MATH 6121 lecture notes

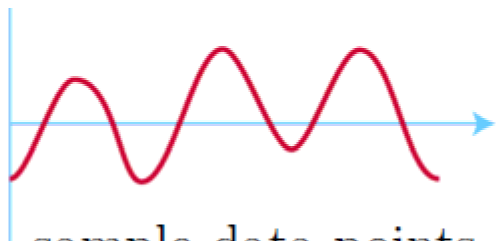
Transcribed and formatted by John M. Campbell
jmaxwellcampbell@gmail.com

10 October 11 lecture

10.1 Applications of representation theory

A nice idea for a presentation topic for MATH 6121 may involve applications of representation theory.

There is an Apple voice message application which involves a representation on the cyclic group.



sample data points

Given a sample $[d_1, d_2, \ldots, d_n]$ of data points, we construct a function from the multiplicative cyclic group

$$C_n = \{1, a, a^2, \ldots, a^{n-1}\}$$

to $\{d_1, d_2, \ldots, d_n\}$, and then decompose corresponding irreducible representations using the **discrete Fourier transform**.

The **continuous Fourier transform** is often used for solving partial differential equations. The discrete Fourier transform may be used for the decomposition of data into frequencies and amplitudes.

The above example applies to cyclic groups only. This provides a very easy example of breaking down a representation into 1-dimensional irreducible representations.

How are representation theory and the Fourier transform linked?

How does analyzing linear data generalize to more complicated groups?

10.2 Maschke's theorem

The symmetric group S_n acts on a set $\{x_1, x_2, \ldots, x_n\}$ of variables in a natural way, with $\sigma(x_i) = x_{\sigma(i)}$ for all indices i, and $\sigma \in S_n$.

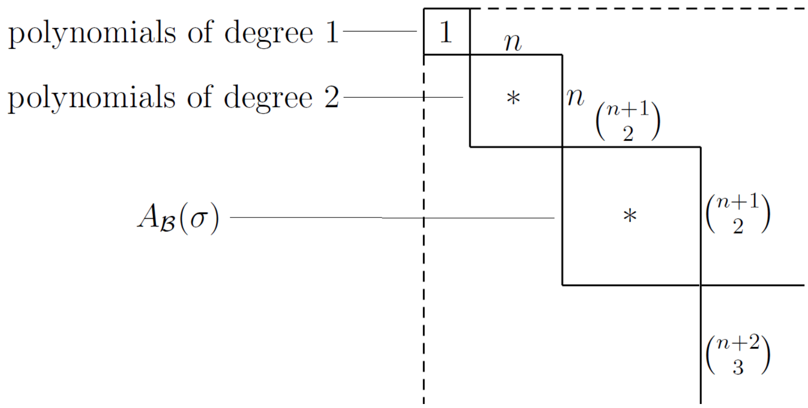
Let $V = \mathbb{C}[x_1, x_2, \ldots, x_n]$.

How does this decompose into irreducibles?

$$\sigma(x_{i_1} x_{i_2} \cdots x_{i_\ell}) = x_{\sigma(i_1)} x_{\sigma(i_2)} \cdots x_{\sigma(i_\ell)}.$$

Rewrite V as follows:

$$\begin{aligned} V &= \mathbb{C}[x_1, x_2, \ldots, x_n] \\ &= \bigoplus_{\ell \geq 0} (\text{polynomials of homogeneous degree } \ell \text{ in } x_1, x_2, \ldots, x_n.) \end{aligned}$$



In the case whereby $n = 2$, we have that $\{x_1^2, x_1x_2, x_2^2\}$ is a basis.

In the case whereby $n = 3$, we have that $\{x_1^2, x_1x_2, x_1x_3, x_2^2, x_2x_3, x_3^2\}$ is a basis.

If we choose $\{x_1^2, x_1x_2, x_1x_3, x_2^2, x_2x_3, x_3^2\}$ as a basis, there will be a corresponding permutation matrix in the third "block" along the main diagonal.

How do these block matrices break down into irreducible components? Each one of these sub-pieces somehow breaks down.

Let $C_3 = \{e, a, a^2\}$ denote the multiplicative cyclic group of order 3.

Define $V = \mathscr{L}_{\mathbb{C}}\{e_1, e_2, e_3\}$.

