	0		0	0
2		0		0
3		0		0
4	0	0	0	
5	0	0	0	

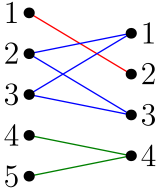
The primary decomposition of $\mathcal{I}$

Theorem (Fink '11); conjecture (Cartwright, Engström)

$\mathcal{I}$ has the primary decomposition $\mathcal{I} = \bigcap_G P_G$ *running over* *admissible graphs G.*

Each P_G is prime, so $\mathcal{I}$ is radical.

A bipartite graph on $\Omega_2 \amalg \Omega_3$ is *admissible* if adding any edge unites two connected components.



$$P_G = (p_{i_1 j_1 k_1} p_{i_2 j_2 k_2} - p_{i_2 j_1 k_1} p_{i_1 j_2 k_2} : (j_1, k_1) \text{ and } (j_2, k_2) \in G \text{ connected}) + (p_{ijk} : (j,k) \notin G)$$

	1	2	3	4
1	0		0	0
2		0		0
3		0		0
4	0	0	0	
5	0	0	0	

Right: the tensor (p_{ijk}) viewed along the i direction.

Theorem (Fink '11); conjecture (Cartwright, Engström)

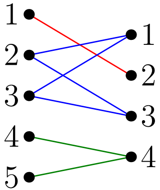
$\mathcal{I}$ has the primary decomposition $\mathcal{I} = \bigcap_G P_G$ *running over admissible graphs G.*

Each P_G is prime, so $\mathcal{I}$ is radical.

A bipartite graph on $\Omega_2 \amalg \Omega_3$ is *admissible* if adding any edge unites two connected components.

$$P_G = (p_{i_1 j_1 k_1} p_{i_2 j_2 k_2} - p_{i_2 j_1 k_1} p_{i_1 j_2 k_2} : (j_1, k_1) \text{ and } (j_2, k_2) \in G \text{ connected}) + (p_{ijk} : (j,k) \notin G)$$

Right: the tensor (p_{ijk}) viewed along the i direction.



	1	2	3	4
1	0		0	0
2		0		0
3		0		0
4	0	0	0	
5	0	0	0	

Proof

Theorem

$$\mathcal{I} = \bigcap_{G \text{ admissible}} P_G$$

- $\mathcal{I} \subseteq$ each P_G ✓
- For $\supseteq$: Let $\deg p_{ijk} = e_{jk}$. Let $G(d) =$ support of $d \in \mathbb{N}^{\Omega_2 \times \Omega_3}$.

Key fact about connectedness

Let f be a monomial multiple of $p_{i_1 j_1 k_1} p_{i_2 j_2 k_2} - p_{i_2 j_1 k_1} p_{i_1 j_2 k_2}$.
Then $f \in \mathcal{I} \iff (j_1, k_1)$ and (j_2, k_2) are connected in $G(\deg f)$.

Let $\overline{G(d)}$ be an "admissible closure" of $G(d)$.

Claim. $P_{\overline{G(d)}}$ has the smallest multidegree d piece of any P_G.

$$(\mathcal{I})_d \overset{?}{\supseteq} (P_{\overline{G(d)}})_d \overset{?}{\supseteq} (\bigcap_G P_G)_d$$

Proof

Theorem

$$\mathcal{I} = \bigcap_{G \text{ admissible}} P_G$$

- $\mathcal{I} \subseteq$ each P_G ✓
- For $\supseteq$: Let $\deg p_{ijk} = e_{jk}$. Let $G(d) =$ support of $d \in \mathbb{N}^{\Omega_2 \times \Omega_3}$.

Key fact about connectedness

Let f be a monomial multiple of $p_{i_1 j_1 k_1} p_{i_2 j_2 k_2} - p_{i_2 j_1 k_1} p_{i_1 j_2 k_2}$.
Then $f \in \mathcal{I} \iff (j_1, k_1)$ and (j_2, k_2) are connected in $G(\deg f)$.

Let $\overline{G(d)}$ be an "admissible closure" of $G(d)$.

Claim. $P_{\overline{G(d)}}$ has the smallest multidegree d piece of any P_G.

$$(\mathcal{I})_d \overset{\checkmark}{\supseteq} (P_{\overline{G(d)}})_d \overset{?}{\supseteq} \Big(\bigcap_G P_G\Big)_d$$

Proof

Theorem

$$\mathcal{I} = \bigcap_{G \text{ admissible}} P_G$$

- $\mathcal{I} \subseteq$ each P_G ✓
- For $\supseteq$: Let $\deg p_{ijk} = e_{jk}$. Let $G(d) =$ support of $d \in \mathbb{N}^{\Omega_2 \times \Omega_3}$.

Key fact about connectedness

Let f be a monomial multiple of $p_{i_1 j_1 k_1} p_{i_2 j_2 k_2} - p_{i_2 j_1 k_1} p_{i_1 j_2 k_2}$.

