Varieties of Bayesian Networks on Three Observable Variables and One Hidden Variable

Luis David Garcia

lgarcia@math.vt.edu

Virginia Tech.

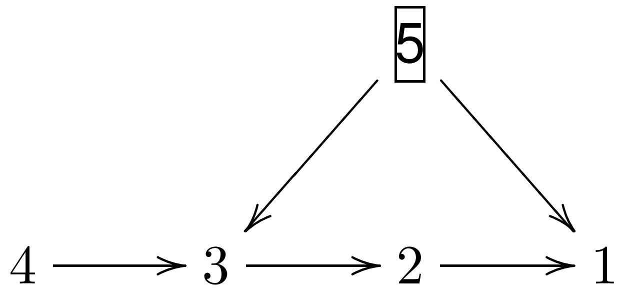
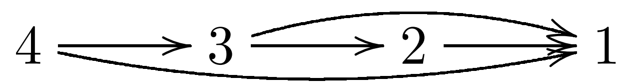
- Bayesian networks can be described implicitly, by a set of independence constraints that the distributions associated with the model must satisfy, or parametrically, by an explicit mapping of a set of parameters to the set of distributions.
- Bayesian networks with hidden variables are usually defined parametrically because the independent constraints on the distribution over the observable variables are not easily established.

> Find the independent constraints on the distributions over the observable variables implied by a Bayesian network with hidden variables.

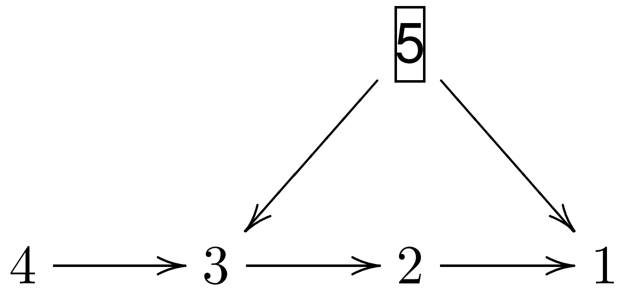
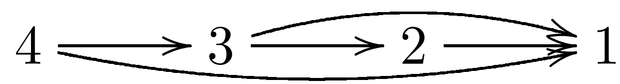
Model Selection

- A Bayesian approach to model selection is to compute $p(\mathrm{Data} \mid \mathrm{Model})$ via integration over all possible parameter values with which the model is compatible and to select a model that maximizes this probability.
- An asymptotic formula for the marginal likelihood known as the Bayesian Information Criteria (BIC) can sometimes be applied.
- Open question to prove the validity of the BIC for selecting models among Bayesian networks with hidden variables.
- Since the independence constraints on the distribution of the observable variables usually vary from one model to another, they can be used to distinguished between models.
- Since these constraints are over the observable variables, their fit to data can be measured directly with some statistical tests.

The P-structure [Geiger-Meek]

	
$2 \perp\!\!\!\perp 4 \mid 3$	
$I_{2 \perp\!\!\!\perp 4 \mid 3} +$ Verma constraints	$I_{2 \perp\!\!\!\perp 4 \mid 3}$
	$\begin{pmatrix} p_{+1k1} & p_{+1k2} \\ p_{+2k1} & p_{+2k2} \end{pmatrix}_{k=1,2}$
No singularities Use BIC to select between the models	

The P-structure [Geiger-Meek]

	
$2 \perp\!\!\!\perp 4 \mid 3$	
$I_{2 \perp\!\!\!\perp 4\|3}$ + Verma constraints	$I_{2 \perp\!\!\!\perp 4\|3}$
	$\begin{pmatrix} p_{+1k1} & p_{+1k2} \\ p_{+2k1} & p_{+2k2} \end{pmatrix}_{k=1,2}$
No singularities Use BIC to select between the models	

Unlike the constraint for the naive Bayes model $2 \longleftarrow \boxed{3} \longrightarrow 1$, the constraints generated by the implicitation procedure for other naive Bayes models did not seem to exhibit such a clear syntactic structure.

