Notes for 190

Spring, 2002

Contents

1 Normal operators

There is a certain view point, connecting the study of linear transformations with matrices, which runs through the notes below. It is not always made explicit, since it would be maddeningly distracting to stop and formulate a precise lemma correlating the two every time this issue reappears; instead the reader is supposed to understand the link and be able to go back and forward whenever necessary.

Every finite dimensional complex vector space is "isomorphic" (which means "looks just like") some $\mathbb{C}^n$. Every linear transformation on $\mathbb{C}^n$ is given by matrix multiplication. In fact, $T : \mathbb{C}^n \to \mathbb{C}^n$ is *linear*[1] if and only if there is some matrix $A_T \in M_n(\mathbb{C})$ with

$$T(v) = vA_T$$

all $v \in \mathbb{C}^n$ (I follow Curtis in multiplying on the right). The operator T is *unitary* if and only if A_T moves *some*[2] orthonormal basis to some other orthonormal basis. T is *invertible* as a linear transformation if and only if A_T is invertible as a linear transformation. Given any basis, $v_1, ..., , v_n$, of $\mathbb{C}^n$, we may find a linear transformation which moves the standard basis, $e_1, e_2, ..., e_n$ to that basis; the transformation is given by the matrix with j^{th} row equal to v_j. Indeed a matrix is invertible if and only if its rows form a basis, if and only if the corresponding linear transformation is invertible.

And so it goes on. We may even end up deciding that there is no real difference between linear transformations and matrices. This is not quite true of course, since linear transformations are these ethereal objects which *transform* vectors, while matrices are literally speaking inert rows of complex numbers, but despite between not quite true the slogan "linear transformations are matrices" does not lead us to far astray.

Definition For $M \in M_n(\mathbb{C})$, we define M^*, the *adjoint of* M, to be the complex conjugate of the transpose of M:

$$M^* = \overline{(M^T)}.$$

Thus if M equals

$$\begin{pmatrix} 1 & 3i & 5 \\ 1-i & 10i & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

then M^* equals

$$\begin{pmatrix} 1 & 1+i & 0 \\ -3i & -10i & 0 \\ 5 & 2 & 1 \end{pmatrix}.$$

At the level of linear transformations one can actually prove a little lemma showing that for any linear $S : \mathbb{C}^n \to \mathbb{C}^n$ there will exist a unique linear $S^* : \mathbb{C}^n \to \mathbb{C}^n$ with

$$\langle S(v), w\rangle = \langle v, S^*(w)\rangle$$

all $v, w \in \mathbb{C}^n$. In fact, if we take the matrix associated to S and take its adjoint and look at the associated linear transformation then one can argue that it does indeed satisfy this above mentioned equality.

That gives us one way of proving that any linear transformation has an *adjoint*. The uniqueness of the adjoint follows from the following easy fact, which I may as well mention now in case it comes up later.

Fact 1.1 *$v, w \in \mathbb{C}^n$ are equal if and only if for all $u \in \mathbb{C}^n$ we have*

$$\langle u, v\rangle = \langle u, w\rangle.$$

[1] I.e. $T(\alpha v + \beta u) = \alpha T(v) + \beta T(u)$ all $\alpha, \beta \in \mathbb{C}$, $v, u \in \mathbb{C}^n$

[2] equivalently, *all*, as it turns out

Here are some basic properties of the adjoint:

Lemma 1.2 *For $A, B \in M_n(\mathbb{C})$*
(i) $A^{**} = A$
(ii) $(AB)^* = A^* B^*$
(iii) $(A^*)^{-1}$ *(if it exists)* $= (A^{-1})^*$
(iv) for any $u, v \in \mathbb{C}^n$, $\langle uA, v\rangle = \langle u, vA^*\rangle$.

Proof (i) and (ii) are direct calculation, while (iii) follows from (ii). (iv) is again a calculation, but one can ease the work by noting that we just have to establish it for $u, v = e_i, e_j$. □

The adjoint operation has appeared already in Curtis under disguise. He in effect defines the unitary group to be the $A \in GL_n(\mathbb{C})$ with

$$A^* = A^{-1}$$

and goes on to show that this definition is equivalent to the one I gave above.

Definition A is *self-adjoint* or *Hermitian* if $A = A^*$. If A has real entries then it self-adjoint if and only if the i, j^{th} entry always equals the j, i^{th}, and we say that A is *symmetric*. A is *normal*[3] if it commutes with its adjoint:

$$AA^* = A^* A.$$

Lemma 1.3 *A is self-adjoint if and only if*

$$\langle vA, u\rangle = \langle v, uA\rangle$$

all $u, v \in \mathbb{C}^n$.

Lemma 1.4 *If A is self-adjoint then*

$$\langle vA, v\rangle \in \mathbb{R}$$

all $v \in \mathbb{C}^n$.

Lemma 1.5 *If A is normal and* $V(\lambda) = \{v \in \mathbb{C}^n : vA = \lambda v\}$, *then*
(i) $vA^* = \bar{\lambda} v$ *all* $v \in V(\lambda)$
(ii) $(V(\lambda))A^* \subset V(\lambda)$
(iii) $(V(\lambda)^\perp)A \subset V(\lambda)^\perp$

Proof For (i) we can calculate that

$$\langle v(A - \lambda I), v(A - \lambda I)\rangle = 0$$

implies

$$\langle v(A^* - \bar{\lambda} I), v(A^* - \bar{\lambda} I)\rangle = 0.$$

(i) gives (ii) and (ii) gives (iii). □

Question Consider the non-normal

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Which parts, if any, of the lemma fail for this matrix?

[3] Roll of drums. Blast of trumpet. This is the central concept for the next while.

Note that in the last lemma, the set $V(\lambda)$ is necessarily a subspace (closed under addition and scalar multiplication) of $\mathbb{C}^n$. If $\lambda \neq \lambda'$ then we necessarily have $V(\lambda) \cap V(\lambda') = \{\vec{0}\}$.

Lemma 1.6 *If A is normal and $V(\lambda)$ as above, $V^*(\bar{\lambda}) = \{v \in \mathbb{C}^n : vA^* = \bar{\lambda}v\}$, then*

(i) $V^*(\bar{\lambda}) = V(\lambda)$
(ii) $(V^*(\bar{\lambda}))A \subset V^*(\bar{\lambda})$
(iii) $(V^*(\bar{\lambda})^{\perp})A^* \subset V^*(\bar{\lambda})^{\perp}$

Theorem 1.7 *Let $A \in M_n(\mathbb{C})$ be normal. Then there is an orthonormal basis of eigenvectors of A.*

Proof By crunching the characteristic polynomial of A, we get some λ_0 with $A - \lambda_0 I$ having zero determinant, hence non-invertible, and hence sending some non-zero vector to the zero vector; in other words, $V(\lambda_0)$ is non-trivial. If $V(\lambda_0) = \mathbb{C}^n$ then we simply take any orthonormal basis of $\mathbb{C}^n$ and finish early.