Then $f \in \mathcal{I} \iff (j_1, k_1)$ and (j_2, k_2) are connected in $G(\deg f)$.

Let $\overline{G(d)}$ be an "admissible closure" of $G(d)$.

Claim. $P_{\overline{G(d)}}$ has the smallest multidegree d piece of any P_G.

$$(\mathcal{I})_d \overset{\checkmark}{\supseteq} (P_{\overline{G(d)}})_d \overset{?}{\supseteq} (\bigcap_G P_G)_d$$

Proof continued: an initial degeneration

By Hilbert function arguments, we may take an initial degeneration.

$$(\mathrm{in}\, P_{\overline{G(d)}})_d \overset{?}{\supseteq} \bigcap_G (\mathrm{in}\, P_G)_d \supseteq \left(\mathrm{in} \bigcap_G P_G\right)_d$$

[Sturmfels '91] on ideals of 2×2 minors:

- For any term order, in P_G is a squarefree monomial ideal.
- Ideals in $P_G \longleftrightarrow$ triangulations of products of simplices.
- For graded revlex order, our generators for P_G are a GB.

□

Corollary

$$\mathrm{in}\, \mathcal{I} = \bigcap \mathrm{in}\, P_G.$$

But this does not produce a Gröbner basis for $\mathcal{I}$.

Proof continued: an initial degeneration

By Hilbert function arguments, we may take an initial degeneration.

$$(\operatorname{in} P_{\overline{G(d)}})_d \overset{?}{\supseteq} \bigcap_G (\operatorname{in} P_G)_d \supseteq \left(\operatorname{in} \bigcap_G P_G\right)_d$$

[Sturmfels '91] on ideals of 2×2 minors:

- For any term order, in P_G is a squarefree monomial ideal.
- Ideals in $P_G \longleftrightarrow$ triangulations of products of simplices.
- For graded revlex order, our generators for P_G are a GB.

■

Corollary

$$\operatorname{in} \mathcal{I} = \bigcap \operatorname{in} P_G.$$

But this does not produce a Gröbner basis for $\mathcal{I}$.

Proof continued: an initial degeneration

By Hilbert function arguments, we may take an initial degeneration.

$$(\mathrm{in}\, P_{\overline{G(d)}})_d \overset{?}{\supseteq} \bigcap_G (\mathrm{in}\, P_G)_d \supseteq \left(\mathrm{in} \bigcap_G P_G\right)_d$$

[Sturmfels '91] on ideals of 2×2 minors:

- For any term order, in P_G is a squarefree monomial ideal.
- Ideals in $P_G \longleftrightarrow$ triangulations of products of simplices.
- For graded revlex order, our generators for P_G are a GB.

∎

Corollary

$$\mathrm{in}\, \mathcal{I} = \bigcap \mathrm{in}\, P_G.$$

But this does not produce a Gröbner basis for $\mathcal{I}$.

Generalisation: binomial edge ideals

The binomial edge ideal of a graph G is

$$J_G = (x_i y_j - x_j y_i : (i,j) \in G) \subseteq \mathbb{K}[x_i, y_i : i \in V(G)].$$

If $|\Omega_1| = 2$, then $\mathcal{I}$ and its components are binomial edge ideals.
So is any CI ideal $X_1 \perp\!\!\!\perp X_B \mid X_{[n]\setminus B\setminus 1}$.

Theorems (Herzog–Hibi–Hreinsdóttir–Kahle–Rauh '10; Ohtani '11)

One can give explicitly

- a decomposition of J_G into prime ideals
- a Gröbner basis for J_G in lex order (sometimes quadratic)
- a sufficient condition for J_G to be Cohen-Macaulay

(Our $\mathcal{I}$ is not CM, and its GB is not quadratic.)

Damadi–Rahmati '16, Banerjee–Núñez-Betancourt '17, de Alba–Hoang 'xx. . .

Generalisation: binomial edge ideals

The binomial edge ideal of a graph G is

$$J_G = (x_i y_j - x_j y_i : (i,j) \in G) \subseteq \mathbb{K}[x_i, y_i : i \in V(G)].$$

If $|\Omega_1| = 2$, then $\mathcal{I}$ and its components are binomial edge ideals.
So is any CI ideal $X_1 \perp\!\!\!\perp X_B \mid X_{[n]\setminus B\setminus 1}$.

Theorems (Herzog–Hibi–Hreinsdóttir–Kahle–Rauh '10; Ohtani '11)

One can give explicitly

- a decomposition of J_G into prime ideals
- a Gröbner basis for J_G in lex order (sometimes quadratic)
- a sufficient condition for J_G to be Cohen-Macaulay

(Our $\mathcal{I}$ is not CM, and its GB is not quadratic.)

Damadi–Rahmati '16, Banerjee–Núñez-Betancourt '17, de Alba–Hoang 'xx. . .

