The Geometry of Tensor Spaces

Jornada de Jóvenes Doctores
en Geometría Algebraica
Barcelona
November 4-5th, 2021

Giorgio Ottaviani

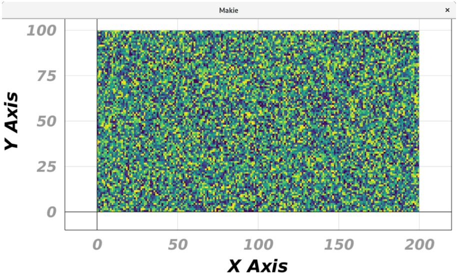
University of Florence, Italy

Books grading make a matrix

Readers grade books from 1 to 5.
Alice rates 5/5 *Don Quixote*, but 2/5 *Les misérables*.

	Anna Karenina		Don Quixote		Les misérables	
Alice	5	3	5	?	2	...
Bob	3	?	4	3	?	...
Mario	?	...	...	...		
Maria						
...						

Typically, only a small percent of the large matrix is known.



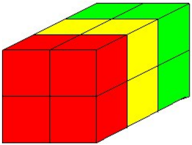
Web services like to recommend you the books that you like. A reasonable assumption is that the large matrix M can be approximated by a small rank matrix, so that

$$\underbrace{M}_{10^5\times 10^4} = \underbrace{A}_{(10^5\times r)} \underbrace{B^t}_{(r\times 10^4)}$$

In many other situations one wishes to decompose $M = \sum_{i=1}^{r} a_i \otimes b_i$ (a_i columns of A, b_i columns of B)

This decomposition is not unique, since $M = (AC)(BC^{-t})^t$ for any invertible C of format $r \times r$.

Tensors encode data with more dimensions, like readers, genres and book titles, which can be encoded as a tensor in $A \otimes B \otimes C$. In the tensor setting the not uniqueness problem disappears.

This is a $2 \times 2 \times 3$ tensor 

Tensor space

$\mathbb{K}$ is a field, typically $\mathbb{R}$ or $\mathbb{C}$.

$\mathbb{K}^{n_1} \otimes \ldots \otimes \mathbb{K}^{n_d} = \mathbb{K}^{n_1 \times \ldots \times n_d}$ is the vector space of tensors with d modes.

In coordinates it is given by $n_1 \cdot \ldots \cdot n_d$ scalars $t_{i_1 \ldots i_d}$.

The case $d = 2$ corresponds to matrices.

Tensor rank

Let

$$T \in \mathbb{K}^{n_1} \otimes \ldots \otimes \mathbb{K}^{n_d}$$

Definition

We say that the tensor T has rank r if

1. T is the sum of r decomposable tensors.

$$T = \sum_{i=1}^{r} v_{i,1} \otimes \ldots \otimes v_{i,d}$$

2. r is the minimal number of summands.

So, by definition, decomposable tensors have rank one.
It is an exercise to show that for matrices ($d = 2$), this definition of rank agrees with the classical one.

Let

$$T \in \mathbb{K}^n \otimes \ldots \otimes \mathbb{K}^n$$

We say that the tensor T is symmetric if $\sigma(T) = T$ for any permutation σ on the modes.
In coordinates, T is symmetric iff

$$t_{i_1 \ldots i_d} = t_{\sigma(i_1) \ldots \sigma(i_d)} \quad \forall \text{ permutation } \sigma$$

Of course, symmetric matrices are exactly the ones we know.

Homogeneous polynomials

Symmetric matrices correspond to quadratic forms, symmetric tensors with d modes correspond to homogeneous polynomials of degree d.

$$\text{Let} \qquad T \in \mathrm{Sym}^d \mathbb{K}^n$$

Definition

We say that the polynomial T has symmetric rank r if

1. T is the sum of r powers of linear forms, $T = \sum_{i=1}^{r} l_i^d$
2. r is the minimal number of summands.

So, by definition, powers of linear forms have symmetric rank one. Again, for symmetric matrices ($d = 2$), this definition of rank agrees with the classical one.

Rank versus symmetric rank

It was believed that rank is equal to symmetric rank for any symmetric tensor (Comon Conjecture).

This conjecture has been disproved in 2018 by Shitov with a complicated counterexample in a 800-dimensional space.

