Stability and moment polytopes of tensors

Harold Nieuwboer

University of Copenhagen

June 12th, 2024

Tensor Ranks & Tensor Invariants

this slide left blank intentionally

Tools for (asymptotic) tensor (sub)rank

Let $T \in V := \mathbb{C}^d \otimes \mathbb{C}^d \otimes \mathbb{C}^d$ be a tensor.

What are the asymptotic rank $\tilde{R}(T)$ and asymp. subrank $\tilde{Q}(T)$?

Need tools to answer this!

Fulvio's talk: $\tilde{R}(T) = \lim_{k\to\infty} \sqrt[k]{R(T^{\boxtimes k})} = \lim_{k\to\infty} \sqrt[k]{\underline{R}(T^{\boxtimes k})}$,
where $\underline{R}$ is the **border rank**.
The set σ_r of tensors $T \in V$ with $\underline{R}(T) \le r$ is an **algebraic variety**.
Let $G = \mathrm{GL}_d \times \mathrm{GL}_d \times \mathrm{GL}_d$. Observe (if $d \ge r$):

$$\sigma_r = \overline{G \cdot \langle r \rangle}$$

where $\langle r \rangle = \sum_{i=1}^{r} e_i \otimes e_i \otimes e_i$ is the rank-r unit tensor.

Christian's talk: $\exists$ homogeneous polynomial F on $\mathbb{C}^4 \otimes \mathbb{C}^4 \otimes \mathbb{C}^4$ s.t.

- $F(T) = 0$ whenever $R(T) \leq 6$, so also whenever $\underline{R}(T) \leq 6$.
- $F(\mathrm{MM}_2) \neq 0$, where $\mathrm{MM}_2 = \langle 2, 2, 2 \rangle$.

Thereby proving $R(\mathrm{MM}_2) \geq \underline{R}(\mathrm{MM}_2) \geq 7$.
F was a **highest weight vector** labelled by the highest weight (λ, μ, ν) with $\lambda = \mu = \nu = (5, 5, 5, 5)$.

To prove lower bounds on $\underline{R}(T)$, can try to separate it from σ_r, i.e., find a homogeneous polynomial f such that $f(T) \neq 0$ and $f|_{\sigma_r} \equiv 0$.

F homogeneous of degree 20 in $4^3 = 64$ variables; very complicated to evaluate!

Without loss of generality: can assume $f \in \mathrm{HWV}_{\lambda,\mu,\nu} \subseteq \mathbb{C}[V]_n$ is a highest weight vector for some (λ, μ, ν).
To separate T from $\overline{G \cdot \langle r \rangle}$, or more generally tensor T from T':

Do there exist (λ, μ, ν) and $f \in \mathrm{HWV}_{\lambda,\mu,\nu}$ with $f|_{G \cdot T} \neq 0$ and $f|_{G \cdot T'} \equiv 0$?

Collect relevant (λ, μ, ν) for T into one object:

Define the **semigroup of representations** $S(T) \subseteq (\mathbb{Z}^d_{\geq 0})^3$ of T: those (λ, μ, ν) where $\lambda, \mu, \nu \vdash_d n$, and

$$\exists f \in \mathrm{HWV}_{\lambda,\mu,\nu} \subseteq \mathbb{C}[V]_n \text{ such that } f|_{G \cdot T} \neq 0.$$

Properties of semigroup of representations

Proposition (Monotonicity)

If $T \trianglerighteq T'$ (i.e., $T' \in \overline{G \cdot T}$) then $S(T) \supseteq S(T')$.

Proof.

If a highest weight vector does not vanish identically on $G \cdot T' \subseteq \overline{G \cdot T}$, then it does not vanish identically on $G \cdot T$ by continuity. □

Proposition (Semigroup property)

If $(\lambda, \mu, \nu), (\lambda', \mu', \nu') \in S(T)$, then $(\lambda + \lambda', \mu + \mu', \nu + \nu') \in S(T)$.

Proof.

If f, f' are highest weight vectors of the above types, then their product $f \cdot f'$ (as polynomials) is a highest weight vector of type $(\lambda + \lambda', \mu + \mu', \nu + \nu')$. It does not vanish on all of $G \cdot T$, since f and f' do not, and $\overline{G \cdot T}$ is irreducible. □

Moment polytope

$S(T)$ is nice, but still really hard to compute!
Define moment polytope:

$$\Delta(T) = \overline{\left\{\frac{1}{n}(\lambda, \mu, \nu) \in S(T) : n = |\lambda| = |\mu| = |\nu|\right\}} \subseteq \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d.$$

In other words, normalize the points in $S(T)$ and take the closure.
Why $\Delta(T)$ easier than $S(T)$?

