Group-theoretic approach to Fast Matrix Multiplication

Jenya Kirshtein

Department of Mathematics
University of Denver

Group-theoretic approach to Fast Matrix Multiplication (DU November 7, 2008)
References


Matrix multiplication

- Standard matrix multiplication takes $2n^3$ arithmetic operations
- Strassen suggested $O(n^{2.81})$ recursive algorithm in 1969

*Question:* What is the smallest number $w$ such that $\forall \epsilon > 0$ matrix multiplication can be performed in at most $O(n^{w+\epsilon})$?

*Question:* Is $w = 2$?

- The best known bound is $w < 2.38$ due to Coppersmith and Winograd (1990)
Outline

- Embed matrices $A, B$ into the elements $\overline{A}, \overline{B}$ of the group algebra $\mathbb{C}[G]$.

- Multiplication of $\overline{A}$ and $\overline{B}$ in the group algebra is carried out in the Fourier domain after performing the Discrete Fourier Transform (DFT) of $\overline{A}$ and $\overline{B}$.

- The product $\overline{A} \overline{B}$ is found by performing the inverse DFT.

- Entries of the matrix $AB$ can be read off from the group algebra product $\overline{A} \overline{B}$.

We’ll consider square matrices and finite groups.
Group Algebra

**Def:** Let $G$ be a finite group and $F$ be a field. A *group algebra* $F[G]$ of $G$ over $F$ is the set of all linear combinations of elements of $G$ with coefficients in $F$, with the natural addition and scalar multiplication, and multiplication defined by

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

Group-theoretic approach to Fast Matrix Multiplication (DU November 7, 2008)
Discrete Fourier Transform (DFT)

Embed polynomials as elements of the group algebra $\mathbb{C}[G]$: Let $G = \langle z \rangle$ be a cyclic group of order $m \geq 2n$. Define

$$\overline{A} = \sum_{i=0}^{n-1} a_i z^i \quad \text{and} \quad \overline{B} = \sum_{i=0}^{n-1} b_i z^i$$

Discrete Fourier Transform is an invertible linear transformation $D : \mathbb{C}[G] \rightarrow \mathbb{C}^{\mid G\mid}$, such that

$$D(\overline{A}) = \left( \sum_{i=0}^{n-1} a_i x_0^i, \sum_{i=0}^{n-1} a_i x_1^i, \ldots, \sum_{i=0}^{n-1} a_i x_{n-1}^i \right),$$

where $x_k = e^{\frac{2\pi i}{n} k}$ is the $n^{th}$ root of unity. Then

$$\overline{A} \overline{B} = D^{-1}(D(\overline{A})D(\overline{B}))$$

*Fast Fourier Transform* algorithm computes DFT in $O(n \log(n))$ arithmetic operations.
Matrices as elements of the group algebra

Let $F$ be a field and $S$, $T$ and $U$ be subsets of $G$.

$$A = (a_{s,t})_{s \in S, t \in T} \quad \text{and} \quad B = (b_{t,u})_{t \in T, u \in U}$$

are $|S| \times |T|$ and $|T| \times |U|$ matrices, indexed by elements of $S$, $T$ and $T$, $U$, respectively. Then embed $A, B$ as elements $\overline{A}, \overline{B} \in F[G]$:

$$\overline{A} = \sum_{s \in S, t \in T} a_{s,t} s^{-1}t \quad \text{and} \quad \overline{B} = \sum_{t \in T, u \in U} b_{t,u} t^{-1}u$$
Triple product property

Let $S \subseteq G$. Let $Q(S)$ denote the right quotient set of $S$, i.e.

$$Q(S) = \{ s_1s_2^{-1} : s_1s_2 \in S \}$$

**Def:** A group $G$ realizes $\langle n_1, n_3, n_3 \rangle$ if $\exists$ subsets $S_1, S_2, S_3 \subseteq G$ s.t. $|S_i| = n_i$ and for $q_i \in Q(S_i)$,

if $q_1q_2q_3 = 1$ then $q_1 = q_2 = q_3 = 1$

(equivalently, if $q_1q_2 = q_3$ then $q_1 = q_2 = q_3 = 1$)

This condition on $S_1, S_2, S_3$ is the **triple product property.**
Theorem [1] Let $F$ be a field. If $G$ realizes $< n, n, n >$, then the number of field operations required to multiply two $n \times n$ matrices over $F$ is at most the number of operations required to multiply two elements of $F[G]$.

Proof. Index $A$ with the sets $S, T$; $B$ with $T, U$ and $AB$ with $S, U$. Consider the product in $F[G]$

$$
\left( \sum_{s \in S, t \in T} a_{st}s^{-1}t \right) \left( \sum_{t' \in T, u \in U} b_{t'u}t'^{-1}u \right)
$$

By triple product property, 

$$
(s^{-1}t) \left( t'^{-1}u \right) = s'^{-1}u' \text{ iff } s = s', t = t', u = u',
$$

so the coefficient of $s^{-1}u$ in the product is

$$
\sum_{t \in T} a_{st}b_{tu} = (AB)_{su}
$$
Characters and character degrees of $G$

**Def:** Let $F$ be $\mathbb{R}$ or $\mathbb{C}$. The set of all invertible $n \times n$ matrices with entries in $F$, under matrix multiplication, forms a group called the *general linear group* $GL(n, F)$.

