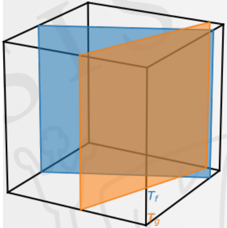
Additive X-rank decomposition

joint work with Ben Lovitz and Alex Casarotti

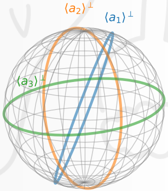
Alexander Blomenhofer (Villum Centre for the Mathematics of Quantum Theory)

Geometry of tensors in Nancy, 03.11.2025

UNIVERSITY OF COPENHAGEN



Alex Casarotti



Benjamin Lovitz

Additive X-rank decomposition

Let $X \subseteq V$ be an algebraic variety, which lives in an affine space V. For a given $T \in \mathbb{C}^N$, the X-rank problem is to find $x_1, \ldots, x_r$ such that

$$T = x_1 + \ldots + x_r,$$

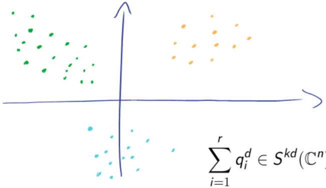
and such that r is minimal. This r is called the X-rank of T, denoted $\operatorname{rank}_X T$.

Examples:

1. Matrix rank: $X = \{ab^T \mid a, b \in \mathbb{C}^n\}$, $V = \mathbb{C}^{n\times n}$.
2. Tensor rank: $X = \{a \otimes b \otimes c \mid a, b, c \in \mathbb{C}^n\}$, $V = \mathbb{C}^{n\times n\times n}$.
3. Skew rank: $X = \{a \wedge b \wedge c \mid a, b, c \in \mathbb{C}^n\}$, $V = \Lambda^3(\mathbb{C}^n)$.
4. Chow rank: $X = \{abc \mid a, b, c \in \mathbb{C}^n\}$, $V = S^3(\mathbb{C}^n)$.

Here: $uv = \frac{1}{2}(u \otimes v + v \otimes u)$ and $u \wedge v = \frac{1}{2}(u \otimes v - v \otimes u)$.

Why should we care about additive X-rank?

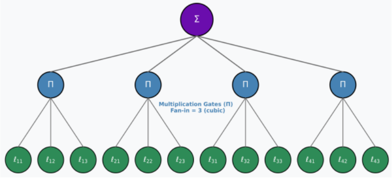


$$\sum_{i=1}^{r} q_i^d \in S^{kd}(\mathbb{C}^n)$$

Circuit Complexity

- Chow ranks
- Waring / Powering circuits

$$\sum_{i=1}^{r} a_{i1} a_{i2} \cdots a_{id} \in S^d(\mathbb{C}^n)$$



Statistics

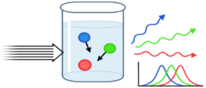
- Gaussian Mixture Estimation
- Subspace Learning

Additive X-rank decompositions

Quantum

- Bosonic Entanglement (sym. rank)
- Fermionic Entanglement (skew rank)
- Spinor varieties
- W-States

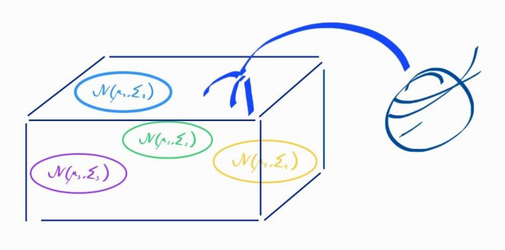
$$\frac{1}{r}\sum_{i=1}^{r}(\mu_i^T x)^5 + 10(\mu_i^T x)^3(x^T\Sigma_i x) + 15(\mu_i^T x)(x^T\Sigma_i x)^2$$

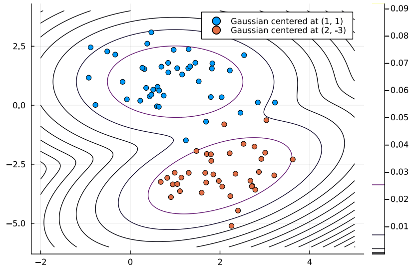


$$\sum_{i=1}^{r} a_{i1} \wedge a_{i2} \wedge \ldots \wedge a_{id} \in \Lambda^d(\mathbb{C}^n)$$

Gaussian mixture models and X-rank

A mixture of r Gaussians $\mathcal{N}(\mu_1, \Sigma_1), \ldots, \mathcal{N}(\mu_r, \Sigma_r)$ is sampled by choosing $i \in \{1, \ldots, r\}$ at random, then sampling the Gaussian with parameters (μ_i, Σ_i).





