Lecture 4

Wednesday, 24 May 2023 15:21

Cohn-Umans approach for matrix multiplication

$$M_{\langle k,m,n\rangle} : \mathbb{C}^{k\times m} \times \mathbb{C}^{m\times n} \longrightarrow \mathbb{C}^{k\times n}$$ (Matrix multiplication map) bilinear

$$A, B \longmapsto AB$$

$$\cap$$

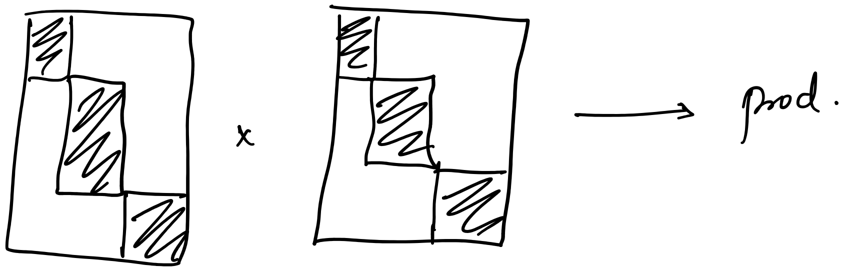
$$(\mathbb{C}^{k\times m})^* \otimes (\mathbb{C}^{m\times n})^* \otimes (\mathbb{C}^{k\times n})$$

General approach so far :- tensor powers of direct sums of matrix mult. tensors, recurse

Is there an abstract approach that gives a good perspective of the various different approaches?

Idea Embed $M_{\langle n,n,n\rangle}$ into Semisimple algebras

Informal defn [Semisimple algebras] Algebra in which multiplication is isomorphic to block-diagonal matrix mult.



* Hope that the algebra has a nice structure so that questions about w reduce to group-theoretic questions.

Defn [Semisimple Algebras] An associative Artinian algebras (over a field) that have a trivial Jacobson radical.

e.g. $\frac{1}{x} \in \mathbb{C}[x][\frac{1}{x}]$

$$D_2 = \frac{\mathbb{C}\langle x, \partial x\rangle}{\langle \partial x x - x\partial x - 1\rangle}$$ (Weyl algebra) Contains polynomial linear combs of differential operator. ; non-comm.

$$x^2 + x\partial x - x + 7 \in D_2$$

$$\partial x . x = x\partial x + 1$$

. $\mathbb{C}$ ∂ [illegible] an elem in $\mathbb{C}[x][^1/_x]$; action is just differentiation

$$\partial x . x = x \partial x + 1$$

apply $f \in D_2$ to any elem in $\mathbb{C}[x][1/x]$; action is just differentiation

$$\partial x \circ \frac{1}{x} = \frac{-1}{x^2}$$

$$(x\partial x + 1) \circ \frac{1}{x} = \frac{-1}{x} + \frac{1}{x} = 0 \Rightarrow (x\partial x + 1) \text{ is an annihilator of } \frac{1}{x}$$

<u>Thm</u> [Wedderburn's Theorem] Any finite dim. semisimple algebra is isomorphic to a finite product $\prod M_{n_i}(D_i)$

$n_i \times n_i$ matrices over D_i

division algebras over the field

[Read "Wedderburn-Artin Ring Theory" in Knapp's Advanced Alg]

Example of a Semi-Simple Alg.

G- finite group. $\mathbb{C}[G]$ - group algebra (formal linear combs of elements of the group)

$$\sum_{g \in G} a_g \cdot g \qquad a_g \in \mathbb{C}$$

$$\left(\sum_{g \in G} a_g g\right) + \left(\sum_{g \in G} b_g g\right)$$

$$= \sum_{g \in G} (a_g + b_g) g$$

$$\left(\sum_{g \in G} a_g g\right)\left(\sum_{h \in G} b_h h\right) = \sum_{f \in G} \sum_{\substack{g,h \in G \\ g+h=f}} (a_g + b_h) f$$

$\mathbb{C}[G]$ is a semi simple algebra

✱ Notice if $G = C_n$, and g is a generator, then

$$\left(\sum_{i=0}^{n-1} a_i g^i\right) \times \left(\sum_{i=0}^{n-1} b_i g^i\right) = \sum_{i=0}^{n-1} \left(\sum_{\substack{j,k \\ j+k \equiv i \bmod n}} a_j b_k\right) g^i,$$

multiplication in $\mathbb{C}[C_n]$ is a cyclic convolution

<u>Observe</u> $\left(\sum_{i=0}^{n-1} a_i x^i\right) * \left(\sum_{i=0}^{n-1} b_i x^i\right)$ is very close to mult. in $\mathbb{C}[C_n]$,

except for the wrap around.

