# Math 524: Linear Algebra

## Notes #5 — Eigenvalues, Eigenvectors, and Invariant Subspaces

Peter Blomgren
⟨blomgren@sdsu.edu⟩

Department of Mathematics and Statistics
Dynamical Systems Group
Computational Sciences Research Center
San Diego State University
San Diego, CA 92182-7720

http://terminus.sdsu.edu/

Fall 2021
(Revised: December 7, 2021)



## Outline







## Student Learning Targets, and Objectives

**Target** Invariant Subspaces

- **Objective** Know how the restriction operator and invariant subspaces are connected
- **Objective** Be familiar with the 1-D "line-type" subspaces and their connection with eigenvalues and eigenvectors.
- **Objective** Know, and be able to use, the fact that eigenvectors corresponding to distinct eigenvalues are linearly independent.

**Target** Eigenvalues, Eigenvectors, and Eigenspaces

- **Objective** Know that every operator on a finite-dimensional, nonzero, complex vector space has an eigenvalue and an upper-triangular matrix with respect to some basis
- **Objective** Know the the definitions of Eigenspaces of operators; and understand the discussion of how "collecting" enough eigenvalues can guarantee invertibility of an operator.





## Introduction

We now turn our attention to one of the cornerstones of Linear Algebra, the study of **Operators on finite-dimensional vector spaces**.

**Rewind (Operator, $\mathcal{L}(V)$)**

- A linear map from a vector space to itself is called an **operator**.
- The notation $\mathcal{L}(V)$ denotes the set of all operators on $V$. In other words, $\mathcal{L}(V) = \mathcal{L}(V, V)$.

We will use our newly acquired abstract understanding of vector spaces and linear maps to the study of eigen-values and eigen-vectors.

Time-Target: $3\times$75-minute lectures.

## Invariant Subspaces

Let $T \in \mathcal{L}(V)$. If we have [DIVIDE-AND-CONQUER "THEOREM"]

$$V = U_1 \oplus \cdots \oplus U_m$$

where each $U_j$ is a proper subspace of $V$ (*i.e.* $\dim(U_j) < \dim(V)$); then it is sufficient to understand the action of $T$ on each $U_j$.

**Notation (Restriction, $T|_{U_j}$)**

$T|_{U_j}$ is the **restriction** of the linear map $T \in \mathcal{L}(V)$ to the subspace $U_j$.

This only makes sense if $T|_{U_j} : U_j \mapsto U_j$, or if you want $T|_{U_j} \in \mathcal{L}(U_j)$. Such subspaces get their own name...

**Definition (Invariant Subspace)**

Let $T \in \mathcal{L}(V)$. A subspace $U$ of $V$ is called **invariant** under $T$ if $\forall u \in U \Rightarrow T(u) \in U$.





## Invariant Subspaces

**Example (Invariant Subspaces)**

Suppose $T \in \mathcal{L}(V)$, then the following subspace are invariant under $T$

- $\{0\}$
  - If $u \in \{0\}$, then $u = 0$. [LINEARITY] $T(u) = T(0) = 0 \in \{0\}$.
- $V$
  - If $u \in V$, then — since $T \in \mathcal{L}(V)$ — $T(u) \in V$
- $\text{null}(T)$ — (could be $\{0\}$)
  - If $u \in \text{null}(T)$, then $T(u) = 0 \in \text{null}(T)$ [LINEARITY]
- $\text{range}(T)$ — (could be $V$)
  - If $u \in \text{range}(T)$, then $T(u) \in \text{range}(T)$, by definition of range.





## Invariant Subspaces

**Question**

Must an operator $T \in \mathcal{L}(V)$ have any invariant subspaces other than $\{0\}$ and $V$?

— $\text{null}(T)$ and $\text{range}(T)$ do not necessarily provide useful insight.

We will see that the answer is yes, as long as $\dim(V) > 1$ (for $\mathbb{F} = \mathbb{C}$), or $\dim(V) > 2$ (for $\mathbb{F} = \mathbb{R}$).

**Example**

Let $D \in \mathcal{L}(\mathcal{P}(\mathbb{R}))$ is defined by $Dp = p'$, then for any fixed $m$, $\mathcal{P}_m(\mathbb{R})$ is an invariant subspace of $\mathcal{P}(\mathbb{R})$.
In this case:

$\dim(\text{null}(D)) = 1$, and $\dim(\text{range}(D)) = m = (\dim(\mathcal{P}_m(\mathbb{R})) - 1)$.





## Eigenvalues and Eigenvectors :: Eigenvalues

We will look at invariant subspaces in careful detail; first we turn our attention to the case of invariant subspaces with $\dim = 1$.

Consider the 1-dimensional ("line"-type) subspaces: let $v \neq 0 \in V$, and define $U = \{\lambda v : \lambda \in \mathbb{F}\} \equiv \text{span}(v)$.

**If** $U$ is invariant under $T \in \mathcal{L}(V)$, then $T(v) \in U$ ($\forall v \in U$), and hence $\exists \lambda \in \mathbb{F}$ such that

$$T(v) = \lambda v$$

The converse holds: if $T(v) = \lambda v$ for some $v \in V$, and $\lambda \in \mathbb{F}$, then $\text{span}(v)$ is an invariant subspace of $V$ under the linear map $T$.

