An Implicitization Challenge for Binary Factor Analysis

María Angélica Cueto[1] Josephine Yu[2]

[1]Department of Mathematics
University of California, Berkeley

[2]Department of Mathematics
Massachusetts Institute of Technology

MEGA'09

Outline

1. Algebraic Statistics: description of the model.
2. Geometry of the model: First Secants of Segre embeddings and Hadamard products.
3. Tropicalization of the model.
4. Main results.
5. Implicitization Task: build the Newton polytope.

The Statistical model $\mathcal{F}_{4,2}$

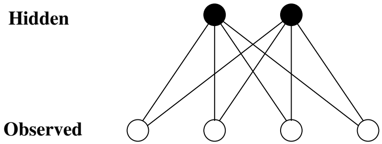


Figure: The undirected graphical model $\mathcal{F}_{4,2}$.

The set of all possible joint probability distributions (X_1, X_2, X_3, X_4) form an algebraic variety $\mathcal{M}$ inside Δ_{15} with expected codimension one and (multi)homogeneous defining equation f.

The Statistical model $\mathcal{F}_{4,2}$

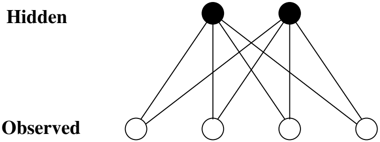


Figure: The undirected graphical model $\mathcal{F}_{4,2}$.

The set of all possible joint probability distributions (X_1, X_2, X_3, X_4) form an algebraic variety $\mathcal{M}$ inside Δ_{15} with expected codimension one and (multi)homogeneous defining equation f.

Problem

Find the degree and the defining polynomial/ Newton polytope of f of $\mathcal{M}$

Geometry of the model

Parameterization of the model: $p : \mathbb{R}^{32} \to \mathbb{R}^{16}$,

$$p_{ijkl} = \sum_{s=0}^{1} \sum_{r=0}^{1} a_{si} b_{sj} c_{sk} d_{sl} e_{ri} f_{rj} g_{rk} h_{rl} \text{ for all } (i,j,k,l) \in \{0,1\}^4.$$

Using homogeneity and the distributive law

$$p : (\mathbb{P}^1 \times \mathbb{P}^1)^8 \to \mathbb{P}^{15} \quad p_{ijkl} = (\sum_{s=0}^{1} a_{si} b_{sj} c_{sk} d_{sl}) \cdot (\sum_{r=0}^{1} e_{ri} f_{rj} g_{rk} h_{rl}).$$

So we have a coordinatewise product of two parameterizations of $\mathcal{F}_{4,1}$: the graphical model corresponding to the 4-claw tree with binary nodes.

Geometry of the model

Parameterization of the model: $p : \mathbb{R}^{32} \to \mathbb{R}^{16}$,

$$p_{ijkl} = \sum_{s=0}^{1} \sum_{r=0}^{1} a_{si} b_{sj} c_{sk} d_{sl} e_{ri} f_{rj} g_{rk} h_{rl} \text{ for all } (i,j,k,l) \in \{0,1\}^4.$$

Using homogeneity and the distributive law

$$p : (\mathbb{P}^1 \times \mathbb{P}^1)^8 \to \mathbb{P}^{15} \quad p_{ijkl} = (\sum_{s=0}^{1} a_{si} b_{sj} c_{sk} d_{sl}) \cdot (\sum_{r=0}^{1} e_{ri} f_{rj} g_{rk} h_{rl}).$$

So we have a coordinatewise product of two parameterizations of $\mathcal{F}_{4,1}$: the graphical model corresponding to the 4-claw tree with binary nodes. **But...**

Geometry of the model

Fact

1. *The binary 4-claw tree model is* $Sec^1(\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1) \subset \mathbb{P}^{15}$.
2. *Coordinatewise product of parameterizations corresponds to Hadamard products of algebraic varieties*

Definition

$X, Y \subset \mathbb{P}^n$, the **Hadamard product** of X and Y is

$$X \centerdot Y = \overline{\{(x_0 y_0 : \ldots : x_n y_n) \,|\, x \in C(X), y \in C(Y), x \centerdot y \neq 0\}} \subset \mathbb{P}^n,$$

Geometry of the model

Proposition

The algebraic variety of the model is $\mathcal{M} = X \star X$ *where* X *is the first secant variety of the Segre embedding* $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \hookrightarrow \mathbb{P}^{15}$.