**Def:** The *trace* of an $n \times n$ matrix $A = (a_{ij})$ is $tr A = \sum_{i=1}^{n} a_{ii}$.

**Def:** A *representation* of $G$ over $F$ is a homomorphism $\rho : G \to GL(n, F)$ for some integer $n$ ($n$ is the degree of $\rho$).

**Def:** Suppose $\rho$ is a representation of $G$. With each $n \times n$ matrix $\rho(g)$ ($g \in G$) we associate the complex number $tr(\rho(g))$, which we call $\chi(g)$. The function $\chi : G \to \mathbb{C}$ is called the *character* of the representation $\rho$ of $G$.

**Def:** If $\chi$ is a character of $G$, then $\chi(1)$ is called a *degree* of $\chi$. 

Group-theoretic approach to Fast Matrix Multiplication (DU November, 7 2008)
Theorem [2] Suppose $G$ realizes $\langle n, n, n \rangle$ and the character degrees of $G$ are $\{d_i\}$. Then
\[ n^w \leq \sum_i d_i^w \]

Theorem [3] Let $G$ be a group with the character degrees $\{d_i\}$. Then $|G| = \sum_i d_i^2$

Corollary [2] Suppose $G$ realizes $\langle n, n, n \rangle$ and has largest character degree $d$. Then
\[ n^w \leq d^{w-2}|G| \]

Question [1] Does there exist a finite group $G$ that realizes $\langle n, n, n \rangle$ and has character degrees $\{d_i\}$ such that
\[ n^3 > \sum_i d_i^3, \]
i.e. $G$ gives a nontrivial bound on $w$. 

Group-theoretic approach to Fast Matrix Multiplication (DU November, 7 2008)
**Example**

**Def:** If $G$ is a group and $X$ is a set, then a (left) **group action** of $G$ on $X$ is a binary function $G \times X \rightarrow X$ denoted $(g, x) \rightarrow g \cdot x$, which satisfies the following two axioms:

1. $(gh) \cdot x = g \cdot (h \cdot x) \quad \forall g, h \in G, \forall x \in X$;
2. $1 \cdot x = x \quad \forall x \in X$ (1 is the identity element of $G$).

**Def:** If $G$ and $H$ are groups with a left action of $G$ on $H$, then the **semidirect product** $H \rtimes G$ is the set $H \times G$ with the multiplication law

$$(h_1, g_1) (h_2, g_2) = (h_1 (g_1 \cdot h_2), g_1 g_2).$$
Example

Denote $Cyc_n$ the cyclic group of order $n$.

Let $H = Cyc_n^3$ and $G = H^2 \times Cyc_2$

$Cyc_2 = \{1, z\}$ acts on $H^2$ by switching the two factors, then $G = \{(a, b) z^i : a, b \in H, i \in \{0, 1\}\}$

Let $H_1, H_2, H_3 \leq H$ be the three factors of $Cyc_n$ in $H$, $H_4 = H_1$.

Define subsets $S_1, S_2, S_3 \subseteq G$ by

$$S_i = \{(a, b) z^j : a \in H_i \setminus \{0\}, b \in H_{i+1}, j \in \{0, 1\}\}$$

$S_1, S_2, S_3$ satisfy the triple product property.
Example

\[ H = \text{Cyc}_n^3 \text{ and } G = H^2 \times \text{Cyc}_2 \]
\[ S_i = \{(a, b) z^j : a \in H_i \setminus \{0\}, b \in H_{i+1}, j \in \{0, 1\}\} \]

**Theorem [4]** Let \( G \) be nonabelian group and \( m \) be an integer. If there exists abelian \( H \triangleleft G \) of index \( m \), then character degrees of \( G \) are 1 and \( m \).

**Theorem** If \( H \leq G \) and \( [G : H] = 2 \), then \( H \triangleleft G \).

Since \( H^2 \) is an abelian subgroup of \( G \) of index \( [G : H] = \frac{|G|}{|H|} = 2 \), character degrees of \( G \) are at most 2.

Then since \( \sum_i d_i^2 = |G|, \quad \sum_i d_i^3 \leq \max \{d_i\} \sum_i d_i^2 \leq 2|G| = 4n^6 \)

Also, \( |S_i| = 2n(n - 1) \) and \( |S_1||S_2||S_3| = 8n^3(n - 1)^3 \),

so we have \( |S_1||S_2||S_3| > \sum_i d_i^3 \) for \( n \geq 5 \).
Example

By Corollary, $|S_1|^w \leq d^{w-2}|G|$ or

$2n(n-1)^w \leq 2^{w-2}2n^6$,

which leads to $w \leq \frac{6\ln n - \ln 2}{\ln(n(n-1))}$.

The best bound on $w$ ($w \leq 2.9088$) is achieved by setting $n = 17$. 
Other results

• *Uniquely Solvable Puzzles* prove $w < 2.48$

• *Local Uniquely Solvable Puzzles* prove $w < 2.41$ and $w < 2.376$

• Conjecture that a *Strong Uniquely Solvable Puzzle* capacity equals $3/2^3$ would imply that $w = 2$

• Conjecture about the existence for any $n$ of an abelian group $G$ with $n$ pairs of subsets $A_i, B_i$ satisfying the *simultaneous double product property* such that $|G| = n^{2+o(1)}$ and $|A_i||B_i| \geq n^{2-o(1)}$ would imply that $w = 2$
Thank you!

Group-theoretic approach to Fast Matrix Multiplication (DU November, 7 2008)