Tutorial: A brief survey on tensor rank and tensor decomposition, from a geometric perspective.

Workshop

Computational nonlinear Algebra

(June 2-6, 2014)

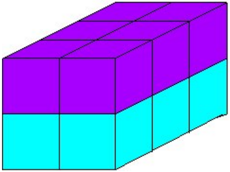
ICERM, Providence

Giorgio Ottaviani

Università di Firenze

Tensors

Let V_i be vector spaces over $K = \mathbb{R}$ or $\mathbb{C}$. A tensor is an element $f \in V_1 \otimes \ldots \otimes V_k$, that is a multilinear map $V_1^\vee \times \ldots \times V_k^\vee \to K$
A tensor can be visualized as a multidimensional matrix.

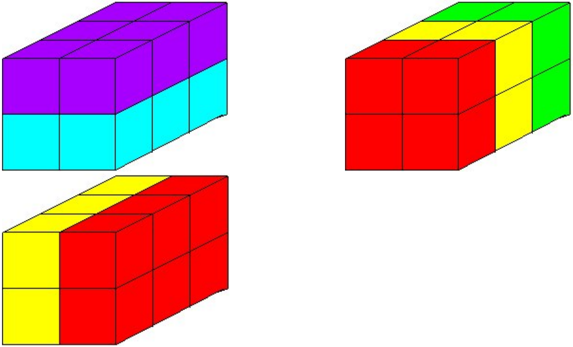


Entries of f are labelled by k indices, as $a_{i_1 \ldots i_k}$
For example, in the case $3 \times 2 \times 2$, with obvious notations, the expression in coordinates of a tensor is

$$a_{000}x_0y_0z_0 + a_{001}x_0y_0z_1 + a_{010}x_0y_1z_0 + a_{011}x_0y_1z_1 +$$
$$a_{100}x_1y_0z_0 + a_{101}x_1y_0z_1 + a_{110}x_1y_1z_0 + a_{111}x_1y_1z_1 +$$
$$a_{200}x_2y_0z_0 + a_{201}x_2y_0z_1 + a_{210}x_2y_1z_0 + a_{211}x_2y_1z_1$$

Slices

Just as matrices can be cutted in rows or in columns, higher dimensional tensors can be cut in slices



The three ways to cut a $3 \times 2 \times 2$ matrix into parallel slices

For a tensor of format $n_1 \times \ldots \times n_k$, there are n_1 slices of format $n_2 \times \ldots \times n_k$.

Multidimensional Gauss elimination

We can operate adding linear combinations of a slice to another slice, just in the case of rows and columns.
This amounts to multiply A of format $n_1 \times \ldots \times n_k$ for $G_1 \in GL(n_1)$, then for $G_i \in GL(n_i)$.

The group acting is quite big
$G = GL(n_1) \times \ldots \times GL(n_k)$.

The group acting, basic computation of dimensions.

The group is big, but not so big...

Let $\dim V_i = n_i$
$\dim V_1 \otimes \ldots \otimes V_k = \prod_{i=1}^k n_i$
$\dim GL(n_1) \times \ldots \times GL(n_k) = \sum_{i=1}^k n_i^2$

For $k \geq 3$, the dimension of the group is in general much less that the dimension of the space where it acts.
This makes a strong difference between the classical case $k = 2$ and the case $k \geq 3$.

Decomposable tensors, of rank one.

We need some "simple" tensors to start with.

Definition

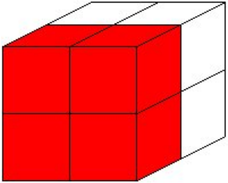
A tensor f is *decomposable* if there exist $x^i \in V_i$ for $i = 1, \ldots, k$ such that $a_{i_1 \ldots i_k} = x^1_{i_1} x^2_{i_2} \ldots x^k_{i_k}$. In equivalent way, $f = x^1 \otimes \ldots \otimes x^k$.

For a (nonzero) usual matrix, decomposable $\iff$ rank one.
Define the rank of a tensor t as

$$\mathrm{rk}(t) := \min\{r | t = \sum_{i=1}^{r} t_i, t_i \text{ are decomposable}\}$$

For matrices, this coincides with usual rank.

Weierstrass Theorem about Tensor Decomposition in $n \times n \times 2$ case



Theorem (Weierstrass)

A general tensor t of format $n \times n \times 2$ has a unique tensor decomposition as a sum of n decomposable tensors

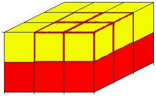
There is a algorithm to actually decompose such tensors. We see how it works in a $3 \times 3 \times 2$ example.

Tensor decomposition in a $3 \times 3 \times 2$ example.

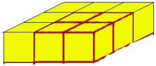
We consider the following "random" real tensor

$$\begin{aligned}
f = & 6x_0y_0z_0 && +2x_1y_0z_0 + 6x_2y_0z_0 \\
& -2014x_0y_1z_0 && +121x_1y_1z_0 - 11x_2y_1z_0 \\
& +48x_0y_2z_0 && -13x_1y_2z_0 - 40x_2y_2z_0 \\
& -31x_0y_0z_1 && +93x_1y_0z_1 + 97x_2y_0z_1 \\
& +63x_0y_1z_1 && +41x_1y_1z_1 - 94x_2y_1z_1 \\
& -3x_0y_2z_1 && +47x_1y_2z_1 + 4x_2y_2z_1
\end{aligned}$$

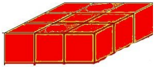
We divide into two 3×3 slices, like in



$\Longrightarrow$ z_0

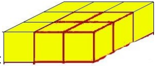
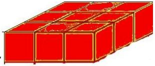


$+z_1$



Two slices

Sum the yellow slice plus t times the red slice.

$f_0 + tf_1 =$  $+t$ 

$$f_0 + tf_1 = \begin{pmatrix} -31t+6 & 63t-2014 & -3t+48 \\ 93t+2 & 41t+121 & 47t-13 \\ 97t+6 & -94t-11 & 4t-40 \end{pmatrix}$$

Singular combination of slices

We compute the determinant, which is a cubic polynomial in t

$$\det(f_0 + tf_1) = 159896t^3 - 8746190t^2 - 5991900t - 69830$$

with roots $t_0 = -.0118594$, $t_1 = -.664996$, $t_2 = 55.3761$.

This computation gives a "guess" about the three summands for z_i, (note the sign change!)

$$f = A_0(.0118594z_0 + z_1) + A_1(.664996z_0 + z_1) + A_2(-55.3761z_0 + z_1)$$

where A_i are 3×3 matrices, that we have to find.
Indeed, we get

$$f_0 + tf_1 = A_0(.0118594 + t) + A_1(.664996 + t) + A_2(-55.3761 + t)$$

and for the three roots $t = t_i$ one summand vanishes, it remains a matrix of rank 2, *with only two colors*, hence with zero determinant.

Finding the three matrices from kernels.

