Moment varieties for mixtures of products

Yulia Alexandr (UC Berkeley)
joint work with Joe Kileel and Bernd Sturmfels

San Francisco State University
March 3, 2023

The (nonparametric) set-up

Consider n independent random variables $X_1, X_2, \ldots, X_n$ on the line $\mathbb{R}$.

Assumptions:

- ⋆ **No** assumptions about X_k, only that moments $\mu_{ki} = \mathbb{E}(X_k^i)$ exist.
- ⋆ The moments μ_{ki} are unknowns.
- ⋆ The only equations we require are $\mu_{k0} = 1$ for $k = 1, 2 \ldots, n$.

We consider a random variable X on $\mathbb{R}^n$ that is the product of these n arbitrary independent random variables on $\mathbb{R}$. By independence, we have

$$\mathbb{E}(X_1^{i_1} X_2^{i_2} \cdots X_n^{i_n}) \;=\; \mathbb{E}(X_1^{i_1}) \cdot \mathbb{E}(X_2^{i_2}) \cdots \mathbb{E}(X_n^{i_n}).$$

This leads us to the *moment variety* $\mathcal{M}_{n,d}$, which has parametrization

$$\boxed{m_{i_1 i_2 \cdots i_n} \;= \mu_{1i_1}\mu_{2i_2} \;\cdots\; \mu_{ni_n} \text{ where } i_1, i_2, \ldots, i_n \geq 0 \text{ and } i_1 + i_2 + \cdots + i_n = d.}$$

The image is a toric variety of dimension at most $nd - 1$ in $\mathbb{P}^{\binom{n+d-1}{d}-1}$.

Example

Consider $\mathcal{M}_{5,3}$ in $\mathbb{P}^{34}$. The solutions to $i_1 + i_2 + i_3 + i_4 + i_5 = 3$ can be grouped into three partitions: $\lambda = (1\ 1\ 1)$, $\lambda = (2\ 1)$, $\lambda = (3)$. Consider the following three toric varieties of dimensions $4, 8, 4$ respectively:

$$\mathcal{M}_{5,(111)} \subset \mathbb{P}^9 : m_{11100} = \mu_{11}\mu_{21}\mu_{31}, \ldots, m_{00111} = \mu_{31}\mu_{41}\mu_{51}.$$
$$\mathcal{M}_{5,(21)} \subset \mathbb{P}^{19} : m_{21000} = \mu_{12}\mu_{21}, m_{12000} = \mu_{11}\mu_{22}, \ldots, m_{00012} = \mu_{41}\mu_{52}.$$
$$\mathcal{M}_{5,(3)} = \mathbb{P}^4 : m_{30000} = \mu_{13}, m_{03000} = \mu_{23}, \ldots, m_{00003} = \mu_{53}.$$

Combining these parametrizations yields the original variety.

We will also study $\mathcal{M}_{n,d}$ under projections $\mathbb{P}^{\binom{n+d-1}{d}-1} \dashrightarrow \mathbb{P}^{|N_\lambda|-1}$ for any partition λ of d with $\leq n$ parts. We denote these toric varieties by $\mathcal{M}_{n,\lambda}$.

T♡ric combinatorics

We are interested in studying the toric varieties $\mathcal{M}_{n,d}$ and $\mathcal{M}_{n,\lambda}$.

Reminders:

- ⋆ Our monomial parametrizations are square-free.
- ⋆ With each our variety we associate a 0-1 *A-matrix* whose columns correspond to the monomials in the parametrization.
- ⋆ Rank(A) is one more than the dimension of the projective toric variety.
- ⋆ The associated polytope is the convex hull of the columns of A.

