Geometry of Hidden-Visible Products of Statistical Models

Guido Montúfar
montufar@mis.mpg.de

Jason Morton
morton@math.psu.edu

ABSTRACT

Dimension of marginals of exponential families
- naive Bayes models / mixtures of products / Segre secants
- restricted Boltzmann machines (RBMs) / Hadamard products of Segre secants
- mixtures of interaction models
- exponential family harmonics

- tropicalization of hidden-visible Kronecker product models
- inference functions (slicings of polytopes by normal-fans)

HIDDEN-VISIBLE PRODUCT MODELS

$$\Delta_X := \left\{(p_v)_{v \in X} \in \mathbb{R}^X : p_v > 0, \sum_{v \in X} p_v = 1\right\}$$

$$\mathcal{E}_{A \otimes B} := \left\{\exp (\theta^T [A \otimes B] - \psi(\theta)) : \theta \in \mathbb{R}^d \right\} \subseteq \Delta_X \times \mathcal{F}$$

$$\mathcal{V}_{A \otimes B} := \left\{\sum_{y \in Y} \exp (\theta^T [A \otimes B] - \psi(\theta)) : \theta \in \mathbb{R}^d \right\} \subseteq \Delta_X$$

JACOBIAN RANK AND TROPICAL MODEL

Jacobian for large parameters

$$\max_{\theta} \text{rank} (J_{\mathcal{V}_{A \otimes B}(\theta)}) + 1 \geq \max_{\theta} \text{rank} (A_x \otimes B_{h_\theta(x)})$$

Inference function

$$h_\theta(x) := \arg \max_y p_{\theta}(y|x) = \arg \max_y \Theta A_x, B_{h_\theta(x)}$$

Tropical morphism

$$\Phi(x; \theta) = \theta^T A_{\theta} = \theta^T [A_x \otimes B_{h_\theta(x)}]$$

By Bieri-Groves

$$\dim(\mathcal{V}_{A \otimes B}) + 1 \geq \max_{\theta} \text{rank} A_{\theta}$$

Slicings

MIXTURES

Segre secants [1, 4]
Mixtures of hierarchical models

(1a) Segre secants [1, 4]
(1b) Mixtures of hierarchical models

PRODUCTS OF MIXTURES

RBM [3, 8]
Products of mixtures of hierarchical models

(2a) RBMs [3, 8]
(2b) Products of mixtures of hierarchical models

PRODUCTS OF RESTRICTED MIXTURES

Exp. family harmonium
- Exponential family harmonium
- Kronecker products of exponential families / toric varieties

$$A_{\theta} = \left(A_{\theta_1} \otimes \cdots \otimes A_{\theta_m}\right)$$

RESULTS

- Combinatorial classification of non-defective cases (extensive but still incomplete)
- Examples of defective cases
- Preprint will be available soon!

OPEN PROBLEMS

- Combinatorics of Kronecker product polytopes
- Markov bases of bipartite graphs (J. Rauh)
- Ranks of Kratochvill polytopes
- Block-wise Kronecker products
- Sub-model dimensions
- Non-defective mixture $$\rightarrow$$ non-defective restricted mixture
- Orthogonal mixed parametrizations
- Deep Models

[1] J. Draisma, A. G. Geramita, and A. Gimigliano. Recent varieties of $$\mathbb{F}^n \times \cdots \times \mathbb{F}^n$$ to monomial sets for $$n \geq 2$$, J. Algebraic Geometry, 18:577–577, 2008.