In order to find A_i, let $a_0 = \begin{pmatrix} -.0589718 & -.964899 & .255916 \end{pmatrix}$, left kernel of $f_0 + t_0 f_1$
let $b_0 = \begin{pmatrix} -.992905 & -.00596967 & -.118765 \end{pmatrix}$, transpose of right kernel of $f_0 + t_0 f_1$.
In the same way, denote
$a_1 =$ left kernel of $f_0 + t_1 f_1$, $a_2 =$ left kernel of $f_0 + t_2 f_1$
$b_1 =$ transpose of right kernel of $f_0 + t_1 f_1$, $b_2 =$ transpose of right kernel of $f_0 + t_2 f_1$,

$$aa = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} -.0589718 & -.964899 & .255916 \\ -.014181 & -.702203 & .711835 \\ .959077 & .0239747 & .282128 \end{pmatrix}$$

$$bb = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} -.992905 & -.00596967 & -.118765 \\ .582076 & -.0122361 & -.813043 \\ .316392 & .294791 & -.901662 \end{pmatrix}$$

Now we invert the two matrices

$$aa^{-1} = \begin{pmatrix} .450492 & -.582772 & 1.06175 \\ -1.43768 & .548689 & -.0802873 \\ -1.40925 & 1.93447 & -.0580488 \end{pmatrix}$$

$$bb^{-1} = \begin{pmatrix} -.923877 & .148851 & -.0125305 \\ -.986098 & -3.43755 & 3.22958 \\ -.646584 & -1.07165 & -.0575754 \end{pmatrix}$$

The first summand A_0 is given by a scalar c_0 multiplied by
$(.450492x_0 - 1.43768x_1 - 1.40925x_2)(-.923877y_0 - .986098y_1 - .646584y_2)$
the same for the other colors.

Decomposition as sum of three terms

By solving a linear system, we get the scalars c_i

$(.450492x_0 - 1.43768x_1 - 1.40925x_2)(-.923877y_0 - .986098y_1 - .646584y_2)(.809777z_0 + 68.2814z_1) +$

$(-.582772x_0 + .548689x_1 + 1.93447x_2)(.148851y_0 - 3.43755y_1 - 1.07165y_2)(18.6866z_0 + 28.1003z_1) +$

$(1.06175x_0 - .0802873x_1 - .0580488x_2)(-.0125305y_0 + 3.22958y_1 - .0575754y_2)(-598.154z_0 + 10.8017z_1)$

and the sum is

$$\begin{array}{rr}
6x_0y_0z_0 + 2x_1y_0z_0 & +6x_2y_0z_0 \\
-2014x_0y_1z_0 + 121x_1y_1z_0 & -11x_2y_1z_0 \\
+48x_0y_2z_0 - 13x_1y_2z_0 & -40x_2y_2z_0 \\
-31x_0y_0z_1 + 93x_1y_0z_1 & +97x_2y_0z_1 \\
+63x_0y_1z_1 + 41x_1y_1z_1 & -94x_2y_1z_1 \\
-3x_0y_2z_1 + 47x_1y_2z_1 & +4x_2y_2z_1
\end{array}$$

The rank of the tensor f is 3, because we have 3 summands, and no less.

Uniqueness of the decomposition

The decomposition we have found is *unique*, up to reordering the summands.
This is a strong difference with the case of matrices, where any decomposition with at least two summands is *never unique*.

For tensors f of rank ≤ 2,the characteristic polynomial vanishes identically.
We understand this phenomenon geometrically, in a while.

Coincident roots, hyperdeterminant

What happens if we have two coincident roots in $\det(f_0 + tf_1)$?

In this case, the discriminant of characteristic polynomial vanishes, the discriminant is an invariant of the tensor, called the *hyperdeterminant*.

The hyperdeterminant of format $n \times n \times 2$ has degree $2n(n-1) = 4\binom{n}{2}$.

References
[Gelfand-Kapranov-Zelevinsky] *Discriminants, resultants and multidimensional determinants*, Birkhauser.
[O] *An introduction to the hyperdeterminant and to the rank of multidimensional matrices.* (book chapter, available on arXiv)

The hyperdeterminant of a general tensor

The hyperdeterminant of a tensor $f \in V_1 \otimes V_2 \otimes V_3$ vanishes if and only if there exist nonzero $x^i \in V_i$ such that

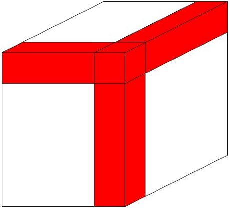
$$f(-, x^2, x^3) = f(x^1, -, x^3) = f(x^1, x^2, -) = 0.$$

It is a codimension 1 condition if the *triangle inequality* holds

$$(\dim V_i - 1) \leq (\dim V_j - 1) + (\dim V_k - 1) \quad \forall i, j, k,$$

which is the assumption for the hyperdeterminant to exist.

A picture is useful. $\mathrm{Det}(f) = 0$ if and only if, after a linear change of coordinates, f is zero on the "red corner".



The generating function for degree of the hyperdeterminant

Let $N(k_0, k_1, k_2)$ be the degree of the hyperdeterminant of format $(k_0 + 1) \times (k_1 + 1) \times (k_2 + 1)$.

Theorem ([GKZ] Thm. XIV 2.4)

$$\sum_{k_0, k_1, k_2 \geq 0} N(k_0, k_1, k_2) z_0^{k_0} z_1^{k_1} z_2^{k_2} = \frac{1}{(1 - (z_0 z_1 + z_0 z_2 + z_1 z_2) - 2 z_0 z_1 z_2)^2}$$

List of degree of hyperdeterminants of format (a, b, c)

format	degree	boundary format
$(1, a, a)$	a	$*$
$(2, 2, 2)$	4	
$(2, 2, 3)$	6	$*$
$(2, 3, 3)$	12	
$(2, 3, 4)$	12	$*$
$(2, 4, 4)$	24	
$(2, 4, 5)$	20	
$(3, 3, 3)$	36	
$(3, 3, 4)$	48	
$(3, 3, 5)$	30	$*$
$(3, 4, 4)$	108	
$(3, 4, 5)$	120	
$(4, 4, 4)$	272	
$(2, b, b)$	$2b(b-1)$	
$(2, b, b+1)$	$b(b+1)$	$*$
$(a, b, a+b-1)$	$\frac{(a+b-1)!}{(a-1)!(b-1)!}$	$*$

Symultaneous diagonalization, Corollary of Weierstrass Theorem

Corollary

For any tensor f of format $n \times n \times 2$, such that $\mathrm{Det}(f) \neq 0$, with slices f_0, f_1, there are invertibles matrix $G, H \in GL(n)$ such that Gf_iH is diagonal for $i = 1, 2$. Gf_0H may be assumed to be the identity.

Expression of hyperdeterminant
If $Gf_0H = \mathrm{Id}_n$, $Gf_1H = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ then

$$\mathrm{Det}(GfH) = \prod_{i<j} (\lambda_i - \lambda_j)^2.$$

The rank may depend on the field.

- What happens if we have a pair of complex imaginary roots ?

