Clay Lecture 3:

Border apolarity in practice

J.M. Landsberg

Texas A&M University

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

Review

$A = \mathbb{C}^{\mathbf{a}}$, $B = \mathbb{C}^{\mathbf{b}}$, $C = \mathbb{C}^{\mathbf{c}}$ on this page $\mathbf{a} = \mathbf{b} = \mathbf{c} = m$

$T \in A \otimes B \otimes C$ has border rank at most r, $\underline{\mathbf{R}}(T) \leq r$ if $\exists$ $T_1(\epsilon), \ldots, T_r(\epsilon)$, $T_j(\epsilon)$ rank one $\forall \epsilon > 0$, $T = \lim_{\epsilon \to 0} \sum_j T_j(\epsilon)$.

Goal: lower bounds on $\underline{\mathbf{R}}(T)$, especially $T = M_{\langle n \rangle}$.

Classical $\underline{\mathbf{R}}(T) \geq m$ via minors of flattening $T : A^* \to B \otimes C$.

Strassen 1983: $\underline{\mathbf{R}}(T) \geq \frac{3}{2}m$ via minors of commutators in space of endomorphisms $T(A^*)T(\alpha)^{-1}$.

L-Ottaviani 2015: $\underline{\mathbf{R}}(T) \geq 2m - 3$, for "good" T
$\underline{\mathbf{R}}(M_{\langle n \rangle}) \geq 2m - \sqrt{m}$, $m = n^2$ via minors of Koszul flattening $\Lambda^p A \otimes B^* \to \Lambda^{p+1} A \otimes C$.

L-Michalek 2019: $\underline{\mathbf{R}}(T) \geq (2.02)m$, $\underline{\mathbf{R}}(M_{\langle n \rangle}) \geq 2m - \log(m) + 1$, $m = n^2$. via Koszul flattenings of $\mathbb{B}_T$-fixed degenerations of T

Complexity Theorists and algebraic geometers: Game essentially over for these techniques.

Buczynska-Bucynski idea

$T = \lim_{\epsilon\to 0}\sum_{j=1}^{r} T_j(\epsilon)$, consider $I_\epsilon \subset Sym(A^*\oplus B^*\oplus C^*)$

$I_\epsilon = \{P \in Sym(A^*\oplus B^*\oplus C^*) \mid P(T_j(\epsilon)) = 0, \forall 1 \le j \le r\}$.

Zero set of I_ϵ considered as subvariety of Segre.

I_ϵ $\mathbb{Z}^3$-graded $(I_\epsilon)_{(s,t,u)} \subset S^sA^* \otimes S^tB^* \otimes S^uC^*$

Taylor series for $T_j(\epsilon)$, only low order terms relevant free to alter higher order terms

Ex. $a_1\otimes b_1\otimes c_2 + a_1\otimes b_2\otimes c_1 + a_2\otimes b_1\otimes c_1 =$
$\lim_{\epsilon\to 0}\frac{1}{\epsilon}[(a_1+\epsilon a_2)\otimes(b_1+\epsilon b_2)\otimes(c_1+\epsilon c_2) - a_1\otimes b_1\otimes c_1]$
$= \lim_{\epsilon\to 0}\frac{1}{\epsilon}$

$[(a_1+\epsilon a_2+\epsilon^2 a_3+\ldots)\otimes(b_1+\epsilon b_2+\epsilon^2 b_3+\ldots)\otimes(c_1+\epsilon c_2+\ldots) -$
$a_1\otimes b_1\otimes c_1]$

$\rightsquigarrow$ WLOG $\epsilon > 0$ points in general position $\Rightarrow$
$\mathrm{codim}((I_\epsilon)_{stu}, S^sA^*\otimes S^tB^*\otimes S^uC^*) = r$ whenever $s+t+u > 1$.

Buczynska-Bucynski idea

$\rightsquigarrow$ curves in Grassmannians of codim r-planes with limits defined $\forall$ s, t, u as $\epsilon \to 0$.

Good News: Limit will be an ideal (Haiman-Sturmfels) but not necessarily saturated

Limit as $\epsilon \to 0$ in Haiman-Sturmfels multi-graded Hilbert scheme.
Good News: Only need finite number of Grassmannians.

Bonus: If T has symmetry, can insist limiting ideal I is Borel fixed.

BB Border Apolarity

If $\underline{\mathbf{R}}(T) \leq r$, there exists a multi-graded ideal I satisfying:

1. I is contained in the annihilator of T. This condition says $I_{110} \subset T(C^*)^\perp$, $I_{101} \subset T(B^*)^\perp$, $I_{011} \subset T(A^*)^\perp$ and $I_{111} \subset T^\perp \subset A^* \otimes B^* \otimes C^*$.

   I.e., $T \in I_{111}^\perp$, i.e., T in limiting r-plane in $A \otimes B \otimes C$, and $T(A^*)$ in limiting r-plane $I_{011}^\perp \subset B \otimes C$ etc...

2. For all (stu) with $s+t+u > 1$, $\operatorname{codim} I_{stu} = r$.

   By general position $\epsilon > 0$ assumption.