Familiar examples

♠ For any n, consider the partition $\lambda = (1^d) = (1\,1\ldots1)$ of $d < n$. The parametrization of $\mathcal{M}_{n,\lambda}$ is then $m_{i_1 i_2 \ldots i_n} \mapsto \mu_{i_{\ell_1}1}\mu_{i_{\ell_2}1}\cdots\mu_{i_{\ell_d}1}$. This is equivalent to

$$\mathbb{R}[m_{\ell_1\ell_2\ldots\ell_d} : 1 \leq \ell_1 < \ell_2 < \ldots < \ell_d \leq n] \to \mathbb{R}[\mu_1, \mu_2, \ldots, \mu_n]$$
$$m_{\ell_1\ell_2\ldots\ell_d} \mapsto \mu_{\ell_1}\mu_{\ell_2}\ldots\mu_{\ell_d}.$$

This is the parametrization of the *hypersimplex*

$$\Delta(n,d) = \mathrm{conv}\left\{ e_{\ell_1} + e_{\ell_2} + \cdots + e_{\ell_d} \ : \ 1 \leq \ell_1 < \ell_2 < \cdots < \ell_d \leq n \right\}$$

and $\mathcal{M}_{n,(1^d)}$ is the associated toric variety of dimension $n-1$ in $\mathbb{P}^{\binom{n}{d}-1}$. Its degree is the Eulerian number $A(n,d)$.

Familiar examples

- Consider the partition $\lambda = (n-1, n-2, \ldots, 2, 1)$. Then the moment coordinates $m_{i_1 i_2 \ldots i_n}$ for $\mathcal{M}_{n,\lambda}$ are indexed by the $n!$ permutations of $\{0, 1, 2, \ldots, n-1\}$.

Proposition (A., Kileel, Sturmfels)

The moment variety $\mathcal{M}_{n,\lambda}$ for the partition $\lambda = (n-1, n-2, \ldots, 2, 1)$ is the toric variety of the Birkhoff polytope, which lives in $\mathbb{P}^{n!-1}$ and has dimension $(n-1)^2$.

T♡ric results

Theorem (A., Kileel, Sturmfels)

The dimension of the moment variety $\mathcal{M}_{n,d}$ is $\min\left\{nd-1, \binom{n+d-1}{d}-1\right\}$.

Given a partition λ of length n, let let $k_0 \geq k_1 \geq \ldots \geq k_s$ be multiplicities of the distinct parts in λ. We define

$$\nu = (\underbrace{s,\ldots,s}_{k_s}, \underbrace{s-1\ldots,s-1}_{k_{s-1}},\ldots,\underbrace{1,\ldots,1}_{k_1},\underbrace{0\ldots,0}_{k_0}).$$

to be the *reduction* of λ. Here $s+1$ is the number of distinct parts of λ.

Ex: both $(8,5,5,4)$ and $(7,7,3,0)$ reduce to $\nu = (2,1,0,0)$, with $s=2$.
Ex: if $\lambda = (1^d)$, we recover the identification $\Delta(n,d)$ with $\Delta(n,n-d)$.

Theorem (A., Kileel, Sturmfels)

The moment variety $\mathcal{M}_{n,\lambda} = \mathcal{M}_{n,\nu}$ has dimension $(n-1)s$.

What about generators?

Example

Consider the variety $\mathcal{M}_{4,4}$ in $\mathbb{P}^{34}$. Its ideal is generated by 52 quadrics and 28 cubics. The subset of the generators which involves the twelve unknowns $m_{2110}, \ldots, m_{0112}$ does not suffice to cut out $\mathcal{M}_{4,(211)}$ in $\mathbb{P}^{11}$.

Theorem (A., Kileel, Sturmfels)

For any partition λ, the ideal of $\mathcal{M}_{n,\lambda}$ is generated by quadrics and cubics.

The ideals for $\mathcal{M}_{n,d}$ are more complicated. We conjecture that there does not exist a uniform degree bound for their generators.

Mixtures

Now we consider the mixtures of r copies of our toric models. Algebraically, these are the *secant varieties* $\sigma_r(\mathcal{M}_{n,d})$ and $\sigma_r(\mathcal{M}_{n,\lambda})$. The first is parametrized by

$$\boxed{m_{i_1 i_2 \cdots i_n} = \sum_{j=1}^{r} \mu_{1i_1}^{(j)} \mu_{2i_2}^{(j)} \cdots \mu_{ni_n}^{(j)} \text{ with } i_1, i_2, \ldots, i_n \geq 0 \text{ and } i_1 + i_2 + \cdots + i_n = d.}$$

These varieties are no longer toric! What can we say about their dimensions, degrees, generators?