Still is not clear if for generic tensors Comon Conjecture remains true. It is true in many small rank cases (as the toy examples we see in next slides).

Main differences between matrix rank and tensor rank, I

Matrix rank does not depend on field extension. Tensor rank may depend on the field.
Indeed

$$2x^3 + 6xy^2 = (x+y)^3 + (x-y)^3 \quad \mathrm{rk}_{\mathbb{R}} = 2$$

$$2x^3 - 6xy^2 = ?$$

$$2x^3 - 6xy^2 = (x+\sqrt{-1}y)^3 + (x-\sqrt{-1}y)^3 \quad \mathrm{rk}_{\mathbb{C}} = 2$$

over $\mathbb{R}$ we have

$$2x^3 - 6xy^2 = 4x^3 - (x+y)^3 - (x-y)^3 \quad \mathrm{rk}_{\mathbb{R}} = 3$$

Main differences between matrix rank and tensor rank, II

If $\mathrm{rk}(M_n) = r$ then $\mathrm{rk}\left(\lim_n M_n\right) \leq r$ (lower semicontinuity of matrix rank)

Tensor rank is not semicontinuous,
$\lim_{t \to 0} \frac{1}{t}\left((x+ty)^3 - x^3\right) = 3x^2y$ which has rank three, indeed

$$3x^2y = 3x^3 + \frac{1}{2}(x+y)^3 + \frac{1}{2}(x-y)^3$$

Even on $\mathbb{C}$ we cannot do better.
Note that 3 is bigger than the dimension of the factor $\mathbb{C}^2$, which again does not happen for matrices !

Main differences between matrix rank and tensor rank, III

Matrix rank is lower or equal than the dimension of the factors.

Tensor rank may exceed the dimension of the factors

Main differences between matrix rank and tensor rank, IV

Matrix rank can be computed efficiently by Gaussian elimination.

Tensor rank is NP-hard to be computed.

Uniqueness of decomposition, the feature that makes tensor decomposition important

A decomposition

$$T = \sum_{i=1}^{r} v_{i,1} \otimes \ldots \otimes v_{i,d}$$

is (*almost always!*) unique for general tensors of *subgeneric* rank, that is when r is smaller than the rank attained by generic tensors. This feature makes tensor decomposition of small rank tensors a central feature for many applications, like book grading discussed at the beginning.
Another striking examples is sound reconstruction.

Geometric view with secant varieties

Symmetric tensors of rank one are powers l^d and fill the d-Veronese variety of $\mathbb{P}^n$.
Recall it is given by the image of the map

$$\begin{array}{ccc} \mathbb{P}(\mathbb{C}^{n+1}) & \rightarrow & \mathbb{P}(\mathrm{Sym}^d\mathbb{C}^{n+1}) \\ v & \mapsto & v^d \end{array}$$

and we call it $v_d(\mathbb{P}^n)$.
Define

$$\sigma_k(v_d(\mathbb{P}^n)) = \overline{\left\{ \sum_{i=1}^{k} l_i^d | \deg l_i = 1 \right\}}$$

which is the Zariski closure of polynomials of symmetric rank k.
For $d = 2$ the varieties $\sigma_k(v_d(\mathbb{P}^n))$ consist of symmetric $d \times d$ matrices of rank $\leq k$ and are well known.

Alexander-Hirschowitz(AH) Theorem on generic symmetric rank

Theorem (Alexander-Hirschowitz 1995)

Let $d \geq 3$. The k-secant variety $\sigma_k(v_d(\mathbb{P}^n))$ has the expected dimension

$$\min\{k(n+1)-1, \binom{n+d}{n} - 1\}$$

unless the following cases

1. $\sigma_5(v_4(\mathbb{P}^2))$ *Clebsch quartics*
2. $\sigma_9(v_4(\mathbb{P}^3))$
3. $\sigma_{14}(v_4(\mathbb{P}^4))$
4. $\sigma_7(v_3(\mathbb{P}^4))$