Otherwise we appeal to earlier lemma to see that $V(\lambda_0)^{\perp}$ is an invariant subspace, and so we can find some other λ_1 which is an eigenvector for the linear transformation associated to A restricted to the subspace $V(\lambda_0)^{\perp}$. Continuing on in this way, and appealing to the fact that $\mathbb{C}^n$ is finite dimensional, we obtain a sequence of eigenspaces, which jointly span $\mathbb{C}^n$. □

Corollary 1.8 *If $A \in M_n(\mathbb{C})$ is normal, then we may find a diagonal matrix D and a unitary U with*

$$A = U^{-1}DU.$$

Actually, one can see that the conclusion of this corollary characterizes normality, and hence since not all matrices are normal we cannot hope to obtain such a strong diagonalizability result in general. About the best we can do for arbitrary matrices is:

Theorem 1.9 Jordan Canonical form *Given an arbitrary linear transformation $T : \mathbb{C}^n \to \mathbb{C}^n$ we can find a sequence of invariant subspaces, $V_1, ..., V_m$, disjoint except for sharing $\vec{0}$, such that on each V_j we have that with respect to some appropriate basis T is represented by the matrix*

$$\begin{pmatrix} \lambda_j & 1 & 0 & 0...0 \\ 0 & \lambda_j & 1 & 0...0 \\ 0 & 0 & \lambda_j & 1...0 \\ .... & & & \\ 0 & 0 & 0 & 0...\lambda_j \end{pmatrix}.$$

We can say something for real matrices and their diagonlizability. In the next theorem we view $\mathbb{R}^n$ as vector space over $\mathbb{R}$ in the usual way.

Theorem 1.10 *If $A \in M_n(\mathbb{R})$ is symmetric, then all the eigenvalues are real and there is an orthonormal basis consisting of eigenvectors.*

Proof Since $A = A^*$ then we have all the eigenvectors are real by (i) 1.5. That much granted, the proof is much the same as before. □

There is a nice corollary of 1.7 which appeared as an exam problem on one of the qualifying exams for the graduate students. (As an aside on departmental gossip I will mention that many of the faculty members where outraged that most of the graduate students were unable to answer this "easy undergraduate linear algebra" problem.)

Exercise Let $H \subset GL_n(\mathbb{C})$ be an abelian group[4] of normal matrices. Show that the elements of H are *simultaneously diagonalizable* – that is to say, there some U with

$$U^{-1}AU$$

diagonal for *all* $A \in H$.

An interesting application of these ideas comes about with *rigid body motion.*

Definition A matrix $A \in O_3(\mathbb{R})$ (the orthogonal[5] 3×3 matrices) is said to be a *rigid body motion* if there is a continuous function

$$t \mapsto A_t$$

$$[0,1] \to O_3(\mathbb{R})$$

with $A_0 = I$ and $A_1 = A$; that is to say, there is a continuous path connecting the to identity to A.

In one of the class presentations the next theorem was credited to Euler.

Theorem 1.11 *Let A be a rigid body motion. Then A is a rotation.*

Proof Since A preserves lengths and angles, it preserves volumes, and hence has determinant one or minus one.[6] The continuous function

$$t \mapsto A_t$$

leads by composition to the continuous function

$$t \mapsto \mathrm{Det}(A_t)$$

$$[0,1] \to \{-1,1\},$$

since each A_t also preserves area. Since $\mathrm{Det}(A_0)$=1, we have by continuity that in fact $\mathrm{Det}(A_t) = 1$ throughout, and in particular A has determinant 1.

Then consider the characteristic polynomial

$$P_A(s) = \mathrm{Det}(sI - A).$$

At $s = 0$ we have $P_A(0) = -1$, whilst as $s \to +\infty$ we have $P_A(s) \to +\infty$. Thus there must be some *real* eigenvalue λ in the interval $(0, +\infty)$; if we let v be a corresponding eigenvector, then it follows from

$$\|v\| = \|vA\|$$

that in fact $\lambda = 1$.

Since A preserves orthogonality, $\{\alpha v : \alpha \in \mathbb{R}\}^\perp$ is A-invariant. Thus we may find a basis of $\mathbb{R}^3$ under which A looks like

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}.$$

[4]That is to say, for all $A, B \in H$ we have $AB = BA$.

[5]The orthogonal matrices are the unitary matrices with real entries; in terms of linear transformations, these correspond to the linear isomorphisms of real euclidean space which preserve angles and distances.

[6]We use a few well known facts about matrices, such as the determinant of a matrx equals the ratio it effects in changing volumes and areas. We could take this as a definition of determinant, though equally one can show it equivalent to one of the usual definitions by first calculating the equivalence for the elementary matrices and then extending to the rest by multiplication.

In fact, we can assume from now on that this is the form of A.

Now consider the behavior of A on the canonical basis $\{e_1, e_2, e_3\}$.

$$e_1 A = e_1.$$

$$e_2 A = (0, a, b).$$

$$e_3 A = (0, c, d).$$

These must again form an orthonormal basis.

From this it rapidly follows that (c, d) is a linear combination of $(a, -b)$ and $a^2 + b^2 = c^2 + d^2 = 1$, and from this we obtain that (c, d) equals either $(a, -b)$ or $(-a, b)$. The determinant being 1 rules out the later possibility and we are left with

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & -b & a \end{pmatrix}.$$

Since $a^2 + b^2 = 1$, we may choose θ with $\cos(\theta) = a$, $\sin(\theta) = b$. Thus with respect to our chosen basis,

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{pmatrix}.$$

In terms of our original basis and eigenvector, A is a rotation of angle θ around the axis v. □

2 Polar decomposition

Lemma 2.1 *For $A \in M_n(\mathbb{C})$, the following are equivalent:*
(i) $A = B^2$ for some Hermitian A;
(ii) A is Hermitian and $\langle vA, v\rangle \in \mathbb{R}^{\geq 0}$ all $v \in V$.

Proof Assuming (i) we obtain

$$A^* = (B^2)^* = B^*B^* = BB = A$$

and thus for any v

$$\langle vA, v\rangle = \langle vB^2, v\rangle = \langle vB, vB^*\rangle = \langle vB, vB\rangle \geq 0$$

since B Hermitian. Conversely assuming (ii) we can diagonalize A and then note that in diagonal form it must have positive reals down the main diagonal and zeros elsewhere. □

Lemma 2.2 *If $A \in M_n(\mathbb{C})$ then we can find V a subspace of $\mathbb{C}^n$ and U unitary such that*
(i) $v \mapsto vUA$ gives a one-to-one and onto linear transformation;
(ii) there is a basis $v_1, v_2, ..., v_m, v_{m+1}, ...v_n$ of V such that
(a) $v_1, ..., v_m$ is a basis of V; and
(b) $v_jUA = \mathbf{0}$ all $j \geq m$.

In (ii) we have essentially said that $v_{m+1}, ..., v_n$ is a basis for the null space of UA.

Theorem 2.3 *For any $A \in M_n(\mathbb{C})$ we may find Hermitian C and unitary U with*

$$A = UC^2.$$

Proof Appealing to 2.2 we may assume that A is invertible. We then note that A^*A is Hermitian and that for any $v \in \mathbb{C}^n$ we have

$$\langle vA^*A, v\rangle = \langle vA^*, vA^*\rangle \geq 0;$$

thus by 2.1 we can find some Hermitian D with

$$A^*A = D^4.$$

Thus it suffices to show that $(A^*)^{-1}D^2$ is unitary. For this we behold:

$$(A^*)^{-1}D^2((A^*)^{-1}D^2)^* = (A^*)^{-1}D^2(D^2)^*((A^*)^{-1})^*$$

$$= (A^*)^{-1}D^2D^2(A^{-1})^{**} = (A^*)^{-1}D^4A^{-1} = (A^*)^{-1}A^*AA^{-1} = I.$$

□

Definition We say that B is *positive* if there is some Hermitian C with $C^2 = B$.

Thus over $M_1(\mathbb{C}) \sim \mathbb{C}$ the positive elements are those in $\mathbb{R}^{\geq 0}$. Then in this simple case the polar decomposition says that any $z \in \mathbb{C}$ can be written in the form

$$ru$$

where $r \geq 0, |u| = 1$, which in turn gives us that $u = e^{i\theta}$ some $\theta \in \mathbb{R}$, and thus

$$z = re^{i\theta} = r\cos\theta + ir\sin\theta,$$

which may be familiar as the polar representation of the complex number z.

