TENSOR DECOMPOSITION AND ALGEBRAIC GEOMETRY

LUKE OEDING

1. Summary

In this lecture I will introduce tensors and tensor rank from an algebraic perspective. I will introduce multilinear rank and tensor rank, and I will discuss the related classical algebraic varieties – subspace varieties and secant varieties. I will give basic tools for computing dimensions, Terracini's lemma and the notion of the abstract secant variety. This will lead to the notion of generic rank. I will briefly disucss implicit equations, which lead to rank tests. Some basic references: [CGO14, Lan12].

In the second lecture I will focus on one special case of tensor decomposition - symmetric tensors and Waring decomposition. I will start by discussing the naive approach, then I will discuss Sylvester's algorithm for binary forms. As a bonus I will show how Sylvester's algorithm for symmetric tensor decomposition also gives a method find the roots of a cubic polynomial in one variable.

I will discuss what to expect regarding generic rank and uniqueness of tensor decomposition. With the remaining time I will discuss the recently defined notion of an eigenvector of a (symmetric) tensor ([Lim05, Qi05]), which leads to a new method (developed with Ottaviani) for exact Waring decomposition ([OO13].

Lecture 1: Tensors and classical Algebraic Geometry

2. Tensors with and without coordinates

The main goal in these lectures is to explain the basics of tensors from an Algebraic Geometry point of view. The main goal is to find exact expressions for specific tensors that are as efficient as possible. The guiding questions are the following:

- Find the space using the least number of variables for whit to represent a tensor.
- Determine the rank of a tensor.
- Find a decomposition as a sum of rank 1 tensors.
- Find equations that determine the rank (or border rank) of a tensor.
- Find the expected rank for a given tensor format.
- Provide an Algebraic-Geometry framework to study these questions.

Notation: A, B, C vector spaces over a field $\mathbb{F}$ (usually $\mathbb{C}$ or $\mathbb{R}$), or when there are many vector spaces $V_1, \ldots, V_n$ are all vector spaces over $\mathbb{F}$. In a somewhat circular manner, we define a tensor as an element of the tensor product of vector spaces.

2.1. Matrices as tensors.

We present two ways to think of matrices:

Without coordinates: The tensor product over $\mathbb{F}$ of vector spaces A and B is the new vector space denoted $A \otimes B$ spanned by all elements of the form $a \otimes b$ with $a \in A$ and $b \in B$.

Date: September 2014.

With coordinates: Suppose $\{a_1, \dots, a_m\}$ is a basis of A and $\{b_1, \dots b_n\}$ is a basis of B, then $A \otimes B$ is the vector space with basis $\{a_i \otimes b_j \mid 1 \leq i \leq m, 1 \leq j \leq n\}$. So every $M \in A \otimes B$ can be written in coordinates as

$$M = \sum_{i,j} M_{i,j} \, (a_i \otimes b_j), \quad \text{with } \; M_{i,j} \in \mathbb{F}.$$

Now an element in $A \otimes B$, (with the given choice of basis) is represented by the data $(M_{i,j})$ and we can think of $(M_{i,j})$ as a matrix,

$$M = \begin{pmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & \dots & M_{2,n} \\ & & & \\ \vdots & & \ddots & \vdots \\ M_{m,1} & M_{m,2} & \dots & M_{m,n} \end{pmatrix}. \tag{1}$$

The dual vector space of a vector space A, denoted A^* is the vector space of linear maps to $\mathbb{F}$. To a basis $\{a_1, \dots, a_m\}$ of A we may associate the dual basis $\{\alpha_1, \dots, \alpha_m\}$, which satisfies the condition

$$\alpha_i(a_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

We can also think of M as a linear mapping from $A^* \to B$. If we evaluate the tensor M on a basis of A^* and express the resulting vector in B as a column vector and put the resulting columns in a matrix, we obtain the transpose of the matrix M in (1). If we write the linear mapping $B^* \to A$ in the same fashion, we obtain the matrix M in (1).

We can vectorize M by choosing an (column-first) ordering on the pairs (i, j) to get

$$M = \begin{pmatrix} M_{1,1} & M_{2,1} & \dots & M_{m,1} & \mid & M_{1,2} & M_{2,2} & \dots & M_{m,2} & \mid & \dots & \mid & M_{1,n} & M_{2,n} & \dots & M_{m,n} \end{pmatrix}.$$

The symbol $\otimes$ is $\mathbb{F}$-bilinear:

$$\lambda \cdot (a \otimes b) = (\lambda \cdot a) \otimes b = a \otimes (\lambda \cdot b)$$
$$(a + a') \otimes b = a \otimes b + a' \otimes b$$
$$a \otimes (b + b') = a \otimes b + a \otimes b'$$

for all $\lambda \in \mathbb{F}$, $a, a' \in A$, $b, b' \in B$.

Note that $\otimes$ is *not commutative*. For instance, if A and B are two different vector spaces $b \otimes a$ might not even make sense in $A \otimes B$.

2.2. **Tensor products of many vector spaces.** The tensor product $\otimes$ is associative, so we can iterate the process and write a general tensor as an element in $A \otimes B \otimes C$ or (when there are many factors) in $V_1 \otimes V_2 \otimes \dots \otimes V_n$.

Without coordinates: The vector space $A \otimes B \otimes C$ is the $\mathbb{F}$-linear span of all $a \otimes b \otimes c$ with $a \in A, \; b \in B, \; c \in C$.

With coordinates: If $\{c_1, \dots, c_p\}$ is a basis of C, and A and B are as above, then $\{a_i \otimes b_j \otimes c_k \mid 1 \leq i \leq m, \; 1 \leq j \leq n, \; 1 \leq k \leq p\}$ is a basis of $A \otimes B \otimes C$.

Let's look at the associative property more closely. If $(A \otimes B) \otimes C = A \otimes (B \otimes C)$, then we delete the parentheses and view $A \otimes B \otimes C$ as a space of (3-rd order) tensors. Note while $\otimes$ is not commutative element-wise, $\otimes$ is actually commutative on the level of vector spaces. So it actually makes sense to write $(A \otimes C) \otimes B = A \otimes B \otimes C$, for instance.

2.3. **Symmetric tensors and homogeneous polynomials.** An important class of tensors are the symmetric tensors. Suppose $F \in V \otimes V \otimes \ldots \otimes V$ (a d-fold tensor product). Let $\{x_1, \ldots, x_n\}$ denote a basis of V and write

$$F = \sum_{i_1,\ldots,i_d \in \{1,\ldots,n} F_{i_1,\ldots,i_d} \, (x_{i_1} \otimes \ldots \otimes x_{i_d}).$$

We say that F is symmetric (or fully symmetric) if

$$F_{i_1,\ldots,i_d} = F_{\sigma(i_1),\ldots,\sigma(i_d)} \quad \text{for all } \sigma \in \mathfrak{S}_d.$$

The coordinates of a symmetric tensor F can be thought of as the coefficients of a degree d homogeneous polynomial on n variables. The vector space of all such symmetric tenors, denoted S^dV, has a basis of monomials,

$$\{x_{i_1} \cdots x_{i_d} \;\mid\; i_j \in \{1, \ldots, n\}\}, \quad \text{with } x_{i_1} \cdots x_{i_d} = \binom{d}{i_1, \ldots, i_d} x_{i_1} \otimes \ldots \otimes x_{i_d}.$$

3. Notions of rank and their associated classical algebraic varieties

All of the different notions of rank we will discuss are unchanged by multiplication by a global non-zero scalar, so it makes sense to work in projective space $\mathbb{P}(A \otimes B \otimes C)$. For $T \in A \otimes B \otimes C$ we will write $[T]$ for the line through T and for the corresponding point in projective space.

