Border Rank
equations and geometry

Fulvio Gesmundo

Recap and plan for today

- Tensor rank: $T \in U \otimes V \otimes W$

$$\mathrm{R}(T) = \min \left\{ r : T = \sum_{i=1}^{r} u_i \otimes v_i \otimes w_i : \begin{array}{l} u_i \in U \\ v_i \in V \\ w_i \in W \end{array} \right\}.$$

- What do we do today?
  - Tensor rank is not (lower) semicontinuous;
  - Border rank and secant varieties;
  - Lower bounds for border rank;
  - Degeneration of tensors;
  - Asymptotic rank is determined by border rank.

Semicontinuity of matrix rank

Recall: The tensor rank of a tensor of order two is its rank as a matrix.

Lemma. The set

$$\sigma_r = \{M : U^* \to V \mid \mathrm{rank}(M) \leq r\} \subseteq U \otimes V$$

is closed (in the Euclidean topology of $U \otimes V$).

Proof. Consider the map

$$\begin{aligned} U \otimes V &\to \mathbb{C}^N \\ M &\mapsto \begin{pmatrix} \text{all size } (r+1) \text{ minors} \\ \text{of } M \text{ (in some fixed basis)} \end{pmatrix}. \end{aligned}$$

This map is continuous because it is given by an N-uple of polynomials (here $N = \binom{\dim U}{r+1}\binom{\dim V}{r+1}$).

The set σ_r is the preimage of $\{0\}$ so it is closed. □

A classical example (essentially due to Sylvester – 1852). Consider

$$\mathbf{w} = |1\rangle \otimes |0\rangle \otimes |0\rangle + |0\rangle \otimes |1\rangle \otimes |0\rangle + |0\rangle \otimes |0\rangle \otimes |1\rangle.$$

Last week, we saw that $\mathrm{R}(|\mathbf{w}\rangle) = 3$.

However

$$\mathbf{w} = \lim_{\varepsilon \to 0} \left[\frac{1}{\varepsilon}(|0\rangle + \varepsilon|1\rangle)^{\otimes 3} - \frac{1}{\varepsilon}|0\rangle^{\otimes 3} \right].$$

So $\mathbf{w} = \lim_{\varepsilon \to 0} T_\varepsilon$ with $\mathrm{R}(T_\varepsilon) = 2$ when $\varepsilon \neq 0$.

This shows that the set

$$\sigma_r^\circ = \{T \in U \otimes V \otimes W : \mathrm{R}(T) \leq r\}$$

is NOT closed in general.

Let $\sigma_r = \overline{\sigma_r^\circ}$ be the closure of the set of tensors of rank at most r.

The **border rank** of a tensor T is

$$\underline{\mathrm{R}}(T) = \min\{r \in \mathbb{N} : T \in \sigma_r\}.$$

Note: $\underline{\mathrm{R}}(T) \leq \mathrm{R}(T)$. We just saw $\underline{\mathrm{R}}(|\mathbf{w}\rangle) \leq 2$.

Important Fact: The set σ_r is an *algebraic variety*: it is the zero set of a collection of polynomial equations on $U \otimes V \otimes W$. [Chevalley's Theorem]

For instance, in the case of matrices, σ_r is the zero set of the collection of minors of size $r + 1$.

The variety $\sigma_1 = \{u \otimes v \otimes w : u \in U, v \in V, w \in W\}$ is called **Segre variety** of rank one tensors.

The variety σ_r is the r-**th secant variety** of the Segre variety.

Given $T \in U \otimes V \otimes W$, how to determine $\underline{\mathrm{R}}(T)$.

- Upper bounds
  - We try to give explicit expressions.
  - We do not have a systematic way to do it.
  - Maybe surprisingly: many good ways to tweak numerical methods.
- Lower bounds
  - We look for equations and use them as obstructions:
    For $F \in \mathbb{C}[U \otimes V \otimes W]$
    $$F|_{\sigma_r} \equiv 0 \text{ and } F(T) \neq 0 \quad \Rightarrow \quad \underline{\mathrm{R}}(T) > r.$$
  - The secant variety $\sigma_r \subseteq U \otimes V \otimes W$ is a variety invariant for the action of $\mathrm{GL}(U) \times \mathrm{GL}(V) \times \mathrm{GL}(W)$
  - Highest weight vectors methods (Christian's talk)
  - Flattening methods

A flattening of $U \otimes V \otimes W$ is a linear maps

$$\mathrm{Flat} : U \otimes V \otimes W \to \mathsf{Hom}(\mathcal{E}, \mathcal{F})$$

for two vector spaces $\mathcal{E}, \mathcal{F}$.

