Ideals Associated to Bayesian Networks

(joint with Luis Garcia, Bernd Sturmfels)

math.AG/0301255

Independence statements

\( X_1, ..., X_n \) discrete random variables

\( X_i \) takes values in \([d_i] = \{1, ..., d_i\}\)

Joint probability distribution \( \Pr(X_1=u_1, ..., X_n=u_n) \)

Think of this as a point in \([0,1]^D \subseteq \mathbb{R}^D \subseteq \mathbb{C}^D\)

\( D = d_1 d_2 ... d_n \)

Let \( R = \mathbb{C}[P_{u_1 u_2 ... u_n} : u_i \in [d_i]] = \mathbb{C}[D] \)

Independence: given \( A, B, C \subseteq \{X_1, ..., X_n\} \) pairwise disjoint, \( A \) and \( B \) are independent given \( C \) (written \( A \perp B | C \))

Geiger-Meek-Sturmfels: translate this into a set of quadratic equations, an ideal

\( I_{A \perp B | C} \subseteq R \)

If \( M = \{ A_i \perp B_i | C_i : \text{some } i \} \), then define

\( I_M = \sum_i I_{A_i \perp B_i | C_i} \)
\[ V_{\geq 0}(I_M) \subseteq V(I_M) \subseteq \mathbb{C}[D] \]

↑ \hspace{1cm} \text{the intersection with the probability simplex}

Example: \( n=3, \) \( d_1 = d_2 = d_3 = 2 \)

8 variables indeterminates \( P_{ijk}, \) \( i,j,k \in \{1,2\} \)

\[ X_1 \perp X_2 \mid X_3 \quad \text{(or} \quad 1 \perp 2 \mid 3 \text{)} \]

\[ \varphi_1 = \begin{pmatrix} 1 & 2 \\ 2 & (P_{111} & P_{121}) \\ P_{211} & P_{221} \end{pmatrix} \quad \varphi_2 = \begin{pmatrix} P_{112} & P_{122} \\ P_{212} & P_{222} \end{pmatrix} \]

\( X_3 = 1 \)

\[ I_{1 \perp 2 \mid 3} = (\det \varphi_1, \det \varphi_2) \]

is a prime ideal.

Thm (Geiger–Meek–Sturmfels): \( I_{A \perp B \mid C} \) is prime, generated by quadrics, and has a quadratic GB.

\[ V_1 = P^1 \times P^1 \hookrightarrow P^3 \]
\[ V_2 = P^1 \times P^1 \hookrightarrow P^3 \]

\[ V_{1 \perp 2 \mid 3} = V(I_{1 \perp 2 \mid 3}) = \mathcal{J}(V_1, V_2) \]

↑ linear join

Stillman

\( \Box \)
Example 6

\[ X_2 \perp\!\!\!\!\!\!\!\!\!\!\perp X_3 \mid \varnothing \]

Let \( P_{tjk} = \sum_{i=1}^{d_1, d_2, d_3} P_{ijk} \)

\[ X_2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \]

Example 6

\[ I_{1 \perp 2 \mid 3, 2 \perp 3} = (\det \varphi_1, \det \varphi_2, \det \varphi) \]

complete \( \cap \), 3 components

Note: change coordinates to \( P_{tjk}, P_{2jk} \) and this becomes binomial.

d_1, d_2, d_3 \text{ arbitrary}
- radical ideal
- 2^{d_3} - 1 components

Bayesian network = \( G \), directed acyclic graph on vertices \( \{1, \ldots, n\} \)

Example:

\[ \begin{align*}
  & 4 \\
  & \downarrow \\
  & 3 \\
  \rightarrow & 2 \\
 \rightarrow & 1
\end{align*} \]

\[ \text{pa}(i) = \text{parents of } i \]

\[ \text{nd}(i) = \text{nondescendants of } i \] (other than parents)

Local Markov conditions:

\[ \text{local}(G) = \left\{ i \perp\!\!\!\!\!\!\!\!\!\!\perp \text{nd}(i) \mid \text{pa}(i) \quad : \quad \text{all vertices } i \right\} \]
In this graph: \( \text{local}(G) = \{1 \perp\!\!\!\!\!\!\perp 4 \mid \{2,3\}\} \)

\[
\text{global}(G) = \left\{ A \perp\!\!\!\!\!\!\perp B \mid C : A, B \text{ are d-separated by } C \right\} \\
A, B, C \text{ disjoint vertex sets}
\]

here: \( I_{\text{local}}(G) = I_{\text{global}}(G) \)

In general: \( I_{\text{local}}(G) \subseteq I_{\text{global}}(G) \)

\underline{Theorem:} If \( G \) is a directed forest (i.e., each node has at most one parent) then \( I_{\text{global}}(G) \) is prime (and has a quadratic GB, etc.)

Continuing with example:

\[
\Pr\left( X_4 = u_4, X_3 = u_3, \ldots \right) = \Pr(X_4 = u_4) \cdot \Pr\left( X_3 = u_3 \mid X_4 = u_4 \right) \cdot \Pr\left( X_2 = u_2 \mid X_3, X_4 \right) \cdot \Pr(X_1 = u_1 | X_4, X_3, X_2)
\]

\[
e.g., \Pr(X_4): 2\left[ \frac{1-a}{a} \right]
\]

\[
\Pr(X_3 \mid X_4) \begin{bmatrix} 1 \cdot b_1 & 1 \cdot b_2 \\ 2 \cdot b_1 & b_2 \end{bmatrix}
\]

\( x_4 \text{ is prime and has a quadratic GB, etc.)
\[ \text{Pr}(X_2 \mid X_4) = \frac{1}{2} \begin{bmatrix} 1 - c_1 & 1 - c_2 \\ c_1 & c_2 \end{bmatrix} \]

\[ \text{Pr}(X_1 \mid X_2, X_3) = \frac{1}{2} \begin{bmatrix} 1 - d_{11} & 1 - d_{12} & 1 - d_{21} & 1 - d_{22} \\ d_{11} & d_{12} & d_{21} & d_{22} \end{bmatrix} \]

(really a 2x2x2 tensor)

\[ p_{i\ldots} = (1 - a_i)(1 - b_i)(1 - c_i)(1 - d_i) \]

etc.

Geometrically:

\[ \mathbb{C}^E \rightarrow \mathbb{C}^D \]

\[ \mathbb{C}[D] \xrightarrow{\Phi} \mathbb{C}[E] \]

**Theorem:** For any \( G \), let

\[ p = \prod_{r \geq 1} \prod_{u_{r+1}, \ldots, u_n} p_{u_{r+1} \ldots u_n}. \]

Then

\[ I_{\text{local}}(G) : p^\infty = \ker \Phi \]

is a prime ideal.

\( I_{\text{global}}(G) : p^\infty \) not generally generated by quadrics.

**Conjecture:**

\[ (\ker \Phi)_{\text{deg} 2} = I_{\text{global}}(G) \]
5 binary random variables
302 Bayesian nets

221

Thm: 221 of these 302 have $I_{local}(6)$ prime
68: radical, not prime
13: not radical

Example:

has 207 minimal primes (all primary) and 34 embedded primes