3.1. **Flattenings, multilinear rank and the subspace variety.** View $A \otimes B$ as a vector space with no extra structure by vectorizing all elements in $A \otimes B$ and view $(A \otimes B) \otimes C$ as a space of matrices (utilizing a basis), or as a space of linear maps

$$(A \otimes B)^* \to C.$$

If $T = (T_{i,j,k}) \in A \otimes B \otimes C$, the corresponding element in $(A \otimes B) \otimes C$ is

$$T = \sum_k \left(\sum_{i,j} T_{i,j,k} \, (a_i \otimes b_j) \right) \otimes c_k,$$

and we can represent T as a matrix, which is called a flattening of T and has block structure: $T_C =$

$$\left(\begin{array}{cccc|cccc|c|cccc} T_{1,1,1} & T_{2,1,1} & \ldots & T_{m,1,1} & T_{1,2,1} & T_{2,2,1} & \ldots & T_{m,2,1} & & T_{1,n,1} & T_{2,n,1} & \ldots & T_{m,n,1} \\ T_{1,1,2} & T_{2,1,2} & \ldots & T_{m,1,2} & T_{1,2,2} & T_{2,2,2} & \ldots & T_{m,2,2} & & T_{1,n,2} & T_{2,n,2} & \ldots & T_{m,n,2} \\ & & \vdots & & & & \vdots & & \ldots & & & \vdots & \\ T_{1,1,p} & T_{2,1,p} & \ldots & T_{m,1,p} & T_{1,2,p} & T_{2,2,p} & \ldots & T_{m,2,p} & & T_{1,n,p} & T_{2,n,p} & \ldots & T_{m,n,p} \end{array}\right).$$

The matrix T_C is called a flattening of T. The other flattenings T_A and T_B are similarly defined.

If we fix one of the indices and let the other two vary and collect the results into matrix we have the so-called slices of the tensor. For instance the m different matrices

$$T_{i,,} = \begin{pmatrix} T_{i,1,1} & T_{i,2,1} & \ldots & T_{i,n,1} \\ T_{i,1,2} & T_{i,2,2} & \ldots & T_{i,n,2} \\ & & \vdots & \\ T_{i,1,p} & T_{i,2,p} & \ldots & T_{i,n,p} \end{pmatrix} : B^* \to C$$

are called the (frontal) slices of T and

Even though we used coordinates to define them, it turns out that the ranks of the flattenings of a tensor are independent of our choice of basis. For this reason it makes sense to define the multilinear rank of T as

$$\mathtt{MR}(T) := (\operatorname{rank}(T_A), \operatorname{rank}(T_B), \operatorname{rank}(T_C)).$$

Example 3.1. Note that the ranks in $\mathtt{MR}(T)$ don't have to be the same. For example, consider

$$T = a_1 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_1.$$

Then

$$T_A = \begin{pmatrix} T_{1,1,1} & T_{1,2,1} & T_{1,1,2} & T_{1,2,2} \\ T_{2,1,1} & T_{2,2,1} & T_{2,1,2} & T_{2,2,2} \end{pmatrix} = \left(\begin{array}{c|c} T_{,,1} & T_{,,2} \end{array} \right),$$

$$T_B = \begin{pmatrix} T_{1,1,1} & T_{1,1,2} & T_{2,1,1} & T_{2,1,2} \\ T_{1,2,1} & T_{1,2,2} & T_{2,2,1} & T_{2,2,2} \end{pmatrix} = \left(\begin{array}{c|c} T_{1,,} & T_{2,,} \end{array} \right),$$

$$T_C = \begin{pmatrix} T_{1,1,1} & T_{2,1,1} & T_{1,2,1} & T_{2,2,1} \\ T_{1,1,2} & T_{2,1,2} & T_{1,2,2} & T_{2,2,2} \end{pmatrix} = \left(\begin{array}{c|c} T_{,1,} & T_{,2,} \end{array} \right).$$

$$T_A = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cccc} 11 & 12 & 21 & 22 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \end{array}, \quad T_B = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cccc} 11 & 21 & 12 & 22 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}, \quad T_C = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cccc} 11 & 21 & 12 & 22 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}.$$

So

$$\mathtt{MR}(T) = (2, 2, 1).$$

Suppose $T \in A \otimes B \otimes C$ and $\mathtt{MR}(T) = (p, q, r)$. Then we can consider the image of each flattening:

$$T_A(B \otimes C) =: \mathbb{C}^p_A \subset A, \quad T_B(A \otimes C) =: \mathbb{C}^q_B \subset B, \quad T_C(A \otimes B) =: \mathbb{C}^r_C \subset C,$$

and by construction $T \in \mathbb{C}^p_A \otimes \mathbb{C}^q_B \otimes \mathbb{C}^r_C$. So computing the multilinear rank of T tells us the smallest tensor space in which T lives.

The set of tensors whose multilinear rank is bounded above by a given triple forms the subspace variety:

$$\operatorname{Sub}_{p,q,r}(A \otimes B \otimes C) := \overline{\{[T] \in \mathbb{P}(A \otimes B \otimes C) \mid \mathtt{MR}(T) = (p, q, r)\}},$$

where the over line indicates the Zariski closure.

In the case of symmetric tensors, we have symmetric flattenings and the symmetric subspace variety, which is

$$\operatorname{Sub}_r(S^d V) := \overline{\{[F] \in \mathbb{P}(S^d V) \mid \mathtt{MR}(T) = (r, r, \ldots, r)\}}.$$

Intuitively, the symmetric subspace variety is the variety of symmetric tensors that can be expressed using fewer variables. Geometrically, if $F \in \operatorname{Sub}_r(S^d V)$, then there exists an r-dimensional subspace $\mathbb{C}^r \subset V$ so that $F \in S^d \mathbb{C}^r$.

3.2. **Generic projection (contraction) rank and the Prank-variety.** We can also view the tensor space $(A\otimes B)\otimes C$ as a vector space of matrices $A^* \to B$, depending on parameters coming from C. With this perspective, we can again write

$$T = \sum_k \left(\sum_{i,j} T_{i,j,k}\ (a_i \otimes b_j) \right) \otimes c_k.$$

but treat the vectors c_k as indeterminants so that

$$T(C) = T_{,,1} \cdot c_1 + \cdots + T_{,,k} \cdot c_k$$

is a matrix depending linearly on the c_k. If we assign values to the c_k we obtain a generic projection in the C-factor. The other generic projections $T(A)$ and $T(B)$ are similarly defined.

Again, while our construction depended on choices of bases, the ranks of the generic projections do not depend on the choice of basis, so it makes sense to define the generic projection ranks of T as

$$\mathtt{PR}(T) := (\operatorname{rank}(T(A)), \operatorname{rank}(T(B)), \operatorname{rank}(T(C))).$$

The set of tensors whose generic projection rank is bounded above by a given triple forms the Prank variety:

$$\operatorname{Prank}_{p,q,r}(A \otimes B \otimes C) := \overline{\{[T] \in \mathbb{P}(A \otimes B \otimes C) \mid \mathtt{PR}(T) = (p,q,r)\}},$$

where the over line indicates the Zariski closure.