Theorem

There exists a map $\mu \colon V \setminus \{0\} \to \mathrm{Herm}_d^3$ such that

$$\Delta(T) = \overline{\mathrm{spec}_{\searrow} \circ \mu(G \cdot T)}.$$

Here, μ is called the **moment map**, *and*

- *Herm_d consists of $d \times d$ Hermitian matrices,*
- *$\mathrm{spec}_{\searrow}(H)$ is the sorted list of eigenvalues $\lambda_1 \geq \ldots \geq \lambda_d$ of H.*

The moment map

How to compute $\mu(T) \in \mathrm{Herm}_d^3$?
First component: reinterpret $T \in U \otimes V \otimes W$ as a linear map $T_1 : U^* \to V \otimes W$, then compute

$$\mu_1(T) = \frac{1}{\|T\|^2} T_1^* T_1.$$

Similar for μ_2, μ_3.

Why is this an equivalent description?

Can also define a moment polytope with respect to $H = \mathrm{SL}_d \times \mathrm{SL}_d \times \mathrm{SL}_d$; then $\Delta^H(T) = \Delta(T) - \frac{1}{d}(\vec{1}, \vec{1}, \vec{1})$ is obtained by shifting.

Equivalence of the descriptions

Special case: $0 \in \Delta^H(X)$.

Theorem (Kempf–Ness)

Let $x = [v] \in \mathbb{P}(V)$. Then the following are equivalent:

- $0 \in \mu(\overline{H \cdot x})$.
- $F_x \colon H \to \mathbb{R}$, $F_x(h) = \log\|h \cdot v\|/\|v\|$ *is bounded from below.*
- $0 \notin \overline{H \cdot v} \subseteq V$.

Key idea: $F_x \colon \mathrm{SU}(d)^3 \backslash H \to \mathbb{R}$ is a **geodesically convex** function on $\mathrm{SU}(d)^3 \backslash H$, a symmetric space of non-positive curvature. Therefore bounded below iff gradient vanishes (asymptotically).

By Mumford's theorem, the third condition is equivalent to "There exists some homogeneous invariant $f \in \mathbb{C}[V]^H$ such that $f(x) \neq 0$, i.e., x is semistable."

i.e., $\mathbb{C}[V]_n$ contains the trivial H-representation for some $n \geq 1$.

Equivalence of the descriptions

What about other points? **Shifting** trick (Mumford, Brion):
The idea is: instead asking $\mu(g \cdot x) = \alpha$, construct a bigger rep $\tilde{V}$ and $\tilde{x}$ such that

$$\tilde{\mu}(\tilde{x}) \sim \mu(x) - \alpha,$$

and ask if $\tilde{\mu}$ can be 0.
Let $\alpha \in \mathfrak{t}^*_+$ be rational. Then $\alpha = \lambda/n$ for some highest weight λ and $n \geq 1$.
Let $w \in V_{\lambda^*}$ be a highest weight vector with weight $-\lambda^*$. Set

$$\tilde{V} = V^{\otimes n} \otimes V_{\lambda^*}, \quad \tilde{v} = (g \cdot v)^{\otimes n} \otimes w.$$

Then $\mu(w) = -\lambda$, and

$$\tilde{\mu}((g \cdot v)^n \otimes w) = n\,\mu(g \cdot v) + \mu(w) = n\,\mu(g \cdot v) - \lambda.$$

$\alpha \in \Delta(\overline{G \cdot x})$ if and only if $0 \in \Delta(\overline{G \cdot ((g \cdot v)^n \otimes w)})$ for some $g \in G$ (generic suffices).

Theorem

$\Delta(T)$ is convex.

Proof.

Use semigroup property of $S(T)$:

Let $\alpha = \lambda/n$, $\beta = \mu/m \in \Delta(T)$ with λ, μ highest weights. Pick highest weight vectors $v_\lambda \in \mathbb{C}[V]_n$ and $v_\mu \in \mathbb{C}[V]_m$, not identically zero on orbit of T. Then $v_\lambda^m \cdot v_\mu^n$ is a highest weight vector in $\mathbb{C}[V]_{2mn}$ with weight $m\lambda + n\mu$, and hence

$$\frac{m\lambda + n\mu}{2mn} = \frac{\alpha + \beta}{2} \in \Delta(T).$$

□

Properties of $\Delta(T)$

Theorem

$\Delta(T)$ is a (rational) polytope, i.e., defined by finitely many linear (rational) inequalities; equivalently, it has finitely many (rational) extreme points.