On complex numbers, we still have rank 3. But on real numbers, the rank becomes 4 [tenBerge, 2000].
On $3 \times 3 \times 2$ case, this is governed by the sign of hyperdeterminant.
Unless a set of measure zero, the following holds

$$\begin{cases} Det(f) > 0 \Longrightarrow \mathrm{rk}_{\mathbb{R}}(f) = 3 \\ Det(f) < 0 \Longrightarrow \mathrm{rk}_{\mathbb{R}}(f) = 4. \end{cases}$$

If the tensor is chosen randomly, according to normal distribution, the probability to get rank 3 is exactly $\frac{1}{2}$ [Bergqvist, 2011].

The rank may depend on the field, in contrast to the matrix case.

Typical real ranks of $n \times n \times 2$ tensors.

Ranks which are attained in subsets of positive measure are called *typical ranks*. On $\mathbb{C}$ there is only one typical rank. On $\mathbb{R}$ there may be several typical ranks, the smallest one coincide with the complex one.

tenBerge proves that for format $n \times n \times 2$, the typical ranks are n or $n + 1$, depending on the characteristic polynomial having n real roots or not (the condition is that the Bezoutian must be positive definite).

Semialgebraic sets and best rank approximation

So in $3 \times 3 \times 2$ case, the hyperdeterminant divides the space in two regions, where the real rank is 3 or 4. But the rank on the hypersurface can be 1,2 or 4, never 3. So for tensors of rank 4, the best rank three approximation *does not exist* on real numbers.

Best rank one approximation

The *distance* of our tensor f of format $3 \times 3 \times 2$ from the three summands of its tensor decomposition , according to the L_2-norm (euclidean), is respectively

2031.02 , 2071.18 , 4427.47.

May we have a smaller distance to other rank one tensors ?

In order to find the best rank one approximation of f we may compute all critical points x for the distance from f to the variety of rank 1 matrices. The condition is that the tangent space at x is orthogonal to the vector $f - x$.

Recall SVD and Eckart-Young theorem

Any matrix A has the SVD decomposition

$$A = U\Sigma V^t$$

where U, V are orthogonal and $\sigma = \mathrm{Diag}(\sigma_1, \sigma_2, \ldots)$, with $\sigma_1 \geq \sigma_2 \geq \ldots$. Decomposing $\Sigma = \mathrm{Diag}(\sigma_1, 0, 0, \ldots) + \mathrm{Diag}(0, \sigma_2, 0, \ldots) + \ldots = \Sigma_1 + \Sigma_2 + \ldots$ we find

$$A = U\Sigma_1 V^t + U\Sigma_2 V^t + \ldots$$

Theorem (Eckart-Young, 1936)

- *$U\Sigma_1 V^t$ is the best rank 1 approximation of A, that is $|A - U\Sigma_1 V^t| \leq |A - X|$ for every rank 1 matrix X.*
- *$U\Sigma_1 V^t + U\Sigma_2 V^t$ is the best rank 2 approximation of A, that is $|A - U\Sigma_1 V^t - U\Sigma_2 V^t| \leq |A - X|$ for every rank ≤ 2 matrix X.*
- *So on, for any rank.*

Best rank approximation and tensor decomposition

Among the infinitely many tensor decompositions available for matrices, Eckart-Young Theorems detects one of them, which is particularly nice in optimization problems.

For tensors we have no choices, because the tensor decomposition is often unique (precise statement later). It is unique in $n \times n \times 2$ case. Does it help in best rank approximation ? The answer is negative, due to a subtle fact we are going to explain.

Eckart-Young revisited.

In the SVD $A = U\Sigma V^t$, the columns u_i of U and v_i of V satisfy the conditions $Av_i = \sigma_i u_i$, $A^t u_i = \sigma_i v_i$.
(u_i, v_i) is called a singular vector pair. They are all the critical points of the distance from A to the variety of rank one matrices.

Theorem (Eckart-Young revisited)

All critical points of the distance from A to the variety of rank $\leq r$ matrices are given by $U\Sigma_{i_1}V^t + \ldots + U\Sigma_{i_r}V^t$, their number is $\binom{n}{r}$.

Lim (2000), variational principle

Looking at critical points of the distance, for tensors of format $m_1 \times \ldots \times m_d$ we get singular vector d-ples, a notion analogous to singular vector pairs for matrices.

Theorem (Lim)

The critical points of the distance from $f \in \mathbb{R}^n \otimes \mathbb{R}^n \otimes \mathbb{R}^2$ *to the variety of rank* 1 *tensors are given by triples* $(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^2$ *such that*

- $f \cdot (x \otimes y) = \lambda z,$
- $f \cdot (y \otimes z) = \lambda x,$
- $f \cdot (z \otimes x) = \lambda y.$

$(x \otimes y \otimes z)$ in Lim Theorem is called a *singular vector triple* (defined independently by Qi). λ is called a *singular value*.

15 singular vector triples for our tensor f of format $3 \times 3 \times 2$

We may compute all singular vector triples for

$$\begin{aligned} f = & 6x_0y_0z_0 & +2x_1y_0z_0 + 6x_2y_0z_0 \\ & -2014x_0y_1z_0 & +121x_1y_1z_0 - 11x_2y_1z_0 \\ & +48x_0y_2z_0 & -13x_1y_2z_0 - 40x_2y_2z_0 \\ & -31x_0y_0z_1 & +93x_1y_0z_1 + 97x_2y_0z_1 \\ & +63x_0y_1z_1 & +41x_1y_1z_1 - 94x_2y_1z_1 \\ & -3x_0y_2z_1 & +47x_1y_2z_1 + 4x_2y_2z_1 \end{aligned}$$

We find 15 singular vector triples, 9 of them are real, 6 of them make 3 conjugate pairs.
The minimum distance is 184.038, and the best rank one approximation is given by the singular vector triple
$(x_0 - .0595538x_1 + .00358519x_2)(y_0 - 289.637y_1 + 6.98717)(6.95378z_0 - .2079687z_1)$.
It is unrelated to the three summands of tensor decomposition, in contrast with Eckart-Young Theorem for matrices.

Why Eckart-Young theorem does not hold for tensors ?

The way Eckart-Young generalizes to tensors is more subtle.

Theorem (Draisma-Horobet-O-Sturmfels-Thomas)

The 15 *critical points* p_i *satisfy*

$$\mathrm{Det}(f - p_i) = 0$$

It is part of a more general theory about critical points *(after coffee break!)*.

The phenomenon of the Theorem was first found by [Stegeman-Comon] in $2 \times 2 \times 2$ case, where they showed by examples that *subtracting the best rank* 1 *approximation, may increase the tensor rank !*

In case $n \times n \times 2$, there are $\binom{2n}{2} = n(2n-1)$ critical values,

Why geometry?

Corrado Segre in XIX century understood the tensor decomposition involved in Weierstrass Theorem in terms of projective geometry.
The tensor t is a point of the space $\mathbb{P}(\mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^2)$.
The decomposable tensors make the "Segre variety"

$$X = \mathbb{P}(\mathbb{C}^3) \times \mathbb{P}(\mathbb{C}^3) \times \mathbb{P}(\mathbb{C}^2) \to \mathbb{P}(\mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^2)$$

From f there is a unique secant plane meeting X in three points.
This point of view is extremely useful also today.