Remark

The model is highly symmetric. Invariant under relabeling of observed nodes and by changing role of two states (0 or 1). Therefore, we have an action of the group $B_4 = \mathbb{S}_4 \ltimes (\mathbb{S}_2)^4$, *the group of symmetries of the 4-cube.*

Geometry of the model

Proposition

The algebraic variety of the model is $\mathcal{M} = X \mathbin{.} X$ *where* X *is the first secant variety of the Segre embedding* $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \hookrightarrow \mathbb{P}^{15}$.

Remark

The model is highly symmetric. Invariant under relabeling of observed nodes and by changing role of two states (0 or 1). Therefore, we have an action of the group $B_4 = \mathbb{S}_4 \ltimes (\mathbb{S}_2)^4$, *the group of symmetries of the 4-cube.*

Useful facts about X:

1. The ideal $I(X)$ is a well-studied object: it is the 9-dim *irreducible* subvariety of all $2 \times 2 \times 2 \times 2$-tensors of tensor rank at most 2.
2. Known set of generators for $I(X)$: 3×3-minors of all three 4×4-flattenings of these tensors.

Tropicalizing the model

Definition

For an algebraic variety $X \subset \mathbb{C}^n$ with defining ideal $I = I(X) \subset K[x_1, \ldots, x_n]$, the tropicalization of X or I is defined as:

$$\mathcal{T}(X) = \mathcal{T}(I) = \{w \in \mathbb{R}^{n+1} \,|\, \mathrm{in}_w(I) \text{ contains no monomial}\}$$

where $\mathrm{in}_w(I) = \langle \mathrm{in}_w(f) : f \in I \rangle$, and $\mathrm{in}_w(f)$ is the sum of all *nonzero* terms of $f = \sum_\alpha c_\alpha x^\alpha$ such that $\alpha \cdot w$ is **maximum**.

Tropicalizing the model

Definition

For an algebraic variety $X \subset \mathbb{C}^n$ with defining ideal $I = I(X) \subset K[x_1, \ldots, x_n]$, the tropicalization of X or I is defined as:

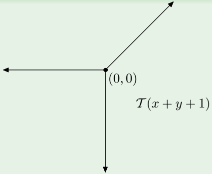
$$\mathcal{T}(X) = \mathcal{T}(I) = \{w \in \mathbb{R}^{n+1} \mid \mathrm{in}_w(I) \text{ contains no monomial}\}$$

where $\mathrm{in}_w(I) = \langle \mathrm{in}_w(f) : f \in I \rangle$, and $\mathrm{in}_w(f)$ is the sum of all *nonzero* terms of $f = \sum_\alpha c_\alpha x^\alpha$ such that $\alpha \cdot w$ is **maximum**.

Example

$L = (x + y + 1 = 0) \subset \mathbb{C}^2$

gives the well-known picture:



Tropicalizing the model

Remark

Basic features of $\mathcal{T}(X)$ for $X \subset \mathbb{P}^n$ with homogeneous ideal $I = I(X)$:

1. *It is a rational polyhedral subfan of the Gröbner fan of I.*
2. *If I is prime, then $\mathcal{T}(X)$ is pure of the same dimension as X (Bieri-Groves Thm) and it is connected in codimension one.*
3. *Maximal cones have canonical multiplicities attached to them. With these multiplicities, $\mathcal{T}(X)$ satisfies the balancing condition.*
4. *The lineality space of the fan $\mathcal{T}(X)$ is the set*

$$L = \{w \in \mathcal{T}(X) : in_w(I) = I\}.$$

   It describes action of the maximal torus acting on X (diagonal action by the lattice $L \cap \mathbb{Z}^{n+1}$.)

5. *Morphisms can be tropicalized and monomial maps have very nice tropicalizations.*

Tropicalizing the model

Theorem (S-T-Y)

Let $A \in \mathbb{Z}^{d\times n}$, *defining a monomial map* $\alpha : (\mathbb{C}^*)^n \to (\mathbb{C}^*)^d$ *and a canonical linear map* $A : \mathbb{R}^n \to \mathbb{R}^d$.
Let $V \subset (\mathbb{C}^*)^n$ *be a subvariety. Then*

$$\mathcal{T}(\alpha(V)) \;=\; A(\mathcal{T}(V)).$$

Moreover, if α *induces a generically finite morphism on* V*, we have an explicit formula to push-forward the multiplicities of* $\mathcal{T}(V)$ *to multiplicities of* $\mathcal{T}(\alpha(V))$.