3 Group representations

Definition A *representation* of a group G is a homomorphism

$$\rho : G \to GL_n(\mathbb{C}).$$

It is *unitary* if its range is included in the unitary matrices.

In other words, a *unitary representation* of the group G is a homomorphism[7] from G to $U_n(\mathbb{C})$. Mostly in this section and next I will work with unitary representations, which allow us to make use of the diagonizability. Some of the results actually go through in wider generality, providing we are willing to avail ourselves of the stated but unproven Jordan canonical form. For finite groups one can show that every representation is *equivalent* to a unitary representation.

Example Let $G = \langle a | a^3 = 1 \rangle \cong \mathbb{Z}_3$. Define $\rho : G \to U_n(\mathbb{C})$ by

$$\rho(a^m) = \begin{pmatrix} (-\frac{1}{2} - \frac{i\sqrt{3}}{2})^m & 0 \\ 0 & (-\frac{1}{2} + \frac{i\sqrt{3}}{2})^m \end{pmatrix}.$$

Exercise Given $|G| = n$, with an enumeration $\{g_j\}_{j \leq n}$ of G, show that we can define a unitary representation

$$\rho : G \to U_n(\mathbb{C})$$

by letting $\rho(g)$ be the matrix which has mostly zeros, except 1 at the i, j^{th} spot when $g \cdot g_j = g_i$.

Definition Two representations $\rho, \theta : G \to GL_n(\mathbb{C})$ are *equivalent* if there is some $A \in GL_n(\mathbb{C})$ with

$$\rho(g) = A^{-1}\theta(g)A$$

all $g \in G$. A not necessarily invertible $m \times n$ matrix A is said to *intertwine* $\rho : G \to GL_n(\mathbb{C})$ and $\theta : G \to GL_m(\mathbb{C})$ if at every g we have

$$A\rho(g) = \theta(g)A.$$

Definition Given a representation

$$\rho : G \to GL_n(\mathbb{C}),$$

we say that a subspace $V \subset \mathbb{C}^n$ is *ρ-invariant* if at every $g \in G$ we have

$$V\rho(g) \subset V.$$

We say that ρ is *irreducible* if the only invariant subspaces are $\mathbb{C}^n$ and $\{\mathbf{0}\}$.

Lemma 3.1 *Let $\rho : G \to U_n(\mathbb{C})$ be a unitary representation and $V \subset \mathbb{C}^n$ a ρ-invariant subspace. Then we may find a ρ-invariant W with*

$$\mathbb{C}^n = V \bigoplus W;$$

that is to say, $V \cap W = \{\mathbf{0}\}$ and every $u \in \mathbb{C}^n$ can be written in the form $u = v + w$ for some $v \in V, w \in W$.

[7] There is an annoying detail, bequeathed to us by Curtis' voraciously disturbing notation. If the assignment $g \mapsto \rho(g)$ is a homomorphism to $GL_n(\mathbb{C})$, then that won't necessarily guarantee that the assignment to g of the *linear transformation associated* to $\rho(g)$ gives us a homomorphism of G to the group of linear transformations on $\mathbb{C}^n$. It should, but it does not. If for each $A \in GL_n(\mathbb{C})$ we let φ_A be the associated linear transformation, $v \mapsto vA$, then alas $\varphi_A \circ \varphi_B = \varphi_{B \circ A}$. In actually fact, I am going to simply try to ignore this annoying detail, keep up the facade of identifying linear transformations with matrices, and make no distinction between a homomorphism from G to $GL_n(\mathbb{C})$ and a homomorphism from G to the group of linear transformation on $\mathbb{C}^n$. One can actually pass between them, with suitable care; given a representation $\rho : G \to GL_n(\mathbb{C})$ one defines a homomorphism into the group of linear transformations $\hat{\rho} : g \mapsto \varphi_{(\rho(g))^{-1}}$.
This was a pain. Enough said.

Proof Note that $V^\perp = \{w \in \mathbb{C}^n : \forall v \in V(\langle v, w\rangle = 0)\}$ is also ρ-invariant. $\square$

Thus by continually applying the above lemma and appealing to the finite dimensionality, we may eventually break any representation down into a finite sum of irreducible representations. Of course one can also go in the opposite direction: Given two representations ρ, θ, mapping G into the linear groups on V and W, one can form the space

$$V \bigoplus W = \{(v, w) : v \in V, w \in W\},$$

and associate to each g the linear transformation on $V \bigoplus W$ which moves $(v, w) \mapsto (S_{\rho(g)}(v), T_{\theta(g)}(w))$, where $S_{\rho(g)}$ is the transformation provided by ρ and $T_{\theta(g)}(w)$ is the transformation provided by θ.

Theorem 3.2 *Let $\rho : G \to GL_n(\mathbb{C})$ be irreducible and let A intertwine ρ with itself. Then there is some $\lambda \in \mathbb{C}$ with $A = \lambda I$.*

Proof Let λ be an eigenvalue for A and $V(\lambda)$ be the corresponding eigenspace. Then for any $g \in G$, $v \in V(\lambda)$

$$v\rho(g)A = vA\rho(g) = (\lambda v)\rho(g) = \lambda(v(\rho(g))$$

and hence $v\rho(g)$ is still in $V(\lambda)$. Thus $V(\lambda)$ is ρ-invariant, contains a non-zero vector, and hence by assumption on ρ must be equal to all of $\mathbb{C}^n$. $\square$

This theorem or the next easy proposition is sometimes called *Schur's lemma.*

Proposition 3.3 *Let ρ, θ be two irreducible representations on $\mathbb{C}^n, \mathbb{C}^m$ which are intertwined by A, and $n \times m$ matrix. Then either A is trivial, in the sense of consisting of nothing but zeros, or $m = n$ and A is invertible.*

Proof Consider the null space of A, $\{v \in \mathbb{C}^n : vA = \mathbf{0}\}$. This is ρ-invariant subspace of $\mathbb{C}^n$, and hence equal to either $\mathbb{C}^n$ or $\{\mathbf{0}\}$. In the latter case, we can then consider the range space of A, $\{vA : v \in \mathbb{C}^n\}$ and use the same argument to conclude that it must be all of $\mathbb{C}^m$. $\square$

Definition Let $\rho : G \to GL_n(\mathbb{C})$ be a representation. We say that an irreducible representation θ is included in ρ if we may write

$$\mathbb{C}^n = V \bigoplus W$$

where the representation ρ restricted to V (i.e. for each $g \in G$ just consider the linear transformation $\rho(g)$ restricted to V) is equivalent to θ.

Lemma 3.4 *Given a representation $\rho : G \to GL_n(\mathbb{C})$ and an irreducible representation θ, we may write*

$$\mathbb{C}^n = U \bigoplus W$$

where ρ restricted to U is equivalent to a finite direct sum of θ and θ is not included in ρ restricted to W.

Given an irreducibility representation ρ, we have that there are no non-trivial intertwiners. The situation with finite sums of irreducible representations is a bit more subtle; we can say something, but we need the language of vector spaces to do so. Note here that the collection of intertwiners between two representations is closed under addition as well as scalar multiplication, and in this sense forms a vector space in its own right.

Lemma 3.5 *Given a representation $\rho : G \to GL_n(\mathbb{C})$ and an irreducible representation θ, with*

$$\mathbb{C}^n = U \bigoplus W$$

where $\rho|_U$ is equivalent to a direct sum of k copies θ, and θ is not included in $\rho|_W$, then the dimension of the intertwining operators between θ and ρ equals k.

Definition Given a representation $\rho : G \to GL_n(\mathbb{C})$, $p : \mathbb{C}^n \to \mathbb{C}^n$ is said to be a *projection* for ρ if $p^2 = p$ and for all $g \in G$, $p\rho(g) = \rho(g)p$.