J.M. Landsberg, *Tensors: Geometry and Applications*, AMS 2012

Secant varieties

Secant varieties give basic interpretation of rank of tensors in Geometry.

Let $X \subset \mathbb{P}V$ be irreducible variety.

$$\sigma_k(X) := \overline{\bigcup_{x_1,\ldots,x_k \in X} < x_1, \ldots, x_k >}$$

where $< x_1, \ldots, x_k >$ is the projective span.

There is a filtration $X = \sigma_1(X) \subset \sigma_2(X) \subset \ldots$

This ascending chain stabilizes when it fills the ambient space.

So $\min\{k | \sigma_k(X) = \mathbb{P}V\}$ is called the generic X-rank.

Terracini Lemma

Terracini Lemma describes the tangent space at a secant variety

Lemma

Terracini Let $z \in < x_1, \ldots, x_k >$ *be general. Then* $T_z\sigma_k(X) =< T_{x_1}X, \ldots, T_{x_k}X >$

$X = \mathbb{P}V \times \mathbb{P}W$
Then $\sigma_k(X)$ parametrizes linear maps $V^\vee \to W$ of rank $\leq k$.
In this case the Zariski closure is not necessary, the union is already closed.

Eckart-Young Theorem may be understood in this setting.

Dual varieties

If $X \subset \mathbb{P}V$ then

$$X^\vee := \overline{\{H \in \mathbb{P}V^\vee | \exists \text{ smooth point } x \in X \text{ s.t. } T_xX \subset H\}}$$

is called the *dual variety* of X. So $X^\vee$ consists of hyperplanes tangent at some smooth point of X.
By Terracini Lemma

$$\sigma_k(X)^\vee = \{H \in PV^\vee | H \supset T_{x_1}X, \ldots, T_{x_k}X \text{ for smooth points } x_1, \ldots, x_k\}$$

namely, $\sigma_k(X)^\vee$ consists of hyperplanes tangent at $\geq k$ smooth points of X.

Duality in euclidean setting

In euclidean setting, duality may be understood in terms of orthogonality.

Considering the affine cone of a projective variety X, the dual variety consists of the cone of all vectors which are orthogonal to some tangent space to X.

Basic dual varieties

The dual variety of $m \times n$ matrices of rank r is given by $m \times n$ matrices of corank r. In particular the dual of the Segre variety of matrices of rank 1 is the determinant hypersurface.
The determinant can be defined by means of projective geometry!
The dual variety of tensors of format $(m_0 + 1) \times (m_1 + 1) \times (m_2 + 1)$ is the hyperdeterminant hypersurface, whenever $m_i \leq m_j + m_k \ \forall i, j, k$.

Expected dimension for secant varieties

Let $X \subset \mathbb{P}^N$ be an irreducible variety.
The naive dimensional count says that

$$\dim \sigma_k(X) + 1 \leq k(\dim X + 1)$$

When $\dim \sigma_k(X) = \min\{N, k(\dim X + 1) - 1\}$ then we say that $\sigma_k(X)$ has the expected dimension. Otherwise we say that X is k-defective.
Correspondingly, the expected value for the general X-rank is

$$\lceil \frac{N+1}{\dim X + 1} \rceil$$

In defective cases, the general X-rank can be bigger than the expected one.

Basic dimensional computation

If $\sigma_k(X)$ has the virtual dimension $k(\dim X + 1) - 1\}$, then the general tensor of rank k has only *finitely many* decompositions.

This assumption is never satisfied for matrices, when $k \geq 2$. It is likely satisfied for many interesting classes of tensors.

Symmetric tensors = homogeneous polynomials

In the case $V_1 = \ldots = V_k = V$ we may consider symmetric tensors $f \in S^d V$.
Elements of $S^d V$ can be considered as homogeneous polynomials of degree d in $x_0, \ldots x_n$, basis of V.
So polynomials have rank (as all tensors) and also symmetric rank (next slides).

Symmetric Tensor Decomposition (Waring)

A *Waring decomposition* of $f \in S^d V$ is

$$f = \sum_{i=1}^{r} c_i (l_i)^d \qquad \text{with } l_i \in V$$

with minimal r

Example: $7x^3 - 30x^2y + 42xy^2 - 19y^3 = (-x + 2y)^3 + (2x - 3y)^3$
$\mathrm{rk}\left(7x^3 - 30x^2y + 42xy^2 - 19y^3\right) = 2$

Symmetric case: the Alexander-Hirschowitz Theorem

Theorem (Campbell, [1891] Terracini, [1916] Alexander-Hirschowitz [1995])

The general $f \in S^d\mathbb{C}^{n+1}$ $(d \geq 3)$ has rank

$$\lceil \frac{\binom{n+d}{d}}{n+1} \rceil$$

which is called the generic rank, with the only exceptions

- $S^4\mathbb{C}^{n+1}$, $\quad 2 \leq n \leq 4$, *where the generic rank is* $\binom{n+2}{2}$
- $S^3\mathbb{C}^5$, *where the generic rank is* 8, *sporadic case*

Toward an Alexander-Hirschowitz Theorem in the non symmetric case

Defective examples

dim $V_i = n_i + 1$, $n_1 \leq \ldots \leq n_k$

Only known examples where the general $f \in V_1 \otimes \ldots \otimes V_k$ $(k \geq 3)$ has rank different from the *generic rank*

$$\lceil \frac{\prod (n_i + 1)}{\sum n_i + 1} \rceil$$

are

- unbalanced case, where $n_k \geq \prod_{i=1}^{k-1}(n_i + 1) - \left(\sum_{i=1}^{k-1} n_i\right) + 1$, note that for $k = 3$ it is simply $n_3 \geq n_1 n_2 + 2$
- $k = 3$, $(n_1, n_2, n_3) = (2, m, m)$ with m even [Strassen],
- $k = 3$, $(n_1, n_2, n_3) = (2, 3, 3)$, sporadic case [Abo-O-Peterson]
- $k = 4$, $(n_1, n_2, n_3, n_4) = (1, 1, n, n)$

Results in the general case

Theorem (Strassen-Lickteig)

there are no exceptions (no defective cases) $\mathbb{P}^n \times \mathbb{P}^n \times \mathbb{P}^n$ *beyond the variety* $\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2$

Theorem

- *The unbalanced case is completely understood [Catalisano-Geramita-Gimigliano].*
- *The exceptions listed in the previous slide are the only ones in the cases:*
  (i) $k = 3$ *and* $n_i \leq 9$
  (ii) $s \leq 6$ *[Abo-O-Peterson]*
  (iii) $\forall k,\ n_i = 1$ *(deep result, [Catalisano-Geramita-Gimigliano])*

Proof uses an inductive technique, developed first for $k = 3$ in [Bürgisser-Claussen-Shokrollai].

Asymptotical behaviour

[Abo-O-Peterson]

Asymptotically ($n \to \infty$), the general rank for tensors in $\mathbb{C}^{n+1} \otimes \ldots \otimes \mathbb{C}^{n+1}$ (k times) tends to

$$\frac{(n+1)^k}{nk+1}$$

as expected.

