New lower bounds for matrix multiplication

A. Conner, A. Harper and J.M. Landsberg

Texas A&M University

Landsberg supported by NSF grant CCF-1814254

Astounding conjecture

Strassen (1968) wrote an explicit algorithm to multiply $n \times n$ matrices with $O(n^{2.81}) < O(n^3)$ arithmetic operations.

Bini 1978, Schönhage 1983, Strassen 1987, Coppersmith-Winograd 1988 $\rightsquigarrow O(n^{2.378})$ arithmetic operations.

Astounding Conjecture

For all $\epsilon > 0$, matrices can be multiplied using $O(n^{2+\epsilon})$ arithmetic operations.

1988-2011 no progress, 2011-14 Stouthers,Williams,LeGall $O(n^{2.373})$ arithmetic operations.

Tensor formulation of conjecture

Set $N = n^2$.
Matrix multiplication is a *bilinear* map

$$M_{\langle n\rangle} : \mathbb{C}^N \times \mathbb{C}^N \to \mathbb{C}^N,$$

Bilinear maps $\mathbb{C}^N \times \mathbb{C}^N \to \mathbb{C}^N$ may also be viewed as trilinear

maps $\mathbb{C}^N \times \mathbb{C}^N \times \mathbb{C}^{N*} \to \mathbb{C}$.

In other words

$$M_{\langle n\rangle} \in \mathbb{C}^{N*} \otimes \mathbb{C}^{N*} \otimes \mathbb{C}^N.$$

Exercise: As a trilinear map, $M_{\langle n\rangle}(X, Y, Z) = \mathrm{trace}(XYZ)$.

Tensor formulation of conjecture

A tensor $T \in \mathbb{C}^N \otimes \mathbb{C}^N \otimes \mathbb{C}^N =: A \otimes B \otimes C$ has *rank one* if it is of the form $T = a \otimes b \otimes c$, with $a \in A$, $b \in B$, $c \in C$. $\sim$ bilinear maps that can be computed using one scalar multiplication.

The *rank* of a tensor T, $\mathbf{R}(T)$, is the smallest r such that T may be written as a sum of r rank one tensors. $\sim$ number of scalar multiplications needed to compute the corresponding bilinear map.

Tensor formulation of conjecture

Theorem (Strassen): $M_{\langle n\rangle}$ can be computed using $O(n^\tau)$ arithmetic operations $\Leftrightarrow$ $\mathbf{R}(M_{\langle n\rangle}) = O(n^\tau)$

Let $\omega := \inf_\tau\{\mathbf{R}(M_{\langle n\rangle}) = O(n^\tau)\}$

exponent of matrix multiplication.
Astounding conjecture: $\omega = 2$

border rank of $T \in A{\otimes}B{\otimes}C$ $\underline{\mathbf{R}}(T)$ denotes the smallest r such that T is a limit of tensors of rank r. I.e., smallest r such that $[T] \in \sigma_r := \sigma_r(Seg(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C))$, where, for $X \subset \mathbb{P}V$

$$\sigma_r(X) := \overline{\cup_{x_1,\dots,x_r\in X}\mathrm{span}\{x_1,\dots,x_r\}}$$

Theorem (Bini 1980) border rank is also a legitimate complexity measure: $\underline{\mathbf{R}}(M_{\langle n\rangle}) = O(n^\omega)$.

How to *disprove* astounding conjecture?

Find a polynomial P (in N^3 variables) in the ideal of σ_r,

Show that $P(M_{\langle n\rangle}) \neq 0$.

Embarassing (?): had not been known even for $M_{\langle 2\rangle}$, i.e., for σ_6 when $N = 4$.

Why did I think this would be easy?: Representation Theory

Matrices of rank at most r: zero set of size $r+1$ minors.

Ideal of Segre generated by size 2 minors of flattenings tensors to matrices: $A \otimes B \otimes C = (A \otimes B) \otimes C$.

Ideal of σ_2 generated by degree 3 polynomials.

Representation theory: systematic way to search for polynomials.

