THE UNIVERSITY OF WARWICK

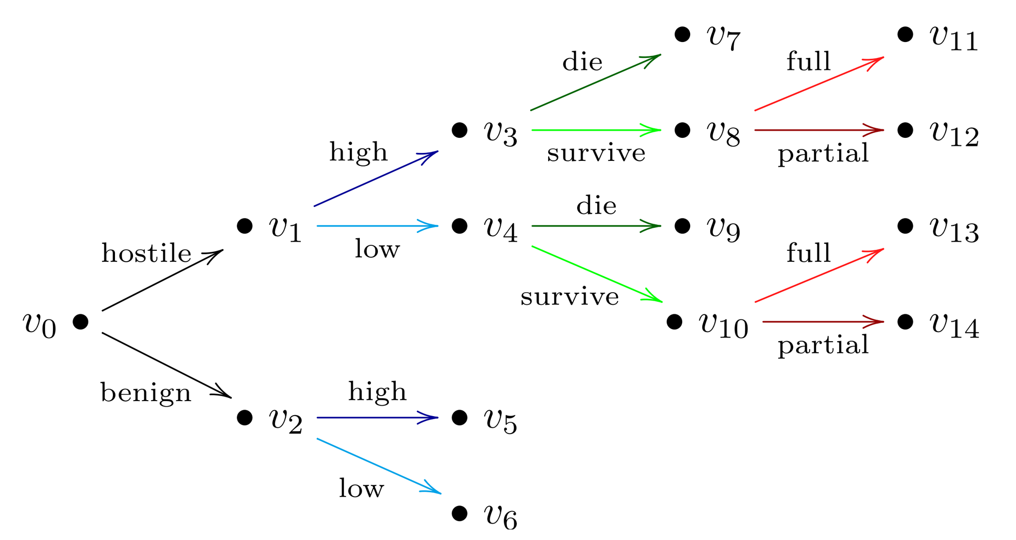
Algebraic Methods and Interpolating Polynomials for Chain Event Graphs

Christiane Görgen and Jim Q. Smith
Department of Statistics
University of Warwick, Coventry CV4 7AL, UK
c.gorgen@warwick.ac.uk j.q.smith@warwick.ac.uk

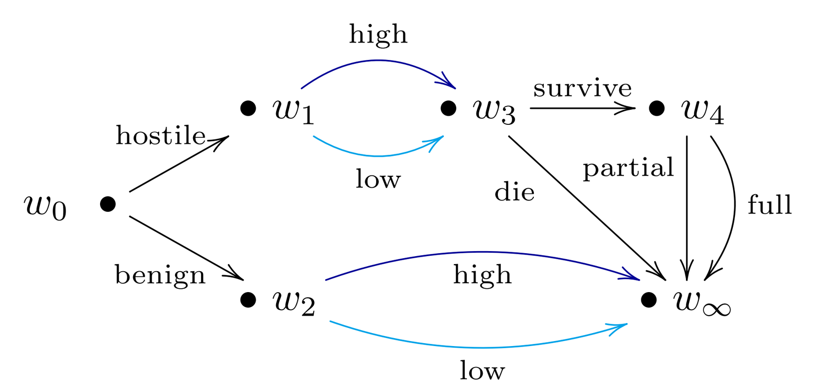
ABSTRACT: Chain Event Graphs are a recently developed graphical model based on trees (Smith and Anderson, 2008). I am currently fully characterising their statistical model equivalence classes. Graphical characterisations have proved challenging and rather inelegant (Thwaites and Smith, 2011), I therefore address these known difficulties using algebraic geometry. Here, I exploit some links between graphs and objects in polynomial rings and establish a useful vocabulary in terms of commutative algebra.

The Chain Event Graph model

Assume we are given a *staged* probabilitiy tree $\mathcal{T}$ describing the unfoldings of events in a cell culture



where colours indicate equality of edge probabilities. Vertices are in the same *position* if their emanating subtrees are isomorphic.
Construct the CEG $\mathcal{C}(\mathcal{T})$



with vertex set = set of positions, and edge set = edges between position representatives from the tree.

The characteristic polynomial

The *interpolating polynomial* of a CEG $\mathcal{C}$ is

$$c(\theta; k, \mathcal{C}) := \sum_{\lambda \in \Lambda(\mathcal{C})} k(\lambda)\pi(\lambda) = \sum_{\lambda \in \Lambda(\mathcal{C})} k(\lambda) \prod_{e \in E_\lambda} \theta(e)$$

where $\Lambda(\mathcal{C})$ is the set of root-to-sink paths, k is a complex function, and $\pi(\lambda)$ a path's associated probability, a monomial in the primitive (edge) probabilities Θ. This is in analogy to (Pistone, Riccomagno and Wynn, 2001).
If k is an indicator function we call c the *characteristic polynomial*.