Symmetric Rank and Comon Conjecture

The minimum number of summands in a Waring decomposition is called the symmetric rank

Comon Conjecture

Let t be a symmetric tensor. Are the rank and the symmetric rank of t equal ? Comon conjecture gives affirmative answer.

Known to be true when $t \in S^d\mathbb{C}^{n+1}$, $n = 1$ or $d = 2$ and few other cases.

The problem of counting singular d-ples

How many are the singular d-ples of a general tensor?

In the format $(2,2,2)$ they are 6, in the format $(3,3,3)$ they are 37. Note they are more than the dimension of the factors, and even more than the dimension of the ambient space.

The number of singular vector d-ples

Theorem (Friedland-O)

The number of singular d-ples of a general tensor $t \in \mathbb{P}(\mathbb{R}^{m_1}) \times \ldots \times \mathbb{P}(\mathbb{R}^{m_d})$ over $\mathbb{C}$ of format $(m_1, \ldots, m_d)$ is equal to the coefficient of $\prod_{i=1}^{d} t_i^{m_i - 1}$ in the polynomial

$$\prod_{i=1}^{d} \frac{\hat{t}_i^{\,m_i} - t_i^{m_i}}{\hat{t}_i - t_i}$$

where $\hat{t}_i = \sum_{j \neq i} t_j$

Amazingly, for $d = 2$ this formula gives the expected value $\min(m_1, m_2)$.

Interpretation with vector bundles

For the proof, we express the d-ples of singular vectors as zero loci of sections of a suitable vector bundle on the Segre variety.
Precisely, let $X = \mathbb{P}(\mathbb{C}^{m_1}) \times \ldots \times \mathbb{P}(\mathbb{C}^{m_d})$ and let $\pi_i \colon X \to \mathbb{P}(\mathbb{C}^{m_i})$ be the projection on the i-th factor. Let $\mathcal{O}(\underbrace{1,\ldots,1}_{d})$ be the very ample line bundle which gives the Segre embedding.

Then the bundle is $\oplus_{i=1}^{d}(\pi_i^* Q)(1, 1, \ldots, 1, \underset{\substack{\uparrow \\ i}}{0}, 1, \ldots, 1)$

We may conclude with a Chern class computation.

In the format $(\underbrace{2,\ldots,2}_{d})$ the number of singular d-ples is $d!$.

List for tensors of order 3

List of the number of singular triples in the format (d_1, d_2, d_3)

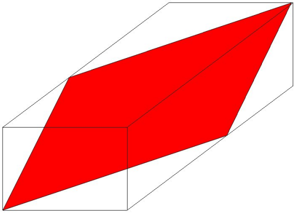
d_1, d_2, d_3	$c(d_1, d_2, d_3)$	
$2, 2, 2$	6	
$2, 2, n$	8	$n \geq 3$
$2, 3, 3$	15	
$2, 3, n$	18	$n \geq 4$
$3, 3, 3$	37	
$3, 3, 4$	55	
$3, 3, n$	61	$n \geq 5$
$3, 4, 4$	104	
$3, 4, 5$	138	
$3, 4, n$	148	$n \geq 6$

The stabilization property

The output stabilizes for (a, b, c) with $c \geq a + b - 1$.
For a tensor of size $2 \times 2 \times n$ there are 6 singular vector triples for $n = 2$ and 8 singular vector triples for $n > 2$.
The format $(a, b, a + b - 1)$ is the *boundary format*, well known in hyperdeterminant theory [Gelfand-Kapranov-Zelevinsky]. It generalizes the square case, a equality holds in triangle inequality.

The diagonal is well defined in the boundary format case.

In the boundary format it is well defined a unique "diagonal" given by the elements $a_{i_1 \ldots i_d}$ which satisfy $i_1 = \sum_{j=2}^{d} i_j$



(indices start from zero).

The symmetric case

Theorem (Cartwright-Sturmfels)

In the symmetric case, a tensor in $S^d(\mathbb{C}^m)$ has

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

singular vectors (which can be called eigenvectors).

For $d = m = 3$ the number of eigenvectors is 7. In general we compute [Oeding-O]

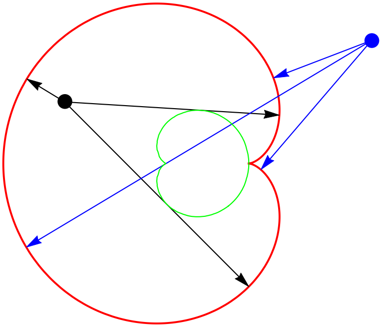
$$c_{m-1}(T\mathbb{P}^{m-1}(d-2)) = \frac{(d-1)^m - 1}{d-2}$$

The first proof of the formula in the symmetric case has been given by [Cartwright-Sturmfels] through the computation of a toric volume. It counts the number of eigenvectors of a symmetric tensor.

We have the same geometric interpretation with the Veronese

Euclidean Distance Degree

The construction of critical points of the distance from a point u, can be generalized to any affine (real) algebraic variety.
We call Euclidean Distance Degree (shortly *ED degree*) the number of critical points of $d_u = d(u, -)\colon X \to \mathbb{R}$. As before, the number of critical points does not depend on u, provided u is generic.



Look at Wikipedia animation on "evolute".

Duality for ED

Theorem (Draisma-Horobet-O-Sturmfels-Thomas)

There is a canonical bijection between

- *critical points of the distance from p to rank ≤ 1*
- *critical points of the distance from p to hyperdeterminant hypersurface.*

Correspondence is $x \mapsto p - x$

In particular from the 15 critical points for the distance from our $3 \times 3 \times 2$ tensor f to the variety of rank one matrices, we may recover the 15 critical points for the distance from f to hyperdeterminant hypersurface.

Duality for ED, in generality

Theorem (Draisma-Horobet-O-Sturmfels-Thomas)

There is a canonical bijection between

- *critical points of the distance from p to a projective variety X*
- *critical points of the distance from p to the dual variety $X^\vee$.*

Correspondence is $x \mapsto p - x$. In particular
$\mathrm{EDdegree}(X) = \mathrm{EDdegree}(X^\vee)$

The Catanese-Trifogli formula

There is a formula, due to Catanese and Trifogli, for ED degree in terms of Chern classes, provided X is transversal to the quadric $\sum x_i^2 = 0$ of isotropic vectors.
Applying this formula to $n \times n$ matrices of rank 1, $n \geq 2$ we get $4, 13, 40, 121, \ldots$ instead of $2, 3, 4, 5, \ldots$. Why ?
Applying this formula to tensors of rank one and format $2 \times 2 \times 2$ we get 34 instead of the expected 6. Why ?
The reason is that the transversality with respect to the quadric is NOT satisfied. ED degree is invariant by orthogonal transformations, but not by general linear projective transformations.
So the approach considered in [O-Friedland] has to be considered counting critical points for tensors.

