Burnside's Theorem on Matrix Algebras

Joel H. Shapiro

August 28, 2014

This note describes Burnside's Theorem on matrix algebras, and its application to triangularization of commutative families of matrices.

The English mathematician William Burnside published a paper in 1905[1] proving that if, for a group G of $n \times n$ (necessarily invertible) complex matrices, there's no subspace of $\mathbb{C}^n$ (other than $\{0\}$ and $\mathbb{C}^n$) that every member of G maps into itself, then G must contain n^2 linearly independent matrices, and so spans the full algebra of $n \times n$ complex matrices.

[1] *On the condition of reducibility of any group of linear substitutions,* Proc. London Math.Soc. 3 (1905) 430-434.

For an example of how one might use Burnside's Theorem, suppose $n > 1$ and that G is *abelian* (i.e. commutative). Then the linear span of G is a commutative algebra of matrices, and so (because $n > 1$) can't be the full algebra. Burnside's Theorem thus guarantees a nontrivial subspace of $\mathbb{C}^n$ that's taken into itself by every member of G. If this "G-invariant subspace"—call it M—has dimension > 1 we can restrict G to M and apply Burnside's Theorem to the resulting group of restrictions, thus obtaining a further G-invariant subspace of strictly smaller dimension. Upon repeating the argument, if necessary, we eventually arrive at a one dimensional G-invariant subspace, i.e.

Here "nontrivial" means "neither $\{0\}$ nor $\mathbb{C}^n$."

Every commutative group of $n \times n$ matrices has a common eigenvector.

Below you'll find a short, accessible proof of a modern version of Burnside's Theorem, which shows, for example, that in the above application it's not necessary for our commutative family of matrices to be a group. We'll take this example further, identifying it as the fundamental step in extending Schur's famous triangularization theorem to commuting families of matrices.

1 *Notation and terminology*

Vector spaces. V will always denote a vector space over the field $\mathbb{C}$ of complex numbers. M will always denote a (linear) subspace of V. A subspace M is said to be *nontrivial* if it's neither the zero-subspace nor the whole space.

Invariance. To say a subset of V is invariant under a linear transformation on V means that the subset gets taken by that transformation into itself. To say our subset is invariant under a *collection* of linear transformations means that it's invariant for every transformation in the collection.

Algebras. $\mathscr{L}(V)$ will denote the collection of all linear transformations $V \to V$. It's an *algebra* over $\mathbb{C}$, i.e. a vector space with multiplication (composition of transformations) that behaves "properly" with respect to scalar multiplication.

A *subalgebra* of $\mathscr{L}(V)$ is just a vector subspace $\mathscr{A}$ of $\mathscr{L}(V)$ that's closed under operator multiplication (i.e. composition). We'll always assume here that our subalgebras are *nontrival*, i.e. not just the zero-transformation.

Irreducibility. To say a subalgebra $\mathscr{A}$ of $\mathscr{L}(V)$ is *irreducible* means that it has no nontrivial invariant subspace.

Transitivity. To say a subalgebra $\mathscr{A}$ of $\mathscr{L}(V)$ is *transitive* means that for every vector $v \in V\backslash\{0\}$,

$$\mathscr{A}v := \{Av : A \in \mathscr{A}\} = V.$$

2 *Modern version of Burnside's Theorem*

HERE'S THE RESULT generally known these days as " Burnside's theorem on matrix algebras."

Theorem. *For a finite dimensional complex vector space V, the only irreducible subalgebra of $\mathscr{L}(V)$ is $\mathscr{L}(V)$ itself.*

2.1 *Remarks on the Theorem*

(a) THE MODERN THEOREM IMPLIES THE ORIGINAL ONE.

Proof. If G is a group of linear transformations of V, let $\mathscr{A}$ be the linear span of G; it's a subalgebra of $\mathscr{L}(V)$. A subspace M of V is invariant for $\mathscr{A}$ if and only if it's invariant for G, so the hypothesis of Burnside's original theorem is that $\mathscr{A}$ is irreducible, hence by Theorem 2 it's equal to $\mathscr{L}(V)$. In other words: G spans $\mathscr{L}(V)$, so must contain a basis for $\mathscr{L}(V)$. □

Note that this argument shows—as promised above—that Burnside's original theorem only requires G to be a collection of matrices that's closed under matrix multiplication.

(b) $\mathscr{L}(V)$ IS IRREDUCIBLE.