The order-5 moments of a (uniform) mixture Y of Gaussians $\mathcal{N}(\mu_1, \Sigma_1), \ldots, \mathcal{N}(\mu_r, \Sigma_r)$ can be collected either in a symmetric tensor $\mathbb{E}_{\sim Y}[Y^{\otimes 5}] \in \mathbb{C}^{n \times \cdots \times n}$ or in the coefficients of the quintic homogeneous polynomial $\mathbb{E}_{\sim Y}[(Y^T x)^5]$ in variables $x = (x_1, \ldots, x_n)$.

Application: Gaussian mixtures

The order-5 moments of a (uniform) mixture Y of Gaussians $\mathcal{N}(\mu_1, \Sigma_1), \ldots, \mathcal{N}(\mu_r, \Sigma_r)$ are the coefficients of the quintic form [1]

$$\mathbb{E}_{\sim Y}[(Y^T x)^5] = \frac{c_5}{r} \sum_{i=1}^{r} (\mu_i^T x)^5 + 10(\mu_i^T x)^3 (x^T \Sigma_i x) + 15(\mu_i^T x)(x^T \Sigma_i x)^2$$

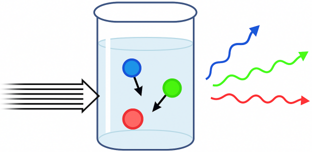
Theorem (B.-Casarotti)

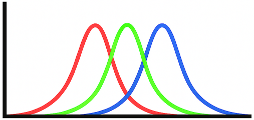
The parameters of a general Gaussian mixture model are identifiable from the fifth order moments, if the rank r of the mixture is bounded by
$r \leq \frac{1}{\binom{n+1}{2}+n} \binom{n+d-1}{d} - \binom{n+1}{2} - n - 1 = \mathcal{O}(n^3)$.

Under more restrictive assumptions, there exist efficient algorithms to recover the parameters from moments.

[1] NB: Generally, the moments can be obtained from the identity of characteristic functions $\phi_Y(x) = \mathbb{E}_{\sim Y}[\exp(\mathring{\imath} Y^T x)] = \sum_{i=1}^{r} \exp(\mathring{\imath} x^T \mu_i - \frac{1}{2} x^T \Sigma_i x)$.

Nick's Example





If a fluorophore is hit by white light, the emitted Fourier spectrum will be a sum of Voigt profiles[2] centered around its emission frequencies.

A Voigt profile $V_{\sigma,\lambda}(x) = \int G_\sigma(\tau) L_\lambda(x-\tau)$ is a convolution of a Gaussian and a Cauchy-Lorentz distribution.

In the limit $\lambda \to 0$, we get a Gaussian mixture.

Distributions $Y \to$ Moments $\mathbb{E}_{\sim Y}[Y_{i_1} \cdots Y_{i_d}] \to$ Symmetric tensors $\mathbb{E}_{\sim Y}[Y^{\otimes d}]$.

[2] W. Demtröder, Laser Spectroscopy 1

Central questions on X-rank

(A) How many terms from X do we need to express general tensors T? (generic rank)

(B) When are additive rank decompositions unique? (nondefectivity/identifiability)

(C) When can we solve additive rank problems? (algorithms)

If X is "nice" (for instance, if X is a GL_n-invariant subvariety of the symmetric or alternating tensor space), then (A) and (B) have simple approximate answers.

Coarse answers to (A) and (B)

Theorem (B.-Casarotti, 2023)

Let X be an (irreducible) affine cone in a space V, where V is also an irreducible G-module. Assume that X is G-invariant ($GX \subseteq X$). Then,

(A) The generic X-rank is at most $\frac{\dim V}{\dim X} + \dim X$.

(B) General forms of X-rank $r \leq \frac{\dim V}{\dim X} - \dim X - 1$ have finitely many minimum rank decompositions.

