JOINT DIAGONALIZABILITY

MANUEL W. LUETHI

Abstract. We give a genuinely linear proof that a commuting family of diagonalizable matrices is jointly diagonalizable.

In 2017, during my PhD studies at ETH, I was organizing a first year course in Linear Algebra for mathematicians and physicists. There, I faced the following problem. The lecturer (and I) wanted to discuss linear algebra over arbitrary fields where sensible, but without delving too much into algebra. After all it should be a first year course. In particular, there was the issue that we did not discuss algebraic closures and factorization of polynomials except for $\mathbb{R}$ and $\mathbb{C}$. Moreover, due to time constraints we ended up not discussion the minimal polynomial of a matrix. Still I wanted to discuss the fact that any commuting family of diagonalizable operators is jointly diagonalizable, and thus I came up with the following "genuinely linear" proof, which was given to the students as a (guided) exercise. As it was previously unknown to me and as I like the fact that it is purely linear, I decided to write up a clean version of it.

In what follows, V is a finite dimensional vector space over a field $\mathbb{K}$, and $\mathcal{T} \subset \operatorname{End}(V)$ is a commuting family of diagonalizable operators, i.e. $ST = TS$ for all $S, T \in \mathcal{T}$ and for every $T \in \mathcal{T}$ there is a basis of V consisting of eigenvectors of T. Given $T \in \operatorname{End}(V)$, we denote by $\sigma(T)$ the set of eigenvalues of T and for all $\lambda \in \mathbb{K}$ we denote by $E_\lambda(T)$ the subspace

$$E_\lambda(T) = \{v \in V : Tv = \lambda v\}.$$

Note that $E_\lambda(T) \neq \{0\}$ if and only if $\lambda \in \sigma(T)$ and for diagonalizable $T \in \operatorname{End}(V)$ we have

$$V = \bigoplus_{\lambda \in \mathbb{K}} E_\lambda(T) = \bigoplus_{\lambda \in \sigma(T)} E_\lambda(T).$$

Theorem 1. *$\mathcal{T}$ is* jointly diagonalizable, *i.e. there is a basis $\mathcal{B}$ of V consisting of eigenvectors of $\mathcal{T}$ in the sense that for all $v \in \mathcal{B}$ and for all $T \in \mathcal{T}$ there is some $\lambda_T \in \mathbb{K}$ such that $Tv = \lambda_T v$.*

For the sake of readability, we split the proof into several lemmata.

Lemma 2. *Let $T \in \operatorname{End}(V)$ and assume $W \subseteq V$ is T-invariant, i.e. $T(W) \subseteq W$. Then*

$$\overline{T} : V/W \to V/W, \overline{T}(v + W) = T(v) + W$$

is well-defined, linear. If $v \in V$ is an eigenvector of T with eigenvalue λ, then

$$\overline{T}(v + W) = \lambda(v + W).$$

Moreover, if T is diagonalizable, then so is $\overline{T}$.

Date: January 12, 2019.

Proof of Lemma 2. Let $v_1, v_2 \in V$ satisfy $v_1 - v_2 \in W$, then

$$T(v_1) + W = \big(T(v_2) + \underbrace{T(v_1 - v_2)}_{\in W}\big) + W = T(v_2) + W.$$

Hence $\overline{T}$ well-defined. Linearity is a routine check. Let $v_1, v_2 \in V$ and $\mu \in \mathbb{K}$, then

$$\begin{aligned}\overline{T}\big((v_1 + W) + \mu(v_2 + W)\big) &= \overline{T}\big((v_1 + \mu v_2) + W\big)\\ &= T(v_1 + \mu v_2) + W\\ &= \big(T(v_1) + \mu T(v_2)\big) + W\\ &= \big(T(v_1) + W\big) + \mu\big(T(v_2) + W\big)\\ &= \overline{T}(v_1 + W) + \mu\overline{T}(v_2 + W).\end{aligned}$$

If $v \in V$ is an eigenvector of T, then by definition

$$\overline{T}(v + W) = T(v) + W = \lambda v + W = \lambda(v + W). \tag{1}$$

Assume now that T is diagonalizable. We show that $\overline{T}$ is diagonalizable. Note that whenever $\{v_1, \ldots, v_n\}$ is a basis of V, then the set $\{v_1 + W, \ldots, v_n + W\}$ of cosets is a generating set of V/W and thus contains a basis of V/W. As V by assumption admits a basis consisting of eigenvectors of T, (1) shows that there is a basis of V/W consisting of eigenvectors of $\overline{T}$.[1] □