Example 3.2. For example, consider again

$$T = a_1 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_1.$$

Then

$$T(A) = \begin{pmatrix} a_1 & 0 \\ a_2 & 0 \end{pmatrix}, \quad T(B) = \begin{pmatrix} b_1 & b_2 \\ 0 & 0 \end{pmatrix}, \quad T(C) = \begin{pmatrix} c_1 & 0 \\ 0 & c_1 \end{pmatrix}.$$

So

$$\mathtt{PR}(T) = (1,1,2).$$

3.3. **Tensor rank, border rank and secant varieties.** As a vector space $A \otimes B \otimes C$ is spanned by elements of the form $a \otimes b \otimes c$. We say that an indecomposable tensor $a \otimes b \otimes c$ has rank 1. The set of tensors of rank 1 forms the Segre variety

$$\operatorname{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C) := \{[T] \in \mathbb{P}(A \otimes B \otimes C) \mid \mathtt{R}(T) = 1\} \subset \mathbb{P}(A \otimes B \otimes C).$$

In coordinates, the points $[T]$ on the Segre variety have the form $T = x \otimes y \otimes z$, or

$$T_{i,j,k} = x_i \cdot y_j \cdot z_k.$$

An element $T \in A \otimes B \otimes C$ has an expression as linear combination of rank one elements. The minimum number r of rank one elements needed to express T is called the rank or tensor rank of T:

$$\mathtt{R}(T) := \min_r \left\{ T = \sum_{s=1}^{r} \lambda_s\ (a_s \otimes b_s \otimes c_s) \quad \middle| \quad \lambda_s \in \mathbb{F},\ a_s \in A,\ b_s \in B,\ c_s \in C \right\}.$$

We may be interested in approximating T by a family of tensors T_ϵ of (possibly) lower rank. This leads to the notion of border rank, which is

$$\mathtt{BR}(T) := \min_r \left\{ \exists \{T_\epsilon \mid \mathtt{R}(T_\epsilon) = r\ \ \forall\ \epsilon > 0\} \quad \text{and} \quad \lim_{\epsilon \to 0} T_\epsilon = T \right\}.$$

If $X \subset \mathbb{P}^N$ is any algebraic variety, the r-th secant variety of X, denoted $\sigma_r(X)$ is the Zariski closure of all points in $\mathbb{P}^N$ which are a linear combination of r points from X. In the tensor case we have

$$\sigma_r(\mathrm{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)) := \overline{\{[T] \in \mathbb{P}(A \otimes B \otimes C) \mid \mathtt{BR}(T) = r\}}.$$

Note that in the Zariski closure the rank can go up in the limit and this causes the ill-posedness of many low rank approximation problems, see [dSL08].

Similarly, in the symmetric case we define the symmetric rank and symmetric border rank of a symmetric tensor $F \in S^dV$ as

$$\mathtt{R}_S(F) := \min_r \left\{ F = \sum_{s=1}^{r} \lambda_s \, (\ell_s) \quad \middle| \quad \lambda_s \in \mathbb{F},\ \ell \in V \right\},$$

and

$$\mathtt{BR}_S(F) := \min_r \left\{ \exists \{F_\epsilon \mid \mathtt{R}_S(F_\epsilon) = r \ \ \forall\, \epsilon > 0\} \quad \text{and} \quad \lim_{\epsilon \to 0} F_\epsilon = F \right\}.$$

The set of symmetric tensors of (symmetric) rank 1 forms the Veronese variety

$$\nu_d(\mathbb{P}V) := \{[F] \in \mathbb{P}(S^dV) \mid \mathtt{R}(F) = 1\}.$$

The points $[F]$ on the Veronese variety have the form $F = \ell^d$, for $\ell \in V$, or in coordinates

$$F_{i_1,\ldots,i_d} = \ell_{i_1} \cdots \ell_{i_d}, \quad \text{with } \ell = (\ell_1, \ldots, \ell_n) \in V.$$

The Zariski closure of the points $[F] \in \mathbb{P}S^dV$ of symmetric border rank r form the r-th secant variety of the Veronese variety:

$$\sigma_r(\nu_d(\mathbb{P}V)) := \overline{\{[F] \in \mathbb{P}(S^dV) \mid \mathtt{BR}_S(F) = r\}}.$$

Example 3.3. Consider the point $[x^{d-1}y] \in \nu_d(\mathbb{P}^1)$ for $d \geq 2$. Suppose for contradiction that $x^{d-1}y$ were a pure power.

Expand

$$x^{d-1}y = (ax + bb)^d$$

and consider the system of equations gotten by comparing coefficients on both sides:

$$\begin{array}{rcl} 0 & = & a^d \\ 1 & = & d \cdot a^{d-1}b \\ 0 & = & \binom{d}{2} \cdot a^{d-2}b^2. \\ \vdots & & \\ 0 & = & b^d \end{array}$$

Since this system of equations is inconsistent (look at the first, last and second equations), $x^{d-1}y$ is not a pure power. On the other hand

$$\lim_{\epsilon \to 0} \frac{x^d - (x + \epsilon y)^d}{d\epsilon} = x^{d-1}y,$$

which shows that $x^{d-1}y \in \sigma_2\nu_d(\mathbb{P}^1)$, and hence has border rank 2. More work shows that $x^{d-1}y$ has rank d.

4. Dimension, Terracini's lemma and expected behavior

There are many ways to compute the dimension of an algebraic variety. For our purpose, we prefer the differential-geometric approach.

4.1. **Computing tangent spaces.** If $X \subset \mathbb{P}V$ is an algebraic variety, let $\widehat{X} \subset V$ denote the cone over X. The affine tangent space to X at a smooth point $[x] \in X$, denoted $T_xX \subset V$, is the set of all tangent vectors to $\widehat{X}$ at x. An effective way to compute T_xX is as the collection of derivatives of all smooth curves $\gamma\colon [0,1] \to X$ that pass through x:

$$T_xX = \left\{ \frac{\partial \gamma}{\partial t}(0) \mid \ \gamma(t) \in X \ \forall t \in [0,1],\ \gamma(0) = x \right\}.$$

Example 4.1. Consider $X = \mathrm{Seg}(\mathbb{P}A \times \mathbb{P}B)$, and $x = [a \otimes b] \in X$. Let $\gamma(t) = a(t) \otimes b(t)$ be an arbitrary curve in X through $a \otimes b$, that is, we take $a(t)$ and $b(t)$ to respectively be arbitrary curves in A and B with $a(0) = a$ and $b(0) = b$.

First note that all points of the Segre product look like $[a \otimes b]$ for $a \in A$ and $b \in B$ arbitrary. We apply the product rule to the tensor product and compute:

$$\gamma'(0) = a'(0) \otimes b(0) + a(0) \otimes b'(0) = a'(0) \otimes b + a \otimes b'(0).$$

Now since $a(t)$ and $b(t)$ were arbitrary curves through a and b respectively, we can take $a'(0)$ to be any direction vector in A and we can take $b'(0)$ to be any direction vector in B. Therefore

$$T_{a\otimes b}\,\mathrm{Seg}(\mathbb{P}A \times \mathbb{P}B) = A \otimes b + a \otimes B$$

The two vector spaces $A \otimes b$ and $a \otimes B$ overlap precisely at $\langle a \otimes b \rangle$, so we can write the tangent space as a direct sum

$$T_{a\otimes b}\,\mathrm{Seg}(\mathbb{P}A \times \mathbb{P}B) = \langle a \otimes b \rangle \ \oplus\ (A/a) \otimes \langle b \rangle \ \oplus\ \langle a \rangle \otimes (B/b).$$

So if A and B are respectively m and n-dimensional, then the Segre product is a $(m-1)+(n-1)$-dimensional subvariety of $\mathbb{P}^{mm-1}$ (the dimension of the cone is one more).