We use flattenings to *translate* membership in σ_r into membership into a set of low rank matrices.

Let $r_0 = \max\{\mathsf{rank}(\mathrm{Flat}(T)) : T \in \sigma_1\}$.

Proposition. [Landsberg-Ottaviani]
Let $T \in U \otimes V \otimes W$.

$$\text{If } T \in \sigma_r \text{ then } \mathsf{rank}(\mathrm{Flat}(T)) \leq r \cdot r_0.$$

Equivalently, if $\mathsf{rank}(\mathrm{Flat}(T)) > R$, then $\underline{\mathrm{R}}(T) > R/r_0$.

Proof. If $\mathrm{R}(T) \leq r$, then the statement is clear by linearity.

But on the matrix side, $\mathsf{rank}(\mathrm{Flat}(T)) \leq r_0 r$ is a closed condition, so it also holds at the limit.

Flattening methods: standard flattenings

A tensor $T \in U \otimes V \otimes W$, defines a linear map

$$T_U : U^* \to V \otimes W$$
$$\alpha \mapsto \alpha(T)$$

via contraction on the first factor. Similarly T_V, T_W.

Claim. If $T = u \otimes v \otimes w$, then $\text{rank}(T_U) = r_0 = 1$.

Indeed: $\mathrm{image}(T_U) = \{\alpha(u) \cdot v \otimes w : \alpha \in U^*\} = \mathrm{span}(v \otimes w)$.

Consequence. $\underline{\mathrm{R}}(T) \geq \text{rank}(T_U)$.

Example. We show $\underline{\mathrm{R}}(\mathbf{I}_r) = r$ where

$$\mathbf{I}_r = |1\rangle \otimes |1\rangle \otimes |1\rangle + \cdots + |r\rangle \otimes |r\rangle \otimes |r\rangle \quad \in \quad U \otimes V \otimes W.$$

We have

$$T_U(\langle k|_U) = |k\rangle_V \otimes |k\rangle_W$$

therefore $\mathrm{image}(T_U) = \mathrm{span}(|1,1\rangle_{UV}, \ldots, |r,r\rangle_{UV})$, which has dimension r.

Flattening methods: Koszul flattenings

Let $T \in U \otimes V \otimes W$. Let $\mathrm{Flat}(T)$ be the composition

$$U^* \otimes W \xrightarrow{T_U \otimes \mathrm{Id}_W} V \otimes W \otimes W \xrightarrow{\text{skew}} V \otimes \Lambda^2 W$$

where $\text{skew}(w_1 \otimes w_2) = \frac{1}{2}(w_1 \otimes w_2 - w_2 \otimes w_1)$.

Claim. If $T = u \otimes v \otimes w$, then $\text{rank}(\mathrm{Flat}(T)) = r_0 = \dim W - 1$.

Indeed

$$\alpha \otimes w' \mapsto$$
$$\alpha(u) \cdot v \otimes w \otimes w' \mapsto$$
$$\alpha(u) \cdot v \otimes (w \otimes w' - w' \otimes w)$$

The image is (canonically) isomorphic to $\mathrm{span}(v) \otimes W/\mathrm{span}(w)$.
It has dimension equal to $\dim W - 1$.

An Example

Let dim U = dim V = dim W = 3. So $r_0 = 2$.

Let $T = |1,1,1\rangle_{UVW} + |2,2,2\rangle_{UVW} + |3,3,3\rangle_{UVW} + |1,2,3\rangle_{UVW}$.

We show $\underline{\mathrm{R}}(T) \geq 4$ by showing rank$(\mathrm{Flat}(T)) \geq 7$.