Main results

In our case $\mathcal{M} = X \centerdot X = \alpha(X \times X)$ where α is the monomial map associated to matrix $(Id_{16} \mid Id_{16})$.

Main results

In our case $\mathcal{M} = X \,\mathbf{.}\, X = \alpha(X \times X)$ where α is the monomial map associated to matrix $(Id_{16} \mid Id_{16})$. In general...

Main results

In our case $\mathcal{M} = X \centerdot X = \alpha(X \times X)$ where α is the monomial map associated to matrix $(Id_{16} \mid Id_{16})$. In general...

Theorem (—, Yu)

Given $X, Y \subset \mathbb{P}^n$ two projective irreducible varieties none of which is contained in a proper coordinate hyperplane, we can consider the associated projective variety $X \centerdot Y \subset \mathbb{P}^n$. Then as **sets***:*

$$\mathcal{T}(X \centerdot Y) = \mathcal{T}(X) + \mathcal{T}(Y).$$

Main results

In our case $\mathcal{M} = X \centerdot X = \alpha(X \times X)$ where α is the monomial map associated to matrix $(Id_{16} \mid Id_{16})$. In general...

Theorem (—, Yu)

Given $X, Y \subset \mathbb{P}^n$ two projective irreducible varieties none of which is contained in a proper coordinate hyperplane, we can consider the associated projective variety $X \centerdot Y \subset \mathbb{P}^n$. Then as **sets***:*

$$\mathcal{T}(X \centerdot Y) = \mathcal{T}(X) + \mathcal{T}(Y).$$

$\mathcal{T}(X)$ can be computed with `Gfan`, so we know $\mathcal{T}(\mathcal{M})$ as a set! BUT we want more...

Main results

In our case $\mathcal{M} = X \,\boldsymbol{.}\, X = \alpha(X \times X)$ where α is the monomial map associated to matrix $(Id_{16} \mid Id_{16})$. In general...

Theorem (—, Yu)

Given $X, Y \subset \mathbb{P}^n$ two projective irreducible varieties none of which is contained in a proper coordinate hyperplane, we can consider the associated projective variety $X \,\boldsymbol{.}\, Y \subset \mathbb{P}^n$. Then as **sets***:*

$$\mathcal{T}(X \,\boldsymbol{.}\, Y) = \mathcal{T}(X) + \mathcal{T}(Y).$$

$\mathcal{T}(X)$ can be computed with `Gfan`, so we know $\mathcal{T}(\mathcal{M})$ as a set! BUT we want more...

We want to compute multiplicities at *regular points* of $\mathcal{T}(\mathcal{M})$.

Main results

In our case $\mathcal{M} = X \,.\, X = \alpha(X \times X)$ where α is the monomial map associated to matrix $(Id_{16} \mid Id_{16})$. In general...

Theorem (—, Yu)

Given $X, Y \subset \mathbb{P}^n$ two projective irreducible varieties none of which is contained in a proper coordinate hyperplane, we can consider the associated projective variety $X \,.\, Y \subset \mathbb{P}^n$. Then as **sets***:*

$$\mathcal{T}(X \,.\, Y) = \mathcal{T}(X) + \mathcal{T}(Y).$$

$\mathcal{T}(X)$ can be computed with `Gfan`, so we know $\mathcal{T}(\mathcal{M})$ as a set! BUT we want more...
We want to compute multiplicities at *regular points* of $\mathcal{T}(\mathcal{M})$.

Our map α is monomial BUT NOT generically finite but very close to being gen. finite. We generalize the previous theorem to obtain multiplicities in $\mathcal{T}(\mathcal{M})$...

Main results

$$\begin{array}{ccc} \mathbb{T}^n \supseteq V & \xrightarrow{\alpha} & W \subseteq \mathbb{T}^d \\ \pi \downarrow & & \downarrow \pi \\ V' = V/H & \xrightarrow{\bar{\alpha}} & W/\alpha(H). \end{array}$$

Theorem (—, Yu)

Let $V \subset (\mathbb{C}^*)^n$ be a subvariety with torus action given by a lattice L and take the quotient by this action $V' = V/H$. Then,

$$\mathcal{T}(\bar{\alpha}(V')) \;=\; A'(\mathcal{T}(V')).$$

Moreover, if $L' = A(L)$ is a primitive sublattice of $\mathbb{Z}^d$ and if $\bar{\alpha}$ induces a generically finite morphism on V', we have an explicit formula to push-forward the multiplicities of $\mathcal{T}(V)$ to $\mathcal{T}(\alpha(V))$.