Under a technical condition (Gauß map of X nondegenerate), (B) can be strengthened to "unique minimum rank decomposition" (Massarenti-Mella, 2024).

Examples: For general (sym.) 3-tensors of Chow rank-r, the minimum decomposition is unique, if $r \leq \mathcal{O}(n^2)$. Analogous results for skew rank, tensor rank etc.

A sneak peak into nondefectivity proofs

Theorem (B.-Casarotti, 2023)

Let X be an irreducible affine cone in a space V, where V is also an irreducible G-module. Assume that X is G-invariant ($GX \subseteq X$). Then,

(A) The generic X-rank is at most $\frac{\dim V}{\dim X} + \dim X$.

(B) General forms of X-rank $r \leq \frac{\dim V}{\dim X} - \dim X$ have finitely many minimum rank decompositions.

Proof sketch.

Step 1: Consider the dimension a_r of the space $(T_{x_1}X + \ldots + T_{x_r}X) \cap T_yX$, where $x_1, \ldots, x_r, y$ are generic points of X.
Step 2: Show that either $a_r = 0$ or $a_r = \dim X$ or $a_r < a_{r+1}$.
Step 3: Conclude there are at most $\dim X$ possible values of r with $a_r \notin \{0, \dim X\}$. $\square$

Trick in Step 2: If $a_r = a_{r+1}$, then V contains an irreducible G-module of dimension $\geq a_r$.

Some new answers to question (C)

Computing decompositions is much harder than showing uniqueness, even for general tensors. Still, if X has nice invariance properties, we can do something for small ranks.

Theorem (Vannieuwenhoven, 2024, "Chiseling")

There is a time-$\mathcal{O}(n^7)$ algorithm which computes the unique minimum skew rank decomposition of a concise alternating tensor $T \in \Lambda^3(\mathbb{C}^{3n}) \subseteq \mathbb{C}^{3n \times 3n \times 3n}$ of skew rank n.

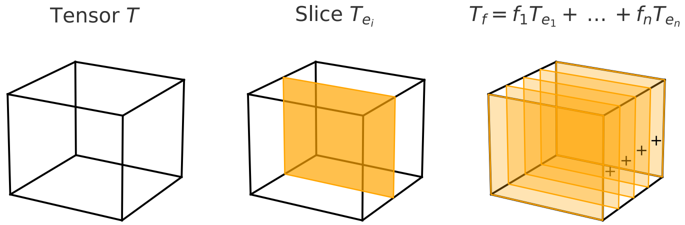
Theorem (B.-Lovitz, 2025, Contraction varieties)

There is a linear-time algorithm which computes the unique minimum Chow rank decomposition of a concise symmetric tensor $T \in S^3(\mathbb{C}^{3n}) \subseteq \mathbb{C}^{3n \times 3n \times 3n}$ of Chow rank n.

There is a subquadratic time algorithm which computes the unique minimum Chow rank decomposition of a general symmetric tensor $T \in S^{2d+1}(\mathbb{C}^{3n})$ of Chow rank $r \leq \frac{1}{c_d}\binom{n+d-1}{d} - c_d \approx n^d$, where $c_d = \binom{2d+1}{d}$ and $n \geq 2d+1$.

The contraction variety

Let $T \in \mathbb{C}^{n \times n \times n}$ be a 3-tensor. For $f \in \mathbb{C}^n$, we write T_f for the contraction of T by f. So $T_f = f_1 T_{e_1} + \ldots + f_n T_{e_n}$, where T_{e_i} are the slices of T. Visualized:



The contraction variety of T is defined as

$$Y_T = \{f \in \mathbb{C}^n \mid \det(T_f) = 0\}.$$

Some contraction varieties

(A) Tensor Rank If $T = \sum_{i=1}^n a_i^{\otimes 3}$ and $a_1, \dots, a_n$ are linearly independent, then

$$Y_T = \bigcup_{i=1}^n \langle a_i \rangle^\perp$$

(B) Chow Rank For $n = 3r$, if $T = \sum_{i=1}^r a_{i1} a_{i2} a_{i3}$ and $\{a_{ij}\}_{i=1,\dots,r, j=1,2,3}$ is linearly independent, then

$$Y_T = \bigcup_{i=1}^r \langle a_{i1} \rangle^\perp \cup \langle a_{i2} \rangle^\perp \cup \langle a_{i3} \rangle^\perp.$$