- ◇ Consider the secant variety $\sigma_2(\mathcal{M}_{5,2})$. The parametrization is given as

$$m_{20000} = \mu_{12}^{(1)} + \mu_{12}^{(2)} \ , \ \ldots \ , \ m_{11000} = \mu_{11}^{(1)} \mu_{21}^{(1)} + \mu_{11}^{(2)} \mu_{21}^{(2)}, \ \ldots .$$

Example (continued)

Note $\mathcal{M}_{5,2} = \mathcal{M}_{5,(2)} \star \mathcal{M}_{5,(11)} = \mathbb{P}^4 \star \mathcal{M}_{5,(11)}$, since

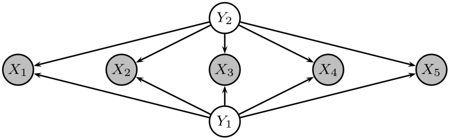
$$\begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} & m_{15} \\ m_{12} & m_{22} & m_{23} & m_{24} & m_{25} \\ m_{13} & m_{23} & m_{33} & m_{34} & m_{35} \\ m_{14} & m_{24} & m_{34} & m_{44} & m_{45} \\ m_{15} & m_{25} & m_{35} & m_{45} & m_{55} \end{bmatrix} = \begin{bmatrix} \mu_{12} & \mu_{11}\mu_{21} & \mu_{11}\mu_{31} & \mu_{11}\mu_{41} & \mu_{11}\mu_{51} \\ \mu_{11}\mu_{21} & \mu_{22} & \mu_{21}\mu_{31} & \mu_{21}\mu_{41} & \mu_{21}\mu_{51} \\ \mu_{11}\mu_{31} & \mu_{21}\mu_{31} & \mu_{32} & \mu_{31}\mu_{41} & \mu_{31}\mu_{51} \\ \mu_{11}\mu_{41} & \mu_{21}\mu_{41} & \mu_{31}\mu_{41} & \mu_{42} & \mu_{41}\mu_{51} \\ \mu_{11}\mu_{51} & \mu_{21}\mu_{51} & \mu_{31}\mu_{51} & \mu_{41}\mu_{51} & \mu_{52} \end{bmatrix}.$$

$$\sigma_2(\mathcal{M}_{5,2}) = \sigma_2\big(\mathbb{P}^4 \star \mathcal{M}_{5,(11)}\big) = \mathbb{P}^4 \star \boxed{\sigma_2(\mathcal{M}_{5,(11)})} \subset \mathbb{P}^4 \star \mathbb{P}^9 = \mathbb{P}^{14}.$$

The ideal of $\sigma_2(\mathcal{M}_{5,(11)})$ is principal, generated by the pentad

$$\begin{aligned} &m_{12}m_{13}m_{24}m_{35}m_{45} - m_{12}m_{13}m_{25}m_{34}m_{45} - m_{12}m_{14}m_{23}m_{35}m_{45} + m_{12}m_{14}m_{25}m_{34}m_{35} \\ &+ m_{12}m_{15}m_{23}m_{34}m_{45} - m_{12}m_{15}m_{24}m_{34}m_{35} + m_{13}m_{14}m_{23}m_{25}m_{45} - m_{13}m_{14}m_{24}m_{25}m_{35} \\ &- m_{13}m_{15}m_{23}m_{24}m_{45} + m_{13}m_{15}m_{24}m_{25}m_{34} + m_{14}m_{15}m_{23}m_{24}m_{35} - m_{14}m_{15}m_{23}m_{25}m_{34}. \end{aligned}$$

This is the factor analysis model $\mathbf{F}_{5,2}$.



