M. Stillman

Ideals Associated to Bayesian Networks

(joint with Luis García, Bernd Sturmfels)

math.AG/0301255

Independence statements

$X_1, \dots, X_n$ discrete random variables

X_i takes values in $[d_i] = \{1, \dots, d_i\}$

Joint probability distribution $\Pr(X_1 = u_1, \dots, X_n = u_n)$
$u_i \in [d_i]$

Think of this as a point in $[0,1]^D \subseteq \mathbb{R}^D \subseteq \mathbb{C}^D$,
$D = d_1 d_2 \cdots d_n$

Let $R = \mathbb{C}[p_{u_1 u_2 \cdots u_n} : u_i \in [d_i]] = \mathbb{C}[D]$

Independence: given $A, B, C \subseteq \{X_1, \dots, X_n\}$ pairwise disjoint, A and B are <u>independent given C</u> (written $A \perp\!\!\!\perp B \mid C$)

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

$$I_{A \perp\!\!\!\perp B \mid C} \subseteq R.$$

If $M = \{A_i \perp\!\!\!\perp B_i \mid C_i \ : \text{some } i\}$, then define

$$I_M = \sum_i I_{A_i \perp\!\!\!\perp B_i \mid C_i} .$$

$$V_{\geq 0}(I_M) \subseteq V(I_M) \subseteq \mathbb{C}[D]$$

the intersection with the probability simplex

$\boxed{\text{Example (a)}}$ $n=3$, $d_1 = d_2 = d_3 = 2$

8 ~~variables~~ indeterminates

p_{ijk}, $i,j,k \in \{1,2\}$

$X_1 \perp\!\!\!\perp X_2 \mid X_3$ (or $1 \perp\!\!\!\perp 2 \mid 3$)

$$\varphi_1 = \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} p_{111} & p_{121} \\ p_{211} & p_{221} \end{pmatrix} \end{array} \qquad \varphi_2 = \begin{pmatrix} p_{112} & p_{122} \\ p_{212} & p_{222} \end{pmatrix}$$

$X_3 = 1$

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

is a prime ideal.

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

$$V_1 = \mathbb{P}^1 \times \mathbb{P}^1 \hookrightarrow \mathbb{P}^3$$
$$V_2 = \mathbb{P}^1 \times \mathbb{P}^1 \hookrightarrow \mathbb{P}^3$$

$$V_{1 \perp\!\!\!\perp 2 \mid 3} = V(I_{1 \perp\!\!\!\perp 2 \mid 3}) = J(V_1, V_2)$$

linear join

Stillman 2

$\boxed{\text{Example (b)}}$ $X_2 \perp\!\!\!\perp X_3 \,(\mid \emptyset)$

Let $P_{+jk} = \sum_{i=1}^{d_1} P_{ijk}$

$$\begin{array}{cc} & X_3 \\ X_2 & \begin{array}{cc} & \begin{array}{cc} \phantom{P_{+11}} & \phantom{P_{+12}} \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} P_{+11} & P_{+12} \\ P_{+21} & P_{+22} \end{pmatrix} \end{array} \end{array} = \varphi$$

$\boxed{\text{Example (c)}}$ $I_{1 \perp\!\!\!\perp 2 | 3,\ 2 \perp\!\!\!\perp 3} = (\det \varphi_1, \det \varphi_2, \det \varphi)$

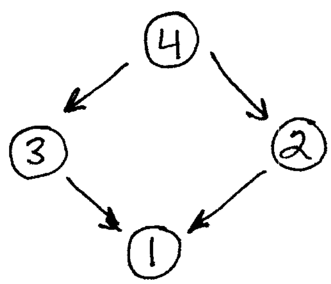
complete $\cap$, 3 components

Note: change coordinates to P_{+jk}, P_{2jk}
and this becomes binomial.