The case of Clebsch quartics

Consider differential operators P of degree 2 on $\mathbb{C}^{n+1}$. They are polynomials of degree 2 in $\partial_i = \frac{\partial}{\partial x_i}$ for $i = 0, \ldots, n$.
Note $P(\sum_{i=1}^k l_i^4)$ is a linear combination of $l_1^2, \ldots, l_k^2$. In particular the map

$$\begin{array}{rccc} C_f\colon & \mathrm{Sym}^2\mathbb{C}^{3\vee} & \to & \mathrm{Sym}^2\mathbb{C}^3 \\ & P & \mapsto & P(f) \end{array}$$

has image contained in $\langle l_1^2, \ldots, l_5^2\rangle$ for $f = \sum_{i=1}^5 l_i^4$ so it has rank ≤ 5 for $f \in \sigma_5(v_4(\mathbb{P}^2))$.Since the map is given by a 6×6 matrix C_f, it follows that $\det C_f$ is the equation of hypersurface of Clebsch quartics in $\mathbb{P}(\mathrm{Sym}4\mathbb{C}^3) = \mathbb{P}^{14}$!

A conjecture toward an analog of AH-Theorem for generic (unsymmetric) rank

In the (unsymmetric) general case, rank one tensors make the Segre variety.
There is a conjectured list of exceptions (Abo-O-Peterson) for an analog of AH-theorem about the dimension of k-secants.
One of them is $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$, (four qubits) where the 3-secant is expected to be an hyperurface, indeed it has codimension 2.

Uniqueness of decomposition for subgeneric cases

Theorem (Chiantini-Ciliberto, Mella, Ballico, Chiantini-O-Vannieuwenhoven)

The general $f \in \sigma_k(v_d\mathbb{P}^n)$ *when* $\sigma_k(v_d\mathbb{P}^n)$ *is properly contained in the ambient space it has a unique decomposition unless the following cases*

1. *The defective cases of AH Theorem*
2. $\sigma_9(v_6(\mathbb{P}^2))$ 2 *decompositions, AH Thm implies finitely many.*
3. $\sigma_8(v_4(\mathbb{P}^3))$ 2 *decompositions*
4. $\sigma_9(v_3(\mathbb{P}^5))$ 2 *decompositions*

The last three items have the common property that an elliptic normal curve passes through:

2. 9 *points in* $\mathbb{P}^2$
3. 8 *points in* $\mathbb{P}^3$
4. 9 *points in* $\mathbb{P}^5$

Uniqueness of decomposition for generic symmetric cases

The generic tensors in $\mathbb{P}(\mathrm{Sym}^d\mathbb{C}^{n+1})$ has NOT a unique decomposition unless the following cases

1. $\mathrm{Sym}^d\mathbb{C}^2$ for odd d (odd binary forms)
2. $\mathrm{Sym}^5\mathbb{C}^3$ plane quintics
3. $\mathrm{Sym}^3\mathbb{C}^4$ cubic surfaces (Sylvester Pentahedral Theorem)

Theorem (Sylvester Pentahedral Theorem)

The general cubic surface can be written in a unique way as a sum $\sum_{i=1}^{5} l_i^3$ of 5 cubes of linear forms.

Note the shape of the Theorem is opposite from previous cases. The number of decompositions (when finitely many) is unknown even on $\mathbb{P}^2$ for $d \geq 10$.

Uniqueness of decomposition for (unsymmetric) cases

The expected shape of the results is analogous to the symmetric case.

In the subgeneric case it is expected uniqueness of decomposition unless a list of exceptions. Here there are again exceptions related to elliptic curves with 2 decompositions (among them $\sigma_5(\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1)$ (five qubits) and an additional *mysterious* exception given by $\sigma_8(\mathbb{P}^2 \times \mathbb{P}^5 \times \mathbb{P}^5)$ [Chiantini-Mella-O 2014] with 6 decompositions.

The $3 \times 4 \times 5$ case

In the generic case it is expected NOT uniqueness of decomposition unless a list of exceptions. There is a nice exception which is given by $3 \times 4 \times 5$ Theorem.

Theorem (Hauenstein-Oeding-O-Sommese)

The general tensor in $\mathbb{C}^3 \otimes \mathbb{C}^4 \otimes \mathbb{C}^5$ has a unique decomposition as a sum of 6 decomposable tensors.

The case of qubits

The case of qubits is almost completely understood.

