Border rank lower bounds part 1:

Representation theory: the systematic utilization of symmetry in (multi-)linear algebra

J.M. Landsberg

Texas A&M University

Supported by NSF grant CCF-1814254 and Clay foundation (Clay senior scholar)

Goals

- Exhibit the utility of representation theory in the study of tensors.

- Explain border rank lower bounds for matrix multiplication.

Tensors

V complex vector space $\dim(V) = \mathbf{v}$.

$V^{\otimes d} := \{T : V^* \times \cdots \times V^* \to \mathbb{C} \text{ multilinear}\}$

$\mathfrak{S}_d$: permutation group acts on $V^{\otimes d}$:
$\sigma \cdot (v_1 \otimes \cdots \otimes v_d) := v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(d)}$

Special subspaces:

$$S^d V := \{T \in V^{\otimes d} \mid \sigma \cdot T = T \ \forall \sigma \in \mathfrak{S}_d\}$$

symmetric tensors, homogeneous degree d polynomials on V^*.
Write $v_1 v_2 \cdots v_d := \sum_{\sigma \in \mathfrak{S}_n} v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(d)} \in S^d V$.

$$\Lambda^d V := \{T \in V^{\otimes d} \mid \sigma \cdot T = \mathrm{sgn}(\sigma) T \ \forall \sigma \in \mathfrak{S}_d\}$$

skew-symmetric tensors. Write
$v_1 \wedge \cdots \wedge v_d := \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(d)} \in \Lambda^d V$.

What else in $V^{\otimes d}$?

Projective space, projective varieties

Recall: $\mathbf{Q}, \ldots, \underline{\mathbf{R}}, \mathbf{R}$ notions of rank: all *invariant* under rescaling $T \mapsto \lambda T$, $\lambda \in \mathbb{C}^*$.

Recall: want to take limits

$\rightsquigarrow$ convenient to work in projective space $\mathbb{P}V := V \backslash \{0\} / \mathbb{C}^*$, $T \in V$, write $[T] \in \mathbb{P}V$, $X \subset \mathbb{P}V$, write $\hat{X} \subset V$.

$X \subset \mathbb{P}V$ zero set of collection of homogeneous polynomials called *algebraic variety* (or *projective variety*)

Examples of projective varieties

Ex. $X = Seg(\mathbb{P}A \times \mathbb{P}B) \subset \mathbb{P}(A \otimes B)$ *Segre variety* of rank one matrices. Zero set of $\Lambda^2 A^* \otimes \Lambda^2 B^* \subset S^2(A^* \otimes B^*)$.

Ex. $X = Seg(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C) \subset \mathbb{P}(A \otimes B \otimes C)$ *Segre variety* of rank one tensors. Zero set of
$\Lambda^2 A^* \otimes \Lambda^2 (B^* \otimes C^*) + \Lambda^2 B^* \otimes \Lambda^2 (A^* \otimes C^*) + \Lambda^2 C^* \otimes \Lambda^2 (A^* \otimes B^*) \subset S^2(A^* \otimes B^* \otimes C^*)$.

Ex. $X = v_d(\mathbb{P}V) \subset \mathbb{P}S^d V$ *Veronese variety* of rank one deg d polynomials, i.e. $v_d(\mathbb{P}V) = \{[v^d] \mid v \in V\}$ also $= GL(V) \cdot [w^d]$

Ex. $X = G(d, V) = \{[v_1 \wedge \cdots \wedge v_d] \mid v_j \in V\} \subset \mathbb{P}\Lambda^d V$ *Grassmann variety* d-planes through the origin in V:
$[v_1 \wedge \cdots \wedge v_d] = [w_1 \wedge \cdots \wedge w_d]$ iff $\langle v_1, \ldots, v_d \rangle = \langle w_1, \ldots, w_d \rangle$

$X \subset \mathbb{P}V$ any variety, can consider its *secant varieties*:
$\sigma_r(X) := \overline{\{[v] \in \mathbb{P}V \mid v = x_1 + \cdots + x_r,\ x_j \in \hat{X}\}}$

$\sigma_r(Seg(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$: tensors of border rank $\leq r$.