Now compare to the matrix interpretation. Every rank one $m \times n$ matrix M can be represented by the outer product of a m-vector $\vec{a}$ and a n-vector $\vec{b}$. There are $m+n$ parameters, however the line through M is unchanged if $\vec{a}$ and $\vec{b}$ are each rescaled by nonzero scalars. So there are precisely $(m-1)+(n-1)$ free parameters, agreeing with our previous computation.

In a similar fashion, one determines that $\mathrm{Seg}(\mathbb{P}^{m-1} \times \mathbb{P}^{n-1} \times \mathbb{P}^{p-1})$ is a smooth $(m+n+p-3)$-dimensional subvariety of $\mathbb{P}^{mnp-1}$, so almost every tensor has rank > 1.

4.2. **Terracini's lemma.** To find the dimension of secant varieties we have one main tool, namely Terracini's Lemma, whose proof essentially follows from the summation rule for derivatives.

Lemma 4.2 (Terracini's Lemma)**.** *Let $P_1, \ldots, P_r \in X$ be general points and $P \in \langle P_1, \ldots, P_s \rangle \subset \sigma_r(X)$ be a general point. Then the tangent space to $\sigma_r(X)$ in P is*

$$T_P(\sigma_r(X)) = \langle T_{P_1}(X), \ldots, T_{P_r}(X) \rangle.$$

From Terracini's lemma we immediately can determine the expected dimension of a secant variety, which occurs when the collection of tangent spaces at the r points have empty pairwise intersections, and the secant variety doesn't fill the ambient space. In that case we have r independent choices from an $s+1$ dimensional (affine) tangent space for most points on the secant variety. In particular, if we have an s-dimensional variety $X \subset \mathbb{P}^N$, we expect the r-th secant variety to have dimension

$$\mathrm{ExpDim}(\sigma_r(X)) = \min\{N, r \cdot (s+1) - 1\}.$$

So if $\mathrm{ExpDim}(\sigma_r(\mathrm{Seg}(\mathbb{P}^{m-1}\times\mathbb{P}^{n-1}\times\mathbb{P}^{p-1}))) < mnp-1$ then we expect that most tensors in $\mathbb{P}(A\otimes B\otimes C)$ do not have rank r. On the other hand, if the r-th secant variety is expected to fill the ambient space, we expect that most tensors in that space will have rank $\leq r$. When the tangent spaces at general points of X overlap, this causes the dimension of the secant variety to drop, and X will not have the expected dimension, in which case we call X defective, and we set $\delta_r(X) = \mathrm{ExpDim}(\sigma_r(X)) - \dim(\sigma_r(X))$.

Example 4.3. Consider $\sigma_2(\mathbb{P}^1\times\mathbb{P}^1\times\mathbb{P}^1)$ and a genera point $p = a_1\otimes b_1\otimes c_1 + a_2\otimes b_2\otimes c_2$. Terracini's lemma says that

$$T_p\sigma_2(\mathbb{P}^1\times\mathbb{P}^1\times\mathbb{P}^1) = T_{a_1\otimes b_1\otimes c_1}\mathrm{Seg}(\mathbb{P}A\times\mathbb{P}B\times\mathbb{P}C) + T_{a_2\otimes b_2\otimes c_2}\mathrm{Seg}(\mathbb{P}A\times\mathbb{P}B\times\mathbb{P}C)$$
$$= \langle a_1\otimes b_1\otimes c_1\rangle \ \oplus\ (A/a_1)\otimes\langle b_1\rangle\otimes\langle c_1\rangle \ \oplus\ \langle a_1\rangle\otimes(B/b_1)\otimes\langle c_1\rangle \ \oplus\ \langle a_1\rangle\otimes\langle b_1\rangle\otimes(C/c_1)$$
$$+$$
$$\langle a_2\otimes b_2\otimes c_2\rangle \ \oplus\ (A/a_2)\otimes\langle b_2\rangle\otimes\langle c_2\rangle \ \oplus\ \langle a_2\rangle\otimes(B/b_2)\otimes\langle c_2\rangle \ \oplus\ \langle a_2\rangle\otimes\langle b_2\rangle\otimes(C/c_2).$$

First notice that this vector space is the entire $\mathbb{C}^8$, which implies that $\sigma_2(\mathbb{P}^1\times\mathbb{P}^1\times\mathbb{P}^1) = \mathbb{P}^7$ and therefore most $2\times2\times2$ tensors have rank 7 over $\mathbb{C}$.

Another interesting fact is that a spanning set for this tangent space is given by all rank 1 tensors $a_i\otimes b_j\otimes c_k$ such that the indices (i,j,k) come from two disjoint Hamming balls of radius one centered respectively at $(1,1,1)$ and $(2,2,2)$. This type of analysis applies to many other secant varieties of Segre products.

We just noticed that most $2\times2\times2$ tensors have rank 7 over $\mathbb{C}$. Over $\mathbb{R}$ the story is a bit more interesting. There are 2 typical ranks, meaning that real rank 2 and 3 both occur in full dimensional open sets [CtBDLC09]. Comon and Ottaviani [CO12] show that in the case of symmetric tensors of the same format the discriminant of a binary cubic cuts the space of binary cubics into typical ranks 2 and 3 depending on the sign of the discriminant. In the non-symmetric case the space $\mathbb{R}^2\otimes\mathbb{R}^2\otimes\mathbb{R}^2$ is cut by the $2\times2\times2$-hyperdeterminant (which also has degree 4 and whose symmetrization is the discriminant of a binary cubic, see [Oed12].

4.3. **The incidence variety.** One important construction for secant varieties is the notion of an incidence variety, or the abstract secant variety, which provides an alternate construction of the secant variety.

Again, for $X\subset\mathbb{P}^N$ the r-th abstract secant variety of X is

$$\mathcal{I}_r = \{(p, x_1,\dots,x_r) \mid [x_i]\in X,\ p\in\langle x_1,\dots,x_r\rangle\}\subset\mathbb{P}^N\times(\mathbb{P}^N)^{\times r}.$$

The image of the projection to the first factor is the usual r-th secant variety of X. Now $\mathcal{I}_r$ always has dimension $r\cdot(s+1)-1$, which is sometimes called the virtual dimension.

Now we have another geometric interpretation of defectivity. Consider the projection $\mathcal{I}_r \xrightarrow{\pi_1} \sigma_r(X)$, then the defect $\delta_r(X)$ is the dimension of the fiber $\pi_1^{-1}(p)$ over a general point $p\in\sigma_r(X)$.

Remark 4.4. One key insight provided by the incidence variety is that the number of decompositions of a given tensor T (as a sum of r rank one tensors) is completely by the size of the fiber $\pi_1^{-1}(T)$. This is one of the key facts exploited in the recent works of Bocci, Chiantini, Ottaviani, Vannieuwenhoven on determining when identifiability holds for many classes of tensors.

5. First equations and tests for border rank

The ideal of the subspace variety we defined above has a nice description.

Proposition 5.1 (Landsberg-Weyman)**.** *The ideal of the subspace variety* $\mathrm{Sub}_{r,s,\dots,t}(A\otimes B\otimes \dots\otimes C)$ *is generated by all*

- $(r+1)\times(r+1)$*-minors of* T_A*,*
- $(s+1)\times(s+1)$*-minors of* T_B
- $\vdots$
- $(t+1)\times(t+1)$*-minors of* T_C*.*

Landsberg and Weyman actually gave a much more refined statement regarding singularities, ideal generation, etc, but this statement is enough for our purposes.

