Algebra, Geometry and Ranks

Enrico Carlini

Department of Mathematical Sciences
Politecnico di Torino, Italy

Matrix Workshop, Creswick, 6th-7th February 2018

Tensor rank

Let $V_1, \ldots, V_m$ be $\mathbb{C}$-vector spaces of dimension $\dim V_i = n_i + 1$.

A tensor $T \in V = V_1 \otimes \ldots \otimes V_m$ is

$$T = \sum \alpha_{i_1,\ldots,i_m} v_{i_1} \otimes \ldots \otimes v_{i_m}$$

where the coefficients $\alpha_{i_1,\ldots,i_m} \in \mathbb{C}$ and the vectors $v_{i_j} \in V_j$.

There are some distinguished elements in V that we commonly use to represent all the other elements

Elementary tensors

A tensor

$$v_1 \otimes \ldots \otimes v_m \in V$$

with $v_i \in V_i$ is called *elementary tensor*.

Note that using elementary tensors we can construct a basis for V and thus for any $T \in V$ we can write

$$T = \sum_{i=1}^{r} E_i$$

where the E_i are elementary tensors.
We give the following definition

Tensor rank

The *tensor rank* of T is

$$\mathrm{rk}(T) = \min\{r : T = \sum_{i=1}^{r} E_i, E_i \text{ elementary}\}.$$

Tensor rank

Example $V = V_1 \otimes V_2$

In this case $T \in V$ can be written as

$$T = \sum_{i,j} \alpha_{ij} v_1 \otimes v_2.$$

Fixing bases in V_1 and V_2, T corresponds to the $\dim V_1 \times \dim V_2$ matrix

$$A_T = (\alpha_{ij}).$$

Elementary tensors correspond to matrices of **rank one**, thus

$$\mathrm{rk}(T) = \mathrm{rk}(A_T).$$

Tensor rank

By basic properties of the tensor product we know that multilinear operators are tensors. For example, the multiplication of two matrices

$$A \in \mathbb{C}^{n,m}, B \in \mathbb{C}^{m,p}$$

corresponds to a tensor

$\mathbf{M}_{\langle \mathbf{n},\mathbf{m},\mathbf{p} \rangle}$

$$\mathbf{M}_{\langle \mathbf{n},\mathbf{m},\mathbf{p} \rangle} \in \mathbb{C}^{n,m*} \otimes \in \mathbb{C}^{m,p*} \otimes \mathbb{C}^{n,p}$$

is the *matrix multiplication tensor*. If $n = m = p$, that is for square matrices, we just write $\mathbf{M}_{\langle \mathbf{n} \rangle}$.

Knowing $\mathrm{rk}(\mathbf{M}_{\langle \mathbf{n},\mathbf{m},\mathbf{p} \rangle})$ relates to the computational complexity of matrix multiplication.

Tensor rank

It is not difficult to find an upper bound for $\mathrm{rk}(\mathbf{M}_{\langle \mathbf{n},\mathbf{m},\mathbf{p} \rangle})$.

$\mathrm{rk}(\mathbf{M}_{\langle \mathbf{n},\mathbf{m},\mathbf{p} \rangle}) \leq nmp$

Given matrices

$$A = (a_{ij}) \in \mathbb{C}^{n,m}, B = (b_{jl}) \in \mathbb{C}^{m,p}, C = c_{il} \in \mathbb{C}^{n,p}$$

and choosing dual bases $\{\alpha_{ij}\}$ and $\{\beta_{jl}\}$ we get that

$$\mathbf{M}_{\langle \mathbf{n},\mathbf{m},\mathbf{p} \rangle} = \sum_{ijl} \alpha_{ij} \otimes \beta_{jl} \otimes c_{il}$$

and thus the conclusion follows.

For example $\mathrm{rk}(\mathbf{M}_{\langle \mathbf{n} \rangle}) \leq n^3$.

Tensor rank

Strassen's result and $\mathbf{M}_{\langle 2 \rangle}$

The usual matrix multiplication in the case 2×2 is

$$\mathbf{M}_{\langle 2 \rangle} \in \mathbb{C}^{2,2} \otimes \in \mathbb{C}^{2,2} \otimes \mathbb{C}^{2,2}$$

where

$$\mathbf{M}_{\langle 2 \rangle} = \sum_{i=1}^{8} E_i$$

for eight elementary tensors and thus

$$\mathrm{rk}(\mathbf{M}_{\langle 2 \rangle}) \leq 8.$$

But in the '60s Strassen wanted to prove that equality holds and...

