Group-theoretic Algorithms for Matrix Multiplication

- CKSU05

Abstract

- for the first time, use the group-theoretic approach to derive algorithms faster than the standard algorithm
- 2.41
- two conjectures ($\Rightarrow \omega = 2$)

what are we working for?

- goal: the exponent of matrix multiplication, the smallest real number ω for which $n * n$ matrix multiplication can be performed in $O(n^{\omega+\varepsilon})$ operations for each $\varepsilon > 0$.

main work of Cohn and Umans[2003] in their previous paper,denoted as [2] in the paper we are studying

steps of the framework:

- one selects a finite group G satisfying a certain property
- reduce $n * n$ matrix multiplication to multiplication of elements of the *group algebra* C[G]
- via Fourier transform, the latter multiplication is reduced to several smaller matrix multiplication
- the size of those small matrices are the *character degrees of G*
- **Thus we get a recursive algorithm whose running time depends on the character degrees.**
- **Thus the problem of devising matrix multiplication algorithms is imported into the domain of group theory and representation theory.**

the main question raised in [2] is...

- whether the proposed approach could prove nontrivial bounds on ω (that is , to prove $\omega < 3$)
- this was shown to be equivalent to a question in representation theory:
- is there a group G with subsets S_1, S_2, S_3 that satisfy the *triple product property*, and for which $|S_1||S_2||S_3| > \sum_i d_i^3$,where d_i is the set of character degrees of G?
- In our paper we resolve this question in the affirmative.

now comes to our paper,some notations:

- The set $1, 2, ..., k$ is denoted $[k]$.
- The cyclic group of order k is denoted Cyc_k (with addition notation for the group law).
- The symmetric group on a set S is denoted $Sym(S)$ or Sym_n.
- If G is a group and R is a ring, then *R[G]* will denote the group algebra of G with coefficients in R.

some related basic facts in representation theory that will be used

- The group algebra C[G] of a finite group G decomposes as the direct product $C[G] \cong C^{d_1 * d_1} * ... * C^{d_k * d_k}$ of matrix algebras of orders $d_1, ..., d_k$. These orders are the character degrees of G.
- If we compute the dimensions of both sides, we have $|G| = \sum_i d_i^2$.
- If G has an abelian subgroup A, then all the character degrees of G are less than or equal to the index $[G : A]$.

some related basic facts in representation theory that will be used

Theorem (Lemma 1.1)

Let $s_1, s_2, \ldots, s_n$ be nonnegative real numbers, and suppose that for every vector $\mu = (\mu_1, \ldots, \mu_n)$ of nonnegative integers for which $\sum_{i=1}^{n} = N$ we have $\binom{N}{\mu} \prod_{i=1}^{n} s_i^{\mu_i} \leq C^N$. Then $\sum_{i=1}^{n} \leq C$.

summarize the necessary definition and results from [2],their previous paper

- If S is a subset of a group, let $Q(S)$ denote the right quotient set of S,i.e., $Q(S) = s_1 s_2^{-1} : s_1, s_2 \in S$.

Definition (**Definition 1.3([2]).**)

A group realizes $< n_1, n_2, n_3 >$ if there are subsets $S_1, S_2, S_3 \subseteq G$ such that $|S_i| = n_i$, and for $q_i \in Q(S_i)$,if $q_1 q_2 q_3 = 1$ then $q_1 = q_2 = q_3 = 1$. We call this condition on S_1, S_2, S_3 the **triple product property**.

summarize the necessary definition and results from [2]

Theorem (**Lemma 1.4([2]).**)

If G realizes $< n_1, n_2, n_3 >$, then it does so for every permutation of n_1, n_2, n_3.

Theorem (**Lemma 1.5([2]).**)

If $S_1, S_2, S_3 \subseteq G$ and $S_1', S_2', S_3' \subseteq G'$ satisfy the triple product property, then so do the subsets $S_1 \times S_1', S_2 \times S_2', S_3 \times S_3'$.

summarize the necessary definition and results from [2]

Theorem (**Theorem 1.6([2]).**)

Let R be any algebra over C(not necessarily commutative). If G realizes $< n, m, p >$, then the number of ring operations required to multiply $n \times m$ with $m \times p$ matrices over R is at most the number of operations required to multiply two elements of $R[G]$.

summarize the necessary definition and results from [2]

- Let $\triangle_n = (a, b, c) \in Z^3 : a + b + c = n - 1 \; and \; a, b, c \geq 0$.
- For $x \in \triangle_n$, we write $x = (x_1, x_2, x_3)$.
- Let H_1, H_2, H_3 be the subgroups of $Sym(\triangle_n)$ that preserve the first, second and third coordinates, respectively.
- Specifically, $H_i = \pi \in Sym(\triangle_n) : (\pi(x))_i = x_i forallx \in \triangle_n$.