Let C_3 act on V as indicated below:

$$\begin{aligned} a(e_1) &= e_2 & \qquad a^2(e_1) &= e_3 \\ a(e_2) &= e_3 & a^2(e_2) &= e_1 \\ a(e_3) &= e_1 & a^2(e_3) &= e_2 \end{aligned}$$

Observe that $\mathscr{L}\{e_1, e_2, e_3\}$ is reducible, with:

$$\mathscr{L}\{e_1 + e_2 + e_3\} \subseteq \mathscr{L}\{e_1, e_2, e_3\}.$$

So, by Maschke's theorem, we thus find that V is decomposible, with:

$$V \cong \mathscr{L}\{e_1 + e_2 + e_3\} \oplus \mathscr{L}\{b_2, b_3\}.$$

If we let $\mathcal{B} = \{e_1, e_2, e_3\}$, then we find that $A_{\mathcal{B}}(a)$ is equal to the following binary matrix:

$$A_{\mathcal{B}}(a) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Now, letting $\mathcal{B}' = \{e_1 + e_2 + e_3, e_2, e_3\}$, we have that:

$$A_{\mathcal{B}}(a) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}.$$

Think of the subspace acting on itself, sending basis elements to basis elements.

In order to find a basis $\{b_2, b_3\}$ as above for $\mathscr{L}\{b_2, b_3\}$, we begin by constructing an inner product on V, following the proof of Maschke's theorem.

More specifically, we are interested in a Hermitian scalar product:

$$\langle \vec{v}, \vec{u} \rangle = \overline{\langle \vec{u}, \vec{v} \rangle}.$$

Our inner product must be linear in the first term:

$$\langle c_1\vec{v} + c_2\vec{v}', \vec{u} \rangle = c_1\langle \vec{v}, \vec{u} \rangle + c_2\langle \vec{v}, \vec{u} \rangle.$$

This inner product must be linear in the second term as well, i.e., it must be bilinear.

It must be positive-definite: $\langle \vec{v}, \vec{v} \rangle \geq 0$, with $\langle \vec{v}, \vec{v} \rangle = 0 \implies \vec{v} = \vec{0}$.

Define the scalar product $\langle \cdot, \cdot \rangle$ as follows:

$$\langle \vec{e}_i, \vec{e}_j \rangle = \begin{cases} 1 & \text{if } i = j; \\ 0 & \text{otherwise.} \end{cases}$$

Following the proof of Maschke's theorem, we define the scalar product $[\cdot, \cdot]$ by "averaging over" the entire group C_3 as follows:

$$[x, y] = \frac{1}{3}\langle x, y \rangle + \frac{1}{3}\langle a(x), a(y) \rangle + \frac{1}{3}\langle a^2(x), a^2(y) \rangle.$$

Now, choose a basis $\mathcal{B}$ for V, so that:

$$\langle \vec{v}, \vec{u} \rangle := \overline{[\vec{v}]_{\mathcal{B}}}\, [\vec{u}]_{\mathcal{B}}\, .$$

It suffices to construct a basis for the orthogonal complement to $\mathscr{L}\{\vec{e}_1 + \vec{e}_2 + \vec{e}_3\}$.

We are interested in constructing a vector space of the form $W \oplus W^{\perp}$. Apply the Gram-Schmidt algorithm in the following manner:

$$\begin{aligned} b_2 &= e_2 - \frac{[e_2, e_1 + e_2 + e_3]}{[e_1 + e_2 + e_3, e_1 + e_2 + e_3]}(e_1 + e_2 + e_3), \\ b_3 &= e_3 - \frac{[e_3, e_1 + e_2 + e_3]}{[e_1 + e_2 + e_3, e_1 + e_2 + e_3]}(e_1 + e_2 + e_3). \end{aligned}$$

We thus have that:

$$\begin{aligned} [b_2, e_1 + e_2 + e_3] &= 0, \\ [b_3, e_1 + e_2 + e_3] &= 0. \end{aligned}$$

Exercise 10.1. Check that the above equalities hold.

Therefore,

$$b_2 = -\frac{1}{3}e_1 + \frac{2}{3}e_2 - \frac{1}{3}e_3 = e_2 - \frac{1}{3}(e_1 + e_2 + e_3)$$
$$b_3 = -\frac{1}{3}e_1 - \frac{1}{3}e_2 + \frac{2}{3}e_3 = e_3 - \frac{1}{3}(e_1 + e_2 + e_3).$$

Exercise 10.2. Check that the above equalities hold.

Therefore,

$$a(b_2) = -\frac{1}{3}e_2 + \frac{2}{3}e_3 - \frac{1}{3}e_1 = b_3,$$
$$a(b_3) = -\frac{1}{3}e_2 - \frac{1}{3}e_3 + \frac{2}{3}e_1 = -b_2 - b_3.$$

Question 10.3. Why is it that when we act on $W^\perp$ by our group do we get something which is still in this space? Use Maschke's theorem to answer this question.