Varieties and group actions

Say $X \subset \mathbb{P}V$ is a *G-variety* for some $G \subset GL(V)$, if $g \cdot x \in X$ for all $x \in X$ and all $g \in G$.

$Seg(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)$ is a $G = GL(A) \times GL(B) \times GL(C) \subset GL(A \otimes B \otimes C)$ variety.

X is a G-variety $\Rightarrow \sigma_r(X)$ is a G- variety.

Essential observation: $Z \subset \mathbb{P}V$ is a G-variety $\Rightarrow I_Z \subset Sym(V^*)$ is a *G-module*

Representation Theory

A *representation* of a group G is a group homomorphism $\mu : G \to GL(V)$. Say V is a G-module.

$W \subset V$ is a *submodule* if $\forall\ g \in G$, $w \in W$, $\mu(g)w \in W$.

V is *irreducible* if it has no proper submodules.

Non-Example: action of $\mathfrak{S}_d$ on $\mathbb{C}^d$ is not irreducible.

Examples: $\Lambda^d V$, $S^d V \subset V^{\otimes d}$ are irreducible $GL(V)$-modules.

Lemma (Schur's Lemma)

Let V, W be irreducible G-modules and $f : V \to W$ a G-module map. Then either $f = 0$ or f is an isomorphism. If furthermore $V = W$, then $f = \lambda\,\mathrm{Id}$ for some constant λ.

Decomposition of $V^{\otimes 3}$ as a $GL(V)$-module

Saw $S^3V \oplus \Lambda^3 V \subset V^{\otimes 3}$ What else?

What about partially symmetric? $S^2V{\otimes}V \subset V^{\otimes 3}$ not irred., $S^2V{\otimes}V \supset S^3V$. Consider kernel of symmetrization map $S^2V{\otimes}V \to S^3V$. A $GL(V)$-submodule!

Ex. $v_1{\otimes}v_2{\otimes}v_3 + v_2{\otimes}v_1{\otimes}v_3 - v_1{\otimes}v_3{\otimes}v_2 - v_3{\otimes}v_1{\otimes}v_2$ is in kernel.

Representations of $GL(V)$ in $V^{\otimes d}$

Let π: partition of d. Have associated *Young diagram*

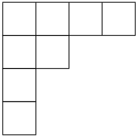

Figure: Young diagram for $\pi = (4, 2, 1, 1)$

Label boxes $\{1, \ldots, d\}$, called a *Young tableau without repetitions*. For example

3	7	2	5
1	4		
6			
8			

.

Representations of $GL(V)$ in $V^{\otimes d}$

default Young tableau

$$\begin{array}{|c|c|c|c|}\hline 1 & 2 & 3 & 4 \\ \hline 5 & 6 \\ \cline{1-2} 7 \\ \cline{1-1} 8 \\ \cline{1-1} \end{array} .$$

Given any Young tableau without repetitions, write
$V^{\otimes d} = V_1 \otimes V_2 \otimes \cdots \otimes V_d$,

for each row symmetrize the corresponding copies of V,

e.g., for $\pi = (4, 2, 1, 1)$ after the symmetrization, one obtains an element of $S^4 V \otimes S^2 V \otimes V \otimes V \subset V^{\otimes 8}$. Next, skew-symmetrize along the columns.

Example, default Young tableau for $\pi = (2, 1)$, get
$v_i \otimes v_j \otimes v_k \mapsto v_i \otimes v_j \otimes v_k + v_j \otimes v_i \otimes v_k \mapsto$
$v_i \otimes v_j \otimes v_k - v_k \otimes v_j \otimes v_i + v_j \otimes v_i \otimes v_k - v_k \otimes v_i \otimes v_j$.

