Conditional independence ideals with hidden variables

Fatemeh Mohammadi (IST Austria)
Johannes Rauh (York University)

June 10, 2015

Conditional independence

Consider three discrete random variables X_1, X_2, X_3 with finite ranges $[r_1], [r_2], [r_3]$ and the joint state space: $\mathcal{X} = [r_1] \times [r_2] \times [r_3]$.

X_1 is (conditionally) independent of X_2 given X_3

$$P(X_1 = x_1, X_2 = x_2 | X_3 = x_3) = P(X_1 = x_1 | X_3 = x_3) P(X_2 = x_2 | X_3 = x_3)$$

$$p_{x_1 x_2 x_3} = p_{x_1 + x_3} p_{+ x_2 x_3}$$

$$X_1 \perp\!\!\!\perp X_2 \mid X_3$$

Lemma

The following equivalent conditions holds:

1. $X_1 \perp\!\!\!\perp X_2 \mid X_3$
2. *For each $x_3 \in [r_3]$, the matrix $(p_{x_1, x_2, x_3})_{x_1, x_2}$ has rank one.*
3. $p_{x_1 x_2 x_3} p_{x_1' x_2' x_3} = p_{x_1 x_2' x_3} p_{x_1' x_2 x_3}$ *for all $x_1, x_1' \in [r_1], x_2, x_2' \in [r_2], x_3 \in [r_3]$.*

Conditional independence models

Consider n random variables $X_1, \ldots, X_n$, taking values in the finite sets $[r_1], \ldots, [r_n]$. For any $A \subseteq [n]$ let X_A be the random vector $(X_i)_{i \in A}$. For disjoint subsets $A, B, C \subset [n]$, a CI statements has the form

$$X_A \perp\!\!\!\perp X_B \mid X_C, \qquad \text{or in short:} \quad A \perp\!\!\!\perp B \mid C$$

Consider the joint distribution P of $X_1, \ldots, X_n$ as an n-tensor $P = (p_{x_1,\ldots,x_n})_{x_i \in [r_i]}$. The statement $A \perp\!\!\!\perp B \mid C$ says:

1. Take the marginal over $[n] \setminus (A \cup B \cup C)$.
2. For any fixed value of X_C take the slice with constant $(x_k)_{k \in C}$.
3. Flatten this slice to a matrix, with rows indexed by $(x_i)_{i \in A}$, columns indexed by $(x_j)_{j \in B}$.

The resulting matrix has rank one.

An example

Short notation:
$X_A \perp\!\!\!\perp X_B \mid X_C \iff (p_{x_A x_B x_C +})_{x_A, x_B}$ has rank one for each x_C.

Example

Let $n = 5$ and $\mathcal{X}_i = \{0, 1\}$. The statement $\{1, 2\} \perp\!\!\!\perp \{3\} \mid \{4\}$ holds if and only if the two matrices

$$\begin{pmatrix} p_{00000} + p_{00001} & p_{00100} + p_{00101} \\ p_{01000} + p_{01001} & p_{01100} + p_{01101} \\ p_{10000} + p_{10001} & p_{10100} + p_{10101} \\ p_{11000} + p_{11001} & p_{11100} + p_{11101} \end{pmatrix}, \begin{pmatrix} p_{00010} + p_{00011} & p_{00110} + p_{00111} \\ p_{01010} + p_{01011} & p_{01110} + p_{01111} \\ p_{10010} + p_{10011} & p_{10110} + p_{10111} \\ p_{11010} + p_{11011} & p_{11110} + p_{11111} \end{pmatrix}$$

have rank one.

Saturated CI statements

Definition

A CI statement $A \perp\!\!\!\perp B \mid C$ is saturated, if it involves all random variables, i.e. $A \cup B \cup C = [n]$.

Observation: Saturated CI statements lead to binomial ideals.

- $R = \mathbb{C}[p_{x_1 x_2 \cdots x_n} : \ \ x_1 \in [r_1], \ldots, x_n \in [r_n]]$
- $A \perp\!\!\!\perp B \mid C \iff (p_{x_A x_B x_C +})_{x_A, x_B}$ has rank one for each x_C
- $I_{A \perp\!\!\!\perp B \mid C}$ is generated by all 2-minors of $(p_{x_A x_B x_C +})_{x_A, x_B}$

Theorem (Eisenbud, Sturmfels '96)

Binomial ideals have a binomial primary decomposition.