Lemma 3. *Let $T \in \mathrm{End}(V)$ and let $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ denote pairwise distinct eigenvalues of T and let $v_i \in E_{\lambda_i}(T)$ $(1 \leq i \leq k)$. If $v_1 + \cdots + v_k = 0$, then $v_1 = \cdots = v_k = 0$.*

Proof of Lemma 3. It suffices to show that any finite collection of eigenvectors for pairwise distinct eigenvalues is linearly independent. Hence assume that the v_i are eigenvectors (i.e. none equal 0), and assume that $\mu_1, \ldots, \mu_k \in \mathbb{K}$ are scalars such that

$$\mu_1 v_1 + \cdots + \mu_k v_k = 0.$$

We show that this relation is trivial by induction. This is clear in case $k = 1$. Note that the base for induction did not use T. Hence assume now that $k \geq 2$ and that given an arbitrary linear map on a vector space the statement is true for sums of up to $k - 1$ eigenvectors with pairwise distinct eigenvalues. We note that $W = E_{\lambda_r}(T)$ is a T-invariant subspace. Consider the map $\overline{T} : V/W \to V/W$ constructed in Lemma 2 and note that

$$0 = \mu_1\overline{T}(v_1) + \cdots + \mu_{k-1}\overline{T}(v_{k-1}).$$

Note that $v_i \notin W$ for all $1 \leq i < k$, hence $\overline{T}(v_1), \ldots, \overline{T}(v_{k-1})$ form a collection of eigenvectors of $\overline{T}$ for pairwise distinct eigenvalues by Lemma 2. The induction assumption implies that $\mu_1 = \cdots = \mu_{k-1} = 0$ and hence the claim. □

Lemma 4. *Assume that $T \in \mathrm{End}(V)$ is diagonalizable and $W \subset V$ is a T-invariant subspace, then the restriction $T|_W \in \mathrm{End}(W)$ is diagonalizable and the eigenspace decomposition of W with respect to $T|_W$ is given by*

$$W = \bigoplus_{\lambda \in \sigma(T)} \big(W \cap E_\lambda(T)\big).$$

[1] One should convince oneself that this makes sense in case $V = W$.

Remark. The decomposition of W in Lemma 4 relies on the T-invariance of W. In general it is not true that a subspace is the sum of the intersections of the subspace with all the direct summands. The counterexample is $W = \mathbb{R}(e_1 + e_2) \subseteq \mathbb{R}^2$, where the decomposition of the ambient vector space is given by $\mathbb{R}^2 = \mathbb{R}e_1 \oplus \mathbb{R}e_2$. Here e_1, e_2 denotes a basis of $\mathbb{R}^2$.

Proof of Lemma 4. Choose an enumeration $\sigma(T) = \{\lambda_1, \dots, \lambda_k\}$ of the elements of $\sigma(T)$, such that $\lambda_i \neq \lambda_j$ whenever $1 \leq i < j \leq k$. As T is assumed to be diagonalizable, we have

$$V = \bigoplus_{i=1}^{k} E_{\lambda_i}(T).$$

In what follows, we write E_i for $E_{\lambda_i}(T)$. We want to show that

$$W = \bigoplus_{i=1}^{k} (W \cap E_i).$$

Let $1 \leq i \leq k$, then $(W \cap E_i) \subset E_i$ and $\sum_{j\neq i}(W \cap E_j) \subset \sum_{j\neq i} E_j$ and in particular

$$\{0\} = (W \cap E_i) \bigcap \sum_{j\neq i} (W \cap E_j),$$

implying that indeed

$$\sum_{i=1}^{k} (W \cap E_i) = \bigoplus_{i=1}^{k} (W \cap E_i).$$

We know furthermore that $\bigoplus_{i=1}^{k}(W \cap E_i) \subset W$. It remains to prove the opposite inclusion. Let $w \in W$ and using diagonalizability of T assume that $v_1, \dots, v_k \in V$ satisfy $v_i \in E_i$ and $w = v_1 + \cdots v_k$. Then

$$0 = w + W = (v_1 + \cdots + v_k) + W = (v_1 + W) + \cdots + (v_k + W).$$

Every element in $\{v_i + W \mid 1 \leq i \leq k\}$ distinct from 0 is an eigenvektor of $\overline{T}$ and for any pair of such elements the eigenvalues are distinct. It follows that $v_i + W = 0$ for all $1 \leq i \leq k$, and hence $v_i \in W$ for each $1 \leq i \leq k$. This shows that indeed $w \in \sum_{i=1}^{k}(W \cap E_i)$.