3. each I_{stu} is Borel-fixed.

4. I is an ideal, so the multiplication maps $I_{s-1,t,u} \otimes A^* \oplus I_{s,t-1,u} \otimes B^* \oplus I_{s,t,u-1} \otimes C^* \to S^s A^* \otimes S^t B^* \otimes S^u C^*$ have image contained in I_{stu}.

Border Apolarity in practice

4. I is an ideal, so the multiplication maps $I_{s-1,t,u} \otimes A^* \oplus I_{s,t-1,u} \otimes B^* \oplus I_{s,t,u-1} \otimes C^* \to S^s A^* \otimes S^t B^* \otimes S^u C^*$ have image contained in I_{stu}.

In particular codim of image of $I_{s-1,t,u} \otimes A^* \oplus I_{s,t-1,u} \otimes B^* \oplus I_{s,t,u-1} \otimes C^* \to S^s A^* \otimes S^t B^* \otimes S^u C^*$ is at least r. Rank condition!

After fixing choice of Borel fixed subspaces, have **polynomial** necessary conditions!

Border Apolarity in practice

Given T, to prove $\underline{\mathbf{R}}(T) > r$, prove can't have I satisfying above.

1. determine all codimension r Borel fixed subspaces of $A^* \otimes B^*$ annihilating $T(C^*) \subset A \otimes B$. get all candidates for I_{110}. Do same for candidate $I_{101} \subset A^* \otimes C^*$ and $I_{011} \subset B^* \otimes C^*$.

2. Compute the rank of $I_{110} \otimes A^* \to S^2 A^* \otimes B^*$. If too large (image has codim $< r$) REJECT! "(210)-test" ditto rank of $I_{110} \otimes B^* \to A^* \otimes S^2 B^*$ Do same for all candidates and other spaces.

3. For each so far ok triple, compute rank of $I_{110} \otimes C^* \oplus I_{101} \otimes B^* \oplus I_{011} \otimes A^* \to A^* \otimes B^* \otimes C^*$. If too large (image has codim $< r$) REJECT! "(111)-test"

continue all cases so far win already with 1-3.

Matrix multiplication and border apolarity

Here $A = U^* {\otimes} V$, $B = V^* {\otimes} W$, $C = W^* {\otimes} U$,

$M_{\langle n \rangle}$ reordering of $\mathrm{Id}_U \otimes \mathrm{Id}_V \otimes Id_W$, $\mathrm{Id}_U \in U^* {\otimes} U$.

$$M_{\langle n \rangle}(C^*) = U^* {\otimes} \mathrm{Id}_V \otimes W$$

$$\subset A {\otimes} B = (U^* {\otimes} V) {\otimes} (V^* {\otimes} W) = M_{\langle n \rangle}(C^*) \oplus [U^* {\otimes} \mathfrak{sl}(V) {\otimes} W]$$

Need to understand Borel fixed subspaces in $U^* {\otimes} \mathfrak{sl}(V) {\otimes} W$.

Borel: upper triangular invertible matrices in
$SL(U) \times SL(V) \times SL(W) = SL_n \times SL_n \times SL_n$.

Borel fixed subspaces for $U^*\otimes\mathfrak{sl}(V)\otimes W$

Candidate I_{110} codim$=r$ Equivalently, $I_{110}^{\perp}$, dim$=r$ containing $T(C^*)=U^*\otimes \mathrm{Id}_V\otimes W$ need to add $r-n^2$ dimensional Borel fixed subspace Case $M_{\langle 2\rangle}$: $r=6$, $n^2=4$, $r-n^2=2$

$$
\begin{tikzcd}
 & x_1^2\otimes y_1^2 \arrow[d] \arrow[ddl] \arrow[ddr] & \\
 & x_1^2\otimes y_1^1 - x_2^2\otimes y_1^2 \arrow[dd] \arrow[dddl] \arrow[dddr] & \\
x_1^2\otimes y_2^2 \arrow[dd] \arrow[dddr] & & x_1^1\otimes y_1^2 \arrow[dd] \arrow[dddl] \\
 & x_2^2\otimes y_1^1 \arrow[dddl] \arrow[dddr] & \\
x_1^2\otimes y_2^1 - x_2^2\otimes y_2^2 \arrow[dd] \arrow[dddr] & & x_1^1\otimes y_1^1 - x_2^1\otimes y_1^2 \arrow[dd] \arrow[dddl] \\
 & x_1^1\otimes y_2^2 \arrow[dd] & \\
x_2^2\otimes y_2^1 \arrow[ddr] & & x_2^1\otimes y_1^1 \arrow[ddl] \\
 & x_1^1\otimes y_2^1 - x_2^1\otimes y_2^2 \arrow[d] & \\
 & x_2^1\otimes y_2^1 &
\end{tikzcd}
$$

$x_j^i=u^i\otimes v_j$ etc.. three choices

Matrix multiplication

To show $\underline{\mathbf{R}}(M_{\langle 2\rangle}) > 6$: none of three choices of I_{110} passes both (210) and (120) tests. Explicitly, just had to compute ranks of sparse 24×40 matrices with entries $\{0, \pm 1\}$ and show over $18 = 24 - 6$. In homework, shortcuts to make calculation easier, even hand checkable.