2004 L-Manivel: No polynomials in ideal of σ_6 of degree less than 12

2013 Hauenstein-Ikenmeyer-L: No polynomials in ideal of σ_6 of degree less than 19. However there are polynomials of degree 19. Caveat: too complicated to evaluate on $M_{\langle 2 \rangle}$. Good news: easier polynomial of degree 20 (trivial representation) $\rightsquigarrow$ (L 2006, Hauenstein-Ikenmeyer-L 2013) $\underline{\mathbf{R}}(M_{\langle 2 \rangle}) = 7$.

Polynomials via a retreat to linear algebra

$T \in A{\otimes}B{\otimes}C = \mathbb{C}^N{\otimes}\mathbb{C}^N{\otimes}\mathbb{C}^N$ may be recovered up to isom. from the linear space $T(C^*) \subset A{\otimes}B$.
tensors up to changes of bases $\sim$ linear subspaces of spaces of matrices up to changes of bases.
Even better than linear maps are endomorphisms. Assume $T(C^*) \subset A{\otimes}B$ contains an element of full rank. Use it to obtain an isomorphism $A{\otimes}B \simeq \mathrm{End}(A) \rightsquigarrow$ space of endomorphisms.
$\mathbf{R}(T) = N \Leftrightarrow N$-dimensional space of simultaneously diagonalizable matrices
$\underline{\mathbf{R}}(T) \leq N \Leftrightarrow$ limits of N-dimensional spaces of simultaneously diagonalizable matrices

Good News: Classical linear algebra!

Bad News: Open question.

Retreat to linear algebra, cont'd

Simultaneously diagonalizable matrices $\Rightarrow$ commuting matrices

Good news: Easy to Test.

Better news (Strassen): Can upgrade to tests for higher border rank than N: $\underline{\mathbf{R}}(T) \geq N + \frac{1}{2}$(rank of commutator)

$\rightsquigarrow$ (Strassen 1983) $\underline{\mathbf{R}}(M_{\langle n \rangle}) \geq \frac{3}{2}n^2$

Variant: (Lickteig 1985) $\underline{\mathbf{R}}(M_{\langle n \rangle}) \geq \frac{3}{2}n^2 + \frac{n}{2} - 1$

1985-2012: no further progress other than for $M_{\langle 2 \rangle}$.

Retreat to linear algebra, cont'd

Perspective: Strassen mapped space of tensors to space of matrices, found equations by taking minors.

Classical trick in algebraic geometry to find equations via minors.
$\rightsquigarrow$ (L-Ottaviani 2013) $\underline{\mathbf{R}}(M_{\langle n \rangle}) \geq 2n^2 - n$

Found via a $G = GL(A) \times GL(B) \times GL(C)$ module map from $A \otimes B \otimes C$ to a space of matrices (systematic search possible).

Explicitly: $A \otimes B \otimes C \to \mathrm{Hom}(\Lambda^p A^* \otimes B, \Lambda^{p+1} A \otimes C)$.
Polynomials: minors of matrix.

Punch line: Found modules of determinantal equations by exploiting symmetry of σ_r.

Note: Only gives good bounds if $\dim B \sim \dim C \geq \dim A$.
Example: says nothing new for $M_{\langle 2nn \rangle}$, $M_{\langle 3nn \rangle}$ for $n > 3$.

Bad News: Barriers

Theorem (Bernardi-Ranestad, Buczyńska-Buczyński-Galcazka, Efremenko-Garg-Oliviera-Wigderson): Game (almost) over for determinantal methods.

Variety of zero dimensional schemes of length r is not irreducible $r > 13$.

Determinantal methods detect zero dimensional schemes (want zero dimensional smoothable schemes).

$$\sigma_r(X) := \overline{\bigcup \{\langle R \rangle \mid \mathrm{length}(R) = r,\ \mathrm{support}(R) \subset X,\ R : \mathrm{smoothable}\}}$$

secant variety.

$$\kappa_r(X) := \overline{\bigcup \{\langle R \rangle \mid \mathrm{length}(R) = r,\ \mathrm{support}(R) \subset X\}}$$

cactus variety.