Statistical equivalence

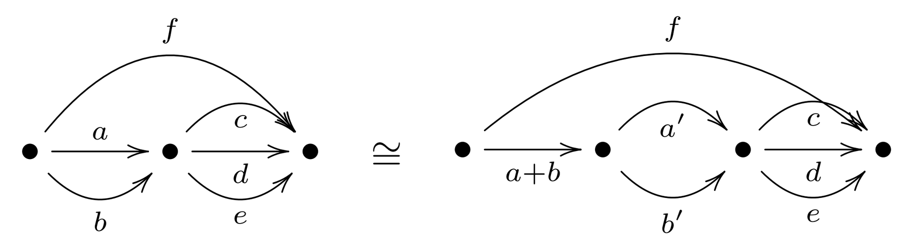
Two graphical statistical models are *equivalent* if they act as an index of the same class of probability distributions. A given CEG $\mathcal{C}$ is a parametric model in the primitive probabilities. Two such models are *statistically equivalent* if there is a bijection between their parameter sets.

We call two CEGs *algebraically equivalent* if they have the 'same' characteristic polynomial.

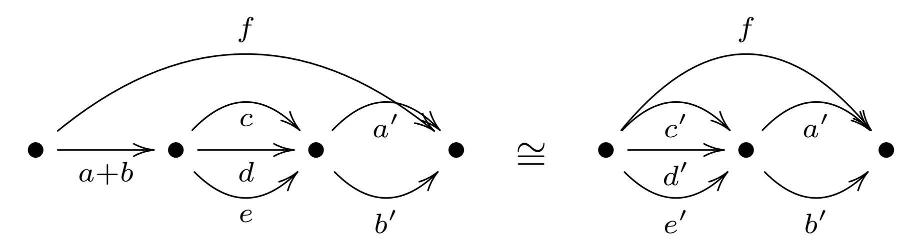
We note that for example the three Markov-equivalent Bayesian networks $X \to Z \to Y$, $X \leftarrow Z \leftarrow Y$ and $X \leftarrow Z \to Y$ give rise to equivalent Chain Event Graphs.

The model from an algebraic viewpoint

Consider the equivalent Chain Event Graphs



and, still equivalent,



where a, b, c, d, e, f and a', b', c', d', e' are primitive probabilities and the latter are functions of the former.
Thus, two equivalent CEGs can have quite different topologies, unlike Bayesian networks which must always share the same skeleton.

However, the characteristic polynomials of the first and last CEG above are

$$\begin{aligned} c(\cdot; \mathbb{1}_{\Lambda_1}, \mathcal{C}_1) &= ac + ad + ae + bc + bd + be + f \\ &= (a+b)(c+d+e) + f, \\ c(\cdot; \mathbb{1}_{\Lambda_2}, \mathcal{C}_2) &= c'a' + c'b' + d'a' + d'b' + e'a' + e'b' + f \\ &= (c' + d' + e')(a' + b') + f. \end{aligned}$$

We notice:

- The *floret* structure of the CEG (and its underlying tree) is recovered in the algebraic description via factorising terms.
- The two characteristic polynomials above are said to be *structurally equivalent*, that is they have the same number of summands and same degree structure and one is a reparametrisation of the other.

Minimal representations

The first and last CEG above are in *minimal form*, that is they are those representatives of their equivalence class which have the lowest number of vertices and edges:

If the tree underlying a CEG is saturated, we transform it into a *star*, a single vertex whose edges correspond to the former root-to-leaf paths.
Geometrically, events represented by a star correspond to projections on a subset of axes in a coordinate system spanned by the model's atomic probabilities.

In the case of staged trees, we retain the root, leaves and those florets belonging to staged vertices, while shortening all other subpaths.
Then events λ of the type 'a unit passes through w' are linked to monomial ideals via $\pi(\lambda) \in \langle \theta(w) \rangle$.

Theorems

We deduce the following statements:

- The characteristic polynomial associated to the minimal form of a Chain Event Graph is of *minimal degree* among all polynomials associated to equivalent CEG models.
- We can draw statistical inference from the characteristic polynomial alone. It
  1. shows us the length and number of paths (atoms) in the graphical model,
  2. lets us identify the root (up to independence) and identify the leaves,
  3. gives some information about the number of edges per floret and nesting (partial order), and
  4. provides information about *cuts* and hence about the conditional independence structure of the model.
- Two CEGs with the same characteristic polynomial are statistically equivalent. Two minimal CEGs with structurally different characteristic polynomials are not statistically equivalent.

Hence, this algebraic object encrypts the tree's topology and therefore most of the model information. What we deduce is a statement of the type 'algebraically equivalent = statistically equivalent'.

Future Work

We plan to establish a whole dictionary of interrelations between the characteristic polynomial's structure and our CEG model. These will then help to characterise the model equivalence class.

REFERENCES used in our work include the following articles:
Jim Q. Smith and Paul E. Anderson, *Conditional independence and Chain Event Graphs*, Journal of Artifical Intelligence 172, 42-68, 2008.
Peter A. Thwaites and Jim Q. Smith, *Separation theorems for Chain Event Graphs*, CRiSM 11-09, 2011.
Robert G. Cowell and Jim Q. Smith, *Causal discovery through MAP selection of stratified Chain Event Graphs*, Electronic Journal of Statistics 8, 965-997, 2014.
Giovanni Pistone and Eva Riccomagno and Henry P. Wynn, *Gröbner bases and factorisation in discrete probability and Bayes*, Statistics and Computing 11, 37-46, 2001.

Workshop on ALGEBRAIC STATISTICS, Prague Stochastics 2014, August 25-29, Prague, Czech Republic.
This research (with grant number A.STAA.1314.CXG) is funded by EPSRC Engineering and Physical Sciences Research Council