Dimensions of mixtures

Proposition (A., Kileel, Sturmfels)

The dimension of the moment variety satisfies the upper bound

$$\dim(\sigma_r(\mathcal{M}_{n,d})) \;\leq\; \min\{\, rnd - rn + n - 1,\ \binom{n+d-1}{d} - 1 \,\}. \qquad (1)$$

Consider $\sigma_2(\mathcal{M}_{5,3}) = \sigma_2(\mathcal{M}_{5,(3)} \star \widetilde{\mathcal{M}}_{5,3}) = \mathbb{P}^4 \star \sigma_2(\widetilde{\mathcal{M}}_{5,3})$ in $\mathbb{P}^{34}$. Note:

- $\star$ we know $\dim(\widetilde{\mathcal{M}}_{5,3}) \leq (5 \cdot 3 - 5) - 1 = 9$,
- $\star$ so $\dim(\sigma_2(\mathcal{M}_{5,3})) \leq 4 + 1 + (2 \cdot 9 + 1) = 24$.

Theorem (A., Kileel, Sturmfels)

The dimension $\sigma_r(\mathcal{M}_{n,d})$ is bounded above by the optimal value of

$$\boxed{\text{maximize } c_1 + c_2 + \cdots + c_d - 1 \qquad \begin{array}{l} \text{subject to} \quad 0 \leq c_i \leq nr \quad \text{for } i \in [d] \\ \text{and } \sum_{i \in S} c_i \leq \sum_{\lambda \cap S \neq \varnothing} |N_\lambda| \quad \text{for } S \subseteq [d]. \end{array}}$$

The last sum ranges over partitions $\lambda \vdash d$ of length $\leq n$ having nonempty intersection with S.

Conjecture: this bound is tight for $d \geq 3$!

Implicitization

Solving the implicitization problem is difficult!

Remark: Our initial parametrization is not one-to-one. If ω is a primitive dth root of unity then we can replace μ_{ki} by $\mu_{ki}\,\omega^i$ without changing $m_{i_1 i_2 \cdots i_n}$. We parameterize to make degree computations in **Julia** faster.

Consider the variety $\mathcal{M}_{6,(111)}$. Its ideal is given by the 2×2 minors of

$$\begin{bmatrix} \star & \star & \star & \star & \star & m_{123} & m_{124} & m_{125} & m_{126} & m_{134} & m_{135} & m_{136} & m_{145} & m_{146} & m_{156} \\ \star & m_{123} & m_{124} & m_{125} & m_{126} & \star & \star & \star & \star & m_{234} & m_{235} & m_{236} & m_{245} & m_{246} & m_{256} \\ m_{123} & \star & m_{134} & m_{135} & m_{136} & \star & m_{234} & m_{235} & m_{236} & \star & \star & \star & m_{345} & m_{346} & m_{356} \\ m_{124} & m_{134} & \star & m_{145} & m_{146} & m_{234} & \star & m_{245} & m_{246} & \star & m_{345} & m_{346} & \star & \star & m_{456} \\ m_{125} & m_{135} & m_{145} & \star & m_{156} & m_{235} & m_{245} & \star & m_{256} & m_{345} & \star & m_{356} & \star & m_{456} & \star \\ m_{126} & m_{136} & m_{146} & m_{156} & \star & m_{236} & m_{246} & m_{256} & \star & m_{346} & m_{356} & \star & m_{456} & \star & \star \end{bmatrix}$$

The ideal of $\sigma_2(\mathcal{M}_{6,(111)})$ is generated by 20 cubics and 12 quintics. The ideal of $\sigma_3(\mathcal{M}_{6,(111)})$ has no quadrics or cubics, but contains a unique quartic. Computations in **Julia** reveal:

$$\deg(\sigma_2(\mathcal{M}_{6,(111)})) \;=\; 465 \quad \text{and} \quad \deg(\sigma_3(\mathcal{M}_{6,(111)})) \;=\; 80.$$

More implicitization

Proposition (A., Kileel, Sturmfels)