Remark 5.2. There are many more types of flattenings when there are more than 3 factors, such as

$$T_{A\otimes B}\colon (C\otimes D)^*\to A\otimes B.$$

One question from the lecture was about how much is understood about the intersections of those types of subspace varieties. For instance the defectivity of $\sigma_3(\mathbb{P}^1\times\mathbb{P}^1\times\mathbb{P}^1\times\mathbb{P}^1)$ can be explained by noting that in the case all the vector spaces are 2 dimensional, the determinant of $T_{A\otimes B}$ and the determinant of $T_{A\otimes C}$ are algebraically independent and vanish on $\sigma_3(\mathbb{P}^1\times\mathbb{P}^1\times\mathbb{P}^1\times\mathbb{P}^1)$, causing its dimension to be one less than the expected dimension.

This explains some of the interesting features of flattenings to more than one factor, and it seems an interesting avenue to pursue.

Consider a general point $[T]\in\sigma_r(\mathrm{Seg}(\mathbb{P}A\times\mathbb{P}B\times\mathbb{P}C))$ and write

$$T=\sum_{s=1}^{r} a_s\otimes b_s\otimes c_s.$$

Now we can consider separately the span of the vectors appearing in the decomposition

$$\begin{aligned} A' &:= \langle a_1,\dots,a_s\rangle,\\ B' &:= \langle b_1,\dots,b_s\rangle,\\ C' &:= \langle c_1,\dots,c_s\rangle. \end{aligned}$$

Then $T\in A'\otimes B'\otimes C'$. Since everything happens on an open set, we have the following:

Proposition 5.3. *The following containment holds*

$$\sigma_r(\mathrm{Seg}(\mathbb{P}A\times\mathbb{P}B\times\mathbb{P}C))\subset \mathrm{Sub}_{r,r,r}(A\otimes B\otimes C).$$

In particular if $T\in\sigma_r(\mathrm{Seg}(\mathbb{P}A\times\mathbb{P}B\times\mathbb{P}C))$*, the* $(r+1)\times(r+1)$ *minors of the flattenings* T_A, T_B *and* T_C *must all vanish.*

Usually the containment is strict. For instance, $\sigma_4(\mathbb{P}^2\times\mathbb{P}^2\times\mathbb{P}^2)$ is a degree 9 hypersurface in $\mathbb{P}^{26}$, but all the flattenings have size 3×9, which in particular, have no 5×5 minors.

On the other hand, minors of flattenings are enough to detect border rank 1 (classical), and border rank 2, which was conjectured by Garcia, Stillman and Sturmfels:

Theorem 5.4 (Landsberg,Manivel,Weyman 2004–08, Raicu2012)**.** *The ideal of* $\sigma_2(\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_d})$ *is generated by all* 3×3 *minors of flattenings.*

The most refined statement over $\mathbb{C}$ is by Raicu. Michalek-Oeding
For border rank 3 the following result was recently obtained:

Theorem 5.5 (Yang Qi 2013)**.** *As a set $\sigma_3(\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_d})$ is defined by all 4×4 minors of flattenings and Strassen's degree 4 commutation conditions.*

For border rank 4 much less is known in general, but for 3 factors we have the following result, which is known as the "salmon prize problem."

Theorem 5.6 (Friedland2010, Bates-Oeding2011, Friedland-Gross2012)**.** *As a set $\sigma_4(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)$ is defined by*

- *all 5×5 minors of flattenings,*
- *Strassen's degree 5 commutation conditions,*
- *Landsberg and Manivel's degree 6 equations,*
- *Strassen's degree 9 commutation conditions.*

For higher secants and for more factors many sporadic cases are known, but little is known in general. Landsberg and Ottaviani's work on what they call Young flattenings provides a vast generalization of the usual flattenings that accounts for most of the known equations of secant varieties. Much more work is needed to find equations of secant varieties.

Lecture 2: Algorithms for tensor decomposition and identifiability

6. Symmetric tensor decomposition

Now we come to the question of how to actually find a decomposition of a tensor. For this we will focus on the case of symmetric tensors. Waring's classic problem is to write a number as a sum of powers of whole numbers using as few terms as possible. Waring's problem for polynomials is: write a given polynomial as a sum of powers of linear forms using as few terms as possible. Such a decomposition always exists over $\mathbb{C}$, but may be much more subtle over other fields. For instance in characteristic 2, the monomial xy has no decomposition as the sum of squares. Because of this, we assume we are working in characteristic 0 for what follows.

Let $\{x_1, \dots, x_n\}$ be a basis of V, with dual basis $y_i := \frac{\partial}{\partial x_i}$. The induced basis of S^pV is given by monomials $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, with $|\alpha| = p$, and the induced basis on the dual is given by the monomials in partial derivatives $\frac{\partial}{\partial x^\beta} := \frac{\partial}{\partial x_1^{\beta_1}} \cdots \frac{\partial}{\partial x_n^{\beta_n}}$ with $|\beta| = q$.

The naive formulation of the problem is the following: Given $F = \sum F_\alpha x^\alpha$

$$\text{minimize } \ r \text{ so that } \quad F = \sum_{s=1}^{r} \lambda_s \ell_s^d, \quad \text{with } \ell_s \in V.$$

If we expand both sides of the equation and compare coefficients, we have the following problem

$$\text{minimize } \ r \text{ so that the system } \quad F_\alpha = \sum_{s=1}^{r} \lambda_s \sum_{|\beta_s|=d} \ell_{1,s}^{\beta_{1,s}} \cdots \ell_r^{\beta_{r,s}}, \quad \text{with } \ell_{s,n} \in \mathbb{C} \text{ has a solution.}$$

Finally, one can take a brute force approach to finding the tensor rank of F and the tensor decomposition of F. That is arbitrarily pick an r, and try to solve the following optimization problem (for some appropriate norm $||\cdot||$).

$$\text{minimize } \quad ||F - \sum_{s=1}^{r} \lambda_s \ell_s^d||, \quad \text{for } \lambda \in \mathbb{F} \text{ and } \ell_s \in V.$$

While such an approach is guaranteed to eventually find a solution (given enough time, computational power and memory), the approach can quickly become infeasible even for plane quintics.

6.1. **Symmetric flattenings (Catalecticants).** Let's return to flattenings, but now in the symmetric case. Since S^dV has the interpretation as a d-mode symmetric tensor, we can relax some of the symmetry to obtain an inclusion $S^dV \subset S^pV \otimes S^qV$ for any positive integers p, q such that $p + q = d$. Let $C^p(F)$ denote the image of F in $S^pV \otimes S^qV$. Like usual flattenings $C^p(F)$ can be viewed as a linear mapping:

$$C^p(F) \colon S^pV^* \to S^qV$$

and can be represented by the following matrix

$$C^p(F) = \left(\frac{\partial F}{\partial x^{\alpha+\beta}}\right)_{\alpha,\beta},$$

with $|\alpha| = p$ and $|\beta| = q$. The matrix $C^p(F)$ is often called a catalecticant matrix, moment matrix, or Hankel matrix.

Symmetric flattenings all satisfy the following:

- The rank of $C^p(x_1^d)$ is one, and so is the rank of $C^p(\ell^d)$ for all linear forms $\ell \in V$.
- The construction is linear in the F-argument: $C^p(\lambda F + F') = \lambda C^p(F) + C^p(F')$.
- Matrix rank is sub-additive, so if F has rank r, then $\operatorname{rank}(C^p(F)) \leq r$.