If $i = 2, 3$:

$$\langle i|_U \otimes |j\rangle_W \xrightarrow{T_U \otimes \mathrm{id}_W}$$

$$|i,i\rangle_{VW} \otimes |j\rangle_W = |i\rangle_V \otimes |i\rangle_W \otimes |j\rangle_W \xrightarrow{\text{skew}}$$

$$|i\rangle_V \otimes (|i,j\rangle_{WW} - |j,i\rangle_{WW}).$$

Get 4 linearly independent elements in $\mathrm{image}(\mathrm{Flat}(T))$.

If $i = 1$:

$$\langle 1|_U \otimes |j\rangle_W \xrightarrow{T_U \otimes \mathrm{id}_W}$$

$$(|1,1\rangle_{VW} + |2,3\rangle_{VW}) \otimes |j\rangle_W \xrightarrow{\text{skew}}$$

$$|1\rangle_V \otimes (|1,j\rangle_{WW} - |j,1\rangle_{WW}) + |2\rangle_V \otimes (|3,j\rangle_{WW} - |j,3\rangle_{WW})$$

Get 3 more linearly independent elements in $\mathrm{image}(\mathrm{Flat}(T))$.

More general flattenings

- Young flattenings built on other representations of $\mathrm{GL} \times \mathrm{GL} \times \mathrm{GL}$.
- We can use representation theory to compute ranks of the flattenings.
- Standard flattenings give *all* equations of σ_r for $r = 1, 2$.
  [classical] + [Landsberg-Manivel]
- Koszul flattenings give *all* equations of σ_r for $r = 3$.
  [Strassen] + [Qi]
- Barriers:
  no flattening gives equations for $r \geq 6n$ if $\dim U = \dim V = \dim W = n$.
  [Galazka] + [Efremenko-Garg-Oliveira-Wigderson]
- Other methods for lower bounds:
  Apolarity, border apolarity, border substitution.
  [Iarrobino] + [Buczyńska-Buczyński] + [Landsberg-Michałek]
  They go further, but we do not really know how far.

Given two tensors $T, S \otimes U \otimes V \otimes W$, we say that T restricts to S if there exist linear maps

$$A : U \to U \qquad B : V \to V \qquad C : W \to W$$

such that

$$(A \otimes B \otimes C)(T) = S.$$

We say that T **degenerates** to S if there exist linear maps

$$A(\varepsilon) : U \to U, \quad B(\varepsilon) : V \to V, \quad C(\varepsilon) : W \to W$$

depending polynomially on a formal variable ε such that

$$(A(\varepsilon) \otimes B(\varepsilon) \otimes C(\varepsilon))(T) = \varepsilon^a S_a + \varepsilon^{a+1} S_{a+1} + \cdots + \varepsilon^e S_e$$

with $S = S_a$.

We have

$$T \text{ restricts to } S \Rightarrow T \text{ degenerates to } S.$$

Geometrically: A degeneration is a limit of restrictions along a curve a degree e.

Example. Consider

$$
\begin{aligned}
A(\varepsilon) = B(\varepsilon) = C(\varepsilon) : \mathbb{C}^2 &\to \mathbb{C}^2 \\
|0\rangle &\mapsto |0\rangle + \varepsilon |1\rangle \\
|1\rangle &\mapsto -|0\rangle
\end{aligned}
$$

As 2×2 matrices

$$A(\varepsilon) = B(\varepsilon) = C(\varepsilon) = \begin{pmatrix} 1 & -1 \\ \varepsilon & 0 \end{pmatrix}$$

Recall the second unit tensor:

$$\mathbf{I}_2 = |0\rangle \otimes |0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle \otimes |1\rangle .$$

Then

$$A(\varepsilon) \otimes B(\varepsilon) \otimes C(\varepsilon)(\mathbf{I}_2) = \varepsilon \mathbf{w} + \varepsilon^2 S_1 + \varepsilon^3 S_2$$

Therefore $\mathbf{I}_2$ degenerates to $\mathbf{w}$.

Theorem.
$\underline{\mathrm{R}}(T) \leq r$ if and only if $\mathbf{I}_r$ degenerates to T.