Proof. This is trivial if $\dim V = 1$, so suppose V has dimension > 1. Without loss of generality we can suppose $V = \mathbb{C}^n$ for some $n > 1$. Let M be a nontrivial subspace of V, so there's a vector $v \neq 0$ that belongs to V, but not to M. Fix $m \in M$ with $\|m\| = 1$, and let S be the linear transformation defined on V by

We use the usual Euclidean norm and inner product on $V = \mathbb{C}^n$.

$$Sx = \langle x, m \rangle v \qquad (x \in V).$$

The usual notation for this operator is $S := v \otimes m$.

Then

$$S(m) = \langle m, m \rangle v = v$$

so $\{0\} \neq S(M) \subset S(V) = \text{span}\{v\}$, hence $S(M) = \text{span}\{v\}$, a one dimensional subspace that intersects M only at the origin. So M is not invariant for S, and hence not for $\mathscr{L}(V)$. □

(c) Transitivity is equivalent to irreducibility.

Proof. This is trivial for $\dim V = 1$, so suppose V has dimension > 1.

Suppose $\mathscr{A}$ is irreducible. For $0 \neq x \in V$ the set $\mathscr{A}x$ is a subspace of V (because $\mathscr{A}$ is a vector space) and it's invariant for $\mathscr{A}$ (because $\mathscr{A}$ is an algebra). It can't be the zero-subspace; if it were, then $K := \cap_{A \in \mathscr{A}} \ker A$ would be an invariant subspace for $\mathscr{A}$ that contains x, and so isn't the zero-subspace. By irreducibility we'd have $K = V$, i.e. $\mathscr{A} = \{0\}$, a contradiction to our standing assumption that subalgebras are nontrivial. Thus $\mathscr{A}x \neq \{0\}$, so it must equal V, else it would be a nontrivial invariant subspace for $\mathscr{A}$, contradicting the assumed irreducibility of that subalgebra.

Conversely, suppose $\mathscr{A}$ is transitive, i.e. that $\mathscr{A}x = V$ for every $0 \neq x \in V$. Suppose M is a subspace of V that's $\neq \{0\}$ and is invariant for $\mathscr{A}$. Choose $0 \neq x \in M$. Then $V = \mathscr{A}x \subset M$, i.e $M = V$. Thus $\mathscr{A}$ has no nontrivial invariant subspace. □

2.2 *Application to Schur Triangularization*

One of the fundamental results of matrix theory is Schur's Triangularization Theorem, which asserts that every complex matrix is unitarily equivalent to one in upper triangular form. Stated in terms of linear transformations, this result asserts that for every $A \in \mathscr{L}(V)$ there's an orthonormal basis for V with respect to which the matrix of A is upper triangular. Thanks to Burnside's theorem this result can easily be extended to assert that:

> *For any commuting subset $\mathscr{F}$ of $\mathscr{L}(V)$ there's an orthonormal basis of V with respect to which every operator in $\mathscr{F}$ has an upper triangular matrix.*

Idea of proof: The result is trivial if $\dim V = 1$, so suppose V has dimension $n > 1$. The key is to find, for all the transformations in $\mathscr{F}$, a common eigenvector, i.e. a one dimensional invariant subspace. We may as well assume $\mathscr{F}$ is an algebra—otherwise replace it by the algebra it generates (the linear span of all finite products of elements of $\mathscr{F}$), the point being that $\mathscr{F}$ and the algebra it generates have the same invariant subspaces. Since our algebra $\mathscr{F}$ is *commutative*, it's not all of $\mathscr{L}(V)$ (remember: $\dim V > 1$), hence Burnside's Theorem guarantees for $\mathscr{F}$ a nontrivial invariant subspace M. If $\dim M = 1$ we've got what we want. Otherwise restrict $\mathscr{F}$ to M and apply Burnside's Theorem again. Repeat if necessary. Eventually—thanks to the finite dimensionality of V—you arrive at a one dimensional invariant subspace for $\mathscr{F}$.

So far we know that $\mathscr{F}$ has a common eigenvector; call it v. Thus with respect to the orthogonal direct sum decomposition $V = \text{span}\{v\} \oplus \text{span}\{v\}^\perp$, each of the transformations in $\mathscr{F}$ has matrix that's "block upper-triangular" in the sense that it looks like $\left(\begin{smallmatrix}\lambda & * \\ 0 & B\end{smallmatrix}\right)$ where $\lambda \in \mathbb{C}$ and B is an $(n-1) \times (n-1)$ matrix.