Question: Describe the implications among a collection of CI statements (preferably in terms of prime components of $I_{\mathcal{C}}$).

$$I_{\mathcal{C}} = \bigcap I_{\mathcal{C}_i}$$

Binomial edge ideals

- $\{X_0 \perp\!\!\!\perp X_1 \mid X_2, X_0 \perp\!\!\!\perp X_2 \mid X_1\}$ [Fink]
- $\{X_0 \perp\!\!\!\perp X_A \mid X_{[n]\setminus A} : \ \textit{various subsets} \ \ A\}$ [HHHKR, Ay-Rauh]
  - $|\mathcal{X}_0| = 2$: binomial edge ideals [HHHKR]
  - $|\mathcal{X}_0| > 2$: generalized binomial edge ideals [Ay-Rauh]
- $\{X_i \perp\!\!\!\perp X_j \mid X_{[n]\setminus\{i,j\}} : i < j\}$ [Swanson-Taylor]

Theorem [Fink]

Let $\{X_0 \perp\!\!\!\perp X_1 \mid X_2, X_0 \perp\!\!\!\perp X_2 \mid X_1\}$ with $\mathcal{X}_0 = [2]$, $\mathcal{X}_1 = [n_1]$, $\mathcal{X}_2 = [n_2]$. Let $P = (p_{x_0 x_1 x_2})$ be a vanishing point of $I_{\mathcal{C}}$. Then $I_{\mathcal{C}}$ is the binomial edge ideal of the bipartite graph G with vertex set $[n_1] \cup [n_2]$ and the edge set

$$\{(x_1, x_2) : \ p_{x_0 x_1 x_2} \neq 0 \textit{ for some } x_0\}.$$

Binomial edge ideals

Definition

Let G be a graph on the vertex set $[n]$ and $R = \mathbb{C}[p_{1x}, p_{2x} : x \in [n]]$. The binomial edge ideal $I_G \subset R$ is generated by the binomials

$$p_{1x}p_{2y} - p_{1y}p_{2x} \quad \textit{for all edges } \{x, y\} \textit{ of } G.$$

Known facts about binomial edge ideals:

- radical ideal
- nice description for Gröbner bases of I_G
- combinatorial description for primary decomposition.
- For $W \subset [n]$, let $m_W = \langle p_{1x}, p_{2x} : x \in W\rangle$ then

$$(I_G + m_W) : (\prod_{x \notin W} p_{1x}, p_{2x})^{\infty}$$

  is a prime component of I_G.

CI statements with hidden variables

- X_0, X_1: visible random variables, H: hidden random variable
- joint probability distribution: $P = (p_{i,x,h})_{i\in\mathcal{X}_0, x\in\mathcal{X}_1, h\in\mathcal{H}}$.
- $X_0 \perp\!\!\!\perp X_1 | H$ iff each slice $P_h := (p_{i,x})_{i\in\mathcal{X}_0, x\in\mathcal{X}_1}$ has rank one.
- The marginal distribution of X_0 and X_1 is $P^{X_0,X_1} = \sum_{h\in\mathcal{H}} P_h$.
- Therefore P^{X_0,X_1} has rank at most $|\mathcal{H}|$.

- Find all matrices $P = (p_{i,x})_{i\in\mathcal{X}_0, x\in\mathcal{X}_1}$ of non-negative rank at most $|\mathcal{H}|$ with the normalization condition $\sum_{i,x} p_{i,x} = 1$.

- The set of matrices of given non-negative rank at most r is a semi-algebraic set whose semi-algebraic condition is not known for general r. However, it is known that its Zariski closure equals the set of all rank r matrices, and it is described by

the determinantal ideal of all $(r+1)\times(r+1)$-minors of P.

CI statements with hidden variables

Question

Let $\mathcal{C} = \{ X_0 \perp\!\!\!\perp X_1 \,|\{X_2, H_1\} \,,\, X_0 \perp\!\!\!\perp X_2 \,|\{X_1, H_2\} \}$. Describe $I_{\mathcal{C}}$ and its primary decomposition combinatorially.

Theorem

Let $\mathcal{C} = \{ X_0 \perp\!\!\!\perp X_1 \,|\{X_2, H_1\} \,,\, X_0 \perp\!\!\!\perp X_2 \,|\{X_1, H_2\} \}$ with

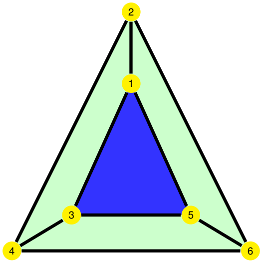
- $\mathcal{X}_0 = [d]$, $\mathcal{X}_1 = [n_1]$ and $\mathcal{X}_2 = [n_2]$
- $\mathcal{H}_1 = [r_1]$ and $\mathcal{H}_2 = [r_2]$
- $\Delta^{n_1,0} = \Big\{\{(i,1),(i,2),\dots,(i,n_2)\} : i \in [n_1]\Big\}$
- $\Delta^{0,n_2} = \Big\{\{(1,j),(2,j),\dots,(n_1,j)\} : j \in [n_2]\Big\}$.