6.2. **Apolarity.** A key tool for this study is apolarity. That is, given $\ell = \alpha x + \beta y \in \mathbb{C}^2$ we define its polar as $\ell^\perp = -\beta\partial_x + \alpha\partial_y \in (\mathbb{C}^2)*$. Notice that $\ell^\perp(\ell) = 0$ and also

$$\ell^\perp(\ell^d) = 0.$$

By the Fundamental Theorem of Algebra, elements in $S^d(\mathbb{C}^2)^*$ can be factored as $\ell_1^\perp \cdots \ell_d^\perp$ for some $\ell_i \in \mathbb{C}^2$.

The following proposition is the key to Sylvester's algorithm:

Proposition 6.1. *Let ℓ_i be distinct for $i = 1, \ldots, e$. Then there exists $\lambda_i \in \mathbb{C}$ such that*

$$f = \sum_{i=1}^{e} \lambda_i (\ell_i)^d$$

if and only if

$$\ell_1^\perp \cdots \ell_e^\perp f = 0.$$

Proof. Since we have $\ell^\perp(\ell^d) = 0$ this implies the forward direction and shows that $\langle \ell_1^d, \ldots, \ell_e^d \rangle \subset \mathtt{Ker}(\ell_1^\perp \cdots \ell_e^\perp)$. The other inclusion comes from noticing that both linear spaces have dimension e so equality holds, and this proves the other direction of the proposition. □

6.3. **Sylvester's algorithm by example.** Consider the following cubic polynomial in one variable

$$f(x) = 9x^3 - 3x^2 + 81x - 124.$$

For convenience we prefer work with the homogenization

$$f(x, y) = 9x^3 - 3x^2y + 81xy^2 - 124y^3.$$

We're interested in knowing how to write $f(x, y)$ as a sum of cubes of linear forms. We will do this via Sylvester's algorithm. As a consequence we'll be able to solve the cubic equation $f(x, y) = 0$.

Let $\{x, y\}$ be a basis of $\mathbb{C}^2$, so we can consider $f \in S^3\mathbb{C}^2$.

Step 1 is to construct the symmetric flattening:

$$C_f \colon S^2(\mathbb{C}^2)^* \longrightarrow \mathbb{C}^2 \ ,$$

$$C_f = \begin{array}{c} \\ \partial_x \\ \partial_y \end{array} \begin{array}{c} \begin{array}{ccc} \partial_x^2 & \partial_x\partial_y & \partial_y^2 \end{array} \\ \begin{pmatrix} 54 & -6 & 162 \\ -6 & 162 & -744 \end{pmatrix} \end{array} = 6 \cdot \begin{pmatrix} 9 & -1 & 27 \\ -1 & 27 & -124 \end{pmatrix}$$

Now compute the kernel:

$$\mathtt{ker}(C_f) = \left\{ \begin{pmatrix} -5 \\ 9 \\ 2 \end{pmatrix} \right\} \subset S^2(\mathbb{C}^2)^*$$

Reinterpret the kernel as a vector space (a line) of polynomials:

$$P = -5\partial_x^2 + 9\partial_x\partial_y + 2\partial_y^2$$

This is a quadratic polynomial in two homogeneous variables, so it factors (over $\mathbb{C}$):

$$P = (-\partial_x + 2\partial_y) \cdot (5\partial_x + \partial_y)$$

(I picked a nice example so that it actually factors over $\mathbb{Z}$.)

Now notice that if $(\alpha\partial_x + \beta\partial_y)\ell = 0$ for some $\ell \in \mathbb{C}^2$ then also $(\alpha\partial_x + \beta\partial_y)\ell^d = 0$ for all $d > 0$.

So the polar forms are $2x + y$ and $(x - 5y)$. We then expand and solve the system defined by:

$$f(x, y) = 9x^3 - 3x^2y + 81xy^2 - 124y^3 = \lambda_1(2x + y)^3 + \lambda_2(x - 5y)^3,$$

and find that (miraculously) $\lambda_1 = \lambda_2 = 1$.

Solving cubic equations As a side benefit, Sylvester's algorithm also allows us to solve the (dehomogenized) cubic equation

$$0 = f(x) = 9x^3 - 3x^2 + 81x - 124.$$

The decomposition we computed implies that

$$0 = (2x + 1)^3 + (x - 5)^3,$$

so

$$1 = \left(\frac{2x + 1}{-x + 5}\right)^3$$

and if $\omega = e^{2\pi i/3}$ then we have three solutions (for $j = 0, 1, 2$):

$$x = \frac{5\omega^j - 3}{\omega^j + 2}.$$

7. Eigenvectors of tensors

7.1. **Eigenvectors of matrices.** Recall that if M represents a linear map from V to V, we say that $v \in V$ is an eigenvector of M associated to an eigenvalue $\lambda \in \mathbb{F}$ if

$$Mv = \lambda v.$$

Since v and λv point in the same direction, we can write the eigenvector equation in another way as

$$(Mv) \wedge v = 0,$$

where the wedge product $\wedge$ is defined as $v \wedge w = \frac{1}{2}(v \otimes w - w \otimes v)$.

7.2. **Eigenvectors of symmetric tensors (polynomials).** Now we consider how to extend the definition of an eigenvector of a matrix to an eigenvector of a tensor. We work in the symmetric case first. Suppose $P \in S^dV$, and write the first symmetric flattening

$$C^1_P \colon S^{d-1}V^* \to V$$

We can think of this as a linear mapping that is multilinear and symmetric:

$$C^1_P \colon V^* \times \cdots \times V^* \to V.$$

With this perspective we will say that $v \in V$ is an eigenvector of the tensor $P \in S^dV$ if

$$C^1_P(v, \ldots, v) = \lambda v, \quad \text{for some } \lambda \in \mathbb{C},$$

or if

$$C^1_P(v, \ldots, v) \wedge v = 0.$$

This essential definition appeared in Lim '05 and Qi '05.

A general $m \times m$ matrix will have m linearly independent eigenvectors. Cartwright and Sturmfels computed the number of eigenvectors of a general symmetric tensor $P \in S^d\mathbb{C}^n$, which is

$$\frac{(d-1)^n - 1}{d-2}.$$

Example 7.1. Consider $xyz \in S^3V$. The first symmetric flattening of xyz is

$$C^1_{xyz} = \begin{array}{c} \\ \partial_x \\ \partial_y \\ \partial_z \end{array} \begin{array}{c} \begin{array}{cccccc} \partial_x^2 & \partial_x\partial_y & \partial_x\partial_z & \partial_y^2 & \partial_y\partial_z & \partial_z^2 \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}.$$

Now consider the system of equations:

$$C^1_{xyz}((ax+by+cz)^2) = \lambda(ax+by+cz)$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} a^2 \\ 2ab \\ 2ac \\ b^2 \\ 2bc \\ c^2 \end{pmatrix} = \lambda \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

$$\begin{aligned} 2bc &= \lambda a \\ 2ac &= \lambda b. \\ 2ab &= \lambda c \end{aligned}$$

The seven eigenvalues of xyz are the following solutions of this system. If $\lambda = 0$ there are 3 solutions corresponding to two of the variables a, b, c being zero, and one solution corresponding to all 3 variables equal to zero (which we exclude). The other set of solutions is when $\lambda = 1$. Here we have one solution $a = b = c = \frac{1}{2}$, and 3 more solutions with $a, b, c = \pm\frac{1}{2}$ with precisely two of the three negative.