(C) Skew Rank For $n = 3r$, if $T = \sum_{i=1}^r a_{i1} \wedge a_{i2} \wedge a_{i3}$ and $\{a_{ij}\}_{i=1,\dots,r, j=1,2,3}$ is linearly independent, then

$$Y_T = \bigcup_{i=1}^r \langle a_{i1}, a_{i2}, a_{i3} \rangle^\perp.$$

The contraction variety II

Proposition

Let $T \in \mathbb{C}^{n\times n\times n}$ be a concise symmetric tensor of symmetric rank n. Then, Y_T is a union of hyperplanes.

Proof.

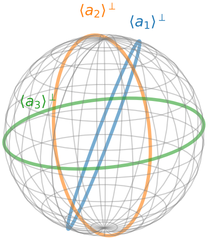
By the rank assumption, we have $T = \sum_{i=1}^n a_i^{\otimes 3}$ for some $a_i \in \mathbb{C}^n$. By conciseness, we know that $a_1, \dots, a_n \in \mathbb{C}^n$ are linearly independent. A contraction has the form

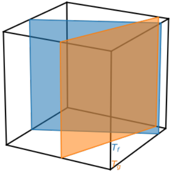
$$T_f = \sum_{i=1}^{n} \langle a_i, f\rangle a_i a_i^T.$$

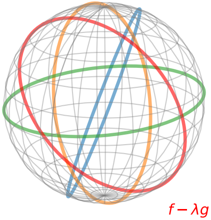
Obviously, this has full rank if and only if $\langle a_i, f\rangle \neq 0$ for all $i = 1, \dots, n$. Therefore,

$$Y_T = \bigcup_{i=1}^{n} \langle a_i\rangle^{\perp}.$$

Spectral algorithm







$$\det(T_{f-\lambda g}) = \det(T_f - \lambda T_g) = 0 \iff \det(T_g^{-1} T_f - \lambda I_n) = 0$$

This is a (generalized) eigenvalue problem and can be solved fast (time $\mathcal{O}(n^3) = \mathcal{O}(|T|)$).
The eigenvectors x_i satisfy $T_f x_i = \text{const} \cdot a_i$.

Spectral algorithm

Input: A concise symmetric tensor T of symmetric rank n.
Output: The unique minimum rank decomposition $T = \sum_{i=1}^{n} a_i^{\otimes 3}$ of the tensor.

(1) Pick generic $f, g \in \mathbb{C}^n$ and consider the general line $\mathcal{L}\colon \lambda \mapsto f - \lambda g$.
(2) Compute the intersection points $h_i = f - \lambda_i g$ of the line with Y_T, $i = 1, \ldots, n$.
(3) Compute corresponding kernel elements $x_i \in \ker T_{h_i}$.
(4) Output $\{T_f x_i\}_{i=1,\ldots,r}$, which equals $\{a_i\}_{i=1,\ldots,n}$ up to scalar multiples.

Correctness: Since $Y_T = \bigcup_{i=1}^{n} \langle a_i \rangle^{\perp}$, a general line intersects Y_T in n simple points h_i, one for each irreducible component. W.l.o.g., let $h_i \in \langle a_i \rangle^{\perp}$. Then, if $x_i \in \ker T_{h_i}$, we see that $x_i \perp a_j$ for each $j \neq i$. Thus, $T_f x_i = \langle a_i, f \rangle \langle a_i, x_i \rangle a_i$ is a nonzero multiple of a_i.

Summary

1. Additive X-rank decompositions well-behaved when X has good invariance properties.
2. Useful to extract information from data that is "mixed together".
3. Identifiability of Gaussian mixtures from moments of order 5 or higher. Algorithms for rank $\leq n-1$ available in special cases (homoscedastic or centered (B.)) or for very low rank $r = \sqrt{n}$ (Ge-Huang-Kakade, 2015).
4. Chow decompositions of cubic forms can be computed in linear time. For higher odd-order forms in subquadratic time. (B.-Lovitz)

Chow decompositions (left) and nondefectivity of reducible X-rank decompositions (right).

Merci pour votre attention !