Why is this scalar product invariant under the group? When we act on $W^\perp$, we obtain an element of $W^\perp$. Use the proof of Maschke's theorem to explain this. Intuitively, it's because we averaged over the whole group.

If $g \in G$, then:

$$[\phi(g)(x), \phi(g)(y)] = [x, y],$$
$$[\phi(g)(x), y] = [x, \phi(g^{-1})(y)].$$

Recall that by definition of the scalar product $[\cdot,\cdot]$, we have that:

$$[x, y] = \frac{1}{|G|}\sum_{g\in G}\langle \phi(g)(x), \phi(g)(y)\rangle.$$

We thus have that $[\phi(h)(x), \phi(h)(y)] = [x, y]$.

If $x \in W$ and $y \in W^\perp$, then:

$$\begin{aligned} \phi(g)(x) \in W \implies [\phi(g)(x), y] &= 0 \\ &= [x, \phi(g^{-1})(y)]. \end{aligned}$$

So $[x, \phi(g^{-1})(y)] = 0$.

So, $\phi(g^{-1})(y) \in W^\perp$.

Therefore, $W^\perp$ is a submodule too, and:

$$V \cong W \oplus W^\perp.$$

Given bases of $\mathcal{B}_1$ and $\mathcal{B}_2$ of W and $W^\perp$, we obtain a basis $\mathcal{B}$ of V.

So, we find that any vector $\vec{v}$ can be written as $\vec{u} \in W$ plus $\vec{u}' \in W^\perp$.

More precisely, since $V \cong W \oplus W^{\perp}$, we find that $\vec{v}$ can be written in the form suggested below:

$$\vec{v} = \vec{u} + \vec{u}' \cong \left(\vec{u}, \vec{0}\right) + \left(\vec{0}, \vec{u}'\right) = (\vec{u}, \vec{u}')\,.$$

There are implicit isomorphisms involved above. Observe that $A_{\mathcal{B}}(a)$ has a block matrix structure:

$$A_{\mathcal{B}}(a) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}$$

In general, what can be said about matrices of the form $A_{\mathcal{B}'}(g)$? Are matrices of this form block-triangular? Will it not be block-triangular for all g at the same time?

What can be said about block-diagonal matrices of the form $A_{\mathcal{B}''}(g)$?

first part is a submodule

$$A_{\mathcal{B}''}(g) = \left[\begin{array}{c|c} * & 0 \\ \hline 0 & * \end{array}\right]$$ — second part is an orthogonal submodule

10.3 Schur's lemma

Schur's lemma is a very important result in representation theory. This result was formulated in the following manner in class:

Schur's Lemma (1905): If $\theta\colon M \to N$ is a G-morphism, and M and N are irreducible modules, then $\theta \equiv 0$ or θ is an isomorphism.

What is a G-morphism? We begin by reviewing some terminology related to the above lemma.

Definition 10.4. Let G be a group. A **left G-module** consists of an abelian group M together with a left group action $\bullet\colon G \times M \to M$ such that

$$g \bullet (a + b) = g \bullet a + g \bullet b$$

for all $g \in G$ and all $a, b \in M$.

Definition 10.5. A function $f\colon M \to N$ is a **morphism of G-modules** or a **G-linear map** or a **G-homomorphism** or a **G-morphism** if f is a group homomorphism and

$$f(g \bullet m) = g \bullet f(m)$$

for all $g \in G$ and all $m \in M$.

Let $M = \mathscr{L}\{\vec{e}_1, \vec{e}_2\}$, and let $G = \{e, a\} \cong C_2$. Let $a(\vec{e}_1) = \vec{e}_1$, and let $a(\vec{e}_2) = \vec{e}_2$. We thus obtain the trivial action on M.

But in this case, M is not irreducible: we can break it down into irreducible submodules.

Now, define T so that $T(\vec{e}_1) = \vec{e}_2$ and $T(\vec{e}_2) = \vec{e}_1$. The mapping T commutes with the above action. But this is an isomorphism.

The mapping T given by the matrix

$$_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

is not an isomorphism. But it is a linear transformation from M to M that commutes with the action of G. Since

$$A_{\mathcal{B}}(a) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2,$$

we find that:

$$_{\mathcal{B}}[T]_{\mathcal{B}} A_{\mathcal{B}}(a) = A_{\mathcal{B}}(a) {_{\mathcal{B}}[T]_{\mathcal{B}}}.$$

Lemma 10.6. *If $\theta\colon M \to N$ is a G-morphism, then* $\ker(\theta) \leq M$ *and* $\operatorname{im}(\theta) \leq N$.