Definition For G a finite group, let $\mathbb{C}[G]$ consist of all functions $f : G \to \mathbb{C}$; note that this space is naturally isomorphic to $\mathbb{C}^{|G|}$. We let define the *regular left representation* of G to be the representation which associates to each g the linear transformation λ_g:

$$f \mapsto \lambda_g(f),$$

where $\lambda_g(f)$ is defined by

$$(\lambda_g(f))(h) = f(g^{-1}h)$$

for all $h \in G$; one routinely checks that this indeed gives us a homomorphism into the group of linear transformations on $\mathbb{C}[G]$.

Lemma 3.6 *For G finite, there is a non-trivial intertwining operator from $\mathbb{C}[G]$ to any non-trivial representation.*

Proof Let $g \mapsto S_g$ be the homomorphism from G to the group of linear transformations of some vector space V; we can assume that it is an irreducible representation; we fix some non-zero $v \in V$. We can then define

$$T : \mathbb{C}[G] \to V$$

by

$$f \mapsto \sum_{g \in G} f(g) S_g(v).$$

$\square$

Note then that if the space V has dimension n, then we essentially have dimension n choices for T: we could have chosen v to be any element of V, and conversely given the intertwiner T, T is completely determined by $T(\delta_e)$ (where δ_e is the function which assumes the value 1 at the identity of G and zero elsewhere).

So now suppose we apply 3.4 to write

$$\mathbb{C}[G] = \bigoplus_{i \leq \ell} (\bigoplus_{n_i} V_i)$$

where the $\lambda|_{V_i}$'s are mutually non-equivalent irreducible representations, appearing as suggested by the notation n_i many times inside the regular left representations. Then it follows from 3.5 that the intertwiners between λ and each $\lambda|_{V_i}$ is exactly n_i; which in turn by the observation of the last paragraph must equal the dimension of V_i. Finally note that the dimension of $\mathbb{C}[G]$ equals the size of G, which in turn must equal the sums of the dimensions of the various $\bigoplus_{n_i} V_i$.

Thus with simple homespun methods we have established the following startling theorem:

Theorem 3.7 *Suppose G is finite group and we enumerate the irreducible representations*

$$\rho_i : G \to GL_{n_i}(\mathbb{C}),$$

each ρ_i having dimension n_i. Then

$$|G| = \sum (n_i)^2.$$

4 Character and trace

Definition If $M \in M_n(\mathbb{C})$, we let $\tau(M)$, the *trace of* M, be the sum of the diagonal entries.

Example Thus the trace of

$$\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1+i & 0 \\ 4 & 0 & 6i \end{pmatrix}$$

is $2+7i$.

Lemma 4.1 *For $A, B \in M_n(\mathbb{C})$ we have*

$$\tau(AB) = \tau(BA).$$

Proof A brute calculation. □

Thus we have $\tau(A) = \tau(C^{-1}AC)$ for any invertible C.

Lemma 4.2 *If $A \in U_n(\mathbb{C})$ has all its eigenvalues with absolute value 1, then*

$$\tau(A^{-1}) = \overline{\tau(A)}.$$

Proof We may assume that A is a diagonal matrix, and then note that in this form it will have complex numbers of modulus 1 down the diagonal and zeros elsewhere. □

Definition $\chi : G \to \mathbb{C}$ is a *character* if there is a representation

$$\rho : G \to GL_n(\mathbb{C})$$

with $\chi = \tau \circ \rho$ – that is to say, $\chi(g)$ is the trace of $\rho(g)$. We say that a character is *irreducible* if it arises from an irreducible representation.

Remark In the case $n = 1$ any character will be a homomorphism; however if $\rho : \mathbb{Z} \to GL_n(\mathbb{C})$ is defined by

$$\rho(n) = \begin{pmatrix} e^{i2\sqrt{2}n\pi} & 0 \\ 0 & 1 \end{pmatrix}$$

then the induced trace is *not* a homomorphism.

Lemma 4.3 *Let G be a finite group, $\rho : G \to U_n(\mathbb{C})$ a unitary representation, $\chi : G \to \mathbb{C}$ be the induced character. Then for any $g \in G$*

(i) all the eigenvalues of $\rho(g)$ have absolute value 1;
(ii) $\chi(g^{-1}) = \overline{\chi(g)}$;
(iii) if $\chi(g) = \chi(1)$ then $\rho(g) = I$.

Proof Fix g; we may assume $\rho(g)$ is diagonal. Then (i) follows since g^m is the identity for some $m \in \mathbb{N}$. (ii) then follows from 4.2. For (iii) we can observe that a sequence of n complex numbers sum to n if and only if they all equal 1. □

Example Let $G = D_8$ = group of symmetries of the square. So $G = \langle a, b; a^4 = 1, b^2 = 1, bab^{-1} = a^{-1} \rangle$.
Etc.

g	1	a	a^2	b	ba
$\rho(g)$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$
$\chi(g)$	2	0	-2	0	0

5 Characters for abelian groups

The character theory for finite abelian groups is especially satisfying. Every irreducible representation will be one dimensional and every irreducible character will be a homomorphism from G to $\mathbb{C}$.

The irreducible characters will form an orthonormal basis for $\mathbb{C}[G]$. In an appropriate group structure, the characters will actually form a group themselves, called the dual group, $\hat{G}$, of G. The dual group of the dual group of G,

$$\hat{\hat{G}},$$

equals the original group G.

I should make a general remark about terminology. For abelian groups it can be shown that the *irreducible* characters are already sufficient to *separate points*: that is to say, for any $g \neq h$ in an abelian group G there is an irreducible character $\chi : G \to \mathbb{C}$ with $\chi(g) \neq \chi(h)$. For this reason people frequently restrict themselves entirely to irreducible characters of abelian groups and use the term *character* to only refer to *irreducible character*. For this section only I will adopt that terminology: Everywhere I say *character* you should read *irreducible character.*

Lemma 5.1 *If G is abelian and $\chi : G \to \mathbb{C}$ is a character, then χ is a homomorphism from G into the multiplicative group of the non-zero complex numbers.*

Proof It is easily seen that if T and S are commuting linear transformations ($TS = ST$) and $V_{\lambda,T}$ is the space of eigenspace for T and eigenvalue λ, then $S[V_{\lambda,T}] \subset V_{\lambda,T}$. Then by restricting ourselves to $S|_{V_{\lambda,T}}$ we may find a common eigenvector. Continuing in this fashion we can show that any collection of mutually commuting linear transformations has a common eigenvector, and hence an abelian group only has one dimensional irreducible representations.[8]

Thus, if $\chi : G \to \mathbb{C}$ is an (irreducible) character, then there is a homomorphism

$$\rho : G \to GL_1(\mathbb{C})$$

with $\chi(g)$ equal to the trace of $\rho(g)$ for all $g \in G$; under the natural identification of $GL_1(\mathbb{C})$ with the multiplicative group of $\mathbb{C} \setminus \{0\}$ one has that the trace of every complex number is just that very same complex number. □

Lemma 5.2 *Let G be a finite abelian group. Then for all characters $\chi_1, \chi_2 : G \to \mathbb{C} \setminus \{0\}$ we have*

$$\sum_{g \in G} \chi_1(g)\overline{\chi_2(g)} = 0.$$

[8]Here I am implicitly assuming that we start off considering finite dimensional vector spaces; it actually turns out that the same result is true if we contemplate the wider perspective of representations of abelian groups on separable and possibly infinite dimensional Hilbert spaces, but the proof is far more subtle. The proof becomes more subtle because a normal operator on an infinite dimensional Hilbert space may have *no* eigenvectors.