Observe $\left(\sum_{i=0} a_i x\right) * \left(\sum_{i=0} b_i\right)$ is

except for the wrap around.

If we took C_m, $m \geq 2n$, then polynomial mult. is the same as mult. in $\mathbb{C}[C_m]$.

Thm [Fast Fourier Transform Alg] There is an invertible linear transformation $\to D : \mathbb{C}[G] \to \mathbb{C}^{|G|}$ that turns mult. in $\mathbb{C}[G]$ into pointwise mult. in $\mathbb{C}^{|G|}$. There is a very efficient algorithm to compute the transformation. & the inverse.

→ So what we do is embed the polynomials into $\mathbb{C}[C_m]$ to get $\sum a_i g^i$, $\sum b_i g^i$, compute their Discrete Fourier transform, Compute pointwise mult of their DFT's, and compute the inverse DFT

* Turns out using $\simeq m \log m \simeq n \log n$ mults, we can compute products of polynomials.

* The Cohn-Umans approach is to embed matrix mult. into group algebra mult. in an analogous way.

(Vague Plan)

appropriately (cleverly) chosen

① Mat Mult $\hookrightarrow \mathbb{C}[G] \xrightarrow[\text{DFT}]{} \mathbb{C}^{|G|}$

② Do pointwise mult in $\mathbb{C}^{|G|}$ and come back.

Defn [Right Quotient] S is a subset of a finite group. Define

$$Q(S) = \{ st^{-1} \mid s, t \in S \}$$

→ if S is a sub group, then $Q(S) = S$

Defn [Triple product property] Subsets X, Y, Z of G satisfy TPP if

$\forall\, x \in Q(X), y \in Q(Y), z \in Q(Z)$

$$xyz = 1 \implies x = y = z = 1$$

→ if X, Y, Z are subgroups,

$$xyz = 1 \implies x = y = z = 1$$

How TO EMBED?

G - finite group, S, T, U be subsets of G, and

$$\underset{\substack{\uparrow \\ |S|\times|T| \text{ matrix}}}{A} = (a_{s,t})_{s\in S,\, t\in T}, \quad \underset{\substack{\uparrow \\ |T|\times|U| \text{ matrix}}}{B} = (b_{t,u})_{t\in T,\, u\in U}$$

Define $\underset{\substack{\cap \\ \mathbb{C}[G]}}{\bar{A}} = \sum a_{s,t}\, s^{-1}t$, $\bar{B} = \sum b_{t,u}\, t^{-1}u$.

Turns out if S, T, U satisfy the triple product property, we can read off entries of AB from $\bar{A}\bar{B} \in \mathbb{C}[G]$

↳ $(AB)_{s,u}$ is the coeff of $s^{-1}u$ in $\bar{A}\bar{B}$

Thm [Wedderburn] $\mathbb{C}[G] \cong \mathbb{C}^{d_1\times d_1} \times \dots \times \mathbb{C}^{d_k \times d_k}$,

k is the no. of conjugacy classes of G.

d_i's are called "character degrees" of G.

$$\left(\Rightarrow \ |G| = \sum_{i=1}^{k} d_i^2 \right)$$

Thus the product of $|S|\times|T|$ matrix times $|T|\times|U|$ matrix reduces to many small matrix multiplications

Defn If you can find G and subsets X, Y, Z satisfying the TPP, then we say G realizes $M_{\langle |X|, |Y|, |Z| \rangle}$.

e.g. $C_k \times C_m \times C_n$ realizes $M_{\langle k,m,n\rangle}$ via the subgroups

$C_k \times \{1\} \times \{1\}$, $\{1\} \times C_m \times \{1\}$, $\{1\} \times \{1\} \times C_n$.