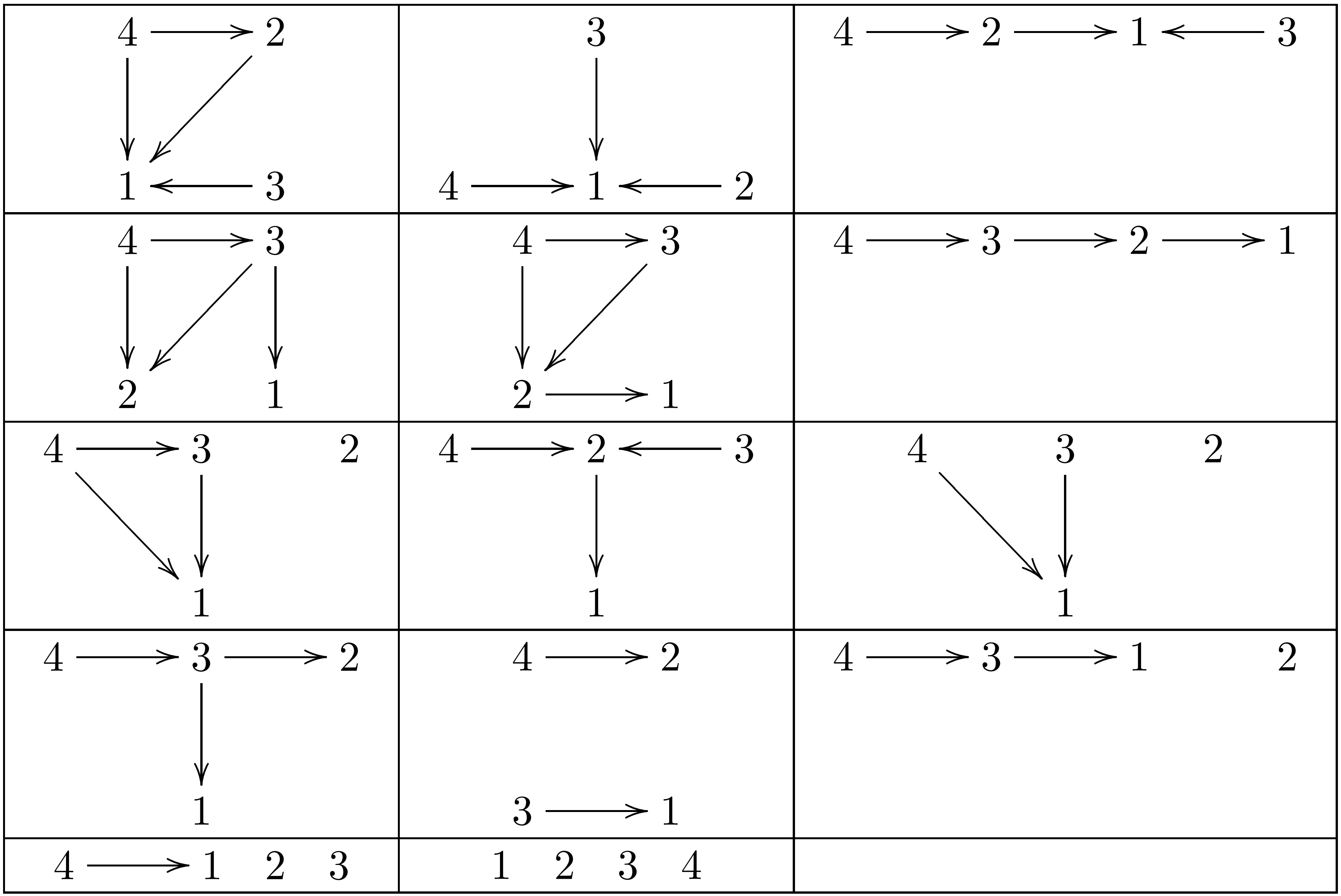
Proposition 19 [Garcia-Stillman-Sturmfels]

- Let G be a Bayesian network on n discrete random variables, where the nodes $r+1, \ldots, n$ correspond to hidden variables.
- Let $\mathbb{R}[D]$ be the ring generated by $p_{i_1 i_2 \cdots i_n}$.
- Let $\mathbb{R}[D']$ be the subring generated by $p_{i_1 \cdots i_r + \cdots +}$.
- Let P_G be the prime ideal of all homogeneous polynomials which vanish on all distributions that factor according to G.

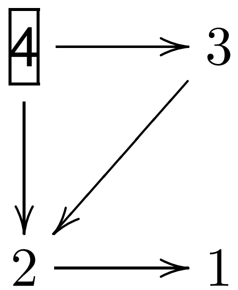
> The set of all polynomial constraints which vanish on the space of observable probability distributions is the prime ideal $Q_G = P_G \cap \mathbb{R}[D']$

Conjecture 21

The prime ideal Q_G of any naive Bayes model G with $r = 2$ classes is generated by the 3×3-subdeterminants of any two-dimensional table obtained by flattening the n-dimensional table $(p_{i_1 i_2 \cdots i_n})$.


