Staged tree models

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 Staged tree models generalise discrete Bayesian networks (BNs). They are particularly useful where BNs unnecessarily assume an underlying state space to have a product structure. Centrally, staged trees do not rely on an a priori set of problem variables and are particularly strong when a model is specified in terms of relationships between a collection of events. This poster gives an overview about recent developments in the algebraic analysis of these models.

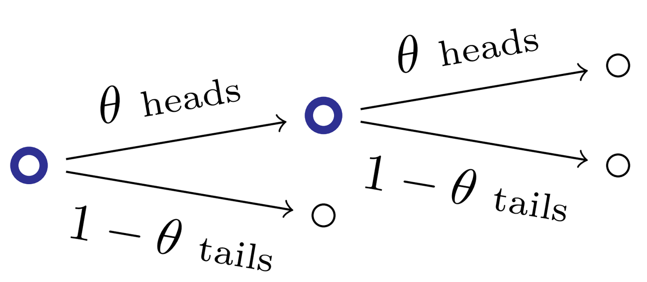
Staged tree models

A *probability tree* is an event tree graph $\mathcal{T} = (V, E)$ with edge labels $\theta(e) \in (0,1)$, $e \in E$. If $\sum_{e \in E(v)} \theta(e) = 1$ for all $v \in V$, the product

$$\pi_{\boldsymbol{\theta}}(\lambda) = \prod_{e \in E(\lambda)} \theta(e)$$

of labels along root-to-leaf paths λ defines a probability distribution on $\mathcal{T}$.

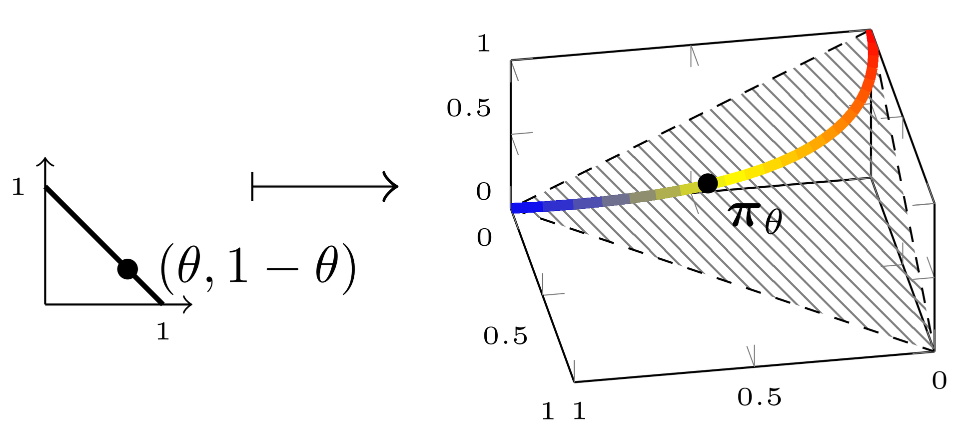
A *staged tree* is a probability tree where the emanating labels of certain vertices can be identified.



Every (staged) probability tree represents a discrete statistical model

$$\mathbb{P}_{\mathcal{T}} = \left\{ \left(\theta^2, \theta(1-\theta), 1-\theta\right) \mid \theta \in (0,1) \right\}$$

which can be parametrised based on its tree graph.

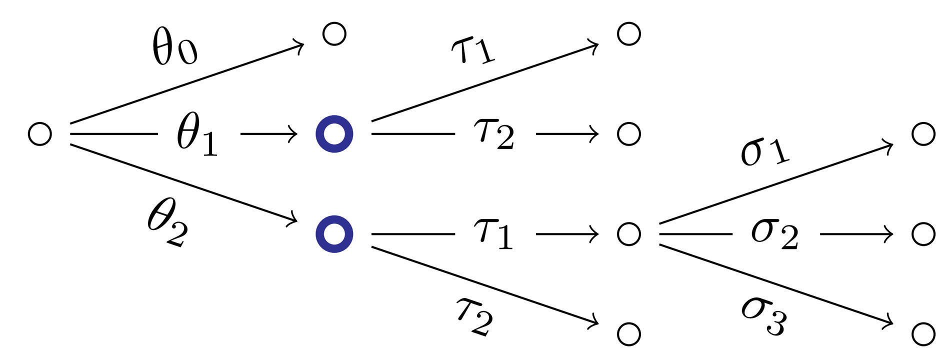


- Staged tree models include all discrete (and context-specific) BNs as a special case.
- They are always *faithful* and all atomic probabilities are *strictly positive*.
- (Staged) Tree graphs are efficient in modelling *asymmetric problems*.

Statistical equivalence

- Can we characterise which different staged trees represent the same model?
- Can we find operators to move around this class?