Determinantal equations are equations for the cactus variety.

Punch line: **Barrier** to progress.

How to go further?

So far, lower bounds via symmetry of σ_r.

The matrix multiplication tensor also has symmetry:

$T \in A \otimes B \otimes C$, define *symmetry group of T*
$G_T := \{g \in GL(A) \times GL(B) \times GL(C) \mid g \cdot T = T\}$

$GL_n^{\times 3} \subset G_{M_{\langle n \rangle}} \subset GL_{n^2}^{\times 3} = GL(A) \times GL(B) \times GL(C)$:

Proof: $(g_1, g_2, g_3) \in GL_n^{\times 3}$

$$\operatorname{trace}(XYZ) = \operatorname{trace}((g_1 X {g_2}^{-1})(g_2 Y {g_3}^{-1})(g_3 Z {g_1}^{-1}))$$

How to exploit G_T?

Given $T \in A \otimes B \otimes C$
$\underline{\mathbf{R}}(T) \leq r \Leftrightarrow \exists$ curve $E_t \subset G(r, A \otimes B \otimes C)$ such that
i) For $t \neq 0$, E_t is spanned by r rank one elements.
ii) $T \in E_0$.

For all $g \in G_T$, gE_t also works.
$\rightsquigarrow$ can insist on normalized curves (for $M_{\langle n \rangle}$, those with E_0 Borel fixed).

$\rightsquigarrow$ (L-Michalek 2017) $\underline{\mathbf{R}}(M_{\langle n \rangle}) \geq 2n^2 - \log_2 n - 1$

More bad news: this method cannot go much further.

New idea: Buczyńska-Buczyński (review of last week)

Use more algebraic geometry: Consider not just curve of r points, but the curve of **ideals** $I_t \in Sym(A^* \oplus B^* \oplus C^*)$ it gives rise to: *border apolarity method*

$T = \lim_{t \to 0} \sum_{j=1}^{r} T_{j,t}$

I_t ideal of $[T_{1,t}] \cup \cdots \cup [T_{r,t}] \subset \mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C$

Can insist that limiting ideal I_0 is Borel fixed: reduces to small search in each multi-degree.

Instead of single curve $E_t \subset G(r, A \otimes B \otimes C)$ limiting to Borel fixed point, for each (i, j, k) get curve $\{I_{ijk,t}^{\perp}\} \subset G(r, S^i A \otimes S^j B \otimes S^k C)$, each limiting to Borel fixed point *and* satisfying compatibility conditions.

Upshot: algorithm that either produces all normalized candidate I_0's or proves border rank $> r$.

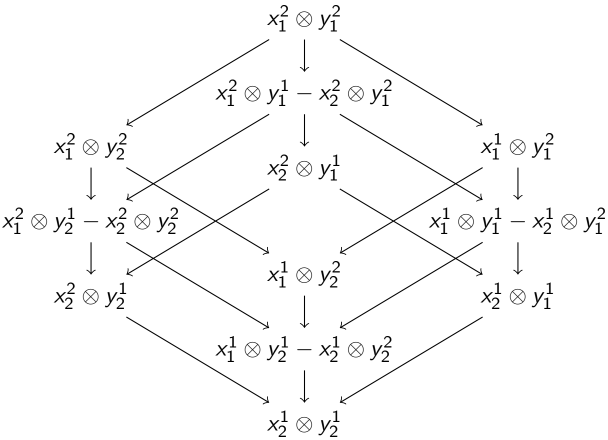
The border apolarity method (Buczyńska-Buczyński)

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 $I_{111} \subset T^{\perp} \subset A^* \otimes B^* \otimes C^*$.
2. For all (ijk) with $i+j+k>1$, $\operatorname{codim} I_{ijk} = r$.
3. each I_{ijk} is Borel-fixed.
4. I is an ideal, so the multiplication maps $I_{i-1,j,k} \otimes A^* \oplus I_{i,j-1,k} \otimes B^* \oplus I_{i,j,k-1} \otimes C^* \to S^i A^* \otimes S^j B^* \otimes S^k C^*$ have image contained in I_{ijk}. Call this the (ijk)-test.