Beyond the binary case

[Rauh–Ay '11] Let $\mathcal{R}$ be any set of CI statements

$$X_1 \perp\!\!\!\perp X_B \mid X_{[n]\setminus B\setminus 1}$$

and $\mathcal{I}_\mathcal{R}$ its ideal.

Application: Robustness. Does output random variable X_1 have unchanged distribution if inputs X_B are "disabled"?

Theorems

- $\mathcal{I}_\mathcal{R}$ is an intersection of primes, one for each subgraph maximal for its connected components. ($\Rightarrow$ moral theorem)
- Explicit reduced GB for $\mathcal{I}_\mathcal{R}$.

Beyond the binary case

[Rauh–Ay '11] Let $\mathcal{R}$ be any set of CI statements

$$X_1 \perp\!\!\!\perp X_B \mid X_{[n]\setminus B\setminus 1}$$

and $\mathcal{I}_\mathcal{R}$ its ideal.

Application: Robustness. Does output random variable X_1 have unchanged distribution if inputs X_B are "disabled"?

Theorems

- $\mathcal{I}_\mathcal{R}$ is an intersection of primes, one for each subgraph maximal for its connected components. ($\Rightarrow$ moral theorem)
- Explicit reduced GB for $\mathcal{I}_\mathcal{R}$.

Another generalisation

[Swanson–Taylor '12] consider the ideal $\mathcal{I}^{(t)}$ of

$$\{X_i \perp\!\!\!\perp X_j \mid X_{[n]\setminus\{i,j\}} : i \leq t, j \leq n\}.$$

Ay–Rauh subsumes $t = 1$. $\mathcal{I}$ is the case $t = 1$, $n = 3$.

Theorems

One can give explicitly

- the minimal primes of $\mathcal{I}^{(t)}$. It is no longer radical!
  The primes are subsets maximal for their connected components.
- Gröbner bases for the binomial parts of the minimal primes.

The full-support component is $\{X_i \perp\!\!\!\perp X_{[n]\setminus i} : i \leq t\}$.

Another generalisation

[Swanson–Taylor '12] consider the ideal $\mathcal{I}^{(t)}$ of

$$\{X_i \perp\!\!\!\perp X_j \mid X_{[n]\setminus\{i,j\}} : i \leq t, j \leq n\}.$$

Ay–Rauh subsumes $t = 1$. $\mathcal{I}$ is the case $t = 1$, $n = 3$.

Theorems

One can give explicitly

- the minimal primes of $\mathcal{I}^{(t)}$. It is no longer radical!
  The primes are subsets maximal for their connected components.
- Gröbner bases for the binomial parts of the minimal primes.

The full-support component is $\{X_i \perp\!\!\!\perp X_{[n]\setminus i} : i \leq t\}$.

Continuous distributions

Let p be a continuous probability density on the metric space Ω.

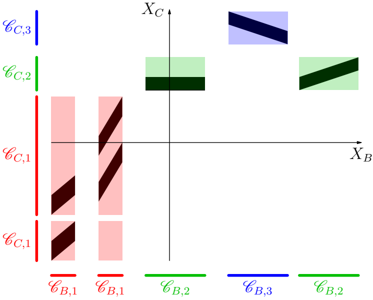
Theorem (Peters '14)

If $X_A \perp\!\!\!\perp X_B \mid X_{C\cup D}$ and $X_A \perp\!\!\!\perp X_C \mid X_{B\cup D}$, then $X_A \perp\!\!\!\perp X_{B\cup C} \mid (X_D, \mathscr{C})$ where $\mathscr{C}$ is the component of $\{(b,c) : p(b,c,d) > 0\}$ containing $x_{B\cup C}$.

Let $\{\mathscr{C}_{B,i}\}_{i=1}^k$ and $\{\mathscr{C}_{C,i}\}_{i=1}^k$ be families of minimal disjoint sets s.t.

$$\{(b,c) : p(b,c,d) > 0\} \subseteq \bigcup_i (\mathscr{C}_{B,i} \times \mathscr{C}_{C,i}).$$

The $\mathscr{C}_{B,i} \times \mathscr{C}_{C,i}$ are the components.



Selected references

- A Fink, *The binomial ideal of the intersection axiom for conditional probabilities*, J. of Algebraic Combinatorics **33** issue 3 (2011), 455–463.
- J Herzog, T Hibi, F Hreinsdóttir, T Kahle, J Rauh, *Binomial edge ideals and conditional independence statements*, Advances in Applied Mathematics **45** no. 3 (2010), 317–333.
- M Ohtani, *Graphs and ideals generated by some 2-minors*, Communications in Algebra **39**, no. 3 (2011), 905–917.
- J Peters, *On the intersection property of conditional independence and its application to causal discovery*, J. of Causal Inference **3** (2015), 97–108.
- J Rauh, N Ay, *Robustness and conditional independence ideals*, arXiv:1110.1338.
- I Swanson, A Taylor, *Minimal primes of ideals arising from conditional independence statements*, J. Algebra **392** (2013), 299–314.

Thanks!