4 2
1 3
3
4 1 2
4 2 1 3
4 3
2 1
4 3
2 1
4 3 2 1
4 3 2
1
4 2 3
1
4 3 2
1
4 3 2
1
4 2
3 1
4 3 1 2
4 1 2 3
1 2 3 4

Cuadratic Constraints (2)



$$P_{G_{14}} = I_{\text{local}(G_{14})} = I_{1 \perp\!\!\!\perp \{3,4\} | 2}$$

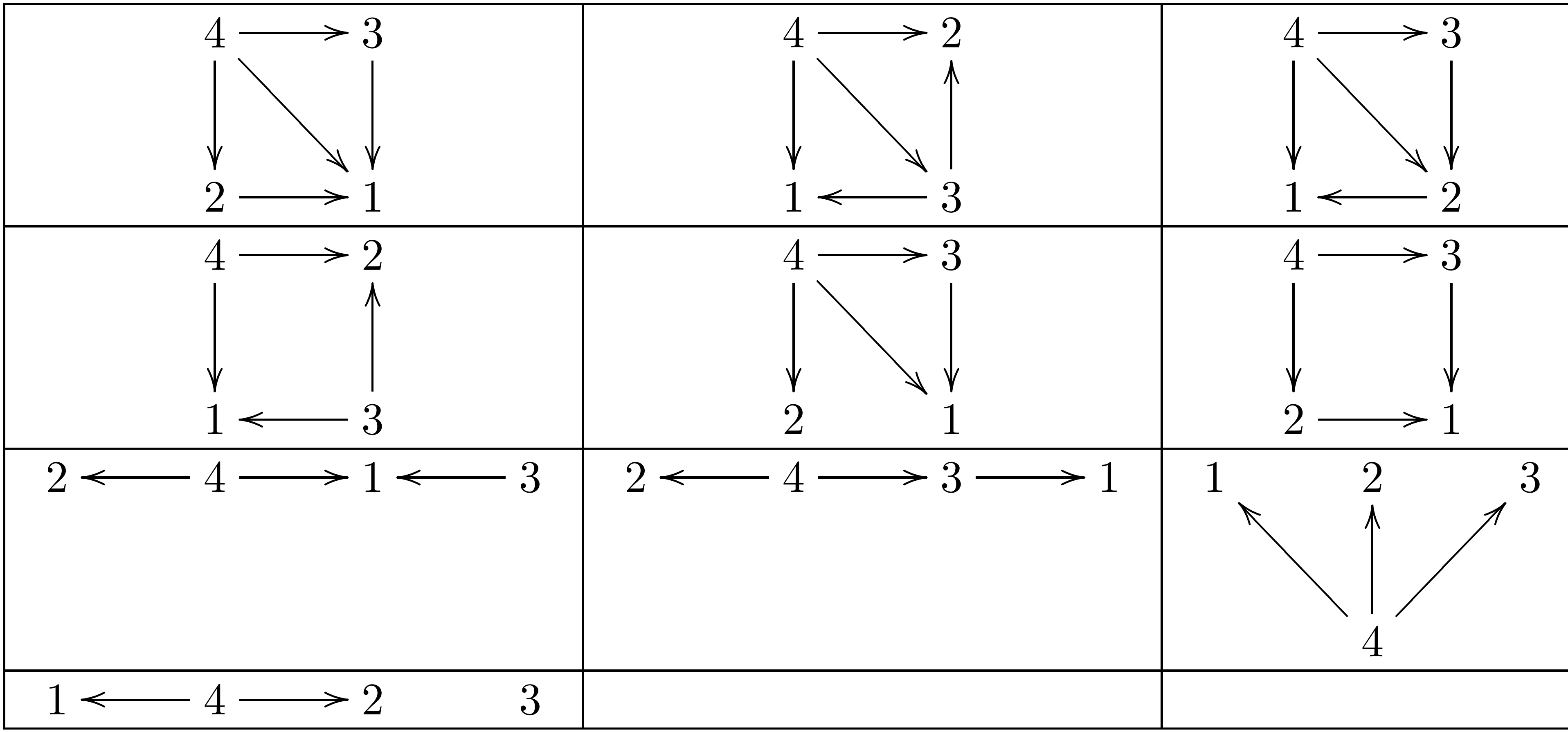
- Generated by the 2×2-minors of the d_2 tables

$$\begin{pmatrix} p_{1j11} & \cdots & p_{1j1d_4} & p_{1j21} & \cdots & p_{1jd_3d_4} \\ \vdots & & \vdots & \vdots & & \vdots \\ p_{d_1j11} & \cdots & p_{d_1j1d_4} & p_{d_1j21} & \cdots & p_{d_1jd_3d_4} \end{pmatrix}$$