Then $I_{\mathcal{C}} = I_{\Delta}$, where Δ is the union of the r_1-skeleton of $\Delta^{n_1,0}$ and the r_2-skeleton of Δ^{0,n_2}, and all its prime components can be read from subcomplexes of Δ.

Example of CI statements with hidden variables

- $\mathcal{C} = \{ X_0 \perp\!\!\!\perp X_1 \,|\{X_2, H_1\} \,,\, X_0 \perp\!\!\!\perp X_2 \,|\{X_1, H_2\} \}$
- $|\mathcal{X}_0| = 3$, $|\mathcal{X}_1| = 2$, $|\mathcal{X}_2| = 3$, $|\mathcal{H}_1| = 3$ and $|\mathcal{H}_2| = 2$.

$$\Delta = \{135, 246, 12, 34, 56\}$$



$$P_{|\mathcal{X}_0|\times|\mathcal{X}_1||\mathcal{X}_2|} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{16} \\ p_{21} & p_{22} & \cdots & p_{26} \\ p_{31} & p_{32} & \cdots & p_{36} \end{pmatrix}$$

The ideal $I_{\mathcal{C}}$ and its prime components

We take all maximal minors of the submatrices of P corresponding to $\Delta = \{135, 246, 12, 34, 56\}$:

$$\begin{pmatrix} p_{11} & p_{13} & p_{15} \\ p_{21} & p_{23} & p_{25} \\ p_{31} & p_{33} & p_{35} \end{pmatrix}, \begin{pmatrix} p_{12} & p_{14} & p_{16} \\ p_{22} & p_{24} & p_{26} \\ p_{32} & p_{34} & p_{36} \end{pmatrix}, \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \\ p_{31} & p_{32} \end{pmatrix}, \begin{pmatrix} p_{13} & p_{14} \\ p_{23} & p_{24} \\ p_{33} & p_{34} \end{pmatrix}, \begin{pmatrix} p_{15} & p_{16} \\ p_{25} & p_{26} \\ p_{35} & p_{36} \end{pmatrix}$$

Then $I_{\mathcal{C}} = I_{\Delta}$ has seven minimal primes associated to the complexes:

$$\begin{aligned} \Delta_{1,4} &= \{1, 4, 56\}, & \Delta_{1,6} &= \{1, 6, 34\}, \\ \Delta_{2,3} &= \{2, 3, 56\}, & \Delta_{2,5} &= \{2, 5, 34\}, \\ \Delta_{3,6} &= \{3, 6, 12\}, & \Delta_{4,5} &= \{4, 5, 12\}, \end{aligned}$$

$$\Delta_0 = \{12, 34, 56, 135, 145, 136, 146, 235, 245, 236, 246\}.$$

CI statements with hidden variables

- X_0, $X_1, \dots, X_k$: visible random variables
- $H_1, \dots, H_l$: hidden random variables
- $\mathcal{C}$: a family of CI statements of the form $X_0 \perp\!\!\!\perp X_A \mid X_B$, where $A, B \subseteq \{X_1, \dots, X_k, H_1, \dots, H_l\}$ are disjoint

We are interested in the set $P_{\mathcal{C}}$ of marginal distributions of $X_0, X_1, \dots, X_k$ of the set of those joint distributions of $X_0, X_1, \dots, X_k, H_1, \dots, H_l$ that satisfy the statements in $\mathcal{C}$.

Question

Whether CI statements with hidden variables can be given an algebraic interpretation? What can we say about the ideal $I_{\mathcal{C}}$?

- Is $I_{\mathcal{C}}$ a radical ideal?
- Give a nice combinatorial primary decomposition for $I_{\mathcal{C}}$.
- Describe a Gröbner basis.

Prime components of $I_{\mathcal{C}}$

Theorem [M.-Rauh]

The minimal primes of $I_{\mathcal{C}}$ are of the form

$$(I_{\mathcal{C}} + m_W) : (\prod_{x \notin W} p_{1x}, p_{2x}, \ldots, p_{dx})^{\infty}$$

where m_W is the ideal associated to a subcomplex of Δ.

Theorem [EHHM 2013], [M. 2012], [M.-Rauh]

The ideal $I_{\mathcal{C}}$ and all its prime components can be read from a simplicial complex associated to $\mathcal{C}$, i.e. these ideals are all determinantal facet ideals studied in [EHHM].

We computed some class of examples, which all are very nice:

- radical ideal,
- nice combinatorial primary decomposition.

References

- Ene, Herzog, Hibi, and Mohammadi: Determinantal facet ideals (Michigan Mathematical Journal, 2013)
- Mohammadi: Prime splittings of determinantal ideals (arXiv:1208.2930, 2012)
- Mohammadi & Rauh: Conditional independence ideals with hidden variables (in preparation)

Thank you!