Exercise 10.7. Prove the above lemma.

We thus proceed to prove Schur's lemma in the following manner. If $\theta\colon M \to N$ is a morphism such that M and N are both irreducible, then $\ker(\theta) = \{0\}$ or $\ker(\theta) = M$, by the above lemma. Therefore, θ is either injective, or is constantly 0. Now, observe that $\operatorname{im}(\theta) = \{0\}$ or $\operatorname{im}(\theta) = N$. We thus find that θ is either constantly equal to 0, or θ is surjective. So, if θ is nontrivial, we thus find that the morphism θ must necessarily be injective and surjective.

Question 10.8. What are some expamples of enumerative results which may be derived from Schur's lemma?

There is a second part of Schur's lemma:

Lemma 10.9. *If $M = N$, and $\theta\colon M \to N$ is an isomorphism, then $\theta = \lambda \cdot \mathrm{id}$ for some scalar λ, where M and N are irreducible.*

For example, consider the module $\mathscr{L}\{\vec{e}_1 + \vec{e}_2 + \vec{e}_3\}$. If

$$_{\mathcal{B}}[T]_{\mathcal{B}} A_{\mathcal{B}}(g) = A_{\mathcal{B}}(g) {_{\mathcal{B}}[T]_{\mathcal{B}}}$$

then it can be shown that the matrix $_{\mathcal{B}}[T]_{\mathcal{B}}$ must be of the form

$$_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{bmatrix} \lambda \end{bmatrix},$$

so that the mapping T is such that $T(\vec{b}_i) = \lambda \vec{b}_i$, with $T(\vec{e}_1 + \vec{e}_2 + \vec{e}_3) = b\vec{e}_1 + b\vec{e}_2 + b\vec{e}_3$.

This illustrates that the only mappings that are going to commute with the action given by a G-morphism are multiples of the identity. Multiplying by λ certainly yields an isomorphism which commutes with the group action. But what's not clear is that this is the only way this could happen.

To prove this, let λ be an eigenvalue of T with eigenvector $\vec{v}$, so that:

$$T(\vec{v}) = \lambda \vec{v}.$$

Equivalently,

$$(T - \lambda \cdot \mathrm{id})(\vec{v}) = \vec{0}.$$

But the mapping $T - \lambda \cdot \mathrm{id}$ is a G-morphism, and it is not an isomorphism, which shows that $T - \lambda \mathrm{id} \equiv 0$.

Given an element T in $\mathrm{Aut}(V)$ such that T is a G-morphism, then:

$$_{\mathcal{B}}[T]_{\mathcal{B}} A_{\mathcal{B}}(g) = A_{\mathcal{B}}(g) {}_{\mathcal{B}}[T]_{\mathcal{B}}.$$

That is, $T\phi(g) = \phi(g)T$. The matrix $_{\mathcal{B}}[T]_{\mathcal{B}}$ is in $\mathrm{Mat}_{n\times n}(\mathbb{C})$ and T is in $\mathrm{Aut}(V)$.

$\mathrm{Com}(A)$ denotes the set of everything that commutes with A.

Recall that an **endomorphism** is a morphism from a structure to itself. An endomorphism may or may not be injective, and an endomorphism may or may not be surjective.

$$\begin{aligned}
&\mathrm{End}_G(V) = \{\theta \colon V \to V \mid \theta \text{ is a } G\text{-morphism}\} \\
&\mathrm{Com}(A) = \{T \in \mathrm{Mat}_{n\times n}(\mathbb{C}) \ : \ TA_{\mathcal{B}}(g) = A_{\mathcal{B}}(g)T\} \\
&Z(\mathrm{Mat}_{n\times n}(\mathbb{C})) = \mathbb{C}^* \mathrm{Id}_{n\times n}.
\end{aligned}$$

Let E_{ij} denote the binary matrix such that the ij-entry of E_{ij} is equal to 1, and each other entry of E_{ij} is equal to 0.

If T commutes with every matrix B, then $TE_{ij} = E_{ij}T$ for all indices i and j. Observe that:

$$E_{ij} \cdot E_{k\ell} = \begin{cases} 0 & \text{if } k \neq j \\ E_{i\ell} & \text{if } j = k \end{cases}.$$

Exercise 10.10. Prove that $Z(\mathrm{Mat}_{n\times n}(\mathbb{C})) = \mathbb{C}^* \mathrm{Id}_{n\times n}$ using matrices of the form E_{ij}.