Representations of the permutation group and $GL(V)$

For default Young tableau, write

$$\mathrm{proj}_{\pi-def} : V^{\otimes d} \to V^{\otimes d}$$

$S_{\pi-def} V := \mathrm{proj}_{\pi-def}(V^{\otimes d})$: image.

$S_{\pi-def} V$ is a $GL(V)$-module.

Write $\ell(\pi)$ for number of parts of π (number of cols. of YD)

Let $\{e_j\}$ basis V and $\ell(\pi) \leq \dim V$. Set

$$v_{\pi-def} := \mathrm{proj}_{\pi-def}(e_1^{\otimes p_1} \otimes e_2^{\otimes p_2} \otimes \cdots \otimes e_\ell^{\otimes p_\ell}).$$

Fact: $\mathfrak{S}_d$-modules indexed by partitions of d. Write $[\pi]$ for corresponding module. The span of the $\mathfrak{S}_d$-orbit of $v_{\pi-def}$ is an irreducible $\mathfrak{S}_d$-module isomorphic to $[\pi]$.

Representations of the permutation group and $GL(V)$

Let $S_\pi V$ denote the isomorphism class of $S_{\pi-def}V$.

Theorem (Schur-Weyl duality)

As a $GL(V) \times \mathfrak{S}_d$-module,

$$V^{\otimes d} = \bigoplus_{|\pi|=d} S_\pi V \otimes [\pi].$$

W, M: G-modules with M irreducible, the *isotypic component* of M in W is the largest submodule of W isomorphic to $M^{\oplus m} = M \otimes \mathbb{C}^m$ for some m (multiplicity). Isotypic component of $S_\pi V$ in $V^{\otimes d}$ is $S_\pi V \otimes [\pi]$ Multiplicity of $S_\pi V$ in $V^{\otimes d}$ is $\dim[\pi]$. Multiplicity of $[\pi]$ in $V^{\otimes d}$ is $\dim S_\pi V$.

Other $GL(V)$ modules: Define $\det^{-1} : GL(V) \to GL(\mathbb{C}^1)$, $g \mapsto \det(g)^{-1}$. All other $GL(V)$- modules are $S_\pi V \otimes (\det^{-1})^{\otimes k}$.

Dimensions

Write $x \in \pi$ for a box in the Young diagram of π.

hook length of x = number of boxes to the right of it in its row, plus the number of boxes below it in its column, plus one,

hook lengths for $(4, 2, 1, 1)$:

7	4	2	1
4	1		
2			
1			

.

content $c(x)$ of x is zero if x is on the main diagonal, j, if it is on the j-th diagonal above the main diagonal, and $-j$ if it is on the j-th diagonal below the main diagonal.

Dimensions cont'd

$$\dim S_\pi \mathbb{C}^m = \prod_{x \in \pi} \frac{m + c(x)}{h(x)}$$

In particular

$$\dim S_\pi \mathbb{C}^m \leq (d+1)^{\binom{m}{2}}.$$

$$\dim[\pi] = \frac{d!}{\prod_{x \in \pi} h(x)}$$

$\dim(\mathbb{C}^m)^{\otimes d} = m^d$ is exponential in d,

Cauchy and Pieri

Cauchy: As $GL(A) \times GL(B)$-module,

$$S^d(A \otimes B) = \oplus_{|\pi|=d, \ell(\pi) \leq \min\{\mathbf{a},\mathbf{b}\}} S_\pi A \otimes S_\pi B$$

$$\Lambda^d(A \otimes B) = \oplus_{|\pi|=d, \ell(\pi) \leq \min\{\mathbf{a},\mathbf{b}\}} S_\pi A \otimes S_{\pi'} B$$

Pieri: $S_{(5,2,1)} V \otimes V = S_{(6,2,1)} V \oplus S_{(5,3,1)} V \oplus S_{(5,2,2)} V \oplus S_{(5,2,1,1)} V$