Theorem (Catalisano-Geramita-Gimigliano 2011)

k-secant varieties to the Segre variety given by n copies of $\mathbb{P}^1$ have the expected dimension unless $\sigma_3(\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1)$

Theorem (Casarotti-Mella, 2019)

The general tensor in the k-secant varieties to the Segre variety given by n copies of $\mathbb{P}^1$ has a unique decomposition for $n \geq 5$ and $k \leq \lfloor \frac{2^n}{n+1} \rfloor - 1$.

Weierstrass canonical form, I

We are ready for the first classification result in Tensor Spaces. From triangle inequality, the only format $2 \times b \times c$ where the hyperdeterminant exists (so that the triangular inequality is satisfied) are $2 \times k \times k$ and $2 \times k \times (k+1)$.
Consider first the $2 \times k \times k$ case.

Theorem (Weierstrass)

Let A be a $2 \times k \times k$ tensor and let A_0, A_1 be the two slices. Assume that $Det(A) \neq 0$. Under the action of $GL(k) \times GL(k)$ A is equivalent to a matrix where A_0 is the identity and $A_1 = diag(\lambda_1, \ldots, \lambda_k)$. In this form the hyperdeterminant of A is equal to $\prod_{i<j}(\lambda_i - \lambda_j)^2$.

Weierstrass canonical form, II

The second case $2 \times k \times (k+1)$ was also solved by Weierstrass.

Theorem (Weierstrass)

All tensors of format $2 \times k \times (k+1)$ with Det $\neq 0$ are $GL(k) \times GL(k+1)$ equivalent to the polynomial multiplication tensor with $\alpha = 1$, $\beta = k-1$.

Kronecker-Weierstrass canonical form, I

It is interesting, and quite unexpected, that the format $2 \times k \times (k+1)$ is a building block for all the other formats $2 \times b \times c$. The canonical form illustrated by the following Theorem is called the Kronecker-Weierstrass canonical form (there is an extension in the degenerate case that we do not pursue here).

Theorem (Kronecker, 1890)

Let $2 \le b < c$. There exist unique $n, m, q \in \mathbb{N}$ satisfying

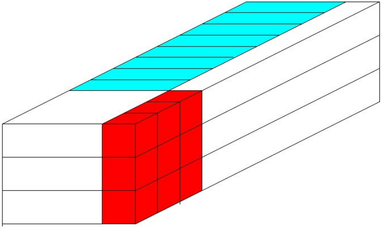
$$\begin{cases} b = nq + m(q+1) \\ c = n(q+1) + m(q+2) \end{cases}$$

such that the general tensor $t \in \mathbb{C}^2 \otimes \mathbb{C}^b \otimes \mathbb{C}^c$ decomposes under the action of $GL(b) \times GL(c)$ as n blocks $2 \times q \times (q+1)$ and m blocks $2 \times (q+1) \times (q+2)$ in Weierstrass form.

Kac Theorem, I

Kac has generalized this statement to the format $2 \leq w \leq s \leq t$ satisfying the inequality $t^2 - wst + s^2 \geq 1$. Note that in these cases the hyperdeterminant does not exist (for $w \geq 3$). The result is interesting because it gives again a canonical form.
Given w, define by the recurrence relation $a_0 = 0$, $a_1 = 1$,
$a_j = wa_{j-1} - a_{j-2}$
For $w = 2$ get $0, 1, 2, \ldots$ and Kronecker's result.
For $w = 3$ get $0, 1, 3, 8, 21, 55, \ldots$ (odd Fibonacci numbers)

Kac Theorem and Fibonacci blocks



Theorem (Kac, 1980)

Let $2 \leq w \leq s \leq t$ satisfying the inequality $t^2 - wst + s^2 \geq 1$. Then there exist unique $n, m, j \in \mathbb{N}$ satisfying

$$\begin{cases} s = & na_j & + \; ma_{j+1} \\ t = & na_{j+1} & + \; ma_{j+2} \end{cases}$$

such that the general tensor $t \in \mathbb{C}^w \otimes \mathbb{C}^s \otimes \mathbb{C}^t$ decomposes under the action of $GL(s) \times GL(t)$as n blocks $w \times a_j \times a_{j+1}$ and m blocks $w \times a_{j+1} \times a_{j+2}$ which are denoted "Fibonacci blocks".

Thanks

Thanks !!