Theorem (**Theorem 1.7([2]).**)

The subgroups H_1, H_2, H_3 defined above satisfy the triple product property.

summarize the necessary definition and results from [2]

Theorem (Theorem 1.8([2]).)

Suppose G realizes $< n, m, p >$ and the character degrees of G are $\{d_i\}$. Then $(nmp)^{\omega/3} \leq \sum_i {d_i}^{\omega}$.

Theorem (Corollary 1.9(2).)

Suppose G realizes $< n, m, p >$ and has largest character degree d. Then $(nmp)^{\omega/3} \leq d^{\omega-2}|G|$.

Proof.

Combine Thm 1.8[2] with the basic fact mentioned before that $|G| = \sum_i d_i^2$, then we have the corollary. $\square$

Beating the sum of the cubes

- Suppose G realizes $< n, m, p >$ and has character degrees $\{d_i\}$.
- Since $\omega \leq 3$,by ruling out the possibility of $\omega = 3$, Thm1.8[2] yields a nontrivial bound on ω if and only if $nmp > \sum_i d_i^3$.
- Then the question is : whether such a group exists?
- In this section we construct one (which shows that our methods do indeed prove nontrivial bounds on ω).

Beating the sum of the cubes

Theorem (Lemma 2.1.)

S_1,S_2, and S_3 satisfy the triple product property.

Proof.

Construct the example and show the proof on the **whiteboard**. $\square$

Uniquely solvable puzzles

Definition (USP)

A *uniquely solvable puzzle(USP) of width* k *is a subset* $U \subseteq 1, 2, 3^k$ *satisfying the following property: For all permutations* $\pi_1, \pi_2, \pi_3 \in Sym(U)$*,either* $\pi_1 = \pi_2 = \pi_3$ *or else there exist* $u \in U$ *and* $i \in [k]$ *such that at least two of* $(\pi_1(u))_i = 1$,$(\pi_2(u))_i = 2$,$(\pi_3(u))_i = 3$ *hold.*

Definition (strong USP)

A *strong USP of width* k *is a subset* $U \subseteq 1, 2, 3^k$ *satisfying the following property: For all permutations* $\pi_1, \pi_2, \pi_3 \in Sym(U)$*,either* $\pi_1 = \pi_2 = \pi_3$ *or else there exist* $u \in U$ *and* $i \in [k]$ *such that exactly two of* $(\pi_1(u))_i = 1$,$(\pi_2(u))_i = 2$,$(\pi_3(u))_i = 3$ *hold.*

Uniquely solvable puzzles

- show the example of a strong USP of size 8 and width 6 on the whiteboard

Theorem (Proposition 3.1)

For each $k \geq 1$, there exists a strong USP of size 2^k and width $2k$.

Proof.

By hand. $\square$

Uniquely solvable puzzles

Definition (the strong USP capacity)

We define the strong USP capacity to ba the largest constant C such that there exist strong USPs of size $(C - o(1))^k$ and width k for infinitely many values of k.
The USP capacity is defined analogously.

Uniquely solvable puzzles

- There is a simple upper bound for the USP capacity, which is of course an upper bound for the strong USP capacity as well.

Theorem (Lemma 3.2.)

The USP capacity is at most $(27/4)^{1/3}$.

Proof.

On the board. □

Uniquely solvable puzzles

- In section 6 of [3] they show implicitly that Lemma 3.2 is sharp.

Theorem (**Theorem 3.3(Coppersmith and Winograd[3]).**)
The USP capacity equals $(27/4)^{1/3}$.

Theorem (**Conjecture 3.4.**)
The strong USP capacity equals $(27/4)^{1/3}$.

- This conjecture would imply that $\omega = 2$.