Thm If G realizes $M_{\langle k,m,n\rangle}$, then $M_{\langle k,m,n\rangle} \lesssim \mathbb{C}[G]$

abuse of notation to denote the tensor corresponding to algebra multiplication

In particular

$$R(M_{\langle k,m,n\rangle}) \le R(\mathbb{C}[G])$$

Proof Just read "HOW TO EMBED" $\boxtimes$

Summary

① G realizes $M_{\langle k,m,n\rangle} \Longrightarrow M_{\langle k,m,n\rangle} \lesssim \mathbb{C}[G]$

② Wedderburn's thm. states that $\mathbb{C}[G]$ is isomorphic to a product of matrix algebras

③ Thus mult. in $\mathbb{C}[G]$ (and more importantly matrix mult.) breaks down into many small matrix mults.

Thm For a non-trivial group G, define

$$\alpha(G) := \min\left\{ \frac{3\log|G|}{\log kmn} \;\middle|\; \begin{array}{l} G \text{ realizes } M_{\langle k,m,n\rangle}, \\ \text{one of } k,m,n > 1 \end{array} \right\}.$$

Then

(1) $2 < \alpha(G) \leq 3$

(2) If G is abelian, $\alpha(G)=3$

(3) If the character degrees of G are $d_1, \dots, d_t$, then

$$|G|^{\omega/\alpha(G)} \leq \sum_{i=1}^{t} d_i^{\omega}.$$

Proof (1a) $\alpha(G) \leq 3$ – trivial: for G, Let $H_1 = H_2 = \{1\}$, $H_3 = G$.

Thus G realizes $M_{\langle |G|,1,1\rangle}$. ✓

(1b) $2 < \alpha(G)$. Let G realize $M_{\langle k,m,n\rangle}$ via S_1, S_2, S_3, where $|Q(S_1)| = k$, $|Q(S_2)| = m$, $|Q(S_3)| = n$. Consider the map

$\rightarrow \; \phi : Q(S_1) \times Q(S_2) \to G$

$(x,y) \longmapsto x^{-1}y$

$x_2 x_1^{-1} y_1 y_2^{-1} \cdot 1 \in Q(S_3)$

– ϕ is injective $\left(x_1^{-1}y_1 = x_2^{-1}y_2 \Rightarrow x_2x_1^{-1}y_1y_2^{-1} = 1 \text{ \& by TPP} \right.$

$\longrightarrow \quad \left. \Rightarrow x_2x_1^{-1} = y_1y_2^{-1} = 1 \Longrightarrow x_1 = x_2 \text{ \& } y_1 = y_2 \right)$

– $\operatorname{Im}(\phi) \cap Q(S_3) = \{1\}$ (Suppose not. Then exists $z \in Q(S_3)$, $z \neq 1$ s.t

$x^{-1}y = z^{\in Q(S_3)} \Longrightarrow x^{-1}yz^{-1} = 1 \Rightarrow x^{-1} = y = z^{-1} = 1 = z$

Contradiction!)

$$\underset{Q(S_1)}{\underset{\in}{x^{-1}}}\,\underset{Q(S_2)}{\underset{\in}{y}} = z \implies x^{-1}yz^{-1} = 1 \Rightarrow x^{-1} = y = z^{-1} = 1 = z$$

Contradiction!)

$*$ $|G| \geq km$ (ineq. is strict unless $n = 1$)

$*$ (Due to symmetry) $|G| \geq mn$ & $|G| \geq km$

$*$ $|G|^3 \geq (kmn)^2$ (with ineq. strict unless $\underbrace{m = k = n = 1}_{\text{not in defn of } \alpha(G)}$)

$$\Rightarrow |G| > (kmn)^{2/3}$$

$$\Rightarrow \alpha(G) > 2. \quad ✓$$

(2) if G abelian, $\alpha(G) = 3$. Take

$$\psi : Q(S_1) \times Q(S_2) \times Q(S_3) \to G$$
$$(a, b, c) \mapsto abc.$$

ψ is injective ($a_1 b_1 c_1 = a_2 b_2 c_2$
$\Rightarrow a_1 a_2^{-1} b_1 b_2^{-1} c_1 c_2^{-1} = 1$
$\Rightarrow a_1 = a_2,\ b_1 = b_2,\ c_1 = c_2$)