Apolarity and Waring decomposition, I

For any $l = \alpha x_0 + \beta x_1 \in \mathbb{C}^2$ we denote $l^\perp = -\beta\partial_0 + \alpha\partial_1 \in \mathbb{C}^{2^\vee}$.
Note that

$$l^\perp(l^d) = 0 \tag{1}$$

so that $l^\perp$ is well defined (without referring to coordinates) up to scalar multiples. Let e be an integer. Any $f \in S^d\mathbb{C}^2$ defines $C_f^e \colon S^e(\mathbb{C}^{2^\vee}) \to S^{d-e}\mathbb{C}^2$
Elements in $S^e(\mathbb{C}^{2^\vee})$ can be decomposed as $(l_1^\perp \circ \ldots \circ l_e^\perp)$ for some $l_i \in \mathbb{C}^2$.

Proposition

Let l_i be distinct for $i = 1, \ldots, e$. There are $c_i \in K$ such that $f = \sum_{i=1}^{e} c_i (l_i)^d$ if and only if $(l_1^\perp \circ \ldots \circ l_e^\perp) f = 0$

Proof: The implication $\Longrightarrow$ is immediate from (1). It can be summarized by the inclusion $< (l_1)^d, \ldots, (l_e)^d > \subseteq \ker(l_1^\perp \circ \ldots \circ l_e^\perp)$. The other inclusion follows by dimensional reasons, because both spaces have dimension e. $\square$

The previous Proposition is the core of the Sylvester algorithm, because the differential operators killing f allow to define the decomposition of f, as we see in the next slide.

Sylvester algorithm for general f Compute the decomposition of a general $f \in S^d U$

- Pick a generator g of $\ker C_f^a$ with $a = \lfloor \frac{d+1}{2} \rfloor$.
- Decompose g as product of linear factors, $g = (l_1^\perp \circ \ldots \circ l_r^\perp)$
- Solve the system $f = \sum_{i=1}^r c_i (l_i)^d$ in the unknowns c_i.

Remark When d is odd the kernel is one-dimensional and the decomposition is unique. When d is even the kernel is two-dimensional and there are infinitely many decompositions.

The catalecticant matrices for two variables

If $f(x,y) = a_0x^4 + 4a_1x^3y + 6a_2x^2y^2 + 4a_3xy^3 + a_4y^4$ then

$$C_f^1 = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 \\ a_1 & a_2 & a_3 & a_4 \end{bmatrix}$$

and

$$C_f^2 = \begin{bmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{bmatrix}$$

The catalecticant algorithm at work

The catalecticant matrix associated to
$f = 7x^3 - 30x^2 + 42x - 19 = 0$ is

$$A_f = \begin{bmatrix} 7 & -10 & 14 \\ -10 & 14 & -19 \end{bmatrix}$$

$\ker A_f$ is spanned by $\begin{bmatrix} 6 \\ 7 \\ 2 \end{bmatrix}$ which corresponds to
$6\partial_x^2 + 7\partial_x\partial_y + 2\partial_y^2 = (2\partial_x + \partial_y)(3\partial_x + 2\partial_y)$

Hence the decomposition

$$7x^3 - 30x^2y + 42xy^2 - 19y^3 = c_1(-x+2y)^3 + c_2(2x-3y)^3$$

Solving the linear system, we get $c_1 = c_2 = 1$

Another example, Waring decomposition of a quintic in three variables, $3 \times 3 \times 3 \times 3 \times 3$ symmetric tensor.

Hilbert, 1888

The general f of order 5 in three variables has a unique decomposition as a sum of seven powers of linear forms.

As an example we pick

$$f = 19x_0^5 + 25x_0^4x_1 + 44x_0^3x_1^2 + 35x_0^2x_1^3 + 30x_0x_1^4 + 36x_1^5 + 38x_0^4x_2 + 50x_0^3x_1x_2 - 20x_0^2x_1^2x_2 + 27x_0x_1^3x_2 + 14x_1^4x_2 - 23x_0^3x_2^2 + 10x_0^2x_1x_2^2 + 45x_0x_1^2x_2^2 - 13x_1^3x_2^2 + 11x_0^2x_2^3 - 29x_0x_1x_2^3 + 29x_1^2x_2^3 + 13x_0x_2^4 - 28x_1x_2^4 + 34x_2^5$$

Question

How to construct explicitly $f = \sum_{i=1}^{7} c_i l_i^5$, with $c_i \in \mathbb{C}$ $l_i = a_i x_0 + b_i x_1 + c_i x_2$?

We answer to this question presenting an algorithm (joint works with Landsberg, Oeding). A related powerful approach is due to Bernardi, Brachat, Comon, Mourrain, Tsigaridas.

The contraction map P_f

$Hom(S^2\mathbb{C}^3, \mathbb{C}^3)$ represents tensors of order 3 partially symmetric in two indices.
We construct the map

$$Hom(S^2\mathbb{C}^3, \mathbb{C}^3) \xrightarrow{P_f} Hom(\mathbb{C}^3, S^2\mathbb{C}^3)$$

if $f = v^5$, $g \in Hom(S^2\mathbb{C}^3, \mathbb{C}^3)$

$$P_{v^5}(g)(w) := \left(g(v^2) \wedge v \wedge w\right) v^2 \qquad (2)$$

and then extended by linearity.
This means $P_{\sum_i c_i v_i^5} = \sum_i c_i P_{v_i^5}$
The formula (2) is the key to understand the connection between tensor decomposition and eigenvectors.

Lemma

$P_{v^5}(M) = 0$ if and only if there exists λ such that $M(v^2) = \lambda v$.

If all v_i are eigenvectors of g then $g \in \ker P_{\sum_i c_i v_i^5}$.
So we have candidates to decompose f: compute the eigenvectors of $\ker P_f$.
Luckily P_f can be computed without knowing the decomposition $\sum_i c_i v_i^5$.
P_f is given by a 18×18 matrix and now we construct it.
We compute the three partials

$\frac{\partial f}{\partial x_0} = 95x_0^4 + 100x_0^3x_1 + 132x_0^2x_1^2 + 70x_0x_1^3 + 30x_1^4 + 152x_0^3x_2 + 150x_0^2x_1x_2 - 40x_0x_1^2x_2 + 27x_1^3x_2 - 69x_0^2x_2^2 + 20x_0x_1x_2^2 + 45x_1^2x_2^2 + 22x_0x_2^3 - 29x_1x_2^3 + 13x_2^4$

$\frac{\partial f}{\partial x_1} = 25x_0^4 + 88x_0^3x_1 + 105x_0^2x_1^2 + 120x_0x_1^3 + 180x_1^4 + 50x_0^3x_2 - 40x_0^2x_1x_2 + 81x_0x_1^2x_2 + 56x_1^3x_2 + 10x_0^2x_2^2 + 90x_0x_1x_2^2 - 39x_1^2x_2^2 - 29x_0x_2^3 + 58x_1x_2^3 - 28x_2^4$

$\frac{\partial f}{\partial x_2} = 38x_0^4 + 50x_0^3x_1 - 20x_0^2x_1^2 + 27x_0x_1^3 + 14x_1^4 - 46x_0^3x_2 + 20x_0^2x_1x_2 + 90x_0x_1^2x_2 - 26x_1^3x_2 + 33x_0^2x_2^2 - 87x_0x_1x_2^2 + 87x_1^2x_2^2 + 52x_0x_2^3 - 112x_1x_2^3 + 170x_2^4$

The catalecticant (Sylvester)

To any quartic we can associate the catalecticant matrix constructed in the following way

$$\begin{array}{c|cccccc} & \partial_{00} & \partial_{01} & \partial_{02} & \partial_{11} & \partial_{12} & \partial_{22} \\ \hline \partial_{00} & & & & & & \\ \partial_{01} & & & & & & \\ \partial_{02} & & & & & & \\ \partial_{11} & & & & & & \\ \partial_{12} & & & & & & \\ \partial_{22} & & & & & & \end{array}$$

$rank(f) = rank(C_f)$ it relates the rank of a tensor with the rank of a usual matrix.

