SIMULTANEOUS COMMUTATIVITY OF OPERATORS

KEITH CONRAD

In this note we work with linear operators on finite-dimensional complex vector spaces. Any such operator has an eigenvector, by the fundamental theorem of algebra. A linear operator is called *diagonalizable* if it has a basis of eigenvectors: there is a basis in which the matrix representation of the operator is a diagonal matrix. We are interested in conditions that make a finite set of linear operators *simultaneously diagonalizable*: there is a basis in which the matrix representation of each operator is diagonal.

Example 1. The matrices $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)$ acting on $\mathbf{C}^2$ are each diagonalizable, but they are not simultaneously diagonalizable: the eigenvectors of A are scalar multiples of $\binom{i}{1}$ and $\binom{-i}{1}$ while the eigenvectors of B are scalar multiples of $\binom{1}{1}$ and $\binom{-1}{1}$.

Example 2. The matrices

$$A = \begin{pmatrix} 7 & -10 & 5 \\ 4 & -5 & 3 \\ -1 & 3 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} 8 & -11 & 6 \\ 5 & -6 & 4 \\ 0 & 2 & 1 \end{pmatrix}$$

acting on $\mathbf{C}^3$ are simultaneously diagonalizable with common eigenbasis

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -3/2 + i/2 \\ -1/2 + i/2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -3/2 - i/2 \\ -1/2 - i/2 \\ 1 \end{pmatrix}.$$

For linear operators to be simultaneously diagonalizable, they at least have to be individually diagonalizable, but more is needed (see Example 1). A further necessary condition is that the operators commute, since diagonal matrices commute. Check the matrices in Example 1 do not commute while those in Example 2 do commute. The following theorem shows that commuting linear operators already have something in common.

Theorem 3. *If $A_1, \ldots, A_r$ are commuting linear operators on a finite-dimensional $\mathbf{C}$-vector space V then they have a common eigenvector in V.*

Proof. We induct on r, the result being clear if $r = 1$ since we work over the complex numbers: every linear operator on a finite-dimensional $\mathbf{C}$-vector space has an eigenvector.

Now assume $r \geq 2$. Let the last operator A_r have an eigenvalue $\lambda \in \mathbf{C}$ and let

$$E_\lambda = \{v \in V : A_r v = \lambda v\}$$

be the λ-eigenspace for A_r. For $v \in E_\lambda$, $A_r(A_i v) = A_i(A_r v) = A_i(\lambda v) = \lambda(A_i v)$, so $A_i v \in E_\lambda$. Thus each A_i *restricts* to a linear operator on the subspace E_λ.

The linear operators $A_1|_{E_\lambda}, \ldots, A_{r-1}|_{E_\lambda}$ commute since the A_i's commute as operators on the larger space V. There are $r - 1$ of these operators, so by induction on r (while quantifying over *all* finite-dimensional $\mathbf{C}$-vector spaces) these operators have a common eigenvector in E_λ. That vector is also an eigenvector of A_r by the definition of E_λ. We're done. □

A common eigenvector for $A_1, \ldots, A_r$ need not have the same eigenvalue for all the operators; the first eigenvector in Example 2 has eigenvalue 2 for A and 3 for B.

Lemma 4. *If $A\colon V \to V$ is a diagonalizable linear operator and W is an A-stable subspace of V then the restriction $A|_W\colon W \to W$ is also diagonalizable.*

Proof. We use a variation on the answer by Zorn at https://math.stackexchange.com/questions/62338/. Letting $\lambda_1, \ldots, \lambda_r$ be the distinct eigenvalues of A acting on V, we have $V = \bigoplus_{i=1}^r E_{\lambda_i}$, where E_{λ_i} is the λ_i-eigenspace of A. For $w \in W$, write $w = v_1 + \cdots + v_r$ where $v_i \in E_{\lambda_i}$. We'll prove each v_i is in W. Then $W = \bigoplus_{i=1}^r (E_{\lambda_i} \cap W)$, so $A|_W$ is diagonalizable.