The proof is then finished off by "compressing" the transformations in $\mathscr{A}$ to the orthogonal complement of $\text{span}\{v\}$ and repeating the process; for the details, see e.g. Sheldon Axler's book[2] (Theorem 5.13, pp.84–85), or Sergei Treil's recent lecture notes[3] (Chapter 6, §1).

[2] *Linear Algebra Done Right* second ed., Springer 1997

[3] *Linear Algebra Done Wrong*. Available online at: `http://www.math.brown.edu/treil/papers/LADW/LADW.html`

3 *Proof of Burnside's Theorem on Matrix Algebras*

In view of §2.1(b) we need only show that if $\mathscr{A}$ is an irreducible algebra then it's equal to $\mathscr{L}(V)$. For this it's enough to prove:

Lemma. *If $\mathscr{A}$ is an irreducible subalgebra of $\mathscr{L}(V)$, then $\mathscr{A}$ contains a rank-one operator.*

Here "operator" means "linear transformation."

THE LEMMA IMPLIES THE THEOREM. We are assuming that there exist nonzero vectors v and w in V such that the irreducible algebra $\mathscr{A}$ contains the rank-one operator $v \oplus w$, i.e. the linear transformation that takes $x \in V$ to the vector $\langle x, w\rangle v$.

We first encountered such an operator at the top of page 3.

It's an easy exercise to prove that for any $T \in \mathscr{L}(V)$:

(1) $\quad T(v \oplus w) = (Tv) \oplus w \quad$ and $\quad (v \oplus w)T = v \oplus (T^* w)$

The first of these equations shows that if $T \in \mathscr{A}$ then so is $(Tv) \oplus w$. But since $\mathscr{A}$ is irreducible, it's transitive, so as T runs through $\mathscr{A}$, the images Tv exhaust V. Thus: $v \oplus w \in \mathscr{A}$ *for any* $v \in V$.

Now use the second equation in (1) above to apply the same reasoning to $T^* w$ as T runs through $\mathscr{A}$. The point here is that the irreducibiity of $\mathscr{A}$ is inherited by the "adjoint algebra"

$\mathscr{A}^* := \{A^* : A \in \mathscr{A}\}$.
Conclusion: $v \oplus w \in \mathscr{A}$ for any $v, w \in V$.

Exercise: The key is that a subspace M is $\mathscr{A}$-invariant iff its orthogonal complement $M^\perp$ is $\mathscr{A}^*$-invariant.

Thus $\mathscr{A}$ contains all rank-one operators on V. Since $\mathscr{L}(V)$ is the span of all such operators, $\mathscr{A} = \mathscr{L}(V)$, as promised. □

PROOF OF THE LEMMA. Let $n = \dim V$. For $n = 1$ the Lemma is trivial, so suppose $n > 1$ and the Lemma's true for all complex vector spaces of dimension between 1 and $n-1$. Suppose $\mathscr{A}$ is a transitive subalgebra of $\mathscr{L}(V)$. We wish to show that $\mathscr{A}$ contains a rank-one operator. We'll start by proving something along these lines, but weaker:

Claim: $\mathscr{A}$ contains a non-invertible transformation S.

Proof of Claim. By convention $\mathscr{A} \neq \{0\}$, and by irreducibility $\mathscr{A}$ doesn't consist entirely of scalar multiples of the identity operator. Thus there's an operator $T \in \mathscr{A}$ that's neither the zero-operator nor a scalar multiple of the identity. If T is not invertible, take $S = T$. Suppose T is invertible. We know T has an eigenvalue λ so $T - \lambda I$ is not invertible, and is not the zero-operator. If $\mathscr{A}$ contains the identity transformation, we're done, with $S = T - \lambda I$. If not, let $S = T(T - \lambda I)$, which is not invertible, and (because T is invertible) is not identically zero, and (because $S = T^2 - \lambda T$) belongs to $\mathscr{A}$. This proves the *Claim.*

Here we use for the first—and only–time the fact that the field $\mathbb{C}$ is algebraically closed!

To finish the proof of the Lemma, let $V_0 = S(V)$, and consider the family $\mathscr{A}_0$ of restrictions to V_0 of the transformations $S\mathscr{A} := \{SA : A \in \mathscr{A}\}$. Then V_0 is a subspace of V which, since S is not invertible, has dimension $< n$. One checks easily that $\mathscr{A}_0$ is a subalgebra of $\mathscr{L}(V_0)$.