Proof.

$S(T)$ is a finitely generated semigroup, see Derksen–Kemper's *Computational Invariant Theory*, section 4.2. □

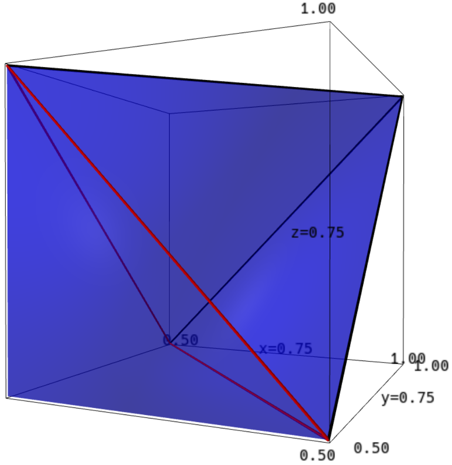
An example: $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$

Let $V = \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ and $G = \mathrm{GL}_2^3$ acting via

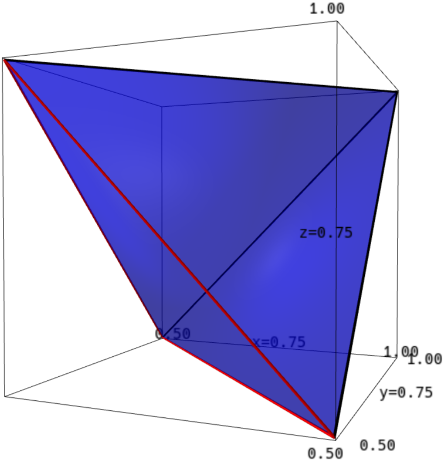
$$(g_1, g_2, g_3) \cdot v = (g_1 \otimes g_2 \otimes g_3)v$$

There are only **finitely** many orbits (6):

- $e_1 \otimes e_1 \otimes e_1$ (product state),
- $e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_2$ and its permutations (maximally entangled pairs),
- $e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_1$ (W-state),
- $e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes e_2$ (GHZ-state, generic),



GHZ-state polytope



W-state polytope

Interpreting generic tensor polytopes

Let $G = \mathrm{GL}_a \times \mathrm{GL}_b \times \mathrm{GL}_c$ act on $V = \mathbb{C}^a \otimes \mathbb{C}^b \otimes \mathbb{C}^b$. Then

$$\Delta(\mathbb{C}^a \otimes \mathbb{C}^b \otimes \mathbb{C}^c) = \left\{ \frac{1}{n}(\lambda, \mu, \nu) : g(\lambda, \mu, \nu) \neq 0 \right\}$$

where $\lambda, \mu, \nu \vdash n$ are partitions of n into at most $a/b/c$ parts respectively,
$g(\lambda, \mu, \nu)$ are the **Kronecker** coefficients for S_n, i.e., S_ν appears as subrep of $S_\lambda \otimes S_\mu$. (See Christian's talk on Schur–Weyl duality!)

So the polytope captures some (asymptotic) non-vanishing Kronecker coefficients, which are notoriously hard to compute!
$\Delta(V)$ described by Klyachko, Berenstein–Sjamaar using Schubert calculus, but hard to make explicit!
Walter–Vergne computed explicitly up to $(a, b, c) = (4, 4, 4)$, $(3, 3, 9)$, $(2, 4, 8)$, $(2, 2, 3, 12)$, using a more practical version of a method due to Ressayre.