Strassen's result and $\mathbf{M}_{\langle 2 \rangle}$

Strassen showed that

$$\mathrm{rk}(\mathbf{M}_{\langle 2 \rangle}) \leq 7,$$

and we now know that equality holds. That is

$$\mathbf{M}_{\langle 2 \rangle} = \sum_{i=1}^{7} F_i$$

for **seven**, and **no fewer**, elementary tensors F_i. Thus one can multiply $n \times n$ matrix with complexity

$$O(n^{\log_2 7}).$$

Tensor rank

$\mathbf{M}_{\langle \mathbf{3} \rangle}$

Clearly

$$\mathrm{rk}(\mathbf{M}_{\langle \mathbf{3} \rangle}) \leq 27,$$

and we know that

$$19 \leq \mathrm{rk}(\mathbf{M}_{\langle \mathbf{3} \rangle}) \leq 23,$$

but we do not know the actual value yet!

Waring rank

Given a vector space $V = \langle x_0, \ldots, x_n \rangle$ we can consider *symmetric tensors* that is elements of $\mathrm{Sym}^{\mathrm{d}} V$ that correspond to *degree d homogeneous polynomials* in the polynomial ring

$$S = \mathbb{C}[x_0, \ldots, x_n].$$

The vector space of degree d forms is usually denoted as S_d.

Elementary symmetric tensors

The elementary tensor

$$v_1 \otimes \ldots \otimes v_d$$

is symmetric iff all the vectors v_i are equal. Thus elementary symmetric tensors are just pure powers

$$L^d \in S_d.$$

Waring rank

Given a homogeneous degree d form $F \in S_d$, we define the *Waring rank*

$$\mathrm{rk}(F) = \min\{r : F = L_1^d + \ldots + L_r^d, L_i \in S_1\}$$

For example

Quadratic forms

If $F \in S_2$, then

$$F(\mathbf{x}) = \mathbf{x} A_F \mathbf{x}^\top$$

for a suitable symmetric matrix A_F. Diagonalizing A_F is equivalent to writing F as a sum of powers of linear forms, thus

$$\mathrm{rk}(F) = \mathrm{rk}(A_F).$$

In char zero, we can find a basis of S_d made by powers

Characteristic zero

If $F \in S_d$, then

$$\mathrm{rk}(F) \leq \binom{n+d}{d}.$$

Positive characteristic

If $xy \in \mathbb{K}[x, y]$ with $\mathrm{char}(\mathbb{K}) = 2$, then

$$\mathrm{rk}(xy) = +\infty$$

since $(ax + by)^2$ cannot contain the monomial xy.

Waring rank

Subadditivity

Clearly
$$\mathrm{rk}(L^d) = 1,$$
and
$$\mathrm{rk}(L_1^d + L_2^d) = 2$$
iff L_1 and L_2 are not proportional. In general we have
$$\mathrm{rk}(L_1^d + \ldots + L_r^d) \leq r$$
and it is not easy to decide whether equality holds or not.

For example

$$\mathrm{rk}((a_1x + b_1y)^2 + (a_2x + b_2y)^2 + (a_3x + b_3y)^2) \leq 2.$$

Waring rank

Since a form F gives rise to a symmetric tensor T it is natural to study the relation between $\mathrm{rk}(F)$ and $\mathrm{rk}(T)$. Note that to compute the former we restrict to elementary *symmetric* tensors and thus

$$\mathrm{rk}(T) \leq \mathrm{rk}(F).$$

Comon's conjecture

The tensor rank and the symmetric tensor rank are equal, that is

$$\mathrm{rk}(T) = \mathrm{rk}(F).$$

August 2017 **counterexample** by J. Shitov.

X-rank

We want to find a uniform setting to deal with ranks. First we note that our rank definition are invariant up to scalar multiplication, thus it is natural to work over the *projective space*.

Projective space

Given a $N+1$ dimensional vector space V, we define

$$\mathbb{P}(V) = \mathbb{P}^N \setminus 0 = V/\mathbb{C}^*$$

and $[v] \in \mathbb{P}(V)$ is the equivalence class $\{\lambda v : \lambda \in \mathbb{C} \setminus \{0\}\}$.

X-rank

We want to work with special subset of the projective space, namely *algebraic varieties*.