Proof Since $\chi_1 \neq \chi_2$ we may find some specific $h_0 \in G$ with $\chi_1(h_0) \neq \chi_2(h_0)$, which amounts to saying that

$$\chi_1(h_0)\frac{1}{\chi_2(h_0)} \neq 1;$$

since $\chi_2 : G \to \mathbb{C} \setminus \{0\}$ is a homomorphism this implies

$$\chi_1(h_0)\chi_2(h_0^{-1}) \neq 1$$

$$\therefore \chi_1(h_0)\overline{\chi_2(h_0)} \neq 1$$

by 4.3(ii).

Now notice that

$$\chi_1(h_0)\overline{\chi_2(h_0)} \sum_{g\in G} \chi_1(g)\overline{\chi_2(g)} = \sum_{g\in G} \chi_1(h_0)\chi_1(g)\overline{\chi_2(h_0)\chi_2(g)}$$

$$= \sum_{g\in G} \chi_1(h_0 g)\overline{\chi_2(h_0 g)},$$

since χ_1, χ_2 are homomorphisms,

$$\therefore = \sum_{g\in G} \chi_1(g)\overline{\chi_2(g)},$$

since $G = \{h_0 g : g \in G\}$. But the only way that we could have $\chi_1(h_0)\overline{\chi_2(h_0)} \neq 1$ and

$$\sum_{g\in G} \chi_1(g)\overline{\chi_2(g)} = \chi_1(h_0)\overline{\chi_2(h_0)} \sum_{g\in G} \chi_1(g)\overline{\chi_2(g)}$$

is if the quantity $\sum_{g\in G} \chi_1(g)\overline{\chi_2(g)}$ equals zero. □

Lemma 5.3 *Let G be a finite abelian group. Then for all characters χ*

$$\sum_{g\in G} \chi(g)\overline{\chi(g)} = |G| (\text{the size of } G).$$

Proof Referring back to 4.3 again we get that at every g

$$\chi(g)\overline{\chi(g)} = \chi(g)\chi(g^{-1})$$

$$= \chi(gg^{-1}) = \chi(0) = 1,$$

since χ is a homomorphism from the abelian group G to the multiplicative group of the non-zero complex numbers. □

The last two lemmas can be reformulated in the language of Hilbert spaces. Let G be a finite abelian group of size n. For any two

$$F_1, F_2 : G \to \mathbb{C}$$

let us define

$$\langle F_1, F_2 \rangle = \frac{1}{n} \sum_{g\in G} F_1(g)\overline{F_2(g)}.$$

This gives us a perfectly good inner product space on the vector space of complex valued functions on G. (Note: This space is in some natural way isomorphic to $\mathbb{C}^n$.) It is a finite dimensional inner product space.

With this terminology paved into the ground, we can reformulate the last two lemmas with the following proposition:

Proposition 5.4 *The characters of a finite abelian group form an orthonormal set.*

So far so good. But of course this proposition only screams out the question of whether they are an *orthonormal basis.* The answer to this is yes indeed, and that is the main result here.

Theorem 5.5 *The characters of a finite abelian group form an orthonormal basis for its space of complex valued functions.*

Proof I am going to cheat somewhat and appeal to the classification theorem for finite abelian groups; you probably won't have seen this unless you have taken math 110a, but perhaps you would be willing to believe it anyway. In actual fact we will give a general argument for arbitrary finite abelian groups in the next section, but this argument is somewhat involved and I want to get out an elementary argument for the abelian case first.

The classification theorem for abelian groups tells us that any finite abelian group can be written as a product of finitely many cyclic groups, each cyclic group necessarily isomorphic to some $\mathbb{Z}_k$ (the natural numbers $\{0, 1, ..., k-1\}$ with addition mod k – that is to say, we let $\ell_1 + \ell_2 = m$ if m is the remainder when we try to divide k into $\ell_1 + \ell_2$). In fact I am going to cheat a little bit more an assume that G is just equal to some such $\mathbb{Z}_n$; that is to say, as a single product. The general case requires some more details, but if you have taken a course in group theory then perhaps those details may be clear to you.

So $G = \mathbb{Z}_n$ and we are trying to show that the characters form an orthonormal basis for $G^{\mathbb{C}}$. We will do this by considering dimension.

$G^{\mathbb{C}}$ has dimension n. If we can show that there are n distinct characters then we will be happy.

Let us consider the n^{th} root of unity

$$\xi = e^{\frac{2\pi}{n}i}.$$

For $\ell = \{1, 2, ..., n\}$ we can define the character

$$\chi_\ell : G \to \mathbb{C}$$

by

$$\chi_\ell(m) = e^{\frac{2m\ell\pi}{n}i}$$

for $m \in \{0, 1, ..., n-1\}$. □

Corollary 5.6 *For G a finite abelian group and $g \neq 0$ in G, there is a character*

$$\chi : G \to \mathbb{C}$$

with $\chi(g) \neq 1$.

Proof Consider the subspace $H \subset G^{\mathbb{C}}$ consisting of all $F : G \to \mathbb{C}$ with $F(g) = F(1)$; it is easily checked that this is a subspace, and so if our orthonormal basis was included in this subspace we would have $H = G^{\mathbb{C}}$, which is obviously untrue. □

In some form the results above go through in a wider context. Given an infinite abelian group G one can still talk about irreducible characters and still hope that the irreducible characters are sufficient to separate points; this much still turns out to be true, but with the usual refrain that the proof is much more subtle.

6 Characters for non-abelian groups

Rather messy.

I want to first of all point out that at least one of the facts we observed for abelian groups passes through to the general case with no trouble at all.

Lemma 6.1 *Let G be a finite group and*

$$g \mapsto A_g$$

a (finite dimensional, unitary[9]) representation and

$$\chi : G \to \mathbb{C}$$

the corresponding character given by $\chi(g) = \text{Trace}(A_g)$. *Then* $\chi(g^{-1}) = \overline{\chi(g)}$.

Proof We observe that the matrices A_g and $A_{g^{-1}}$ commute, are normal, and hence are simultaneously diagonalizable. Without loss of generality we may assume that they are in fact already in diagonal form. (Since trace is conjugation invariant, calculating the character with respect to a different basis introduces no change). Then since the matrices are unitary, A_g consists of complex numbers of absolute value 1 down the diagonal and $A_{g^{-1}}$, as its inverse, must consist of the corresponding complex conjugates in the indicated order down the main diagonal. □

In what follows I wish to increasingly think in terms of a representation as a homomorphism from a group to a group of invertible linear transformations. Morally of course this is much the same thing as a homomorphism into a group of matrices, but in practice one rapidly becomes overwhelmed with the fiddly and fussy details involved in constantly setting up the precise correspondence.

For G a finite[10] group we let $\mathbb{C}[G]$ be the vector space of complex valued functions from G to $\mathbb{C}$; this is of course isomorphic to $\mathbb{C}^n$ where $n = |G|$, it has a basis over $\mathbb{C}$ the functions $\{\delta_g : g \in G\}$, each of which assume the value 1 at the relevant g and zero elsewhere, and it suggest two natural representations of G.

Notation For V a vector space, let $\mathcal{L}(V)$ be the algebra of linear transformation on V; note that we may add two linear transformation and multiply by a scalar, as well as compose, and for this reason we describe the group of objects as an algebra. We let $\mathcal{GL}(V)$ be the collection of invertible linear transformations on V; this is now a group under composition, but we lose the ability to add.