Using strong USPs

Definition

Given a strong USP U of width k, let H be the abelian group of all functions from $U \times [k]$ to the cyclic group Cyc_m(H is a group under pointwise addition).
The symmetric group Sym(U) acts on *(H) via* $\pi(h)(u, i) = h(\pi^{-1}(u), i)$ *for* $\pi \in Sym(U)$,$h \in H$, $u \in U$ *and* $i \in [k]$. *Let* G *be the semidirect product* $H \rtimes Sym(U)$, *and define subsets* S_1, S_2, S_3 *of* G
by letting S_i *consist of all products* π *with* $\pi \in Sym(U)$ *and* $h \in H$ *satisfying* $h(u, j) \neq 0$ *iff* $u_j = i$ *for all* $u \in U$ *and* $j \in [k]$.

Theorem (**Proposition 3.5.**)

If U *is a strong USP, then* S_1,S_2, *and* S_3 *satisfy the triple product property.*

Proof.

On the board. $\square$

Using strong USPs

Theorem (**Corollary 3.6.**)

On the board,with the proof.

- several bounds (on the board):$2.67, 2.48, 2$

The triangle construction

- Suppose $U \subseteq (1,2,3)^k$ is a subset with only two symbols occurring in each coordinate.
  Let H_1 be the subgroup of Sym(U) that preserves the coordinates in which only 1 and 2 occur,
  H_2 the subgroup preserving the coordinates in which only 2 and 3 occur,
  and H_3 the subgroup preserving the coordinates in which only 1 and 3 occur.

Theorem (**Lemma 3.7.**)

The set U is a USP iff H_1,H_2, and H_3 satisfy the triple product property within Sym(U).

Proof.

On the board. □

The triangle construction

Theorem (Proposition 3.8.)

For each $k \geq 1$, there exists a strong USP of size $2^{k-1}(2^k + 1)$ and width $3k$.

Proof.

On the board. □

- It follows that the strong USP capacity is at least $2^(2/3)$
- and $\omega < 2.48$.
- Show the reason on the whiteboard now:

Theorem (Corollary 3.9.)

If U is a USP of width k such that only two symbols occur in each coordinate, then $|U| \leq (2^{2/3} + o(1))^k$.

Proof.

em...?how to prove? □

- The only upper bound on the size of a strong USP is in Lemma 3.2.

The simultaneous double product property

- *simultaneous double product property will be used to modify the underlying group of the combinatorial structure in the algebraic direction.*

Definition (*double product property*)

We say that subsets S_1, S_2 of a group H satisfy the *double product property if*

$q_1 q_2 = 1$ *implies* $q_1 = q_2 = 1$, *where* $q_i \in Q(S_i)$.

The simultaneous double product property

Definition (**Definition 4.1.**)

We say that n pairs of subsets A_i, B_i(for $1 \leq i \leq n$) of a group H satisfy the *simultaneous double product property if*

- *for all i, the pair A_i, B_i satisfies the double product property, and*
- *for all i,j,k, $a_i(a'_j)^{-1}$ implies $i = k$, where $a_i \in A_i, a'_j \in A_j, b_j \in B_j, and b'_k \in B_k$.*

The simultaneous double product property

Theorem (**Lemma 4.2.**)

If n pairs of subsets A_i, B_i satisfy the simultaneous double product property, and n' pairs of subsets $A'_i, B'_i \subseteq H'$ satisfy the simultaneous double product property, then so do the nn' pairs of subsets $A_i \times A'_i, B_j \times B'_j \subseteq H \times H'$.

The simultaneous double product property

- $\Delta_n = \{(a, b, c) \in Z^3 : a + b + c = n - 1 \text{ and } a, b, c \geq 0\}$.
- Given n pairs of subsets A_i, B_i in H for $0 \leq i \leq n - 1$.

Definition

we define triples of subsets in H^3 indexed by $v = (v_1, v_2, v_3) \in \Delta_n$ as follows:
$\widehat{A_v} = A_{v_1} \times \{1\} \times B_{v_3}$
$\widehat{B_v} = B_{v_1} \times A_{v_2} \times \{1\}$
$\widehat{C_v} = \{1\} \times B_{v_2} \times A_{v_3}$

The simultaneous double product property

Theorem (**Theorem 4.3.**)

If n pairs of subsets $A_i, B_i \subseteq H$(with $0 \leq i \leq n-1$) satisfy the simultaneous double product property, then the following subsets S_1, S_2, S_3 of $G = (H^3)^{\Delta_n} \rtimes Sym(\Delta_n)$ satisfy the triple product property:

$S_1 = \widehat{a}\pi : \pi \in Sym(\Delta_n), \widehat{a_v} \in \widehat{A_v}$ *for all* v

$S_2 = \widehat{b}\pi : \pi \in Sym(\Delta_n), \widehat{b_v} \in \widehat{B_v}$ *for all* v

$S_3 = \widehat{c}\pi : \pi \in Sym(\Delta_n), \widehat{c_v} \in \widehat{C_v}$ *for all* v

The simultaneous double product property

Theorem (Theorem 4.4.)