The three catalecticant matrices corresponding to the three partial derivatives $\frac{\partial f}{\partial x_0}$, $\frac{\partial f}{\partial x_1}$, $\frac{\partial f}{\partial x_2}$ are
$$\begin{pmatrix} 2280 & 600 & 912 & 528 & 300 & -276 \\ 600 & 528 & 300 & 420 & -80 & 40 \\ 912 & 300 & -276 & -80 & 40 & 132 \\ 528 & 420 & -80 & 720 & 162 & 180 \\ 300 & -80 & 40 & 162 & 180 & -174 \\ -276 & 40 & 132 & 180 & -174 & 312 \end{pmatrix}$$
$$\begin{pmatrix} 600 & 528 & 300 & 420 & -80 & 40 \\ 528 & 420 & -80 & 720 & 162 & 180 \\ 300 & -80 & 40 & 162 & 180 & -174 \\ 420 & 720 & 162 & 4320 & 336 & -156 \\ -80 & 162 & 180 & 336 & -156 & 348 \\ 40 & 180 & -174 & -156 & 348 & -672 \end{pmatrix}$$
$$\begin{pmatrix} 912 & 300 & -276 & -80 & 40 & 132 \\ 300 & -80 & 40 & 162 & 180 & -174 \\ -276 & 40 & 132 & 180 & -174 & 312 \\ -80 & 162 & 180 & 336 & -156 & 348 \\ 40 & 180 & -174 & -156 & 348 & -672 \\ 132 & -174 & 312 & 348 & -672 & 4080 \end{pmatrix}$$

The construction of P_f

We get, in exact arithmetic, that the 18×18 matrix of P_f is the following block matrix

$$\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 912 & 300 & -276 & -80 & 40 & 132 & -600 & -528 & -300 & -420 & 80 & -40 \\
0 & 0 & 0 & 0 & 0 & 0 & 300 & -80 & 40 & 162 & 180 & -174 & -528 & -420 & 80 & -720 & -162 & -180 \\
0 & 0 & 0 & 0 & 0 & 0 & -276 & 40 & 132 & 180 & -174 & 312 & -300 & 80 & -40 & -162 & -180 & 174 \\
0 & 0 & 0 & 0 & 0 & 0 & -80 & 162 & 180 & 336 & -156 & 348 & -420 & -720 & -162 & -4320 & -336 & 156 \\
0 & 0 & 0 & 0 & 0 & 0 & 40 & 180 & -174 & -156 & 348 & -672 & 80 & -162 & -180 & -336 & 156 & -348 \\
0 & 0 & 0 & 0 & 0 & 0 & 132 & -174 & 312 & 348 & -672 & 4080 & -40 & -180 & 174 & 156 & -348 & 672 \\
-912 & -300 & 276 & 80 & -40 & -132 & 0 & 0 & 0 & 0 & 0 & 0 & 2280 & 600 & 912 & 528 & 300 & -276 \\
-300 & 80 & -40 & -162 & -180 & 174 & 0 & 0 & 0 & 0 & 0 & 0 & 600 & 528 & 300 & 420 & -80 & 40 \\
276 & -40 & -132 & -180 & 174 & -312 & 0 & 0 & 0 & 0 & 0 & 0 & 912 & 300 & -276 & -80 & 40 & 132 \\
80 & -162 & -180 & -336 & 156 & -348 & 0 & 0 & 0 & 0 & 0 & 0 & 528 & 420 & -80 & 720 & 162 & 180 \\
-40 & -180 & 174 & 156 & -348 & 672 & 0 & 0 & 0 & 0 & 0 & 0 & 300 & -80 & 40 & 162 & 180 & -174 \\
-132 & 174 & -312 & -348 & 672 & -4080 & 0 & 0 & 0 & 0 & 0 & 0 & -276 & 40 & 132 & 180 & -174 & 312 \\
600 & 528 & 300 & 420 & -80 & 40 & -2280 & -600 & -912 & -528 & -300 & 276 & 0 & 0 & 0 & 0 & 0 & 0 \\
528 & 420 & -80 & 720 & 162 & 180 & -600 & -528 & -300 & -420 & 80 & -40 & 0 & 0 & 0 & 0 & 0 & 0 \\
300 & -80 & 40 & 162 & 180 & -174 & -912 & -300 & 276 & 80 & -40 & -132 & 0 & 0 & 0 & 0 & 0 & 0 \\
420 & 720 & 162 & 4320 & 336 & -156 & -528 & -420 & 80 & -720 & -162 & -180 & 0 & 0 & 0 & 0 & 0 & 0 \\
-80 & 162 & 180 & 336 & -156 & 348 & -300 & 80 & -40 & -162 & -180 & 174 & 0 & 0 & 0 & 0 & 0 & 0 \\
40 & 180 & -174 & -156 & 348 & -672 & 276 & -40 & -132 & -180 & 174 & -312 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

Theorem

$$r(f) \geq \frac{\mathrm{rank}\,(P_f)}{2}$$

Note that $\mathrm{rank}\,(P_f)$ is even because P_f is skew-symmetric. Equality holds for f general in the variety of tensors with assigned rank.

It relates the rank of the tensor f with the rank of the matrix P_f.

The decomposition of f

We substitute the seven eigenvectors already computed

$$f = c_0(x_0 + 7.97577x_1 + 1.82513x_2)^5 +$$
$$c_1(x_0 + x_1(-6.7325 + 2.91924\sqrt{-1}) + x_2(-3.49842 - 3.27128\sqrt{-1}))^5 +$$
$$c_2(x_0 + x_1(-6.7325 - 2.91924\sqrt{-1}) + x_2(-3.49842 + 3.27128\sqrt{-1}))^5 +$$
$$c_3(x_0 + (.39844)x_1 + (.112957)x_2)^5 +$$
$$c_4(x_0 + x_1(.122478 + .537715\sqrt{-1}) + x_2(-.436832 - .342586\sqrt{-1}))^5 +$$
$$c_5(x_0 + x_1(.122478 - .537715\sqrt{-1}) + x_2(-.436832 + .342586\sqrt{-1}))^5 +$$
$$c_6(x_0 + (-2.94762)x_1 + (12.5538)x_2)^5$$

We need just to solve a square system in the seven unknowns $c_0 \ldots c_6$.