V(I)

Given a *homogeneous ideal* $I \subseteq \mathbb{C}[x_0, \ldots, x_N]$ we define the algebraic variety

$$V(I) = \{p \in \mathbb{P}^n : F(p) = 0 \text{ for each } F \in I\}.$$

Note that to each algebraic variety $X \subseteq \mathbb{P}^N$ corresponds a *radical ideal*

I(X)

$$I(X) = \{F \in S : F(p) = 0 \text{ for each } p \in X\}.$$

Some features of algebraic varieties

- The algebraic variety X is completely determined by the ideal $I(X)$
- All ideal $I \subseteq S$ have a finite number of generators (*Hilbert's basis theorem*)
- For each ideal we can compute a numerical function $HF_{I(X)}(t)$ giving to us several information about X: emptyness, dimension, degree, etc (*Hilbert function*)
- Groebner bases of $I(X)$ are used to study X, for example its projections (*Elimination theory*)
- The image of an algebraic projective variety via a polynomial map is a projective variety
- Algebraic varieties are the closed subset of the Zariski topology

X-rank

Given an algebraic variety $X \subset \mathbb{P}^N$ and a point $p \in \mathbb{P}^N$, we define

X-rank

The X-rank of p with respect to X is

$$X{-}\mathrm{rk}(p) = \min\{r : p \in \langle p_1, \ldots, p_r \rangle, p_i \in X\}$$

where

$$\langle p_1, \ldots, p_r \rangle = \mathbb{P}(\{\lambda_1 v_1 + \ldots + \lambda_r v_r : \lambda_i \in \mathbb{C}\})$$

is the linear span of the points $p_i = [v_i]$'s.

Clearly, $X{-}\mathrm{rk}(p) = 1$ if and only if $p \in X$.

Veronese varieties

Let $S = \mathbb{C}[x_0, \ldots, x_n]$ and consider the map,

$$\mathbb{P}(S_1) \longrightarrow \mathbb{P}(S_d)$$

$$[L] \mapsto [L^d]$$

this is usually denoted as

$$\nu_d : \mathbb{P}^n \longrightarrow \mathbb{P}^N$$

and its image

$$X = \nu_d(\mathbb{P}^n)$$

is an algebraic variety called the *d*-uple *n*-dimensional Veronese variety.

Veronese varieties

Since the Veronese variety $X = \nu_d(\mathbb{P}^n)$ parameterizes pure powers of degree d in $n+1$ variables, it is clear that

$$X-\mathrm{rk}([F]) = \min\{r : [F] \in \langle [L_1^d], \ldots, [L_r^d] \rangle\}$$

and thus the X-rank with respect to the Veronese variety is just the Waring rank.

X-rank

Segre varieties

Given vector spaces $V_1, \ldots, V_t$, we consider the map

$$s : \mathbb{P}(V_1) \times \ldots \times \mathbb{P}(V_t) \longrightarrow \mathbb{P}(V_1 \otimes \ldots \otimes V_t)$$

$$[v_1], \ldots, [v_t] \mapsto [v_1 \otimes \ldots \otimes v_t]$$

this is called the *Segre map* and its image X is called the *Segre product* of the varieties $\mathbb{P}(V_i)$.

X-rank

Segre varieties

Since the Segre variety $X = s(\mathbb{P}(V_1) \times \ldots \times \mathbb{P}(V_t))$ parameterizes elementary tensors in $V_1 \otimes \ldots \otimes V_t$, it is clear that

$$X{-}\mathrm{rk}([T]) = \min\{r : [T] \in \langle [E_1], \ldots, [E_r] \rangle\}$$

and thus the X-rank with respect to the Segre variety is just the (tensor) rank.

X-rank

To study general tensor we can use Segre varieties and to study symmetric tensor we can use Veronese varieties. However, intermediate situations can be of interest. For example,

$$x \otimes y \otimes z \otimes t + y \otimes x \otimes z \otimes t + x \otimes y \otimes t \otimes z + y \otimes x \otimes t \otimes z$$

is a *partially symmetric*.

Segre-Veronese varieties

Segre-Veronese varieties parameterizes tensors with prescribed symmetry, for example the variety

$$s(\nu_2\mathbb{P}(\langle x, y\rangle), \nu_2\mathbb{P}(\langle z, t\rangle))$$

parameterizes elementary objects of the form

$$x \otimes x \otimes z \otimes z.$$

Secant varieties

It is a common process to start from an algebraic variety $X \subseteq \mathbb{P}^N$ and to produce a new one.

Secant varieties

For any non-negative integer i we define the i-th secant variety of X

$$\sigma_i(X) = \overline{\bigcup_{P_1,\ldots,P_i \in X} \langle P_1, \ldots, P_i \rangle}$$

where the bar denotes the Zariski closure.

We note that

there is an open dense subset of $\sigma_i(X)$ formed by points of the type

$$\lambda_1 P_1 + \ldots + \lambda_i P_i,$$

for points P_j's in X and scalars λ_j's.