Since W is A-stable, $A^k(w) \in W$ for all $k \geq 0$. Also $A^k w = \lambda_1^k v_1 + \cdots + \lambda_r^k v_r$. Taking $k = 0, 1, \ldots, r-1$, we have the following equation in V^r:

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_r \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{r-1} & \lambda_2^{r-1} & \cdots & \lambda_r^{r-1} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_r \end{pmatrix} = \begin{pmatrix} w \\ Aw \\ \vdots \\ A^{r-1}w \end{pmatrix}.$$

The vector on the right is in W^r (a subspace of V^r) and the $r \times r$ matrix on the left is invertible (Vandermonde matrix with distinct λ_i). Therefore the vector on the left is in W^r, so each v_i is in W. □

Theorem 5. *If $A_1, \ldots, A_r$ are commuting linear operators on V and each A_i is diagonalizable then they are simultaneously diagonalizable, i.e., there is a basis of V consisting of simultaneous eigenvectors for the A_i.*

Proof. We won't use Theorem 3, but the proof will be essentially the same type of argument as in the proof of Theorem 3; the stronger hypothesis (commutativity and individual diagonalizability) will lead to a stronger conclusion (a basis of simultaneous eigenvectors, not just one simultaneous eigenvector).

The result is clear if $r = 1$, so assume $r \geq 2$. Since the last operator A_r is diagonalizable on V, V is the direct sum of the eigenspaces for A_r. Let λ be an eigenvalue for A_r and E_λ be the λ-eigenspace of A_r in V. As in the proof of Theorem 3, since each A_i commutes with A_r we have $A_i(E_\lambda) \subset E_\lambda$. Thus each A_i *restricts* to a linear operator on the subspace E_λ and the linear operators $A_1|_{E_\lambda}, \ldots, A_{r-1}|_{E_\lambda}$ commute since $A_1, \ldots, A_{r-1}$ commute as operators on V.

By Lemma 4, the restrictions $A_1|_{E_\lambda}, \ldots, A_{r-1}|_{E_\lambda}$ are each diagonalizable on E_λ. Since the number of these operators is less than r, by induction on r there is a basis for E_λ consisting of simultaneous eigenvectors for $A_1|_{E_\lambda}, \ldots, A_{r-1}|_{E_\lambda}$.[1] The elements of this basis for E_λ are eigenvectors for $A_r|_{E_\lambda}$ as well, since *all* nonzero vectors in E_λ are eigenvectors for A_r. Thus $A_1|_{E_\lambda}, \ldots, A_{r-1}|_{E_\lambda}, A_r|_{E_\lambda}$ are all diagonalizable. The vector space V is the direct sum of the eigenspaces E_λ of A_r, so stringing together simultaneous eigenbases of $A_1|_{E_\lambda}, \ldots, A_{r-1}|_{E_\lambda}, A_r|_{E_\lambda}$ as λ runs over the eigenvalues of A_r gives a simultaneous eigenbasis of V for all the $A_1, \ldots, A_r$. □

Remark 6. Theorem 5 is *not* saying commuting operators diagonalize! It says commuting diagonalizable operators simultaneously diagonalize. For example, the matrices $\left(\begin{smallmatrix} 1 & a \\ 0 & 1 \end{smallmatrix}\right)$ for all a commute with each other, but none of them are diagonalizable when $a \neq 0$.

Because we are dealing with operators on finite-dimensional spaces, Theorem 5 extends to a possibly infinite number of commuting operators, as follows.

Corollary 7. *Let $\{A_i\}$ be a set of commuting linear operators on a finite-dimensional $\mathbf{C}$-vector space V. If each A_i is diagonalizable on V then they are simultaneously diagonalizable.*

[1] This choice of basis for E_λ is not made by A_r, but by the other operators together.

Proof. Let U be the subspace of $\mathrm{End}_F(V)$ spanned by the operators A_i's. Since $\mathrm{End}_F(V)$ is finite-dimensional, its subspace U is finite-dimensional, so U is spanned by a finite number of A_i's, say $A_{i_1}, \ldots, A_{i_r}$. By Theorem 5, there is a common eigenbasis of V for $A_{i_1}, \ldots, A_{i_r}$. A common eigenbasis for linear operators is also an eigenbasis for any linear combination of the operators, so this common eigenbasis of $A_{i_1}, \ldots, A_{i_r}$ diagonalizes every element of U, and in particular diagonalizes each A_i. □

Corollary 7 is important in number theory, where it implies the existence of eigenforms for Hecke operators.