Definition We let $\mathcal{GL}(\mathbb{C}[G])$ be the group of invertible linear transformations on $\mathbb{C}[G]$. We define the representation ρ of G given by

$$\rho : G \to \mathcal{GL}(\mathbb{C}[G])$$

$$\rho : g \mapsto \rho_g$$

where for each $g \in G$ and $f : G \to \mathbb{C}$ we define $\rho_g(f)$ by the rule

$$(\rho_g(f))(h) = f(hg);$$

in other words, we multiply on the right. You can check without too much trouble that this defines a representation (the main issue being that $\rho_{g_1 g_2}(f) = \rho_{g_1}(\rho_{g_2}(f))$).

A similar definition works on the left, but now we actually need to be more careful. What works is

$$\lambda : g \mapsto \lambda_g$$

[9]We are in fact making the harmless assumption that all our representations are unitary.

[10]Here and onwards we assume G is finite

where we define each $\lambda_g(f)$ by the rule

$$(\lambda_g(f))(h) = f(g^{-1}h).$$

The choice of g^{-1} may seem strange, but without it we lose the critical identity $\lambda_{g_1 g_2}(f) = \lambda_{g_1}(\lambda_{g_2}(f))$.

These two representations interact nicely. In some sense they commute. This will be the key observation, from which everything else follows.

Definition We let $L(G)$ be the linear transformations in $\mathcal{GL}(\mathbb{C}[G])$ generated by the $\{\lambda_g : g \in G\}$; that is to say, $L(G)$ consists of all linear transformations of the form

$$f \mapsto \sum_{g \in G} c_g \lambda_g(f),$$

for some choice of constants $c_g \in \mathbb{C}$.

Similarly let $R(G)$ be the linear transformations generated by the $\{\rho_g : g \in G\}$.

Lemma 6.2 *Suppose $\theta \in \mathcal{GL}(\mathbb{C}[G])$ commutes with every ρ_g (and hence in fact with every element of $R[G]$). Then $\theta \in L[G]$.*

Proof Recall that the characteristic functions of the form δ_g provide a basis for the vector space. Note that at each $g, h \in G$ we have the slightly unintuitive equation $\rho_g \cdot \delta_h = \delta_{hg^{-1}}$. Consider the behavior of θ at δ_e, the characteristic function of the identity of G.

Say $\theta(\delta_e) = \sum_{g \in G} \alpha_g \delta_g$. Then at every $h \in G$ we have $\delta_h = \rho_{h^{-1}} \delta_e$, and hence

$$\theta(\delta_h) = \rho_{h^{-1}}(\theta(\delta_e)) = \sum_{g \in G} \alpha_g \rho_{h^{-1}} \delta_g = \sum_{g \in G} \alpha_g \delta_{gh},$$

which in turn, if you crunch out the definition, equals

$$\sum_{g \in G} \alpha_g \lambda_g(\delta_h).$$

□

Similarly:

Lemma 6.3 *If $\theta \in \mathcal{GL}(\mathbb{C}[G])$ commutes with every λ_g then $\theta \in R[G]$.*

Recall that we previously saw that we may write a general representation, such as $\rho : G \to \mathbb{C}[G]$, as a direct sum of irreducible representations. I want to go a bit further than that and collect together the irreducible representations of like similarity.

It might be helpful to establish some notation before formulating the next lemma.

Notation For

$$\pi : G \to \mathcal{GL}(V)$$
$$g \mapsto \pi_g$$

a representation, we let $I(\pi, \pi)$ be the collection intertwiners of π to itself: That is to say, those $\theta : V \to V$ such that

$$\theta \circ \pi_g = \pi_g \circ \theta$$

all $g \in G$. We then take the idea one further step and let $Z(\pi, \pi)$ be those $\theta \in I(\pi, \pi)$ which commute with every element in $I(\pi, \pi)$: That is to say, $\varphi : V \to V$ linear is in $Z(\pi, \pi)$ if and only if for every $\theta \in I(\pi, \pi)$ we have $\theta \circ \varphi = \varphi \circ \theta$.

Thus from 6.2 we have $L[G] = I(\rho, \rho)$; we are yet to come to any definite analysis of $Z(\rho, \rho)$, though it does follow from 6.3 that it must actually be included $R[G]$.

Lemma 6.4 $\sum_{g\in G} \alpha_g \rho_g$ *is in* $Z(\rho, \rho)$ *if and only if the function*

$$g \mapsto \alpha_g$$

is constant on conjugacy classes; that is to say, for all $g, h \in G$

$$\alpha_{hgh^{-1}} = \alpha_g.$$

Proof Any such element of $R[G]$ at once commutes with every element of $I(\rho, \rho) = L[G]$ by 6.3, and the trick is to determine when it satisfies further conditions placing it in $Z(\rho, \rho)$.

$\sum_{g\in G} \alpha_g \rho_g$ will be in $Z(\rho, \rho)$ if it commutes with the representation ρ, which is to say it commutes with every element $\{\rho_h : h \in G\}$. Which amounts to saying that at every g and h

$$\alpha_{hg} = \alpha_{gh},$$

which after replacing g by gh^{-1} indeed is as required. □

Lemma 6.5 *Let* $\pi : G \to \mathcal{G}L(V)$ *be a representation. Let*

$$V = \bigoplus_{i\leq k} (\bigoplus_{j\leq \ell} V_{i,j})$$

where each $V_{i,j}$ *is irreducible and* $V_{i,j}$ *is isomorphic to* $V_{i',j'}$ *if and only if* $i = i'$.

Then $Z(\pi, \pi)$ *is generated by the projections onto the subspace*

$$\bigoplus_{j\leq \ell} V_{i,j},$$

for various $i \leq k$*; in particular,* $Z(\pi, \pi)$ *has dimension* k.

Proof For $i \neq i'$ we have that there are *no* intertwiners from $V_{i,j}$ to $V_{i',j'}$, which means that the various projections described above do indeed commute with all the elements of $I(\pi, \pi)$.

On the other hand, an element of $Z(\pi, \pi)$ must behave identically on each $V_{i,j}, V_{i,j'}$ and it must commute with any projection onto a $V_{i,j}$; from this it follows that each $Z(\pi, \pi)$ will be a linear combination of the above described projections. □

Notation Let $CI(G)$ be the collection of conjugation of invariant functions from G to $\mathbb{C}$.

Note that any character (irreducible or not) must be in $CI(G)$: They are all conjugation invariant.

Corollary 6.6 *The dimension of* $CI(G)$ *equals the number of inequivalent irreducible representations of* G.

Proof We may appeal to 3.6 and decompose ρ as

$$\mathbb{C}[G] = \bigoplus_{i\leq k} (\bigoplus_{j\leq \ell} V_{i,j}),$$

as in the last lemma where the $V_{1,1}, V_{2,1}, ..., V_{k,1}$ enumerate *all* the irreducible representations of G. The last lemma tells us that the dimension of $Z(\rho, \rho) = k$; and from 6.4 we have that this in turn equals the dimension of $CI(G)$. □

We can now go ahead and as in the case of abelian groups define an inner product on the conjugacy invariant functions

$$\alpha : G \to \mathbb{C}$$

$$g \mapsto \alpha_g$$

given by

$$\langle \alpha, \beta \rangle = \sum_{g \in G} \alpha_g \overline{\beta_g}.$$

If we have can establish that the characters arising from irreducible representations are orthogonal under this inner product, then it follows from 6.6 and consideration of the dimension that they must span the space $CI(G)$.

Since characters are defined in terms of trace, and trace is defined in terms of matrices, it will be helpful to revert to the standpoint of representations as homomorphisms to groups of matrices.

Lemma 6.7 *Let*

$$\pi_1 : G \to GL_n(\mathbb{C})$$

$$g \mapsto A_g$$

and

$$\pi_2 : G \to GL_m(\mathbb{C})$$

$$g \mapsto B_g$$

be representations of G and let C be any n by m matrix. Then

$$C^\sharp = \sum_{g \in G} A_g C B_{g^{-1}}$$

intertwines the two representations.