Taken from Michael Walter's `moment_polytopes` package:

Vertices

#	V_A	V_B	V_C
1	(1/4, 1/4, 1/4, 1/4)	(1/4, 1/4, 1/4, 1/4)	(1/4, 1/4, 1/4, 1/4)
2	(1/4, 1/4, 1/4, 1/4)	(1/4, 1/4, 1/4, 1/4)	(1/3, 1/3, 1/3, 0)
3	(1/4, 1/4, 1/4, 1/4)	(1/4, 1/4, 1/4, 1/4)	(1/2, 1/2, 0, 0)
4	(1/4, 1/4, 1/4, 1/4)	(1/4, 1/4, 1/4, 1/4)	(1, 0, 0, 0)
5	(1/4, 1/4, 1/4, 1/4)	(1/3, 1/3, 1/3, 0)	(1/3, 1/3, 1/3, 0)
6	(1/4, 1/4, 1/4, 1/4)	(1/3, 1/3, 1/3, 0)	(1/2, 1/2, 0, 0)
7	(1/4, 1/4, 1/4, 1/4)	(1/3, 1/3, 1/3, 0)	(2/3, 1/6, 1/6, 0)
8	(1/4, 1/4, 1/4, 1/4)	(1/3, 1/3, 1/3, 0)	(2/3, 1/4, 1/12, 0)
9	(1/4, 1/4, 1/4, 1/4)	(1/3, 1/3, 1/3, 0)	(3/4, 1/12, 1/12, 1/12)
10	(1/4, 1/4, 1/4, 1/4)	(3/8, 3/8, 1/4, 0)	(5/8, 3/8, 0, 0)
11	(1/4, 1/4, 1/4, 1/4)	(3/8, 3/8, 1/4, 0)	(3/4, 1/8, 1/8, 0)
12	(1/4, 1/4, 1/4, 1/4)	(2/5, 3/10, 3/10, 0)	(7/10, 3/20, 3/20, 0)
13	(1/4, 1/4, 1/4, 1/4)	(5/12, 5/12, 1/6, 0)	(2/3, 1/6, 1/12, 1/12)
14	(1/4, 1/4, 1/4, 1/4)	(1/2, 1/6, 1/6, 1/6)	(1/2, 1/2, 0, 0)
15	(1/4, 1/4, 1/4, 1/4)	(1/2, 1/4, 1/8, 1/8)	(5/8, 3/8, 0, 0)
16	(1/4, 1/4, 1/4, 1/4)	(1/2, 1/4, 1/4, 0)	(1/2, 1/2, 0, 0)
17	(1/4, 1/4, 1/4, 1/4)	(1/2, 1/4, 1/4, 0)	(2/3, 1/6, 1/6, 0)
18	(1/4, 1/4, 1/4, 1/4)	(1/2, 1/4, 1/4, 0)	(3/4, 1/4, 0, 0)
19	(1/4, 1/4, 1/4, 1/4)	(1/2, 3/8, 1/8, 0)	(5/8, 1/8, 1/8, 1/8)
20	(1/4, 1/4, 1/4, 1/4)	(1/2, 1/2, 0, 0)	(1/2, 1/2, 0, 0)
21	(2/7, 2/7, 2/7, 1/7)	(4/7, 1/7, 1/7, 1/7)	(4/7, 3/7, 0, 0)
22	(7/24, 7/24, 5/24, 5/24)	(1/3, 1/3, 1/3, 0)	(3/4, 1/8, 1/8, 0)
23	(3/10, 3/10, 1/5, 1/5)	(2/5, 3/10, 3/10, 0)	(4/5, 1/10, 1/10, 0)
24	(3/10, 3/10, 3/10, 1/10)	(1/2, 1/2, 0, 0)	(11/20, 3/20, 3/20, 3/20)
25	(3/10, 3/10, 3/10, 1/10)	(1/2, 1/2, 0, 0)	(3/5, 1/5, 1/10, 1/10)
26	(1/3, 2/9, 2/9, 2/9)	(1/3, 1/3, 1/3, 0)	(2/3, 1/3, 0, 0)
27	(1/3, 2/9, 2/9, 2/9)	(1/3, 1/3, 1/3, 0)	(7/9, 1/9, 1/9, 0)

How to compute them?

Different approaches:

- Completely understand $S(T)$ (or even the ring of covariants $\mathbb{C}[V]^U$). Carried out for "small" examples around 1900 (Sylvester–Franklin, Hilbert, …) (with some more recent progress), but very impractical!
- Special cases (e.g. Littlewood–Richardson or Kronecker polytope): recursive and combinatorially difficult.
- Ressayre's method: non-Weyl-chamber facets of $\Delta(V)$ correspond to *admissible well-covering pairs*.
- *Scaling:* can test if *specific* $p \in \Delta(T)$. Not always efficient (theoretically), but practically tractable. Unclear how to use to give **complete** description of $\Delta(T)$.
- Combinatorial description due to Franz.