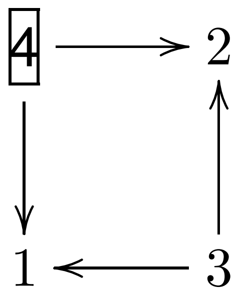
- Q_{14} is generated by the 2×2-minors of the d_2 tables

$$\begin{pmatrix} p_{1j1+} & p_{1j2+} & \cdots & p_{1jd_3+} \\ \vdots & \vdots & & \vdots \\ p_{d_1j1+} & p_{d_1j2+} & \cdots & p_{d_1jd_3+} \end{pmatrix}$$

Cubic Constraints (1)



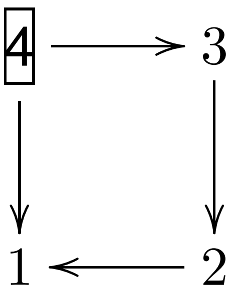
Cubic Constraints (2)



$$P_{G_9} = I_{\text{local}(G_9)} = I_{1 \perp\!\!\!\perp 2 | \{3,4\}} + I_{3 \perp\!\!\!\perp 4}$$

- $I_{3 \perp\!\!\!\perp 4} \cap \mathbb{C}[D'] = 0$.
- $Q_9 = I_{1 \perp\!\!\!\perp 2|\{3,4\}} \cap \mathbb{C}[D']$ if $d_4 = 2$.
- $I = I_{1 \perp\!\!\!\perp 2|\{3,4\}} = \sum_{k_0 \in [d_3]} I_{k_0}$, where I_{k_0} is the ideal generated by the 2×2-minors of the $2\ d_1 \times d_2$-matrices (p_{ijk_0l}).
- $V(I_{k_0} \cap \mathbb{C}[D']) = M_2$ where M_2 is the variety of $d_1 \times d_2$-matrices of rank at most 2, given by the 3×3-subdeterminants of the $d_1 \times d_2$-matrix (p_{ijk_0+}).

Sextic Constraints



- $I_{\mathrm{local}(G_{17})} = I_{\mathrm{global}(G_{17})} = I_{1 \perp\!\!\!\perp 3|\{2,4\}} + I_{2 \perp\!\!\!\perp 4|3}$
- This ideal is not radical in general, so $I_{\mathrm{local}(G_{17})} = P_{G_{17}} \cap L$.
- If $d_4 = 2$, $Q_8 = I_{\mathrm{local}(G_{17})} \cap \mathbb{C}[D'] = I_{1 \perp\!\!\!\perp 3|\{2,4\}} \cap \mathbb{C}[D']$.
- $V(Q_{17})$ is an irreducible proper subvariety of the irreducible variety $V(Q_8)$.

Conjecture $d_1 = d_4 = 2$ **(**$Q_8 = 0$**)** $d_2 = d_3 = 3$

- The ideal Q_{17} is generated by 3 irreducible sextics polynomials.
- Each $j_1, j_2 \in [d_2]$ specify two matrices N_{j_1} and N_{j_2}

$$\begin{pmatrix} p_{1j_11+} & p_{1j_12+} & p_{1j_13+} \\ p_{2j_11+} & p_{2j_12+} & p_{2j_13+} \end{pmatrix} \text{ and } \begin{pmatrix} p_{1j_21+} & p_{1j_22+} & p_{1j_23+} \\ p_{2j_21+} & p_{2j_22+} & p_{2j_23+} \end{pmatrix}$$

- The irreducible sextic polynomial is given by

$$p_{+j_11+}U_1V_1 - p_{+j_12+}U_2V_2 + p_{+j_13+}U_3V_3.$$

Where U_s is the determinant of the 2×2-minor of N_{j_1} obtained by eliminating the s-th column, and V_s is the determinant of the 2×2-matrix N'_{j_2} where the first column of N'_{j_2} equals the s-th column of N_{j_2} and the second column of N'_{j_2} is the product of the remaining two columns of N_{j_2}.