Theorem (—, Yu)

Let $X, Y \subset \mathbb{C}^m$ be two irreducible varieties. Then

$$\mathcal{T}(X \times Y) = \mathcal{T}(X) \times \mathcal{T}(Y)$$

as weighted polyhedral complexes, with $m_{\sigma\times\tau} = m_\sigma m_\tau$ for maximal cones $\sigma \subset \mathcal{T}(X), \tau \subset \mathcal{T}(Y)$, and $\sigma \times \tau \subset \mathcal{T}(X \times Y)$.

The Newton polytope of the implicit equation

If $I = (f)$, we can recover the *Newton polytope of f* from $\mathcal{T}(I)$.

The Newton polytope of the implicit equation

If $I = (f)$, we can recover the *Newton polytope of f* from $\mathcal{T}(I)$. Why?

The Newton polytope of the implicit equation

If $I = (f)$, we can recover the *Newton polytope of f* from $\mathcal{T}(I)$. Why?

1. $\mathcal{T}(I)$ is the union of the codim 1 cones of the *normal fan of* $NP(f)$.
2. multiplicity of a maximal cone is the lattice length of the edge of $NP(f)$ normal to that cone.

The Newton polytope of the implicit equation

If $I = (f)$, we can recover the *Newton polytope of f* from $\mathcal{T}(I)$. Why?

1. $\mathcal{T}(I)$ is the union of the codim 1 cones of the *normal fan of* $NP(f)$.
2. multiplicity of a maximal cone is the lattice length of the edge of $NP(f)$ normal to that cone.

Theorem (D-F-S)

Suppose $w \in \mathbb{R}^n$ is a generic vector so that the ray $(w - \mathbb{R}_{>0}\, e_i)$ intersects $\mathcal{T}(f)$ only at regular points of $\mathcal{T}(f)$, for all i. Let $\mathcal{P}^w$ be the vertex of the polytope $\mathcal{P} = NP(f)$ that attains the maximum of $\{w \cdot x : x \in NP(f)\}$. Then the i^{th} coordinate of $\mathcal{P}^w$ equals

$$\mathcal{P}_i^w = \sum_v m_v \cdot |l_{v,i}|,$$

where the sum is taken over all points $v \in \mathcal{T}(f) \cap (w - \mathbb{R}_{>0} e_i)$, m_v is the multiplicity of v in $\mathcal{T}(f)$, and $l_{v,i}$ is the i^{th} coordinate of the primitive integral normal vector to $\mathcal{T}(f)$ at v.

The Newton polytope of the implicit equation

Theorem

The hypersurface $\mathcal{M}$ has multidegree $(110, 55, 55, 55, 55)$ with respect to the grading defined by the matrix

$$L = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

The Newton polytope of the implicit equation

Theorem

The hypersurface $\mathcal{M}$ *has multidegree* $(110, 55, 55, 55, 55)$ *with respect to the grading defined by the matrix*

$$L = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Bottleneck: Going through the list of all maximal cones supporting $\mathcal{T}(M)$ ($\sim 7\,000\,000$.)

The Newton polytope of the implicit equation

Theorem

The hypersurface $\mathcal{M}$ *has multidegree* $(110, 55, 55, 55, 55)$ *with respect to the grading defined by the matrix*

$$L = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Bottleneck: Going through the list of all maximal cones supporting $\mathcal{T}(M)$ ($\sim 7\,000\,000$.)

We can do better!

The Newton polytope of the implicit equation

Theorem

The hypersurface $\mathcal{M}$ *has multidegree* $(110, 55, 55, 55, 55)$ *with respect to the grading defined by the matrix*

$$L = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Bottleneck: Going through the list of all maximal cones supporting $\mathcal{T}(M)$ ($\sim 7\,000\,000$.)

We can do better!

IDEA: Shoot rays and walk along neighboring chambers.

Theorem

The hypersurface $\mathcal{M}$ has multidegree $(110, 55, 55, 55, 55)$ with respect to the grading defined by the matrix

$$L = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Bottleneck: Going through the list of all maximal cones supporting $\mathcal{T}(M)$ $(\sim 7\,000\,000.)$

We can do better!

IDEA: Shoot rays and walk along neighboring chambers.

Up to now, we have computed 1 155 072 vertices of $NP(f)$ (3 030 orbits.)

Thank you!!!