Proof For any $h \in G$ we wish to show that

$$A_h C^\sharp B_{h^{-1}} = C^\sharp.$$

Writing this out we observe that the equalities

$$A_h C^\sharp B_{h^{-1}} = A_h (\sum_{g \in G} A_g C B_{g^{-1}}) A_{h^{-1}} = \sum_{g \in G} A_h A_g C B_{g^{-1}} B_{h^{-1}},$$

and then using that $g \mapsto A_g, \mapsto B_g$ are homomorphisms, this equals

$$\sum_{g \in G} A_{hg} C B_{(hg)^{-1}},$$

which is simply a different way of summing up

$$\sum_{g \in G} A_g C B_{g^{-1}} = C^\sharp.$$

□

Corollary 6.8 *Assume further that the representations from the last lemma are inequivalent irreducible representations. For each i, j let*

$$a_{i,j} : g \mapsto a_{i,j}(g),$$

$$b_{i,j} : g \mapsto b_{i,j}(g),$$

be the scalar valued functions assigning to each $g \in G$ the i, j^{th} entry of A_g, B_g respectively. Then for any $i, j \leq n$, $k, \ell \leq m$

$$\sum_{g \in G} a_{i,j}(g) b_{k,\ell}(g^{-1}) = 0.$$

Proof We let C be the $n \times m$ matrix with a 1 in the j, k^{th} spot and a solid slate of zeroes everywhere else. It is then not hard to see that

$$C^\sharp = \sum_{g \in G} a_{i,j}(g) b_{k,\ell}(g^{-1}),$$

and then by the previous lemma provides an intertwiner from the first representation to the second; since they are inequivalent irreducible representations, Schur's lemma gives $C^\sharp = 0$. □

Corollary 6.9 *Let $\chi_A, \chi_B : G \to \mathbb{C}$ be irreducible characters. Then*

$$\sum_{g \in G} \chi_A(g) \overline{\chi_B(g)} = 0.$$

Proof We assume that the characters arise from irreducible

$$G \to GL_n(\mathbb{C})$$

$$g \mapsto A_g$$

$$G \to GL_m(\mathbb{C})$$

$$g \mapsto B_g$$

respectively. Recall that $\chi_B(g^{-1}) = \overline{\chi_B(g)}$, so we are actually trying to show

$$\sum_{g \in G} \chi_A(g) \chi_B(g^{-1}) = 0.$$

Then by the last lemma, for any i, j

$$\sum_{g \in G} a_{i,i}(g) b_{j,j}(g^{-1}) = 0,$$

and so certainly

$$\sum_{g \in G} \chi_A(g) \chi_B(g^{-1}) = \sum_{g \in G} \sum_{i \leq n} \sum_{j \leq m} a_{i,i}(g) b_{j,j}(g^{-1}) = \sum_{i \leq n} \sum_{j \leq m} \sum_{g \in G} a_{i,i}(g) b_{j,j}(g^{-1}) = \sum_{g \in G} \chi_A(g) \chi_B(g^{-}1)$$

equals zero, as required. □

Thus, combined with the previous 6.6, we have that the characters form an orthogonal basis for the conjugacy invariant functions on G under the inner product

$$(\chi_1, \chi_2) = \frac{1}{|G|} \sum_{g \in G} \chi_1(g) \overline{\chi_2(g)}.$$

If we have a representation

$$G \to GL_n(\mathbb{C}),$$
$$g \mapsto A_g$$

with associated character $\chi_A : G \to \mathbb{C}$, then we may try to calculate

$$\sum_{g \in G} \chi_A(g) \chi_A(g^{-1}) = \sum_{g \in G} \chi_A(g) \overline{\chi_A(g)}$$

by our usual trick of summing in a different order. For the usual reasons we may assume that A_g is diagonal, and hence that at any i

$$(a_{i,i}(g))^{-1} = \overline{a_{i,i}(g)}.$$

Then it easily comes out

$$\chi_A(g) \overline{\chi_A(g)} = \sum_{i \leq n} a_{i,i}(g)(a_{i,i}(g))^{-1} = n;$$

summing over all g we obtain therefore

$$\sum_{g \in G} \chi_A(g) \overline{\chi_A(g)} = n|G|.$$

Thus if we take the set of functions

$$\{\frac{1}{n}\chi : \chi \text{ arises from an } n\text{-dimensional irreducible representation, for some } n\},$$

then we obtain an orthonormal basis for the conjugacy invariant functions on G. In the case that G is abelian any function on G is conjugacy invariant[11], and thus we subsume the results from the previous section.

There is also a philosophically interesting consequence of these calculations. If π_1, π_2 are irreducible representations then of course the corresponding characters, χ_1, χ_2, are a consequence of the similarity class: Equivalent representations give rise to identical characters. Thus if π_1, π_2 are equivalent representations, then

$$(\chi_1, \chi_2) =_{\text{df}} \frac{1}{|G|} \sum_{g \in G} \chi_1(g) \overline{\chi}_2(g)$$

equals their common dimension. If they are inequivalent then we observed that $(\chi_1, \chi_2) = 0$.

In this manner the slippery question of isomorphism or equivalence or representations has been reduced to the purely mechanical procedure of taking an inner product. Instead of searching high and low for the appropriate isomorphism, we test their equivalence by simply making a calculation.

[11] Since G is commutative exactly when $ghg^{-1} = h$ all g, h.

7 Perron-Frobenius

Definition A *matrice norm* on $M_n(\mathbb{C})$ is a norm under which this collection of n-by-n matrices becomes a Banach space under addition[12] and satisfies the additional inequality

$$\|AB\| \leq \|A\| \cdot \|B\|.$$

Examples For $A = (a_{ij})$, let

$$\|A\|_\infty = \max\{\sum_{j=1}^{j=n} |a_{ij}|\},$$

the maximum of the sum of the absolute value of the row.

If $\|\cdot\|$ is a matrice norm, and S is invertible, then $\|\cdot\|_S$ defined by

$$\|A\|_S = \|S^{-1}AS\|$$

is again a matrice norm.

Definition For $A \in M_{n,n}$ we let $\rho(A)$ be the supremum of $|\lambda|$, as λ ranges over eigenvalues of A.

Lemma 7.1 *$\rho(A)$ equals the infinum of the set of positive reals $\alpha > 0$ such that*

$$\lim_{n\to\infty}(\frac{A}{\alpha})^n \to 0.$$

Proof This amounts to saying that

$$\lim_{n\to\infty} A^n = 0$$

if and only if $\rho(A) < 1$. By considering Jordan canonical form, we may assume that A consists of a sequence of blocks down the main diagonal, each of the form

$$\begin{pmatrix} \lambda & 1 & 0 & \dots \\ 0 & \lambda & 1 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda \end{pmatrix},$$

as λ runs over the eigenvalues. Since these blocks multiply together independently, we may as well simply assume

$$A = \begin{pmatrix} \lambda & 1 & 0 & \dots \\ 0 & \lambda & 1 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda \end{pmatrix},$$

and we want to show that $A^n \to 0$ if and only if $|\lambda| < 1$.

The various n^{th} powers of A have λ^n down the diagonal, and so we certainly have that $|\lambda| < 1$ is necessary for convergence to 0. We are left with showing sufficiency.