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

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

example:



$pa(i)$ = parents of i
$nd(i)$ = nondescendents of i (other than parents)

Local Markov conditions:

$$\text{local}(G) = \{ i \perp\!\!\!\perp nd(i) \mid pa(i) \ : \text{all vertices } i \}$$

Stillman 3

$$\begin{tikzcd} & 4 \arrow[dl] \arrow[dr] & \\ 3 \arrow[dr] & & 2 \arrow[dl] \\ & 1 & \end{tikzcd}$$

In this graph: $\quad \text{local}(G) = \left\{ \begin{array}{l} 1 \perp\!\!\!\perp 4 \mid \{2,3\} \\ 2 \perp\!\!\!\perp 3 \mid 4 \end{array} \right\}$

$$\text{global}(G) = \left\{ A \perp\!\!\!\perp B \mid C \;:\; A, B \text{ are d-separated by } C \right\}$$

A, B, C disjoint vertex sets

"d" stands for "directed"

here: $I_{local(G)} = I_{global(G)}$

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

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

Continuing with example:

$$\Pr(X_4 = u_4, X_3 = u_3, \ldots) = \Pr(X_4 = u_4)$$
$$\cdot \Pr(X_3 = u_3 \mid X_4 = u_4)$$
$$\cdot \Pr(X_2 = u_2 \mid X_4, X_3)$$
$$\cdot \Pr(X_1 = u_1 \mid X_4, X_3, X_2)$$

can delete

e.g., $\Pr(X_4)$: $\begin{array}{c|c} 1 & 1-a \\ 2 & a \end{array}$

$\Pr(X_3 \mid X_4)$

$$\begin{array}{c|cc} & X_4 = 1 & X_4 = 2 \\ \hline 1 & 1-b_1 & 1-b_2 \\ 2 & b_1 & b_2 \end{array}$$

$$
\begin{array}{c|cc}
 & X_3=1 & X_3=2 \\
\hline
\end{array}
$$

$$
\Pr(X_2 \mid X_4) = \begin{array}{c} 1 \\ 2 \end{array}\begin{bmatrix} 1-c_1 & 1-c_2 \\ c_1 & c_2 \end{bmatrix}
$$

$$
\begin{array}{cccc}
X_2=1 & X_2=1 & X_2=2 & X_2=2 \\
X_3=1 & X_3=2 & X_3=1 & X_3=2
\end{array}
$$

$$
\Pr(X_1 \mid X_2, X_3) = \begin{array}{c} 1 \\ 2 \end{array}\begin{bmatrix} 1-d_{11} & 1-d_{12} & 1-d_{21} & 1-d_{22} \\ d_{11} & d_{12} & d_{21} & d_{22} \end{bmatrix}
$$

(really a $2\times 2\times 2$ tensor)

$$
P_{1111} = (1-a_1)(1-b_1)(1-c_1)(1-d_1)
$$

etc.

Geometrically:

$$
\mathbb{C}^E \longrightarrow \mathbb{C}^D
$$

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

Theorem: For any G, let $p = \prod_{r \geq 1} \prod_{u_{r+1}, \dots, u_n} p_{\underbrace{+\cdots+}_{r} u_{r+1} \cdots u_n}$.

Then
$$
I_{local(G)} : p^\infty = I_{global(G)} : p^\infty = \ker \Phi \text{ is a prime ideal.}
$$

("canonical component")
not generally generated by quadrics.

Conjecture:

$$
\left(\ker \Phi\right)_{\deg 2} = I_{global(G)} .
$$

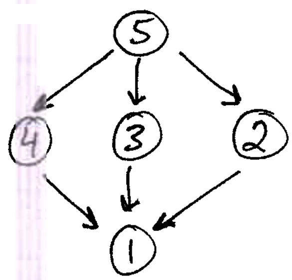
5 binary random variables

302 Bayesian nets

Thm: 221 of these 302 have $I_{global(G)}$ prime

68 : radical, not prime

13 : not radical

e.g.: 

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