Secant varieties

For example

$$\sigma_1(X) = X,$$

and

$\sigma_2(X)$ is the variety of secant lines to X and it is formed by true secant lines and **tangent** lines to X.

Secant varieties

Linear spaces

If X is a linear space, that is $X \simeq \mathbb{P}^m$ for some m, then

$$\sigma_i(X) = X$$

for all $i > 0$.

Hypersurfaces

If $X \subseteq \mathbb{P}^N$ is a hypersurface, not a hyperplane, then

$$\sigma_i(X) = \mathbb{P}^N$$

for all $i > 0$.

Secant varieties

It is clear that, for $X \subset \mathbb{P}^N$, we have

Chain of inclusion

$$X = \sigma_1(X) \subseteq \sigma_2(X) \subseteq \ldots \subseteq \mathbb{P}^N.$$

and also

$\sigma_i(X) = \sigma_{i+1}(X) \Longrightarrow \sigma_i(X)$ is a linear space

For example, for $i = 1$, we note that there is an open dense subset of $\sigma_2(X)$ made by points of the form

$$\lambda_1 P_1 + \lambda_2 P_2$$

for $P_1, P_2 \in X$ and $\lambda_1, \lambda_2 \in \mathbb{C}$. Since

$$\sigma_2(X) = \sigma_1(X) = X$$

we conclude that all lines joining pairs of points in X completely lie in X. Thus X is a linear space.

Secant varieties

If $X \subseteq \mathbb{P}^N$ is non-degenerate, that is X is not contained in any hyperplane, then

$$X = \sigma_1(X) \subset \sigma_2(X) \subset \ldots \subset \sigma_r(X) = \mathbb{P}^N$$

and all inclusions are strict

Thus the natural question to find

$$\min\{r : \sigma_r(X) = \mathbb{P}^N\}.$$

Note that Veronese and Segre varieties are non-degenerate.

Expected dimension

If $X \subseteq \mathbb{P}^N$ and $\dim X = n$, then the *expected dimension* of $\sigma_i(X)$ is

$$\mathrm{expdim}\sigma_i(X) = in + i - 1 = i(n+1) - 1$$

this value comes from a parameter count and it is such that

$$\dim \sigma_i(X) \leq \mathrm{expdim}\sigma_i(X).$$

Note that equality **does not** always hold.

Let's explore the connection between rank and secant varieties

If X-rank(P)=r

then

$$P \in \langle P_1, \ldots, P_r \rangle$$

where the points P_i's are in X. Thus

$$P \in \sigma_r(X).$$

In particular, knowing **some** elements of the ideal

$$I(\sigma_r(X)) = (G_1, \ldots, G_l)$$

gives an effective test to check whether

$$X\mathrm{-rank}(P) \neq r.$$

X-rank and Secant varieties

If $P \in \sigma_r(X)$

then

$$P \in \overline{\bigcup_{P_1,\ldots,P_r \in X} \langle P_1, \ldots, P_r \rangle}$$

but we can have

$$X\mathrm{-rank}(P) < r$$

or

$$X\mathrm{-rank}(P) > r.$$

Border X-rank

The border X-rank of P is the smallest i such that

$$P \in \sigma_i(X).$$

X-rank and Secant varieties

If the border X-rank of P is r

then P is a **limit** of X-rank r elements.

Knowing f the ideal

$$I(\sigma_r(X)) = (G_1, \ldots, G_l)$$

gives an **necessary and sufficient** condition for

the border X -rank(P) to be r.

Example $x^2y \in \mathbb{C}[x, y]$

Let $X = \nu_3(\mathbb{P}^1)$, and thus the X-rank is just the Waring rank. Since

$$x^2y = \lim_{\lambda \to 0} \frac{1}{\lambda} \left[(x + \lambda y)^3 - x^3 \right]$$

we have that $[x^2y] \in \sigma_2(X)$, that is x^2y has border rank two. However, the equation

$$x^2y = (ax + by)^3 + (cx + dy)^3$$

has no solutions in $\mathbb{C}$ and thus

$$\mathrm{rk}(x^2y) \geq 3$$

and actually equality holds.

X-rank and Secant varieties

If $X \subseteq \mathbb{P}^N$ is such that

$$\sigma_r(X) = \mathbb{P}^N,$$

then there is an open and dense subset of $\mathbb{P}^N$ made of elements such that $X-\mathrm{rank} = r$.

Generic X-rank

We say that the *generic* X-rank is r and we write $X-\mathrm{grank} = r$ if

$$\sigma_r(X) = \mathbb{P}^N$$

and $\sigma_{r-1}(X) \neq \mathbb{P}^N$.

Generic element vs **random** element.

Example generic Waring rank for $n = 1$ and $d = 3$

In this case $X = \nu_3(\mathbb{P}^1) \subset \mathbb{P}^3$ and it can be easily checked that

$$\sigma_2(X) = \mathbb{P}^3$$

and thus

$$X\mathrm{-grank} = 2.$$

This means that the **generic** degree 3 element of $\mathbb{C}[x, y]$ has Waring rank 2. But there are elements having Waring rank 3, such as

$$x^2y.$$

Since $\mathrm{expdim}\sigma_i(X) = i(\dim X + 1) - 1$, we get

$$X\mathrm{-expgrank} = \left\lceil \frac{N+1}{\dim X + 1} \right\rceil$$

Waring rank: what do we know

We now consider X to be a Veronese variety and then the X-rank is just the Waring rank of a homogeneous polynomial. We know the *generic Waring rank*:

Alexander and Hirschowitz results $\simeq$ 1990

For degree d forms in $n+1$ variables we have that

$$X - \mathrm{grank}(n, d) = \left\lceil \frac{\binom{n+d}{d}}{n+1} \right\rceil$$

unless $(n, d) =$

$$(n, 2), (2, 4), (3, 4), (4, 3), (4, 4).$$

Waring rank: what do we know

The defective case $(n, d) = (2, 4)$

Let $F \in S_4$ where $S = \mathbb{C}[x, y, z]$ and consider the equation

$$F = \sum_{i=1}^{5} (a_i x + b_i y + c_i z)^5.$$

Since $\dim S_4 = 15$ and since we have 15 variables, we expect to be able to solve for a_i, b_i and c_i. However, it is not difficult to see that, for any choice of a_i, b_i and c_i there exists a partial differential operator ∂ of order 2 such that

$$\partial \circ \sum_{i=1}^{5} (a_i x + b_i y + c_i z)^5 = 0.$$

Waring rank: what do we know

The defective case $(n, d) = (2, 4)$

But for a generic $F \in S_4$ there is no order 2 operator annihilating F, thus

$$\mathrm{rk}(F) \geq 6$$

and actually equality holds.
In other words,

$$\mathrm{expdim}\sigma_5(X) = 14$$

but is is actually 13 and

$$\sigma_5(X) \subset \mathbb{P}(S_4) = \mathbb{P}^{14}.$$

Tensor rank: what do we know

We now consider the Segre variety

$$X = s(\mathbb{P}(V_1) \times \ldots \times \mathbb{P}(V_t)) \subset \mathbb{P}^N$$

and then the X-rank is just the tensor rank.
We can easily write down a formula for

$$X - \text{expgrank} = \left\lceil \frac{1+N}{1-t+\sum_i \dim V_i} \right\rceil$$

but, in general, we do no know whether the formula gives the actual value.

Tensor rank: what do we know

We know the exact value of the X-grank in the following cases:

- $\dim V_i = 2$ for $1 \leq i \leq t$
- $t = 2$, that is only two factors
- $t = 3$, that is only three factors

There are several conjecture giving a list of defective cases and claiming that the list is exhaustive.

Maximal X-rank

Even when we know X-grank, we do not know how big the rank can be for **special** elements.

For example, consider $X = \nu_3(\mathbb{P}^1)$, that is we study the Waring rank of bivariate cubic forms.
By A-H result, we know that

$$X\mathrm{-grank} = 2$$

but we also know that $\mathrm{rk}(x^2y) = 3$, and actually this is the largest possible value.

Maximal rank of binary forms

For degree d binary forms, that is $X = \nu_d(\mathbb{P}^1)$, we have

$$X\mathrm{-mrank} = d + 1.$$

Maximal X-rank

For $X \subset \mathbb{P}^N$, we define:

X-mrank

$$X\mathrm{-mrank} = \max\left\{X\mathrm{-rank}(p) : p \in \mathbb{P}^N\right\}$$

Blekherman-Teitler 2015

$$X\mathrm{-mrank} =\leq 2(X\mathrm{-grank}).$$

Maximal X-rank

Not sharp

Note that the B-T bound is not sharp, even for binary forms, e.g. for binary cubic forms the generic rank is 2 and the maximal rank is 3.

Plane curves

We know the maximal Waring rank for $n = 2$ and $d = 3, 4, 5$ and it is exactly one more than the generic rank, that is $5, 7, 8$.

Form of high rank

We only know few cases of forms having Waring rank larger than the generic rank, for example monomials in 3 variables have this property (2012). But we do **not** have an answer for almost all pairs (n, d).