We note that $T|_{W\cap E_\lambda} = \lambda I_{W\cap E_\lambda}$ and hence the restriction of T to $W \cap E_\lambda$ is diagonalizable. It follows that $W \cap E_\lambda$ admits a basis consisting of eigenvectors of $T|_W$. As for each family of bases $\mathcal{B}_\lambda$ of $W \cap E_\lambda$ the union $\mathcal{B} = \bigcup_{\lambda\in\sigma(T)} \mathcal{B}_\lambda$ is a basis of W, we conclude that W admits a basis consisting of eigenvektors of $T|_W$ and hence $T|_W$ is diagonalizable. $\square$

Proof of Theorem 1. Assume first that $\mathcal{T} = \{T_1, \dots, T_r\} \subset \mathrm{End}(V)$ is a finite family of commuting, diagonalizable operators. We prove the theorem by induction on $r = |\mathcal{T}|$. If $r = 1$, then $\mathcal{T}$ is jointly diagonalizable by assumption, i.e. there is nothing to show. Assume that $r \geq 2$ and assume that the theorem holds for families of operators of cardinality up to $r - 1$. The operator T_r is by assumption diagonalizable, i.e. there is a basis $v_1, \dots, v_n$ of V consisting of eigenvectors of T_r. Given $1 \leq i \leq n$ let $\lambda_i \in \mathbb{K}$ denote the corresponding eigenvalue, i.e. $T_r(v_i) = \lambda v_i$. We choose an enumeration $\{\lambda_1, \dots, \lambda_k\}$ of $\sigma(T_r)$. We denote $E_i = E_{\lambda_i}(T_r)$. Note that for every $S \in \mathrm{End}(V)$ which commutes with T_r, the subspaces E_i are S-invariant. Indeed, if $v \in E_i$, then

$$T_r(Sv) = (T_r S)(v) = (ST_r)(v) = S(T_r v) = S(\lambda_i v) = \lambda_i Sv$$

and thus $Sv \in E_i$. Now consider the family $\mathcal{T}' = \{T_1, \dots, T_{r-1}\}$. As $T_jT_k = T_kT_j$ for all $1 \leq j < r$, it follows that $T_j|_{E_i} \in \mathrm{End}(E_i)$. Lemma 4 implies that $T_j|_{E_i}$ is diagonalizable for every $1 \leq j < r$. As $T_1|_{E_i}, \dots, T_{r-1}|_{E_i}$ is a family of $r-1$ commuting, diagonalizable operators, it is jointly diagonalizable and thus there is a basis of E_i consisting of eigenvectors of $T_1, \dots, T_{r-1}$. As each element in E_i by definition is an eigenvector of T_r, we obtain a basis of E_i consisting of eigenvectors of $T_1, \dots, T_r$. As i was arbitrary and $V = \bigoplus_{i=1}^k E_i$, we have found a basis of V consisting of eigenvectors of $T_1, \dots, T_r$.

In the general case (i.e. $\mathcal{T}$ not finite), let $\mathcal{T}' = \langle \mathcal{T} \rangle \subseteq \mathrm{End}(V)$ be the subspace generated by $\mathcal{T}$. By finite dimension of V, it follows that $\mathcal{T}'$ admits a basis $\mathcal{B}_{\mathcal{T}'} \subseteq \mathrm{End}(V)$ of jointly diagonalizable operators. It is a routine check left to the reader that any basis $\mathcal{B} \subseteq V$ of V consisting of eigenvectors of $\mathcal{B}_{\mathcal{T}'}$ in fact consists of eigenvectors of all the elements in $\mathcal{T}'$, and in particular of $\mathcal{T} \subseteq \mathcal{T}'$. $\square$

Email address: `manuel.luethi@math.ethz.ch`