Since ψ is injective

$$|G| \geq kmn \Rightarrow \alpha(G) \geq 3 \quad ✓$$

(3) Let (k', m', n') be triple responsible for $\alpha(G)$. This means, by defn

$$\alpha(G) = \frac{3 \log |G|}{\log k'm'n'} \Rightarrow (k'm'n')^{\alpha(G)} = |G|^3$$

By defn, G realizes $M_{\langle k', m', n'\rangle}$, so

$$M_{\langle k', m', n'\rangle} \lesssim \mathbb{C}[G] \cong \bigoplus_{i=1}^{t} M_{\langle d_i, d_i, d_i\rangle}$$

Take the lth tensor power

$$M_{\langle (k')^l, (m')^l, (n')^l\rangle} \lesssim \bigoplus_{i=1}^{t} \left(M_{\langle d_i, d_i, d_i\rangle}\right)^{\otimes l}$$
$$= \bigoplus^{t} M_{\langle d_{i_1} d_{i_2} \cdots d_{i_l},\, d_{i_1} d_{i_2} \cdots d_{i_l},\, \cdots\rangle}$$

$$= \bigoplus_{i_1,\dots,i_\ell=1}^{t} M_{\langle d_{i_1}d_{i_2}\cdots d_{i_\ell},\, d_{i_1}d_{i_2}\cdots d_{i_\ell},\, \dots\rangle}$$

Take rank

$$R\left(M_{\langle (k')^\ell, (m')^\ell, (n')^\ell\rangle}\right) \le \sum_{i_1,\dots,i_\ell=1}^{t} R\left(M_{\langle \prod_i d_i,\, \prod_i d_i,\, \prod_i d_i\rangle}\right)$$

$$= c\cdot\left(\sum_{i=1}^{t} d_i^{\omega+\varepsilon}\right)^{\ell}$$

$R(M_{\langle n,n,n\rangle}) = O(n^{\omega+\varepsilon})$ $\forall \varepsilon > 0$

defn of ω

Since $R\left(M_{\langle (k')^\ell, (m')^\ell, (n')^\ell\rangle}\right) \ge (k'm'n')^{\ell\omega/3}$, take ℓ^{th} roots

$$|G|^{\omega/\alpha} = (k'm'n')^{\omega/3} \le \sum_{i=1}^{t} d_i^{\omega+\varepsilon} \qquad ☒$$

APPLICATIONS:-

⊛ $H = C_n^3$, $G = H^2 \rtimes C_2$ ← C_2 acts on H^2 by switching the two factors

Let H_1, H_2, H_3 be the three factors of H viewed as subgroups.
$H_1 = C_n \times \{1\} \times \{1\}$ and so on...

Define subsets

$$S_i = \{(a,b)z^j \mid a \in H_i \setminus \{1\},\ b \in H_{(i\%3+1)},\ C_2 = \langle z\rangle,\ j \in \{0,1\}\}$$

Then G realizes $M_{\langle |S_1|, |S_2|, |S_3|\rangle}$ b coz

S_1, S_2, S_3 satisfy TPP

Setting $n = 17$ gives $\omega \le 2.91$

⊛ Using Wreath product groups $S_n \ltimes A^n$ gives $\omega < 2.41$ (Matches CW bound)

In general, you want $|G| \simeq n^2$, subgroups of size n, and small character degrees

In general, you

and small character degrees

Generalization of all this in the language Commutative Coherent Configuration (Association Schemes)

↳ "Do group theory with groups"

<u>Thm</u> $M_{\langle n,n,n \rangle}$ in a Commutative Coh. Configuration of rank $\simeq n^2$,

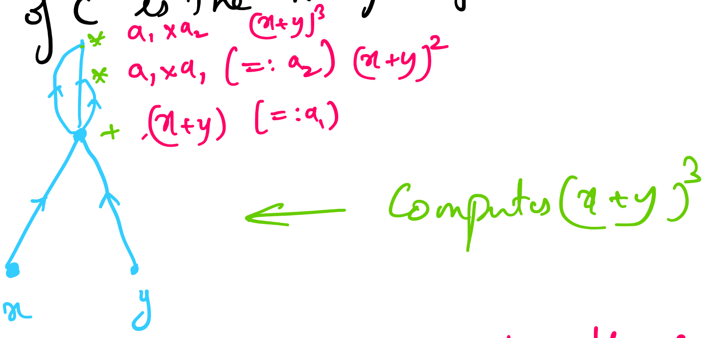
$\omega = 2$.