This is the Waring decomposition of f

$$f = .0011311(x_0 + 7.97577x_1 + 1.82513x_2)^5 +$$
$$(.000199669 + .000111056\sqrt{-1})(x_0 + x_1(-6.7325 + 2.91924\sqrt{-1}) + x_2(-3.49842 - 3.27128\sqrt{-1}))^5 +$$
$$(+.000199669 - .000111056\sqrt{-1})(x_0 + x_1(-6.7325 - 2.91924\sqrt{-1}) + x_2(-3.49842 + 3.27128\sqrt{-1}))^5 +$$
$$(24.25)(x_0 + (.39844)x_1 + (.112957)x_2)^5 +$$
$$(-2.62582 + 3.74206\sqrt{-1})(x_0 + x_1(.122478 + .537715\sqrt{-1}) + x_2(-.436832 - .342586\sqrt{-1}))^5 +$$
$$(-2.62582 - 3.74206\sqrt{-1})(x_0 + x_1(.122478 - .537715\sqrt{-1}) + x_2(-.436832 + .342586\sqrt{-1}))^5 +$$
$$(.000108482)(x_0 + (-2.94762)x_1 + (12.5538)x_2)^5$$

The symmetric case: uniqueness in the subgeneric case

Theorem (Sylvester[1851], Chiantini-Ciliberto, Mella, Ballico, [2002-2005])

The general $f \in S^d\mathbb{C}^{n+1}$ of rank s smaller than the generic one has a unique Waring decomposition, with the only exceptions

- *rank $s = \binom{n+2}{2} - 1$ in $S^4\mathbb{C}^{n+1}$, $2 \leq n \leq 4$, when there are infinitely many decompositions*
- *rank 7 in $S^3\mathbb{C}^5$, when there are infinitely many decompositions*
- *rank 9 in $S^6\mathbb{C}^3$, where there are exactly two decompositions*
- *rank 8 in $S^4\mathbb{C}^4$, where there are exactly two decompositions*
- *rank 9 in $S^3\mathbb{C}^6$, where there are exactly two decompositions*

The cases listed in red are called the *defective cases*.
The cases listed in blue are called the *weakly defective cases*.

Weakly defective examples

Assume for simplicity $k = 3$. Only known examples where the general $f \in V_1 \otimes V_2 \otimes V_3$ ($\dim V_i = n_i + 1$) of subgeneric rank s has a NOT UNIQUE decomposition, besides the defective ones, are

- unbalanced case, rank $s = n_1 n_2 + 1$, $n_3 \geq n_1 n_2 + 1$
- rank 6 $\quad (n_1, n_2, n_3) = (3, 3, 3)$ where there are two decompositions
- rank 8 $\quad (n_1, n_2, n_3) = (2, 5, 5)$, sporadic case [CO], maybe six decompositions

Theorem

- *The unbalanced case is understood [Chiantini-O. [2011]].*
- *There is a unique decomposition for general tensor of rank* s *in* $\mathbb{C}^{n+1} \otimes \mathbb{C}^{n+1} \otimes \mathbb{C}^{n+1}$
  if $s \leq \frac{3n+1}{2}$ *[Kruskal[1977]]*
  if $s \leq \frac{(n+2)^2}{16}$ *[Chiantini-O. [2011]]*
- *The exceptions to uniqueness listed in the previous slide are the only ones in the cases* $\prod n_i \leq 10^4$ *[Chiantini-O-Vannieuwenhoven [2014]]*

Relevance of matrix multiplication algorithm

Many numerical algorithms use matrix multiplication. The complexity of matrix multiplication algorithm is crucial in many numerical routines.

$$M_{m,n} = \text{space of } m \times n \text{ matrices}$$

Matrix multiplication is a bilinear operation

$$\begin{array}{ccc} M_{m,n} \times M_{n,l} & \to & M_{m,l} \\ (A, B) & \mapsto & A \cdot B \end{array}$$

where $A \cdot B = C$ is defined by $c_{ij} = \sum_k a_{ik} b_{kj}$.
This usual way to multiply a $m \times n$ matrix with a $n \times l$ matrix requires mnl multiplications and $ml(n-1)$ additions, so asympotically $2mnl$ elementary operations.
The usual way to multiply two 2×2 matrices requires eight multiplication and four additions.

Rank and complexity

Matrix multiplication can be seen as a tensor
$t_{m,n,l} \in M_{m,n} \otimes M_{n,l} \otimes M_{m,l}$
$t_{m,n,l}(A \otimes B \otimes C) = \sum_{i,j,k} a_{ik} b_{kj} c_{ji} = tr(ABC)$
and the number of multiplications needed coincides with the rank of $t_{m,n,l}$ with respect to the Segre variety $\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C$ of decomposable tensors.
Allowing approximations, the border rank of t is a good measure of the complexity of the algorithm of matrix multiplication.

Strassen result on 2×2 multiplication

Strassen showed explicitly

$$\begin{aligned}
M_{2,2,2} =& a_{11} \otimes b_{11} \otimes c_{11} + a_{12} \otimes b_{21} \otimes c_{11} + a_{21} \otimes b_{11} \otimes c_{21} + a_{22} \otimes b_{21} \otimes c_{21} \\
& + a_{11} \otimes b_{12} \otimes c_{12} + a_{12} \otimes b_{22} \otimes c_{12} + a_{21} \otimes b_{12} \otimes c_{22} + a_{22} \otimes b_{22} \otimes c_{22} \\
=& (a_{11} + a_{22}) \otimes (b_{11} + b_{22}) \otimes (c_{11} + c_{22}) + (a_{21} + a_{22}) \otimes b_{11} \otimes (c_{21} - c_{22}) \\
& + a_{11} \otimes (b_{12} - b_{22}) \otimes (c_{12} + c_{22}) + a_{22} \otimes (-b_{11} + b_{21}) \otimes (c_{21} + c_{11}) \\
& + (a_{11} + a_{12}) \otimes b_{22} \otimes (-c_{11} + c_{12}) + (-a_{11} + a_{21}) \otimes (b_{11} + b_{12}) \otimes c_{22} \\
& + (a_{12} - a_{22}) \otimes (b_{21} + b_{22}) \otimes c_{11} .
\end{aligned} \tag{3}$$

Implementation of Strassen result

Dividing a matrix of size $2^n \times 2^n$ into 4 blocks of size $2^{n-1} \times 2^{n-1}$ one shows inductively that are needed 7^k multiplications and $9 \cdot 2^k + 18 \cdot 7^{k-1}$ additions, so in general $\leq C7^k$ elementary operations.

The number 7 of multiplications needed turns out to be the crucial measure.

The exponent of matrix multiplication ω is defined to be $\underline{\lim}_n \log_n$ of the arithmetic cost to multiply $n \times n$ matrices, or equivalently, $\underline{\lim}_n \log_n$ of the minimal number of multiplications needed.

A consequence of Strassen bound is that $\omega \leq log_2 7 = 2.81\ldots$.

The border rank in case 3×3 is still unknown.

Thanks !!