If H is a finite group with character degrees $\{d_k\}$, and n pairs of subsets $A_i, B_i \subseteq H$ satisfy the simultaneous double product property, then
$\sum_{i=1}^{n}(|A_i||B_i|)^{\omega/2} \leq (\sum_k d_k^{\omega})^{3/2}$.

Proof.
On the board. $\square$

- Using this theorem, the example after Definition 4.1 recovers the trivial bound $\omega \leq 3$ as $k \to \infty$. Show the proof.

The simultaneous double product property

- Now we use two parameters α and β to describe pairs satisfying the simultaneous double product property:
- if there are n pairs,choose α and β so that $|A_i||B_i| \geq n^\alpha$ for all i and $|H| = n^\beta$.
- If H is abelian Theorem 4.4 implies $\omega \leq (3\beta - 2)/\alpha$. show the calculations.

The simultaneous double product property

Theorem (Proposition 4.5.)

For each $m \geq 2$, there is a construction in Cyc_m^{2l} satisfying the simultaneous double product property with $\alpha = log_2(m-1) + o(1)$ and $\beta = log_2 m + o(1)$ as $l \to \infty$.

Proof.

By hand. (kind of disagree with the last part of the proof on the paper) □

- Taking $m = 6$ yields exactly the same bound as in Subsection 3.3 ($\omega \leq 2.48$).

The simultaneous double product property

- The only limitations we know of on the possible values of α and β are the following:

Theorem (**Proposition 4.6.**)

If n pairs of subsets $A_i, B_i \subseteq H$ satisfy the simultaneous double product property, with $|A_i||B_i| \geq n^\alpha$ for all i and $|H| = n^\beta$,then $\alpha \leq \beta$ and $\alpha + 2 \leq 2\beta$.

Proof.

by hand $\square$

The simultaneous double product property

- The most important case is when H is an abelian group. There the bound on ω is $\omega \leq (3\beta - 2)/\alpha$. We've mentioned this.
- Proposition 4.6 shows that the only way to achieve $\omega = 2$ is $\alpha = \beta = 2$. show it by hand.
- and we conjecture that this is possible:

Theorem (Conjecture 4.7.)

For arbitrarily large n, there exists an abelian group H with n pairs of subsets A_i, B_i satisfying the simultaneous double product property such that $|H| = n^{2+o(1)}$ and $|A_i||B_i| \geq n^{2-o(1)}$.

The simultaneous triple product property

- say something on the board
- This apportionment can be viewed as *reducing several independent matrix multiplication problems to a single group algebra multiplication, using triples of subsets satisfying the simultaneous triple product property:*

The simultaneous triple product property

Definition (**Definition 5.1.**)

We say that n triples of subsets A_i, B_i, C_i (for $1 \leq i \leq n$) of a group H satisfy the simultaneous triple product property if
for each i, the three subsets A_i, B_i, C_i satisfy the triple product property, and
for all i,j,k, $a_i(a'_j)^{-1}b_j(b'_k)^{-1}c_k(c'_i)^{-1} = 1$ implies $i = j = k$
for $a_i \in A_i, a'_j \in A_j, b_j \in B_j, b'_k \in B_k, c_k \in C_k$ and $c'_i \in C_i$.
We say that such a group simultaneous realizes
$< |A_i|, |B_i|, |C_i| >$,...,$< |A_n|, |B_n|, |C_n| >$.

The simultaneous triple product property

- Let $H = Cyc_n^3$, and call the three factors H_1, H_2 and H_3. Define the following sets:
- $A_1 = H_1 \setminus \{0\}$, $B_1 = H_2 \setminus \{0\}$, $C_1 = H_3 \setminus \{0\}$
- $A_2 = H_2 \setminus \{0\}$, $B_2 = H_3 \setminus \{0\}$, $C_2 = H_1 \setminus \{0\}$

Theorem (**Proposition 5.2.**)

The two triples A_1, B_1, C_1 and A_2, B_2, C_2 satisfy the simultaneous triple product property.