Picture:

More generally $S_\pi V \otimes S^k V$, $S_\pi V \otimes \Lambda^k V$

Equations for secant varieties of Segre varieties

$T \in A \otimes B \otimes C$, $\underline{\mathbf{R}}(T) \geq \text{rank}(T_A : A^* \to B \otimes C)$ and permuted statements. $\rightsquigarrow$
$\Lambda^{r+1} A^* \otimes \Lambda^{r+1} (B \otimes C)^* \subset I_{r+1}(\sigma_r(Seg(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)))$.

General idea: $X \subset \mathbb{P}M$, $f : M \hookrightarrow U \otimes W$ linear. If $\forall x \in \hat{X}$, $\text{rank}(f(x)) \leq t$, then for all $[m] \in \sigma_r(X)$, $\text{rank}(f(m)) \leq rt$ i.e., size $rt+1$ minors of image spans a subspace of $I_{rt+1}(\sigma_r(X))$

How to find good ones? If M: irred. G-module and X: G-variety, look for G-module inclusions: systematic.

Ex. $A \subset \text{Hom}(\Lambda^p A, \Lambda^{p+1} A) = \Lambda^p A^* \otimes \Lambda^{p+1} A \rightsquigarrow$

$A \otimes B \otimes C \subset \text{Hom}(\Lambda^p A \otimes B^*, \Lambda^{p+1} A \otimes C)$

Advantageous to restrict to $2p+1$ dimensional subspaces.

Matrix multiplication and Koszul flattening

$M_{\langle n\rangle} \in A\otimes B\otimes C = (U^*\otimes V)\otimes(V^*\otimes W)\otimes(W^*\otimes U)$

As a tensor $M_{\langle n\rangle}$ is $\mathrm{Id}_U \otimes \mathrm{Id}_V \otimes \mathrm{Id}_W$ re-ordered.

Koszul flattening map is

$$(M_{\langle U,V,W\rangle})_A^{\wedge p} : \Lambda^p(U^*\otimes V)\otimes(V\otimes W^*) \to \Lambda^{p+1}(U^*\otimes V)\otimes(W^*\otimes U).$$

Presence of $\mathrm{Id}_W = \mathrm{Id}_{W^*} \Rightarrow$ map factors as $(M_{\langle U,V,\mathbb{C}^1\rangle})_A^{\wedge p}\otimes \mathrm{Id}_{W^*}$, where

$$(M_{\langle \mathbf{u},\mathbf{v},1\rangle})_A^{\wedge p} : V\otimes\Lambda^p(U^*\otimes V) \to \Lambda^{p+1}(U^*\otimes V)\otimes U.$$

$$v\otimes(\xi^1\otimes e_1)\wedge\cdots\wedge(\xi^p\otimes e_p) \mapsto \sum_{s=1}^{\mathbf{u}} u_s\otimes(u^s\otimes v)\wedge(\xi^1\otimes e_1)\wedge\cdots\wedge(\xi^p\otimes e_p).$$

Here $u_1,\dots,u_{\mathbf{u}}$: basis of U with dual basis $u^1,\dots,u^{\mathbf{u}}$ of U^*, so $\mathrm{Id}_U = \sum_{s=1}^{\mathbf{u}} u^s\otimes u_s$.

Matrix multiplication and Koszul flattening

Strategic restriction $\Rightarrow \underline{\mathbf{R}}(M_{\langle n \rangle}) \geq 2\mathbf{n}^2 - \mathbf{n}$ (L-Ottaviani 2015)

(previous $\frac{3}{2}\mathbf{n}^2 + \frac{1}{2}\mathbf{n} - 1$ (Lickteig 1985, after Strassen 1983))

Next time: Further uses of representation theory

Thank you for your attention

For more on **tensors**, their geometry and applications, resp. **geometry and complexity**, resp. **recent developments**: