Adapted Bases and Fast Transforms

Michael Orrison

Harvey Mudd College

Joint Work with Michael Hansen and Masanori Koyama

Functions

Let $X = \{x_1, \ldots, x_n\}$ be a finite set, and let $\mathbb{C}X$ be the complex vector space of complex-valued functions defined on X:

$$\mathbb{C}X = \{f : X \to \mathbb{C}\}.$$

We identify $x \in X$ with the function that is 1 on x and 0 on all of other elements.

Using the x_i as a basis, we can then encode $f \in \mathbb{C}X$ as a column vector:

$$f \mapsto \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{bmatrix}.$$

These column vectors represent the data we would like to analyze.

Actions

If G is a finite group acting on X, then G acts on $\mathbb{C}X$ where if $g \in G$ and $f \in \mathbb{C}X$, then

$$(g \cdot f)(x) = f(g^{-1}x).$$

We can then extend this action to the group algebra $\mathbb{C}G$, which makes $\mathbb{C}X$ a $\mathbb{C}G$-permutation module. In the background is a permutation representation

$$\varphi : G \to GL_{|X|}(\mathbb{C})$$

where $\varphi(g)$ is a permutation matrix for all $g \in G$ that encodes the action of G on X:

$$[\varphi(g)]_{ij} = \begin{cases} 1 & \text{if } g \cdot x_j = x_i \\ 0 & \text{otherwise.} \end{cases}$$

The Big Idea

Let G be a finite group acting on X, and let $\mathbb{C}X$ be the resulting permutation module. If we write $\mathbb{C}X$ as a direct sum

$$\mathbb{C}X = U_1 \oplus \cdots \oplus U_m$$

of submodules, then every $f \in \mathbb{C}X$ can be written uniquely as

$$f = f_1 + \cdots + f_m$$

where $f_i \in U_i$. If the U_i are meaningful, then we might be able to better understand f by focusing our attention on the f_i. This leads to generalized spectral analysis, which was pioneered by Diaconis.

Example

If $G = S_3$ and X is the set $\{1, 2, 3\}$, then under the usual action of S_3, we have

$$\mathbb{C}X = \langle \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \rangle \oplus \langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \rangle.$$

For example,

$$\begin{bmatrix} 11 \\ 10 \\ 12 \end{bmatrix} = \begin{bmatrix} 11 \\ 11 \\ 11 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 20 \\ 3 \\ 10 \end{bmatrix} = \begin{bmatrix} 11 \\ 11 \\ 11 \end{bmatrix} + \begin{bmatrix} 9 \\ -8 \\ -1 \end{bmatrix}.$$

Example

If $G = X = \mathbb{Z}/4\mathbb{Z} = \{1, z, z^2, z^3\}$ acts on itself by left multiplication, then we have

$$\mathbb{C}X = \langle \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \rangle \oplus \langle \begin{bmatrix} 1 \\ i \\ -1 \\ -i \end{bmatrix} \rangle \oplus \langle \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \rangle \oplus \langle \begin{bmatrix} 1 \\ -i \\ -1 \\ i \end{bmatrix} \rangle.$$

Note that the associated permutation representation is such that

$$z \mapsto \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Example

If $G = S_n$ and X is the set of k-element subsets of $\{1, \ldots, n\}$ where $k \leq n/2$, then we can write

$$\mathbb{C}X = U_0 \oplus U_1 \oplus \cdots \oplus U_k$$

where U_i corresponds to pure i-th order effects, and $f \in \mathbb{C}X$ is typically viewed as voting data.

Given $f = f_0 + f_1 + \cdots + f_k$, we might ask about the extent to which $||f_i||$ depends on $f_0, \cdots, f_{i-1}$. (See *Algebraic algorithms for sampling from conditional distributions* by Diaconis and Sturmfels in Ann. Statist. Volume 26, Number 1 (1998), 363-397.)

Example

If $G = S_3$ and X is the set $\{1, 2, 3\}$, then under the usual action of S_1, S_2, and S_3, we have

$$\begin{aligned}
\mathbb{C}X &= \langle \begin{bmatrix}1\\0\\0\end{bmatrix} \rangle \oplus \langle \begin{bmatrix}0\\1\\0\end{bmatrix} \rangle \oplus \langle \begin{bmatrix}0\\0\\1\end{bmatrix} \rangle \\
&= \langle \begin{bmatrix}1\\1\\0\end{bmatrix} \rangle \oplus \langle \begin{bmatrix}1\\-1\\0\end{bmatrix} \rangle \oplus \langle \begin{bmatrix}0\\0\\1\end{bmatrix} \rangle \\
&= \langle \begin{bmatrix}1\\1\\1\end{bmatrix} \rangle \oplus \langle \begin{bmatrix}1\\-1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\-2\end{bmatrix} \rangle
\end{aligned}$$

where these decompositions also reflect the different orbits and a certain kind of adaptedness.

Questions

Suppose $\mathbb{C}X = U_1 \oplus \cdots \oplus U_m$ where the U_i are $\mathbb{C}G$-submodules.

1. Given $f \in \mathbb{C}X$, how efficiently can we compute $f_1, \ldots, f_m$? How efficiently can we compute $\|f_1\|, \ldots, \|f_m\|$?

2. Given a basis $\mathcal{B}_i$ for each U_i, how efficiently can we do a change-of-basis from X to the basis $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_m$?

3. How should the U_i and the $\mathcal{B}_i$ be chosen above so as to be meaningful and also computationally helpful?

Machinery

Discrete Fourier Transforms

Every complex group algebra $\mathbb{C}G$ is isomorphic to a direct sum of matrix algebras:

$$\mathbb{C}G \cong \mathbb{C}^{d_1 \times d_1} \oplus \cdots \oplus \mathbb{C}^{d_h \times d_h}.$$

Any associated isomorphism $D = D_1 \oplus \cdots \oplus D_h$ is a (generalized) discrete Fourier transform or DFT.

Note that the D_i form a complete set of irreducible representations for G, and any complete set of irreducible representations for G can be used in this way to construct a DFT for G.

Example

$$D : \mathbb{C}S_3 \to \mathbb{C}^{1\times 1} \oplus \mathbb{C}^{1\times 1} \oplus \mathbb{C}^{2\times 2} = \begin{bmatrix}\cdot\end{bmatrix} \oplus \begin{bmatrix}\cdot\end{bmatrix} \oplus \begin{bmatrix}\cdot & \cdot \\ \cdot & \cdot\end{bmatrix}$$

$$1 \mapsto \begin{bmatrix}1\end{bmatrix} \oplus \begin{bmatrix}1\end{bmatrix} \oplus \begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix}$$

$$b^1_{11} \mapsto \begin{bmatrix}1\end{bmatrix} \oplus \begin{bmatrix}0\end{bmatrix} \oplus \begin{bmatrix}0 & 0 \\ 0 & 0\end{bmatrix}$$

$$b^3_{21} \mapsto \begin{bmatrix}0\end{bmatrix} \oplus \begin{bmatrix}0\end{bmatrix} \oplus \begin{bmatrix}0 & 0 \\ 1 & 0\end{bmatrix}$$

$$b^3_{11} + b^3_{22} \mapsto \begin{bmatrix}0\end{bmatrix} \oplus \begin{bmatrix}0\end{bmatrix} \oplus \begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix}$$

Subgroup-Adapted DFTs

Suppose $H \leq G$. The DFT D for G is subgroup-adapted to the chain $H \leq G$ if for each irreducible representation D_i of G and for all $h \in H$,

1. $D_i(h)$ is block diagonal, where the blocks correspond to irreducible representations of H, and
2. equivalent blocks among all of the D_i are actually equal.

This can also be extended to longer chains subgroups of G, which we'll usually take to have the form

$$\{1\} = G_0 < G_1 < \cdots < G_n = G.$$

Example

$$\mathbb{C}S_2 \cong \mathbb{C}^{1\times 1} \oplus \mathbb{C}^{1\times 1}$$

$$\mathbb{C}S_3 \cong \mathbb{C}^{1\times 1} \oplus \mathbb{C}^{1\times 1} \oplus \mathbb{C}^{2\times 2}$$

$$D : \mathbb{C}S_3 \to \mathbb{C}^{1\times 1} \oplus \mathbb{C}^{1\times 1} \oplus \mathbb{C}^{2\times 2}$$

$$D|_{\mathbb{C}S_2} : \mathbb{C}S_2 \to \begin{bmatrix} \bullet \end{bmatrix} \oplus \begin{bmatrix} \star \end{bmatrix} \oplus \begin{bmatrix} \star & 0 \\ 0 & \bullet \end{bmatrix}$$

Adapted Bases

A basis $\mathcal{B}$ for $\mathbb{C}X$ is adapted to the DFT D if it can be partitioned

$$\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$$

so that each $\mathcal{B}_j$ spans an irreducible submodule of $\mathbb{C}X$, and if this submodule corresponds to D_i, then $[g]_{\mathcal{B}_j} = D_i(g)$ for all $g \in G$.

Example

If $G = S_3$ and X is the set $\{1, 2, 3\}$, then under the usual action of S_1, S_2, and S_3, we have

$$\begin{aligned}
\mathbb{C}X &= \langle \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \rangle \oplus \langle \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \rangle \oplus \langle \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \rangle \\
&= \langle \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \rangle \oplus \langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \rangle \oplus \langle \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \rangle \\
&= \langle \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \rangle \oplus \langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} , \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \rangle.
\end{aligned}$$

Creating Adapted Bases

Let $D : \mathbb{C}G \to \mathbb{C}^{d_1 \times d_1} \oplus \cdots \oplus \mathbb{C}^{d_h \times d_h}$ be a DFT. Let b_{ij}^k be the unique element in $\mathbb{C}G$ such that $D(b_{ij}^k)$ has zeros everywhere except for a 1 in the (i, j) entry of D_k. Then the collection $\{b_{ij}^k\}$ of all such elements forms an adapted basis for $\mathbb{C}G$ called the dual matrix coefficient basis.

More generally, if G acts transitively on X, and $x \in X$, then $\{g \cdot x\}_{g \in G}$ is clearly a spanning set for $\mathbb{C}X$, but

$$\{b_{ij}^k \cdot x\}$$

is spanning set for $\mathbb{C}X$ that contains an adapted basis as a subset. This is how we will create adapted bases, but note that the choice of x matters.

Frequency Subspaces

The elements of the form b^k_{ii} are primitive idempotents for $\mathbb{C}G$, and the subspace

$$b^k_{ii} \cdot \mathbb{C}X$$

is the associated frequency space of $\mathbb{C}X$. Note that

$$(b^k_{11} + \cdots + b^k_{d_k d_k}) \cdot \mathbb{C}X$$

is the isotypic subspace corresponding to the representation D_k.

Key Idea: If D is adapted to the chain $H \leq G$, then the frequency spaces of $\mathbb{C}X$ with respect to $\mathbb{C}H$ are direct sums of the frequency spaces with respect to $\mathbb{C}G$.

Example

If $G = X = \mathbb{Z}/4\mathbb{Z} = \{1, z, z^2, z^3\}$ and $H = 2\mathbb{Z}/4\mathbb{Z}$, then

$$\begin{aligned}\mathbb{C}X &= \langle \begin{bmatrix}1\\0\\1\\0\end{bmatrix} \rangle \oplus \langle \begin{bmatrix}0\\1\\0\\1\end{bmatrix} \rangle \oplus \langle \begin{bmatrix}1\\0\\-1\\0\end{bmatrix} \rangle \oplus \langle \begin{bmatrix}0\\1\\0\\-1\end{bmatrix} \rangle \\ &= \langle \begin{bmatrix}1\\1\\1\\1\end{bmatrix} \rangle \oplus \langle \begin{bmatrix}1\\-1\\1\\-1\end{bmatrix} \rangle \oplus \langle \begin{bmatrix}1\\i\\-1\\-i\end{bmatrix} \rangle \oplus \langle \begin{bmatrix}1\\-i\\-1\\i\end{bmatrix} \rangle.\end{aligned}$$

Fast Transforms

Change-of-Basis

Suppose D is adapted to the chain $H \leq G$. Although G acts on X transitively, the restriction to H need not be transitive, so X might be a union

$$X = X_1 \cup \cdots \cup X_t$$

of orbits of X with respect to H. Suppose you have a basis $\mathcal{B}'$ for $\mathbb{C}X$ that respects this partitition and is adapted to the associated irreducible representations of H.

Questions: How difficult is it to do a change-of-basis from $\mathcal{B}'$ to an adapted basis $\mathcal{B}$ with respect to the action of G? Can we bound the number of nonzero entries in the associated change-of-basis matrix?

Bounds Based on Frequency Subspaces

If the frequency spaces with respect to H have dimensions $\alpha_1, \ldots, \alpha_f$, then we will have no more than

$$\alpha_1^2 + \cdots + \alpha_f^2$$

nonzero entries in the associated change-of-basis matrix.

This can be used to show that if $X = G$ and the X_i are the right cosets of H, then the above sum becomes

$$\sum_{i=1}^{h} \left([G:H]d_i\right)^2 d_i$$

where the sum is over all of the irreducible degrees $d_1, \ldots, d_h$.

Suppose $H \leq G$, and that $d_1, \ldots, d_k$ are the dimensions of the irreducible representations of H, and let $d^3(H) = \sum d_i^3$. Suppose we have

$$\{1\} = G_0 < G_1 < \cdots < G_n = G$$

and set $q_j = [G_j : G_{j-1}]$. Then the step from G_{j-1} to G_j requires no more than the following number of nonzeros:

$$[G_j : G_{j-1}]^2 d^3(G_{j-1})[G : G_j] = q_j^2 q_{j+1} \cdots q_n \; d^3(G_{j-1}).$$

Thus we need no more than

$$\sum_{j=1}^{n} q_j^2 q_{j+1} \cdots q_n \; d^3(G_{j-1})$$

nonzeros, which is also the bound given in Theorem 7.6 of Clausen and Baum's book *Fast Fourier Transforms*.

Two-Sided Attack

If we are dealing with the regular module $\mathbb{C}G$, we can take advantage of the fact that we can act on both the left and the right by subgroups K and H.

Theorem (Mackey's Theorem): As a $(\mathbb{C}K, \mathbb{C}H)$-bimodule,

$$\mathbb{C}G \cong \bigoplus_{g} \mathbb{C}K \otimes_{\mathbb{C}H_g} \mathbb{C}H$$

where the direct sum is taken over a complete set of double coset representatives.

Doubly Adapted Basis for the Symmetric Group

Let $\mathcal{B}_n$ denote the dual matrix coefficient basis for the symmetric group S_n with respect to the **orthogonal form** of the irreducible representations of S_n.

Theorem: If $h \leq k \leq n$, and g is the shortest element in the double coset $S_h g S_k$ in S_n, then the nonzero elements of

$$\{bgb' \mid b \in \mathcal{B}_h \text{ and } b' \in \mathcal{B}_k\}$$

form an orthogonal adapted basis for $\mathbb{C}(S_h g S_k)$ for both the left action of $\mathbb{C}S_h$ and the right action of $\mathbb{C}S_k$.

Fast Fourier Transform

Using such bases with respect to the chain

$$(S_1, S_1) \leq (S_2, S_1) \leq (S_2, S_2) \leq \cdots \leq (S_{n-1}, S_{n-1}) \leq (S_n, S_{n-1})$$

has allowed us to create what we think is a new FFT for the family of symmetric groups.

For $n \leq 18$, the number of nonzero entries required in the matrix factorization of the DFT for S_n is less than $n^2 n!$, which makes our algorithm competitive with Maslen's FFT, which has complexity $O(n^2 n!)$.

The next step is to better understand the orbit/frequency relationship at play in this setup.

Questions

1. What are some groups and sets for which these frequency spaces lead to efficient decomposition algorithms?

2. How do the choices for the x in the spanning sets $\{b_{ij}^k \cdot x\}$ affect the resulting change-of-basis algorithms?

3. Where might we apply these insights and algorithms? (Statistics, Algebraic Statistics, Machine Learning, etc.)