For our induction argument to work we need $\mathscr{A}_0$ to be irreducible. This is true because $\mathscr{A}_0$ is *transitive.* To see why, fix a non-zero vector $w \in V_0$. Then $\mathscr{A}w = V$ by the irreducibility ($\equiv$ transitivity) of $\mathscr{A}$ on V, hence $\mathscr{A}_0 w = S(\mathscr{A}w) = S(V) = v_0$, so $\mathscr{A}_0$ is transitive on V_0, and so irreducible on that subspace.

The induction hypothesis guarantees that $\mathscr{A}_0$ contains a rank-one operator, i.e. a transformation $T \in \mathscr{A}$ for which the restriction of ST to $S(V)$ has rank one. Thus $\dim STS(V) = 1$, i.e. STS, which belongs to $\mathscr{A}$, has rank one. □

4 *Notes*

THE PROOF OF BURNSIDE'S THEOREM given here comes from Lomonosov and Rosenthal's paper: *The simplest proof of Burnside's Theorem on matrix algebras.*[4]. For a different, but equally simple proof,

[4] Lin. Alg. Appl. 383 (2004) 45-47.

along with some supplementary background on Burnside's Theorem, see T.Y. Lam's paper: *A theorem of Burnside on matrix rings.*[5]

[5] American Math. Monthly 105 (1998) 651-653.

If one replaces the Hilbert-space elements of our proof of Burnside's Theorem by some simple duality arguments it becomes clear that the result remains true if the field of complex numbers is replaced by any algebraically closed field.

Burnside's Theorem fails for infinite dimensional complex Hilbert spaces, and for finite dimensional vector spaces over the reals. For the former situation, our argument in §2.1(b) shows that the finite rank operators on infinite dimensional complex Hilbert space furnish an example of an algebra of operators that's irreducible, but not the algebra of all operators.

Here "operator" could also mean "continuous linear transformation."

As for real finite dimensional vector spaces, let $\mathscr{R}$ denote the subalgebra of $\mathscr{L}(\mathbb{R}^2)$ generated by the transformation R of rotation through $\pi/2$ radians (about the origin). Then $\mathscr{R}$ is easily seen to be the linear span of R and the identity on $\mathbb{R}^2$, so it's $\neq \mathscr{L}(\mathbb{R}^2)$. On the other hand $\mathscr{R}$ has the same invariant subspaces as R, namely just the trivial ones. Thus $\mathscr{R}$, although not $\mathscr{L}(\mathbb{R}^2)$, is nevertheless irreducible.

Nevetheless, V. I. Lomonosov[6], in the course of studying the existence of invariant subspaces for (continuous) operators on infinite dimensional Banach spaces, proved for that setting a theorem that, when applied to finite dimensional spaces, reduces to Burnside's theorem. Lomonosov uses his "infinite dimensional Burnside Theorem" to unify and extend several important theorems on invariant subspaces for Banach-space operators.

[6] *An extension of Burnside's Theorem to infinite dimensinal spaces,* Israel J. Math. 75 (1991) 329–339.

For a sketch of Burnside's life and work see the MacTutor article "William Burnside" by J.J. O'Connor and E.F. Robinson[7].

[7] Available online at `www-history.mcs.st-andrews.ac.uk/Biographies/Burnside.html`

Bibliography

Sheldon Axler, *Linear Algebra Done Right* second ed., Springer 1997.

William Burnside, *On the condition of reducibility of any group of linear substitutions,* Proc. London Math.Soc. 3 (1905) 430-434.

T. Y. Lam, *A theorem of Burnside on matrix rings.* American Math. Monthly 105 (1998) 651-653.

Victor Lomonosov and Peter Rosenthal, *The simplest proof of Burnside's Theorem on matrix algebras.* Lin. Alg. Appl. 383 (2004) 45-47.

Victor Lomonosov, *An extension of Burnside's Theorem to infinite dimensinal spaces,* Israel J. Math. 75 (1991) 329–339.

Sergei Treil, *Linear Algebra Done Wrong,* Available online at: http://www.math.brown.edu/treil/papers/LADW/LADW.html

J. J. O'Connor and E. F. Robinson, *William Burnside,*// Available online at www-history.mcs.st-andrews.ac.uk/Biographies/Burnside.html

Fariborz Maseeh Department of Mathematics & Statistics
Portland State University
PO Box 751, Portland, OR 97207-0751
joels314(at)gmail.com
www.joelshapiro.org