7.3. **Eigenvectors of slightly more general tensors.** Let $\bigwedge^a V$ denote the vector space of alternating tensors. If $\{v_1 \ldots, v_n\}$ is a basis of V, then $\{v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_a} \mid i_1 < i_2 < \cdots < i_a\}$ is the natural basis of $\bigwedge^a V$.

Now the eigenvalue-free approach to eigenvectors of tensors allows us to define the notion of an eigenvector of a tensor $T \in S^dV^* \otimes \bigwedge^a V$ (why we need this type of tensor will be apparent later). Consider $T \in S^dV^* \otimes \bigwedge^a V$ as a linear mapping

$$S^dV \to \bigwedge^a V$$

We will say that $v \in V$ is an eigenvector of the tensor $T \in S^dV^* \otimes \bigwedge^a V$ if

$$T(v^d) \wedge v = 0 \in \bigwedge^{a+1} V.$$

The number of such eigenvectors is given by the following:

Proposition 7.2 ([OO13])**.** *Let $T \in T \in S^dV^* \otimes \bigwedge^a V$ be general with $V \cong \mathbb{C}^n$. The number of $e(T)$ of eigenvectors of T is*

$e(M) = m$, *when* $n = 2$ *and* $a \in \{0, 2\}$,	$e(M) = \frac{m^n-1}{m-1}$, *when* $a = 1$ *[CS'10],*
$e(M) = \infty$, *when* $n > 2$ *and* $a \in \{0, n\}$,	$e(M) = 0$, *for* $2 \leq a \leq n-1$,
(classical)	$e(M) = \frac{(m+1)^n+(-1)^{n-1}}{m+2}$, *for* $a = n$.

Our result includes a result of Cartwright-Sturmfels. Our proofs rely on the simple observation that the a Chern class computation for the appropriate vector bundle gives the number of eigenvectors.

(See also [FO14]

8. What to do when flattenings run out

A classical result of Hilbert, Richmond, and Palatini is the following:

Theorem 8.1. *A general $f \in S^5\mathbb{C}^3$ has an essentially unique decomposition as the sum of seven 5-th powers of linear forms.*

Try to find the unique decomposition.

Naively, we can parametrize seven linear forms and solve the system of quintic equations on 21*variables* given by equating the coefficients on each side of the equation

$$f = \sum_{s=1}^{7} (a_s x + b_s y + c_s z)^5.$$

This turns out to be feasible, but computationally intensive. Instead, we seek a solution that is on the same order as linear algebraic computations such as rank and null space computations.

Catalecticants: The most square flattenings are $C_f \colon (S^3\mathbb{C}^3)^* \to S^2\mathbb{C}^3$, which are 6×10 matrices, and can't distinguish rank 6 from rank 7.

The new tool is called an exterior flattening. The following example is part of a more general method described in [OO13]. Consider one of the Koszul maps

$$K \colon \mathbb{C}^3 \xrightarrow{\begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix}} \bigwedge^2 \mathbb{C}^3 \cong (\mathbb{C}^3)^*$$

Now twist the catalecticant by the Koszul matrix K:

$$(K \otimes C)_f \colon \mathbb{C}^3 \otimes (S^2\mathbb{C}^3)^* \xrightarrow{\begin{pmatrix} 0 & -C_{\partial_z f} & C_{\partial_y f} \\ C_{\partial_z f} & 0 & -C_{\partial_x f} \\ -C_{\partial_y f} & C_{\partial_x f} & 0 \end{pmatrix}} (\mathbb{C}^3)^* \otimes S^2\mathbb{C}^3$$

The matrix representing $(K \otimes C)_f$ in this case is 18×18.

Now $K \otimes C$ has the following properties:

- $(K \otimes C)_f$ is skew-symmetric.
- $(K \otimes C)_f$ is linear in the f-argument, that is $(K \otimes C)_{\lambda f + f'} = \lambda (K \otimes C)_f + (K \otimes C)_{f'}$,
- $\operatorname{rank}(K \otimes C)_f = 2$ if $\operatorname{rank}(f) = 1$,
- Subadditivity of matrix rank implies that if $\operatorname{rank} f = r$ then $\operatorname{rank}(K \otimes C)_f \leq 2r$.

Therefore the 16×16 Pfaffians of $(K \otimes C)_f$ vanish if f has rank 7. Moreover, if $R(f) = 7$, then $(K \otimes C)_f$ has a non-trivial kernel. Let M be a general element in $\ker(K \otimes C)_f \subset (S^2\mathbb{C}^3)^* \otimes \mathbb{C}^3$. Note that Cartwright and Sturmfel's count says that M must have 7 eigenvectors.

It turns out that the eigenvectors of M are the linear forms ℓ_s in the decomposition $\sum_{s=1}^{7} \lambda_s \ell_s$ for some $\lambda_s \in \mathbb{C}$.

We can find the eigenverctors of M by solving a (smaller) system of equations implied by $M(v^2) = v$ as we did in Example 7.1.

9. Optimization, and Euclidean Distance Degree

Given $f \in S^d\mathbb{C}^n$ we would like to solve the system of polynomial equations

$$f = \sum_{s=1}^{r} (a_s x + b_s y + \dots c_s z)^d.$$

We might hope for a fast algorithm like in the Hilbert Quintics example, but it turns out that we shouldn't hope for one unless we believe $P = NP$, since the general problem of finding the rank of a polynomial is NP-hard [HL09].

On the other hand, we might be interested in relaxing the exact problem to an optimization problem (as mentioned above):

$$\text{minimize} \quad ||f - \sum_{s=1}^{r} \lambda_s \ell_s^d||, \quad \text{for } \lambda \in \mathbb{F} \text{ and } \ell_s \in V.$$

Just to get an idea of how complicated this problem is, we may try to understand how many optimal solutions we might expect to find. In the case that $||\cdot||$ is the usual norm induced by the euclidean distance, the number of critical points of the function

$$(\vec{\lambda}, \ell_1, \dots, \ell_s) \longmapsto ||f - \sum_{s=1}^{r} \lambda_s \ell_s^d||$$

for a general data point f is called the Euclidean Distance Degree. For more on this topic we invite the reader to consult [DHO$^+$13, Lee14].

10. Exercises

Exercise 10.1. Consider the following $3 \times 3 \times 3$ tensor:

$$T = a_1 \otimes b_2 \otimes c_1 + a_3 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 + a_3 \otimes b_3 \otimes c_3 \in A \otimes B \otimes C.$$

(1) Write the three different flattenings of T and compute the multilinear rank $\mathtt{MR}(T)$.
(2) Write the three different representations of T as a matrix with linear entries and compute the generic projection rank $\mathtt{PR}(T)$.

Exercise 10.2. Consider $T = a_1 \otimes b_1 \otimes c_1 + a_2 \otimes b_1 \otimes c_1 + a_1 \otimes b_2 \otimes c_1 + a_1 \otimes b_1 \otimes c_2$. Show that T has rank 3 but border rank 2.

Exercise 10.3. Show that for $F \in S^dV$, the symmetric rank $\mathtt{R}_S(F) = 1$ if and only if the rank $\mathtt{R}(F) = 1$.

Exercise 10.4.
(1) Show that the variety $\mathrm{Sub}_{p,q,r}$ is irreducible.
(2) Show that the variety $\mathrm{Prank}_{p,q,r}$ is not in general irreducible.
(3) Show that $\mathrm{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)$ is a closed algebraic variety.
(4) Show that $\sigma_r(\mathrm{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$ is irreducible.