Recall: Strassen $\underline{\mathbf{R}}(M_{\langle 3\rangle}) \geq 14$, L-Ottaviani $\underline{\mathbf{R}}(M_{\langle 3\rangle}) \geq 15$, L-Michalek $\underline{\mathbf{R}}(M_{\langle 3\rangle}) \geq 16$.

Conner-Harper-L 2019: $\underline{\mathbf{R}}(M_{\langle 3\rangle}) \geq 17$

Known upper bound is 20 (Smirnov). Why didn't we solve? Have $r = 17$ ideal that passes all tests, in all multi-degrees. But tests are just necess. conditions. Could be ideal from cactus border rank decomp. or could just be garbage, not limit of anything. Work in progress with Warsaw group (BB+ Jelisiejew): winner or not?

Matrix multiplication results cont'd

Recall: so far only $\underline{\mathbf{R}}(M_{\langle 2\rangle})$ known among nontrivial matrix multiplication tensors.

Conner-Harper-L 2019: $\underline{\mathbf{R}}(M_{\langle 223\rangle}) = 10$

Conner-Harper-L 2019: $\underline{\mathbf{R}}(M_{\langle 233\rangle}) = 14$

For tensors where one factor is of much larger dimension than other two, no eqns. beyond flattenings

Conner-Harper-L 2019: For all $\mathbf{n} > 25$,
$\underline{\mathbf{R}}(M_{\langle 2\mathbf{nn}\rangle}) \geq \mathbf{n}^2 + 1.32\mathbf{n} + 1$.

Previously, only $\underline{\mathbf{R}}(M_{\langle 2\mathbf{nn}\rangle}) \geq \mathbf{n}^2 + 1$ (Lickteig).

Conner-Harper-L 2019: For all $\mathbf{n} > 14$, $\underline{\mathbf{R}}(M_{\langle 3\mathbf{nn}\rangle}) \geq \mathbf{n}^2 + 2\mathbf{n}$.

Previously, only $\underline{\mathbf{R}}(M_{\langle 3\mathbf{nn}\rangle}) \geq \mathbf{n}^2 + 2$ (Lickteig).

Other results

Strassen laser method: bound ω indirectly via other tensors.

Prop. (Conner-Gesmundo-L-Ventura) $\det_3$, perm_3 potentially could be used to prove $\omega = 2$.

$\underline{\mathbf{R}}(\det_3) = 17$ (Conner-Harper-L 2019)

$\underline{\mathbf{R}}(\mathrm{perm}_3) = 16$ (Conner-Huang-L 2020)

CGLV Prop. more precisely: $\mathrm{perm}_3 = T_{cw,2}^{\boxtimes 2}$ Known since 1988
$T_{cw,2}^{\boxtimes 2}$ could potentially be used to prove $\omega = 2 \rightsquigarrow$ solves Q open since 1988.

If interested in other tensors for laser method, beam into Berlin on Wed. 6am IPAM time (3pm Berlin time)

Idea of proof for asymptotic results

How to prove lower bounds for all n?

Candidate $I_{110}^{\perp}$:

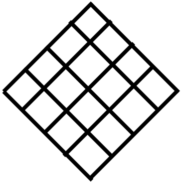
$U^*{\otimes}\,\mathrm{Id}_V\otimes W \subset I_{110}^{\perp} \subset B{\otimes}C = U^*{\otimes}\mathfrak{sl}(V){\otimes}W \oplus U^*{\otimes}\,\mathrm{Id}_V\otimes W$

To prove $\underline{\mathbf{R}}(M_{\langle mnn\rangle}) \geq n^2+\rho$, we show:

$\forall\ E \in G(\rho, U^*{\otimes}\mathfrak{sl}(V){\otimes}W)^{\mathbb{B}}$, (210) or (120) test fails.

Idea of proof for asymptotic results

Set of $U^* \otimes W$ weights of $I_{110}^{\perp}$ "outer structure"



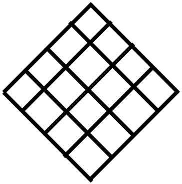
Given $U^* \otimes W$ weight (s, t), set of $\mathfrak{sl}(V)$-weights appearing with it "inner structure" $\mathfrak{sl}(V) = \mathfrak{sl}_2$ or $\mathfrak{sl}_3$

$\rightsquigarrow$ $n \times n$ grid, attach to each vertex a $\mathbb{B}$-closed subspace of $\mathfrak{sl}(V)$.

Split calculation of the kernel into a local and global computation.

Bound local (grid point) contribution to kernel by function of s, t and dimension of subspace of $\mathfrak{sl}(V)$.

Idea of proof for asymptotic results



Solve a nearly convex optimization problem over all possible outer structures.

"Worst case" on boundary.

Show extremal values fail test $\rightsquigarrow$ all choices fail test.

Thank you for your attention

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