Recall the notion of asymptotic rank of a tensor:

$$\underset{\sim}{\mathrm{R}}(T) = \lim_{N\to\infty}[\mathrm{R}(T^{\boxtimes N})]^{1/N}, \quad \overline{\underset{\sim}{\mathrm{R}}}(T) = \lim_{N\to\infty}[\underline{\mathrm{R}}(T^{\boxtimes N})]^{1/N}.$$

Theorem. [Bini-Capovani-Lotti-Romani 1979, Bini 1980]
$\underset{\sim}{\mathrm{R}}(T) = \overline{\underset{\sim}{\mathrm{R}}}(T)$

Proposition.

If $T = \lim(A(\varepsilon) \otimes B(\varepsilon) \otimes C(\varepsilon))\mathbf{I}_r$ with $\deg A(\varepsilon), \deg B(\varepsilon), \deg C(\varepsilon) \leq e$, then

$$\mathrm{R}(T^{\boxtimes N}) \leq r^N(3eN+1).$$

Proposition.

If $T = \lim(A(\varepsilon) \otimes B(\varepsilon) \otimes C(\varepsilon))\mathbf{I}_r$ with $\deg A(\varepsilon), \deg B(\varepsilon), \deg C(\varepsilon) \leq e$, (as functions of ε), then

$$\mathrm{R}(T^{\boxtimes N}) \leq r^N(3eN + 1).$$

Proof.

Let $T_\varepsilon = (A(\varepsilon) \otimes B(\varepsilon) \otimes C(\varepsilon))\mathbf{I}_r$. This is a curve of degree (at most) $3e$ (in ε) in the space of tensors.

Interpolation:

$3e + 1$ generic points on a curve of degree $3e$ have the same span as the whole curve. In particular, they span T.

For $\varepsilon \neq 0$, $\mathrm{R}(T_\varepsilon) \leq r$. Expressing T as linear combination of $3e + 1$ tensors on the curve yields

$$\mathrm{R}(T) \leq r(3e + 1).$$

This is the statement for $N = 1$.

Now:

$$T^{\boxtimes N} = \lim(A(\varepsilon)^{\boxtimes N} \otimes B(\varepsilon)^{\boxtimes N} \otimes C(\varepsilon)^{\boxtimes N})(\mathbf{I}_r)^{\boxtimes N}.$$

Recall $\mathbf{I}_r^{\boxtimes N} = \mathbf{I}_{r^N}$.

Also, if $A(\varepsilon)$ has degree at most e, then $A(\varepsilon)^{\boxtimes N}$ has degree at most eN.

Repeat the argument above.

Theorem.
Let $\underset{\sim}{\mathrm{R}}(T) = \lim_{N\to\infty}[\mathrm{R}(T^{\boxtimes N})]^{1/N}, \quad \overline{\underset{\sim}{\mathrm{R}}}(T) = \lim_{N\to\infty}[\underline{\mathrm{R}}(T^{\boxtimes N})]^{1/N}$. Then

$$\underset{\sim}{\mathrm{R}}(T) = \overline{\underset{\sim}{\mathrm{R}}}(T).$$

Proof.
Since $\underline{\mathrm{R}}(T) \leq \mathrm{R}(T)$, we have $\overline{\underset{\sim}{\mathrm{R}}}(T) \leq \underset{\sim}{\mathrm{R}}(T)$.

Define $r_K = \underline{\mathrm{R}}(T^{\boxtimes K})$. We show $\underset{\sim}{\mathrm{R}}(T) \leq r_K^{1/K}$.

We have

$$\underset{\sim}{\mathrm{R}}(T) \leq [\mathrm{R}(T^{\boxtimes N})]^{1/N} \leq$$

$$\leq [\mathrm{R}((T^{\boxtimes K})^{\boxtimes N/K})]^{1/N} \leq$$

$$\leq [r_K^{N/K}(3e_K \tfrac{N}{K} + 1)]^{1/N} = r_K^{1/K}(3e_K \tfrac{N}{K} + 1)^{1/N}.$$

As $N \to \infty$, we obtain $\underset{\sim}{\mathrm{R}}(T) \leq r_K^{1/K}$.

We conclude

$$\underset{\sim}{\mathrm{R}}(T) \leq \lim_{K\to\infty} r_K^{1/K} = \overline{\underset{\sim}{\mathrm{R}}}(T).$$