I will leave this an exercise. One way to prove it as follows. We show by induction on $j \leq m$ (where A is $m \times m$) that the quantity

$$f(m,j) = \sum_{i\leq m} |a^n_{m-j,i}| \to 0,$$

[12] $\|\alpha A\| = |\alpha|\|A\|$, $\|A + B\| \leq \|A\| + \|B\|$, and the induced metric $d(A,B) = \|A - B\|$ is *complete*, in the sense of all Cauchy sequences having a limit.

where $a^n_{m-j,i}$ equals the i^{th} entry in the $m-j^{\text{th}}$ row of A^n. Once we know that $f(m,j)$ is very small for all $m \geq M$, then at each such m

$$f(m+1, j+1) \leq |\lambda| f(m, j+1) + f(m, j),$$

which, assuming $|\lambda| + f(m,j) < 1$, will start to push $f(m, j+1)$ towards zero. □

Corollary 7.2 *If $\|\cdot\|$ is a matrice norm, then for any A*

$$\rho(A) \leq \|A\|.$$

Notation We write $A \leq B$ if $(A = (a_{ij}), B = (b_{ij})$ and) each $a_{ij} \leq b_{ij}$. Write $0 \leq A$ if each $a_{ij} \geq 0$; $0 < A$ if each $a_{ij} > 0$.

Lemma 7.3 *If $0 \leq A \leq B$ then $\rho(A) \leq \rho(B)$.*

Proof By 7.1. □

Lemma 7.4 *If $0 \leq A$ and $\sum_{j \leq n} a_{ij}$ is constant, then for any i*

$$\rho(A) = \sum_{j \leq n} a_{ij}.$$

Proof For $\vec{x} = (1, 1, ..., 1)$, a vector in $\mathbb{R}^n$ we have that $\vec{x}$ is an eigenvector with eigenvalue $\sum_{j \leq n} a_{ij}$, which in turn equals $\|A\|_\infty$. Thus the lemma follows from 7.2. □

Corollary 7.5 *If $0 \leq A$, then*

$$\min_\ell \sum_{j \leq n} a_{\ell j} \leq \rho(A).$$

Proof Multiplying the rows by the right positive constants, we may find a positive matrix $B = (b_{ij}) \leq A$ with each

$$\sum_{j \leq n} b_{ij} \leq \min_\ell \sum_{j \leq n} a_{\ell j}.$$

□

Corollary 7.6 *If $\vec{x} = (x_1, x_2, ..., x_n)$ is positive (i.e. each $x_i > 0$), and $A \geq 0$, then*

$$\min_i \frac{1}{x_i} \sum_{j \leq n} a_{ij} x_j \leq \rho(A).$$

That is to say, at some $i \leq n$

$$\sum_{j \leq n} a_{ij} x_j \leq x_i \rho(A).$$

Proof Consider

$$X = \begin{pmatrix} x_1 & 0 & \cdot & \cdot \\ 0 & x_2 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & x_n \end{pmatrix}$$

and

$$X^{-1} = \begin{pmatrix} x_1^{-1} & 0 & \cdot & \cdot \\ 0 & x_2^{-1} & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & x_n^{-1} \end{pmatrix}.$$

Then by the last corollary we have

$$\rho(A) = \rho(X^{-1}AX) \geq \min_i \sum_{j \leq n} a_{ij}x_j.$$

□

Corollary 7.7 *If $\vec{x} \geq 0$ and $A \geq 0$ then at some i*

$$\sum_{j \leq n} (A\vec{x})_{ij} \leq \rho(A)\vec{x}_i.$$

Proof We may apply 7.6 to a sequence $(\vec{x}_i)_i$ of strictly positive vectors which approach $\vec{x}$; then the corollary follows by continuity. □

Theorem 7.8 *(Perron) If $A > 0$ then there is an $\vec{x} \geq 0$ with*

$$A\vec{x} = \rho(A) \cdot \vec{x},$$

and hence $\vec{x} > 0$.

Proof Choose λ with $|\lambda| = \rho(A)$ and then choose $\vec{x}$ with

$$A\vec{x} = \lambda \cdot \vec{x}.$$

Then

$$|\lambda||\vec{x}| = |\lambda\vec{x}| = |A\vec{x}| \leq |A||\vec{x}| = A|\vec{x}|.$$

Thus let $y = A|\vec{x}| - |\lambda||\vec{x}|$; we have in the above line that $y \geq 0$, and if $y = 0$ we are done. So suppose for a contradiction that $y \neq 0$.

Then $Ay > 0$

$$\therefore A(A|\vec{x}|) > A|\lambda||\vec{x}| = |\lambda|A|\vec{x}|.$$

Thus if $A|\vec{x}| = \vec{z} = (z_1, z_2, ..., z_n)$ we have $\vec{z} > 0$ and at each $i \leq n$

$$\sum_{j \leq n} a_{ij}z_j > |\lambda|z_i = \rho(A)z_i,$$

contradicting last corollary. □

Corollary 7.9 *If $0 \leq A$ then there is non-zero $\vec{x} \geq 0$ with*

$$\rho(A)\vec{x} = A\vec{x}.$$

Proof We may choose a sequence of positive matrices $(B_i)_i$ with

$$\|B_i\|_\infty \to 0,$$

$$B_{i+1} < B_i,$$

and then we may apply Perron to $A_i = B_i + A$; let $\rho_i = \rho(A_i)$.

At each i we may choose $\vec{x}_i > 0$ of length 1 with

$$A_i \vec{x}_i = \rho_i \cdot \vec{x}_i.$$

By 7.3 we have each $\rho_i > \rho_{i+1} > \rho(A)$, and thus for ρ_∞ the limit point of the monotone sequence ρ_i and $\vec{x} \geq 0$ an accumulation point of the set $\{\vec{x}_i : i \in \mathbb{N}\}$ we have by 7.6

$$\rho(A)\vec{x} \leq \rho_\infty \vec{x} = \lim_i \rho_i \vec{x}_i = \lim_i A_i \vec{x}_i = A\vec{x} \leq \rho(A)\vec{x},$$

and thus there is equality throughout and the lemma is proved. □

Lemma 7.10 *If $A > 0$, $\vec{x} \neq 0$, and $A\vec{x} = \rho(A)\vec{x}$, then $\vec{x} = \alpha\vec{y}$ some $\alpha \in \mathbb{C}$, $\vec{y} > 0$.*

Proof We have by 7.7 that

$$A|\vec{x}| \leq \rho(A)|\vec{x}| = |A\vec{x}| \leq A|\vec{x}|,$$

and hence $|A\vec{x}| \leq A|\vec{x}|$. But then considering the nature of the triangle inequality we must have $\vec{x} = \alpha\vec{y}$ some $\alpha \in \mathbb{C}$ and $\vec{y} = |\vec{x}|$; and then in fact $\vec{y} > 0$ since $A\vec{y} = \rho(A)\vec{y} > 0$. □

Corollary 7.11 *If $A > 0$ then the eigenspace corresponding to $\rho(A)$ has dimension 1.*

Proof Otherwise from the last lemma we obtain independent positive eigenvectors $\vec{z}_1, \vec{z}_2 > 0$ for $\rho(A)$. Then some $k \in \mathbb{R}^+$ gives a new eigenvector $\vec{z} = k\vec{z}_1 - \vec{z}_2 \geq 0$ with a zero entry in some coordinate.

But since $\rho(A)\vec{z} = A\vec{z} > 0$ we have $\vec{z} > 0$ with a contradiction. □

References

[1] M.L. Curtis, **Matrix groups.** Universitext. Springer-Verlag, New York-Heidelberg, 1979.

[2] G.W. Mackey, **Unitary group representations in physics, probability, and number theory.** Mathematics Lecture Note Series, 55. Benjamin/Cummings Publishing Co., Inc., Reading, Mass., 1978.

Aknowledgments: These notes were greatly improved by the input and corrections provided by the class of Math 190, Spring 2002, UCLA.

greg@math.ucla.edu

MSB 6363
UCLA
CA 90095-1555