Proof.

by hand □

- The reason for the strange condition in the definition of the simultaneous triple product property is that it is exactly what is needed to *reduce several independent matrix multiplications to one group algebra multiplication.*

The simultaneous triple product property

Theorem (**Theorem 5.3.**)

Let R be any algebra over $\mathbb{C}$.If H simultaneous realizes $< n_1, m_1, p_1 >, ..., < n_k, m_k, p_k >$, then the number of ring operations required to perform k independent matrix multiplications of sizes $n_1 \times m_1$ by $m_1 \times p_1$,...,$n_k \times m_k$ by $m_k \times p_k$ is at most the number of operations required to multiply two elements of $R[H]$.

Proof.

by hand $\square$

The simultaneous triple product property

Theorem (**Lemma 5.4.**)

If n triples of subsets $A_i, B_i, C_i \subseteq H$ satisfy the simultaneous triple product property, and n' triples of subsets $A'_i, B'_i, C'_i \subseteq H'$ satisfy the simultaneous triple product property,then so do nn' triples of subsets $A_i \times A'_j, B_i \times B'_j, C_i \times C'_j \subseteq H \times H'$.

We will talk about Thm 5.5 and its proof in the last part and show further more that *any bound on ω that can be achieved using the simultaneous triple product property can also be achieved using the ordinary triple product property, but it is an important organizing principle.*

Local strong USPs

In this section we explain how to interpret each of our constructions in this setting.

Definition (local strong USPs)

A local strong USP of width k is a subset $U \subseteq \{1,2,3\}^k$ such that for each ordered triple $(u,v,k) \in U^3$, with u,v,and w not all equal, there exists $i \in [k]$ such that (u_i, v_i, w_i) is an element of $\{(1,2,1),(1,2,2),(1,1,3),(1,3,3),(2,2,3),(3,2,3)\}$.

Theorem (Lemma 6.1.)

Every local strong USP is a strong USP.

Proof.

by hand $\square$

Local strong USPs

Theorem (Theorem 6.2.)

Let U be a local strong USP of width k, and for each $u \in U$ define subsets $A_u, B_u, C_u \subseteq Cyc_l^k$ by
$A_u = x \in Cyc_l^k : x_j \neq 0 \text{ iff } u_j = 1,$
$B_u = x \in Cyc_l^k : x_j \neq 0 \text{ iff } u_j = 2,$ *and*
$C_u = x \in Cyc_l^k : x_j \neq 0 \text{ iff } u_j = 3.$
Then the triples A_u, B_u, C_u satisfy the simultaneous triple product property.

Proof.

by hand. I think there's something wrong in the proof on the paper. $\square$

Local strong USPs

Theorem (Proposition 6.3.)

The strong USP capacity is achieved by local strong USPs. In particular, given any strong USP U of width k, there exists a local strong USP of size $|U|!$ and width $|U|k$.

Proof.

by hand $\square$

Section 6.2 and 6.3 are omitted here in the presentation.

The wreath product construction

- Let H be a group, and define $G = Sym_n \ltimes H^n$, where the symmetric group Sym_n acts on H^n from the right by permuting the coordinates according to $(h^\pi)_i = h_{\pi_i}$. We write elements of G as $h\pi$ with $h \in H^n$ and $\pi \in Sym_n$.

The wreath product construction

Theorem (**Theorem 7.1.**)

If n triples of subsets $A_i, B_i, C_i \subseteq H$ satisfy the simultaneous triple product property, then the following subsets H_1, H_2, H_3 of $G = Sym_n \ltimes H^n$ satisfy the triple product property:
$H_1 = \{h\pi : \pi \in Sym_n, h_i \in A_i \text{ for each } i\}$
$H_2 = \{h\pi : \pi \in Sym_n, h_i \in B_i \text{ for each } i\}$
$H_3 = \{h\pi : \pi \in Sym_n, h_i \in C_i \text{ for each } i\}$

Proof.

by hand $\square$

The wreath product construction

Theorem (**Theorem 5.5.**)

If a group H simultaneously realizes $< a_1, b_1, c_1 >$*,...,*$< a_n, b_n, c_n >$ *and has character degrees* $\{d_k\}$*, then* $\sum_{i=1}^{n}(a_i b_i c_i)^{\omega/3} \leq \sum_k d_k^{\omega}$*.*

Proof.

by hand $\square$

Frequently H will be abelian, in which case $\sum_k d_k^{\omega} = |H|$. That occurs in the example from Prop.5.2, which proves that $\omega < 2.93$ using Theorem 5.5. show the calculations by hand.

any bound that can be derived from Theorem 5.5 can be proved using Theorem 1.8 as well.