The variety $\sigma_2(\mathcal{M}_{5,3})$ *has dimension* 24 *and degree* 3225 *in* $\mathbb{P}^{34}$. *Its prime ideal is generated by* 313 *polynomials, namely* 10 *cubics,* 283 *quintics,* 10 *sextics and* 10 *septics. These are obtained by elimination from the ideal of* 3×3 *minors of the* 5×15 *matrix*

$$\begin{bmatrix} a_{23} & a_{24} & a_{25} & a_{34} & a_{35} & a_{45} & \star & \star & \star & \star & \star & b_{21} & b_{31} & b_{41} & b_{51} \\ a_{13} & a_{14} & a_{15} & \star & \star & \star & a_{34} & a_{35} & a_{45} & \star & b_{12} & \star & b_{32} & b_{42} & b_{52} \\ a_{12} & \star & \star & a_{14} & a_{15} & \star & a_{24} & a_{25} & \star & a_{45} & b_{13} & b_{23} & \star & b_{43} & b_{53} \\ \star & a_{12} & \star & a_{13} & \star & a_{15} & a_{23} & \star & a_{25} & a_{35} & b_{14} & b_{24} & b_{34} & \star & b_{54} \\ \star & \star & a_{12} & \star & a_{13} & a_{14} & \star & a_{23} & a_{24} & a_{34} & b_{15} & b_{25} & b_{35} & b_{45} & \star \end{bmatrix}.$$

Proposition (A., Kileel, Sturmfels)

The variety $\sigma_2(\mathcal{M}_{4,4})$ *has dimension* 27 *and degree* 8650 *in* $\mathbb{P}^{34}$. *Its prime ideal has only three minimal generators in degrees at most six.*

Finiteness up to symmetry

Our ideals satisfy natural inclusions

$$I(\sigma_r(\mathcal{M}_{n,\bullet})) \subset I(\sigma_r(\mathcal{M}_{n+1,\bullet})), \qquad \text{where} \quad \bullet \in \{d, \lambda\},$$

by appending a zero to the indices of every coordinate: $m_{i_1 i_2 \cdots i_n} \mapsto m_{i_1 i_2 \cdots i_n 0}$.
Iterate these inclusions and let the big symmetric group act:

$$\langle S_n I(\sigma_r(\mathcal{M}_{n_0,\bullet}))\rangle \subseteq I(\sigma_r(\mathcal{M}_{n,\bullet})) \quad \text{for } n > n_0.$$

Ideal-theoretic finiteness means $\exists n_0$ such that equality holds for $n > n_0$.

Theorem (A., Kileel, Sturmfels)

Given any partition $\lambda \vdash d$ and integer $r \geq 1$, set-theoretic finiteness holds for the varieties $\sigma_r(\mathcal{M}_{n,d})$ and $\sigma_r(\mathcal{M}_{n,\lambda})$. Ideal-theoretic finiteness holds in the toric case $r = 1$.

Example

The ideal of the variety $\mathcal{M}_{n,(1^d)}$ is generated by quadrics. The indices occurring in each quadratic binomial are 1 in at most $2d$ of the n coordinates. Therefore, ideal-theoretic finiteness holds with $n_0 = 2d$.

If $\lambda = (1\ 1)$, then $n_0 = 4$. Indeed:

$$I(\mathcal{M}_{4,\lambda}) = \langle m_{0101}m_{1010} - m_{0110}m_{1001}, m_{0011}m_{1100} - m_{0110}m_{1001}\rangle$$

$$\begin{aligned} I(\mathcal{M}_{5,\lambda}) = \langle & m_{01001}m_{10100} - m_{01100}m_{10001}, m_{00011}m_{10100} - m_{00110}m_{10001}, \\ & m_{11000}m_{00101} - m_{01100}m_{10001}, m_{10010}m_{00101} - m_{00110}m_{10001}, \\ & m_{10010}m_{01100} - m_{10100}m_{01010}, m_{00011}m_{01100} - m_{00101}m_{01010}, \\ & m_{00110}m_{11000} - m_{10100}m_{01010}, m_{00011}m_{11000} - m_{10001}m_{01010}, \\ & m_{01001}m_{10010} - m_{10001}m_{01010}, m_{00110}m_{01001} - m_{00101}m_{01010}\rangle \end{aligned}$$

Corollary

Fix a partition λ with e nonzero parts, and suppose that n increases. The toric varieties $\mathcal{M}_{n,\lambda}$ satisfy ideal-theoretic finiteness for some $n_0 \leq 3e$.

Thanks!