Franz's description of $\Delta(X)$

Let X be a G-closed subvariety of $\mathbb{P}(V)$.

Write $V = \oplus_{\omega\in\Omega} V_\omega$ for the weight decomposition of V and define for $v \in V$,

$$\operatorname{supp} v = \{\omega \in \Omega : v_\omega \neq 0\}.$$

For $x = [v] \in X$, define the **Borel polytope**

$$P(x) = \bigcap_{u\in U^-} \operatorname{conv}\operatorname{supp}(u \cdot v) \cap \mathfrak{t}^*_+.$$

Theorem (Franz)

$\Delta(X) \supseteq P(x)$, with equality for generic $x \in X$.

Proof relies heavily on the shifting trick.

Franz's description made practical

This yields an **algorithm** for computing $\Delta(X)$!

- Pick a generic $x = [v] \in X$.
- Initialize $\Delta = \operatorname{conv}\operatorname{supp} v \cap \mathfrak{t}^*_+$.
- For every $S \subseteq \Omega$:
  - Determine if there exists some $u \in U^-$ such that $\operatorname{supp}(u \cdot v) \subseteq S$.
  - If yes, update $\Delta = \Delta \cap \operatorname{conv} S$.

Observation

The question "is there $u \in U^-$ such that $\operatorname{supp}(u \cdot v) \subseteq S$" is asking whether the polynomial system

$$(u \cdot v)_\omega = 0, \quad \omega \in \Omega \setminus S$$

has a solution in u! Can be solved using Gröbner bases.

Franz's description made practical

With many tricks, this can be made practical. Naively, need to check $2^{|\Omega|}$ polynomial systems!

Ongoing work w/ van den Berg, Christandl, Lysikov, Walter and Zuiddam (and a lot of CPU hours):
determined complete set of vertices/inequalities of polytopes for

- **all** $v \in \mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3$ under $\mathrm{GL}(3)^3$-action,
- interesting tensors in $\mathbb{C}^4 \otimes \mathbb{C}^4 \otimes \mathbb{C}^4$ under $\mathrm{GL}(4)^3$-action (e.g. the 2×2-matrix multiplication tensor),
- $v \in \mathrm{Sym}^k(\mathbb{C}^d)$ for small k, d,
  e.g. $k = 3$ and $d = 4 \leftrightarrow$ cubic surfaces in $\mathbb{P}^3$,
- 3- and 4-dimensional algebras.

Asymptotic spectrum of tensors

As we saw in Itai's talk:

Theorem (Strassen)

$\tilde{R}(T) = \sup_\varphi \varphi(T)$, $\tilde{Q}(T) = \inf_\varphi \varphi(T)$, where φ are spectral points (points in the asymptotic spectrum of tensors).

Recall that φ are $\mathbb{R}_{\geq 0}$-valued functions on tensors such that

- $\varphi(\langle r \rangle) = r$ (normalization),
- $\varphi(T \oplus S) = \varphi(T) + \varphi(S)$,
- $\varphi(T \boxtimes S) = \varphi(T)\varphi(S)$,
- $\varphi(S) \geq \varphi(T)$ when S restricts to T.

Basic examples: flattening ranks, but these cannot be all.

Quantum functionals

Only other points in asymptotic spectrum we know are due to Christandl–Vrana–Zuiddam: for θ a probability distribution on $\{1,2,3\}$, we define

$$E_\theta(T) = \sup_{(p_1,p_2,p_3)\in\Delta(T)} \sum_{i=1}^{3} \theta_i H(p_i).$$

Then $F_\theta = 2^{E_\theta}$ is a universal spectral point.
Can prove this using properties of moment polytopes, and connection to Kronecker coefficients.

If we had efficient membership testing for $\Delta(T)$, we could compute F_θ efficiently!

Outlook

Interesting open questions:

- Efficient (polynomial-time) algorithms for membership in $\Delta(X)$?
- Efficiently verifiable certificates for inequalities for $\Delta(X)$?
- How big of a representation can compute the polytopes for?
- Interpretation of $\Delta(X)$, e.g. for cubic surfaces? Know they can be quite non-trivial for cubic surfaces with special singularities.
- Asymptotic properties: $\Delta(x^{\boxtimes n})$ for tensors $x \in \mathbb{C}^d \otimes \mathbb{C}^d \otimes \mathbb{C}^d$ and big n?

Thank you!