VP vs VNP, determinantal Complexity.

Wednesday, 24 May 2023 19:01

Defn An arithmetic Circuit C is a finite, directed, acyclic graph with vertices of in-degree 0 or 2, and exactly one vertex of out-degree 0.

- The vertices of in-degree 0 are labelled by elems of $\mathbb{C} \cup \{x_1, \dots, x_n\}$ Called leaves
- Those of in-degree 2 are labelled with + or *, called gates
- If out-degree of a vertex is 0, then it is called output gate
- The size of C is the no. of edges



$*$ $a_1 \times a_2$ $(x+y)^3$
$*$ $a_1 \times a_1$ $(=: a_2)$ $(x+y)^2$
$+$ $(x+y)$ $(=: a_1)$
x
y
Computes $(x+y)^3$

$*$ It is a fact that, upto a polynomial factor, the size of the circuit does not change if the inputs are arbitrary linear transformations on a vector space

Defn [VP] Let $d(n)$, $N(n)$ be polynomials in n, $f_n \in \mathbb{C}[x_1, \dots, x_{N(n)}]$, $\deg(f_n) \le d(n)$ ← seq. of polys. We say the seq. $(f_n) \in VP$ if there exists a sequence of circuits (C_n) of size polynomial in n, computing f_n.

Defn [VNP] A sequence (f_n) is in VNP if there exists a polynomial in n, i.e., $p(n)$, and a sequence $(g_n) \in VP$ s.t

$$f_n(x) = \sum_{e \in \{0,1\}^{p(n)}} g_n(x, e)$$

⊛ Think of sequences in VNP as projections of elements in VP.

Prop $(Per_n) \in VNP$

Proof Define $g_n(x_{1,1} \dots x_{n,n}, y_{1,1} \dots y_{n,n})$

$$:= \underbrace{\underbrace{\left(\prod_{\substack{i,j,l,m \in [n] \\ (i=l) \iff j \neq m}} (1 - y_{i,j} y_{l,m})\right)}_{\alpha_n(Y)} \underbrace{\left(\prod_{i=1}^{n} \sum_{j=1}^{n} y_{i,j}\right)}_{\beta_n(Y)}}_{\gamma_n(Y)} \underbrace{\left(\prod_{i=1}^{n} \sum_{j=1}^{n} x_{i,j} y_{i,j}\right)}_{\mu_n(X,Y)}$$

① $(g_n) \in VP$ (bcoz no of indets $2n^2$, degree of g_n is $O(n^3)$)

② $\gamma_n(e) \neq 0$ iff e is a permutation matrix

⊛ $\alpha_n(e) = 0$ iff there is a row or column with two or more 1's.

⊛ Suppose $\alpha_n(e) \neq 0$. Then $\beta_n \neq 0 \iff$ every row of e contains at least one 1.

⊛ Thus $\gamma_n(e) = \alpha_n(e)\beta_n(e) \neq 0$ iff e is a perm matrix

③ If e is a perm matrix, $\gamma_n(e) = 1$, $\mu_n(X,e) = \prod_{i=1}^{n} x_{i,\sigma(i)}$, where $\underset{S_n}{\overset{\cap}{\sigma}}$ corresponds to the perm e.

④ $$Per_n = \sum_{e \in \{0,1\}^{n^2}} g_n(X, e)$$ ☒

Plan upcoming

define C-complete, C-hard

$(det_n) \in VP$

VP vs VNP ~ P vs NP

non-uniform computation

det. comp (f)

non-uniform computation,

det. comp (f)

$$\frac{n^2}{2} \le dc(\text{perm}_n) \le 2^n - 1$$