The exercises marked (CGO), while classical, also appear (with hints and solutions) in [CGO14].

Exercise 10.5 (CGO)**.** For $X \subset \mathbb{P}^N$, show that, if $\sigma_i(X) = \sigma_{i+1}(X) \neq \mathbb{P}^N$, then $\sigma_i(X)$ is a linear space and hence $\sigma_j(X) = \sigma_i(X)$ for all $j \geq i$.

Exercise 10.6 (CGO)**.** If $X \subseteq \mathbb{P}^N$ is non-degenerate, then there exists an $r \geq 1$ with the property that

$$X = \sigma_1(X) \subsetneq \sigma_2(X) \subsetneq \dots \subsetneq \sigma_r(X) = \mathbb{P}^N.$$

In particular, all inclusions are strict and there is a higher secant variety that coincides with the ambient space.

Exercise 10.7 (CGO)**.** Let $X \subseteq \mathbb{P}^n$ be a curve. Prove that $\sigma_2(X)$ has dimension 3 unless X is contained in a plane. (This is why every curve is isomorphic to a space curve but only birational to a plane curve.)

Exercise 10.8 (CGO)**.** Let M be an $n \times n$ symmetric matrix of rank r. Prove that M is a sum of r symmetric matrices of rank 1.

Exercise 10.9 (CGO)**.** Consider the rational normal curve in $\mathbb{P}^3$, *i.e.* the twisted cubic curve $X = \nu_3(\mathbb{P}(S_1)) \subset \mathbb{P}(S_3)$. We know that $\sigma_2(X)$ fills up all the space. Can we write any binary cubic as the sum of two cubes of linear forms? Try $x_0x_1^2$.

Exercise 10.10 (CGO)**.** Show that $\sigma_5(\nu_4(\mathbb{P}^2))$ is a hypersurface, *i.e.* that it has dimension equal to 13.

Exercise 10.11 (CGO)**.** Let $X = \mathbb{P}V_1 \times \cdots \times \mathbb{P}V_t$ and let $[v] = [v_1 \otimes \cdots \otimes v_t]$ be a point of X. Show that the cone over the tangent space to X at v is the span of the following vector spaces:

$$\begin{gathered} V_1 \otimes v_2 \otimes v_3 \otimes \cdots \otimes v_t, \\ v_1 \otimes V_2 \otimes v_3 \otimes \cdots \otimes v_t, \\ \vdots \\ v_1 \otimes v_2 \otimes \cdots \otimes v_{t-1} \otimes V_t. \end{gathered}$$

Exercise 10.12. Show that $\mathrm{Seg}(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C) = \mathrm{Sub}_{1,1,1}(A \otimes B \otimes C)$.

Exercise 10.13 (CGO)**.** Show that $\sigma_2(\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1) = \mathbb{P}^7$.

Exercise 10.14 (CGO)**.** Use the description of the tangent space of the Segre product and Terracini's lemma to show that $\sigma_3(\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1)$ is a hypersurface in $\mathbb{P}^{15}$ and not the entire ambient space as expected. This shows that the four-factor Segre product of $\mathbb{P}^1$s is defective.

The connection between subspace varieties and secant varieties is the content of the following exercise.

Exercise 10.15 (CGO)**.** Let $X = \mathbb{P}V_1 \times \cdots \times \mathbb{P}V_t$. Show that if $r \leq r_i$ for $1 \leq i \leq t$, then

$$\sigma_r(X) \subset \mathrm{Sub}_{r_1,\ldots,r_t}\,.$$

Notice that for the 2-factor case, $\sigma_r(\mathbb{P}^{a-1} \times \mathbb{P}^{b-1}) = \mathrm{Sub}_{r,r}$.

The following exercise marked (Sturmfels) appeared in Bernd Sturmfels' course on "Introduction to Non-Linear Algebra" at KAIST, Summer 2014:

Exercise 10.16 (Sturmfels)**.** Consider the tensors T whose coordinates are

$$T_{i,j,k} = i + j + k.$$

Give the "simplest" expression of T by choosing a "good" set of coordinates. Compute the rank and border rank of T.

Try to repeat the exercise for $T_{i,j,k} = i^2 + j^2 + k^2$.

Exercise 10.17 (Sturmfels)**.** Give an example of a real $2 \times 2 \times 2$ tensor whose tensor rank over $\mathbb{C}$ is 2 but whose tensor rank over $\mathbb{R}$ is 3.

Exercise 10.18 (Sturmfels)**.** The monomial $\phi = xyz$ is a symmetric $3 \times 3 \times 3$ tensor. Compute the rank and border rank of ϕ (if this is difficult, try to find upper or lower bounds).

Find symmetric tensors of ranks 1 and 2 that best approximate ϕ.

Exercise 10.19 (Sturmfels)**.** The monomial $x_1x_2\cdots x_n$ is a symmetric tensor of format $n \times n \times \cdots \times n$. Find all eigenvectors of this tensor. Start with the familiar case of $n = 2$.

References

[CGO14] E. Carlini, N. Grieve, and L. Oeding. Four lectures on secant varieties. In Susan M. Cooper and Sean Sather-Wagstaff, editors, *Connections Between Algebra, Combinatorics, and Geometry*, volume 76 of *Springer Proceedings in Mathematics & Statistics*, pages 101–146. Springer New York, 2014. arXiv:1309.4145.

[CO12] P. Comon and G. Ottaviani. On the typical rank of real binary forms. *Linear Multilinear Algebra*, 60(6):657–667, 2012.

[CtBDLC09] P. Comon, J.M. F. ten Berge, L. De Lathauwer, and J. Castaing. Generic and typical ranks of multi-way arrays. *Linear Algebra Appl.*, 430(11-12):2997–3007, 2009.

[DHO+13] J. Draisma, E. Horobet, G. Ottaviani, B. Sturmfels, and R. R. Thomas. The Euclidean distance degree of an algebraic variety. *preprint*, August 2013. arXiv:1309.0049.

[dSL08] V. de Silva and L.-H. Lim. Tensor rank and the ill-posedness of the best low-rank approximation problem. *SIAM J. Matrix Anal. Appl.*, 30:1084–1127, 2008.

[FO14] Shmuel Friedland and Giorgio Ottaviani. The number of singular vector tuples and uniqueness of best rank-one approximation of tensors. *Foundations of Computational Mathematics*, pages 1–34, 2014. arXiv:1210.8316.

[HL09] C. J. Hillar and L.-H. Lim. Most tensor problems are NP hard. *arXiv:0911.1393*, 2009.

[Lan12] J.M. Landsberg. *Tensors: geometry and applications*, volume 128 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

[Lee14] H. Lee. The Euclidean Distance Degree of Fermat Hypersurfaces. *preprint*, September 2014. arXiv:1409.0684.

[Lim05] L.-H. Lim. Singular values and eigenvalues of tensors: a variational approach. In *Computational Advances in Multi-Sensor Adaptive Processing, 2005 1st IEEE International Workshop on*, pages 129 –132, 2005.

[Oed12] L. Oeding. Hyperdeterminants of polynomials. *Adv. Math.*, 231(3-4):1308–1326, 2012. arXiv:1107.4659.

[OO13] L. Oeding and G. Ottaviani. Eigenvectors of tensors and algorithms for Waring decomposition. *J. Symbolic Comput.*, 54:9–35, 2013. arXiv:1103.0203.

[Qi05] L. Qi. Eigenvalues of a real supersymmetric tensor. *J. Symbolic Comput.*, 40(6):1302–1324, 2005.