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

$C = W^* \otimes U$ Case $U = V = W = \mathbb{C}^2$. Candidate codim$= r$ I_{110} when $T = M_{\langle 2 \rangle}$ Equivalently, dim$= r$ $I_{110}^{\perp}$ containing $T(C^*) = U^* \otimes \mathrm{Id}_V \otimes W$ $x^i_j = u^i \otimes v_j$ need to add $r - m$ dimensional Borel fixed subspace here $r = 6$, $m = 4$, $r - m = 2$



Border apolarity: results

Conner-Harper-L May 2019:
$\rightsquigarrow$ very easy algebraic proof $\underline{\mathbf{R}}(M_{\langle 2\rangle}) = 7$

$M_{\langle 3\rangle}$?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 June 2019: $\underline{\mathbf{R}}(M_{\langle 3\rangle}) \geq 17$

June 2019 only $\underline{\mathbf{R}}(M_{\langle 2\rangle})$ known among nontrivial matrix multiplication tensors.

Conner-Harper-L August 2019: $\underline{\mathbf{R}}(M_{\langle 223\rangle}) = 10$
Conner-Harper-L August 2019: $\underline{\mathbf{R}}(M_{\langle 233\rangle}) = 14$

All above results only use total degree 3 tests.

Conner-Harper-L Fall 2019: for all $n > 2$, $\underline{\mathbf{R}}(M_{\langle 2nn\rangle}) \geq n^2 + 1.32n$
Previously, only $\underline{\mathbf{R}}(M_{\langle 2nn\rangle}) \geq n^2 + 1$ known.
Conner-Harper-L 2020: for all n, $\underline{\mathbf{R}}(M_{\langle 3nn\rangle}) \geq n^2 + 1.6n$
Previously, only $\underline{\mathbf{R}}(M_{\langle 3nn\rangle}) \geq n^2 + 2$ known.
Just uses (210) and (120) tests!

Idea of proof for asymptotic results

How to prove lower bounds for all n?

Candidate $I_{110}^{\perp}$:
$C = W^* \otimes U$.
$M_{\langle \mathbf{u},\mathbf{v},\mathbf{w} \rangle}(C^*) = U^* \otimes \mathrm{Id}_V \otimes W \subset I_{110}^{\perp} \subset B \otimes C$
$= U^* \otimes \mathfrak{sl}(V) \otimes W \oplus U^* \otimes \mathrm{Id}_V \otimes W$

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

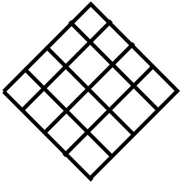
$\forall\ E \in G(\rho, U^* \otimes \mathfrak{sl}(V) \otimes W)^{\mathbb{B}}$, (210) or (120) test fails.

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"

$\rightsquigarrow$ $n \times n$ 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 $\mathfrak{sl}(V)$.



Solve a nearly convex optimization problem over all possible outer structures.
Show extremal values fail test $\rightsquigarrow$ all choices fail test.

What about the barrier?

Bad news: (ijk)-tests are determinantal equations— subject to barrier, i.e., candidate ideals may be candidate cactus border rank decompositions.

More bad news: For *any* tensor $T \in \mathbb{C}^m \otimes \mathbb{C}^m \otimes \mathbb{C}^m$ there exist ideals passing all total degree 3 tests for $\underline{\mathbf{R}}(T) = m + m^{\frac{1}{3}+\epsilon}$, e.g., $m = 9$, $\underline{\mathbf{R}} = 2m$.

How to tell if zero dimensional scheme is smoothable?

In general, hopeless. But: algorithm produces Borel fixed ideals $\rightsquigarrow$ schemes supported at a point.

Here there are recent techniques (Jelisejew).

Spring 2020: full (unsaturated) ideals for $M_{\langle 3 \rangle}$ that pass all tests for border rank 17.
Impostors or Slip?

Stay tuned!

Thank you for your attention

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