Staged Trees and Equivalence Classes of Chain Event Graph Models

Christiane Görgen and Jim Q. Smith

Department of Statistics, University of Warwick, Coventry CV4 7AL, UK.

c.görgen@warwick.ac.uk, j.q.smith@warwick.ac.uk

ABSTRACT We are interested in characterising the class \([\mathcal{T}, \Theta_T]\) of statistically equivalent representations of a probability tree model \(\mathcal{P}(\mathcal{T}, \Theta_T)\). This problem is analogous to the one of finding one essential graph \(\tilde{D}\) which indexes a class \([D]\) of DAG representations for a BN model \(\mathcal{P}_D\). In the case of staged trees (and chain event graphs), no such graphical result is available. We thus present an alternative algebraic characterisation, showing that an interpolating polynomial \(c \in \Theta_T\) uniquely identifies \([\mathcal{T}, \Theta_T]\). This is important for computational reasons, in model selection and for causal discovery.

THE PROBABILITY TREE

Let the graph \(\mathcal{T} = (V, E)\) be an event tree and \(\Theta_T = \{\theta_v : v \in V\}\) a set of parameter vectors \(\theta_v = (\theta(e) | e \in E(v)) \in \Delta_{\#E(v)-1}\). Then,

\[ p_\theta(\lambda) = \prod_{e \in E(h)} \theta(e) \]

defines a strictly positive probability mass function on the root-to-leaf paths \(\lambda \in \Lambda(\mathcal{T})\). We call the pair \((\mathcal{T}, \Theta_T)\) a probability tree.

Such a labelled graph is a picture for a discrete parametric statistical model

\[ \mathcal{P}(\mathcal{T}, \Theta_T) = \{ p_\theta : \theta \in \Lambda(\mathcal{T}) \} \]

Denote by \([\mathcal{T}, \Theta_T]\) the set of possible representations of \(\mathcal{P}(\mathcal{T}, \Theta_T)\). We say that two trees \((\mathcal{T}_1, \Theta_{\mathcal{T}_1})\) and \((\mathcal{T}_2, \Theta_{\mathcal{T}_2})\) are statistically equivalent.

Two vertices \(v, w \in V\) are in the same stage if and only if their parameter vectors coincide, \(\theta_v = \theta_w\). We then write \(v \sim w\).

Think of this as a sort of conditional independence statement, as formalised in Thwaites & Smith (2015).

The stage structure of a tree \((\mathcal{T}, \Theta_T)\) may be captured in the stage ideal

\[ I_T = \sum_{v \sim w} | \theta_{v,i} - \theta_{w,i} | : i = 1, \ldots, \#(E(v)) \]

in a ring \(\mathbb{R}[\Theta_T]\). This object is geometrically very simple but not invariant across \([\mathcal{T}, \Theta_T]\).

• Finite, discrete BN models form a subclass of the class of staged probability trees.
• Staged trees are graphically more complex but much more expressive than BNs.
• These trees have the same algebraic properties as Chain Event Graphs, introduced in Smith & Anderson (2008).

CONCLUSIONS

Our paper analyses discrete and context-specific BNs as well as stratified CEGs as subclasses of staged tree models, and characterises their equivalence classes via interpolating polynomials. We developed an algorithm for the map \(c\), identifying corresponding graphs from a tree-compatible polynomial, and have examined computational aspects of our work in Leonelli et al. (2015). Future research will focus on causal interpretations of these results, based on Cowell & Smith (2015).

POLYNOMIALS IN TREE MODELS

We define an interpolating polynomial of the model \(\mathcal{P}(\mathcal{T}, \Theta_T)\) as

\[ c_T(\theta) = \sum_{\lambda \in \Lambda(\mathcal{T})} \tilde{s}(\lambda) p_\theta(\lambda), \]

where \(g = 1\) or an indicator function. Denote by \([\mathcal{T}, \Theta_T]^c \subseteq [\mathcal{T}, \Theta_T]\) the class of polynomially equivalent tree representations, sharing the same interpolating polynomial.

There is a bijective map \(c : \mathcal{P}(\mathcal{T}, \Theta_T) \to [\mathcal{T}, \Theta_T]^c\), identifying a tree representation from a tree-compatible polynomial \(c\). Interpolating polynomials are tree-compatible. In practice, \(c\) maps a certain order of summation to a corresponding graph, using \(\theta(c) \to c \in E\) (see below).

MAIN RESULTS

PROPOSITION For a finite and discrete BN model with decomposable DAG \(D\) and class of tree representations \([\mathcal{T}, \Theta_T]^c\), we find that \([\mathcal{T}, \Theta_T]^c = [\mathcal{T}, \Theta_T]_D\) for any clique-induced interpolating polynomial \(c\). As a consequence, one interpolating polynomial captures all tree representations of a decomposable BN and thus, in particular, all fatorisations according to \(D\).

THEOREM Let \((S, \Theta_S), (\mathcal{T}, \Theta_T) \in [\mathcal{T}, \Theta_T]\) be statistically equivalent staged trees with interpolating polynomials \(c_S \in \Theta_S\) and \(c_T \in \Theta_T\). Then there is a bijective map \(\Phi : \Theta_S \to \Theta_T\) such that \(c_S(\theta) \mapsto c_T(\theta)\).

Consider a BN with ( clique) parameterisation

\[ p_\theta(x) = \theta(x_1, x_2, x_3) \theta(x_2, x_3, x_4) \theta(x_3, x_4) \theta(x_4, x_5) \]

according to a decomposable DAG \(D\), shown on the right. Define a probability tree \((\mathcal{T}, \Theta_T) \in [\mathcal{T}, \Theta_T]_D\) with \(\theta_v(\lambda) = p(x)\) for all \(x \in X\).

The interpolating polynomial \(c_T(\theta)\) is given below. Each of the summations we see correspond (via the map \(c\) above) to a different tree representation in the class \([\mathcal{T}, \Theta_T]^c = [\mathcal{T}, \Theta_T]_D\). The interpolating polynomial allows us to traverse the whole class of tree representations of the BN, and stratified trees are the ones giving a recursive factorisation of \(p_\theta\) according to \(D\).

\[ c_T(\theta) = \sum_{\lambda \in \Lambda(\mathcal{T})} \tilde{s}(\lambda) p_\theta(\lambda) \]

REFERENCES

Christiane Görgen and Jim Q. Smith, Equivalence Classes of Chain Event Graph Models, in preparation.


ACADEMIC STATISTICS 2015, June 8–11, Dipartimento di Matematica, Universita degli Studi di Genova, Italy.

This research (with grant number EF/1205110/1) is funded by the EPSRC.