Probability trees are closely linked to an associated polynomial:



The sum of all atomic probabilities

$$\begin{aligned} c_{\mathcal{T}} &= \theta_0 + \theta_1\tau_1 + \theta_1\tau_2 \\ &\quad + \theta_2\tau_1\sigma_1 + \theta_2\tau_1\sigma_2 + \theta_2\tau_1\sigma_3 + \theta_2\tau_2 \end{aligned}$$

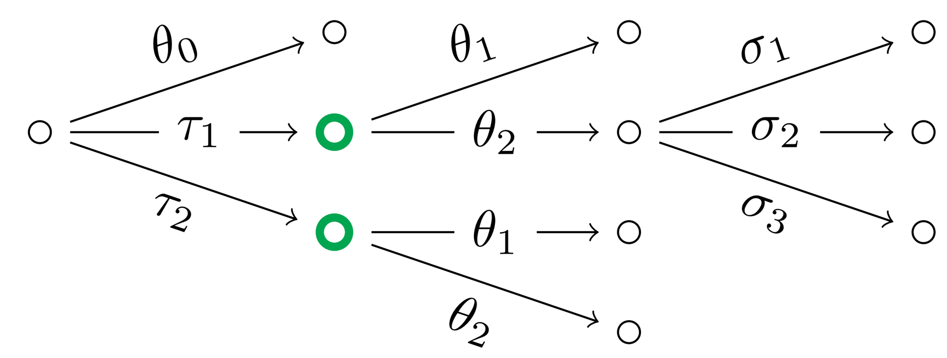
can be written in a bracketed form based on the graph

$$\begin{aligned} &= \theta_0 + \theta_1(\tau_1 + \tau_2) \\ &\quad + \theta_2(\tau_1(\sigma_1 + \sigma_2 + \sigma_3) + \tau_2) \end{aligned}$$

and vice versa a bracketed polynomial as above always belongs to a probability tree.

The swap operator

After projection, we may find factorisations $(\theta_1 + \theta_2)(\tau_1 + \tau_2)$ in the polynomial.

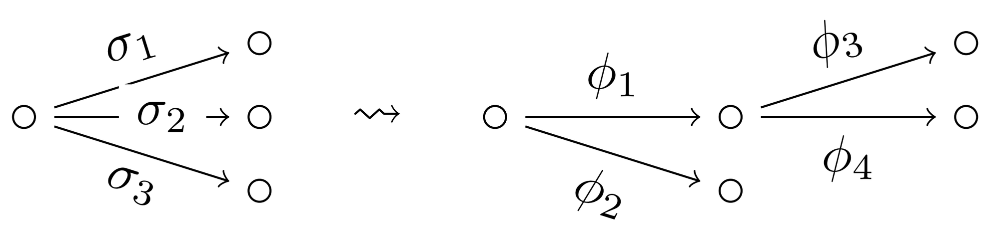


$$\begin{aligned} c_{\mathcal{S}} &= \theta_0 + \tau_1(\theta_1 + \theta_2(\sigma_1 + \sigma_2 + \sigma_3)) \\ &\quad + \tau_2(\theta_1 + \theta_2) \end{aligned}$$

So local independence statements can be reversed in order. The trees $\mathcal{T}$ and $\mathcal{S}$ are *polynomially* and statistically equivalent, $c_{\mathcal{T}} = c_{\mathcal{S}}$.

The resize operator

Reparametrisations (algebraic and graphical) are based on the laws of probability.



Here, $\phi_1 = \sigma_1 + \sigma_2$, $\phi_2 = \sigma_3$, $\phi_3 = \frac{\sigma_1}{\sigma_1 + \sigma_2}$ and $\phi_4 = \frac{\sigma_2}{\sigma_1 + \sigma_2}$.

Causality

- Causal *manipulations* can be performed as operations on the polynomial of a staged tree: in form of projections and differentiations.
- Causal *discovery* is closely linked with the swap operator. We can classify all orders of events depicted within a statistical equivalence class and infer putative causal hypotheses.

Geometry

Staged tree models are solution sets of polynomial equations of the form

$$\Big(\sum_{j \in J} \pi_j\Big)\Big(\sum_{k \in K} \pi_k\Big) = \Big(\sum_{l \in L} \pi_l\Big)\Big(\sum_{i \in I} \pi_i\Big)$$

for some $\pi_t = \pi_{\boldsymbol{\theta}}(\lambda_t)$.

In the coin-toss model, with atomic probabilities $(\pi_1, \pi_2, \pi_3) = \left(\theta^2, \theta(1-\theta), 1-\theta\right)$,

$$\pi_1\pi_3 = (\pi_1 + \pi_2)\pi_1.$$

The model is thus a variety

$$\mathbb{P}_{\mathcal{T}} = V\left(\pi_1\pi_3 - (\pi_1 + \pi_2)\pi_1, \pi_1 + \pi_2 + \pi_3 - 1\right)$$

intersected with the positive cone $\mathbb{R}^3_{>0}$.

Conclusions

Algebraic methods for staged trees provide the tools to traverse the class of all statistically equivalent model representations. They further provide the key to a causal and geometric analysis of these models.
Our results can be generalised to discrete models which have a monomial parametrisation but no tree representation.

REFERENCES Jim Q. Smith and Paul E. Anderson. *Conditional independence and Chain Event Graphs*, Journal of Artifical Intelligence 172, 42–68, 2008.

Christiane Görgen and Jim Q. Smith, *Equivalence Classes of Chain Event Graph Models*, arXiv:1512.00209 [math.ST].

Peter A. Thwaites, Jim Q. Smith, and Eva Riccomagno. *Causal analysis with Chain Event Graphs*. Artificial Intelligence, 174(12-13):889–909, 2010.

Christiane Görgen, Manuele Leonelli and Jim Q. Smith. *A Differential Approach for Staged Trees*, Conference Proceedings of ECSQARU 2015.

ALGEBRAIC STATISTICS SUMMER SCHOOL, June 27–July 2, 2016. Aalto University, Finland.

Funded by EPSRC Engineering and Physical Sciences Research Council