## Eigenvalues and Eigenvectors :: Eigenvalues

In the past, we have surely seen eigenvalues (and eigenvectors) defined for *matrices*, here we generalize the concept to operators on all finite-dimensional subspaces...

**Definition (Eigenvalue)**

Suppose $T \in \mathcal{L}(V)$. A scalar $\lambda \in \mathbb{F}$ is called an **eigenvalue** of $T$ if there exists $v \neq 0 \in V$ such that $T(v) = \lambda v$.





## Eigenvalues and Eigenvectors :: Eigenvalues

**Theorem (Equivalent Conditions to be an Eigenvalue)**

*Suppose $V$ is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{F}$. The following are equivalent:*

(a) *$\lambda$ is an eigenvalue of $T$*

(b) *$T - \lambda I$ is not injective*

(c) *$T - \lambda I$ is not surjective*

(d) *$T - \lambda I$ is not invertible*

**Recall:** $I \in \mathcal{L}(V) : I(v) = v, \forall v \in V$.

(a) $\Leftrightarrow$ (b), by rearranging $T(v) = \lambda v$
(b) $\Leftrightarrow$ (c) $\Leftrightarrow$ (d) by [INJECTIVITY $\Leftrightarrow$ SURJECTIVITY IN FINITE DIMENSIONS (NOTES#3.2)]





## Eigenvalues and Eigenvectors :: Eigenvectors

**Definition (Eigenvector)**

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. A vector $v \in V$ is called an **eigenvector** of $T$ corresponding to the eigenvalue $\lambda$ if $v \neq 0$, and $T(v) = \lambda v$.

**Note:** Eigenvalues can be 0, but
Eigenvectors cannot be the zero-vector.
(!) Eigenvectors corresponding to $\lambda = 0$ come from $\mathrm{null}(T)$.

Since $T(v) = \lambda v$ if and only if $(T - \lambda I)v = 0$, $v \neq 0 \in V$ is an eigenvector of $T$ corresponding to the eigenvalue $\lambda$ if and only if $v \in \mathrm{null}(T - \lambda I)$.





## Eigenvalues and Eigenvectors ::

**Example (Rotation over $\mathbb{R}$ and $\mathbb{C}$)**

Suppose $T \in \mathcal{L}(\mathbb{F}^2)$, is defined by $T(x, y) = (-y, x)$.

$\mathbb{F} = \mathbb{R}$ $T$ is a counterclockwise rotation by $\pi/2$ about the origin in $\mathbb{R}^2$. There is no real scaling of a vector such that $(-y, x) = \lambda(x, y)$. $T$ has no eigenvalue(s) and no eigenvector(s).

$\mathbb{F} = \mathbb{C}$ We are looking for $\lambda \in \mathbb{F}$ such that $(-y, x) = \lambda(x, y)$:

$$\begin{cases} \lambda x &= -y \\ \lambda y &= x \end{cases} \quad \Rightarrow \quad -y = \lambda x = \lambda(\lambda y) = \lambda^2 y$$

EIGENVALUES: $\lambda^2 = -1 \Rightarrow \lambda = \pm i$.

$$\text{EIGENVECTORS:} \begin{cases} (-y, x) = +i(x, y) & \Rightarrow \quad (x, y) = (w, -wi) \\ (-y, x) = -i(x, y) & \Rightarrow \quad (x, y) = (w, +wi) \end{cases}$$

## Eigenvectors Corresponding to Distinct Eigenvalues are Linearly Independent

**Theorem (Linearly Independent Eigenvectors)**

*Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of $T$, and $v_1, \dots, v_m$ are the corresponding eigenvectors; then $v_1, \dots, v_m$ is linearly independent.*

**Proof (Linearly Independent Eigenvectors)**

[By Contradiction] Suppose $v_1, \dots, v_m$ is linearly dependent. Let $k$ be the smallest positive integer such that $v_k \in \mathrm{span}(v_1, \dots, v_{k-1})$. We can find $a_1, \dots, a_{k-1} \in \mathbb{F}$ such that

$$\begin{array}{lrcl} (1) & v_k & = & a_1 v_1 + \cdots + a_{k-1} v_{k-1} \\ & T(v_k) & = & T(a_1 v_1 + \cdots + a_{k-1} v_{k-1}) \\ (2) & \lambda_k v_k & = & a_1 \lambda_1 v_1 + \cdots + a_{k-1} \lambda_{k-1} v_{k-1} \\ \hline \lambda_k(1)-(2) & 0 & = & a_1(\lambda_k - \lambda_1) v_1 + \cdots + a_{k-1}(\lambda_k - \lambda_{k-1}) v_{k-1} \end{array}$$

Since $v_1, \dots, v_{k-1}$ is linearly independent, $\lambda_k$ is magically equal to all the distinct $\lambda_1, \dots, \lambda_{k-1}$. **Contradiction!**





## Number of Eigenvalues $\leq \dim(V)$

**Theorem (Number of Eigenvalues)**

*Suppose $V$ is finite-dimensional. Then each operator on $V$ has at most $\dim(V)$ distinct eigenvalues.*

**Proof (Number of Eigenvalues)**

Let $m = \dim(V)$. We can find at most $m$ linearly independent vectors in $V$; eigenvectors corresponding to distinct eigenvalues are linearly independent (by previous theorem); so we can find at most $m$ eigenvectors; thus at most $m$ distinct eigenvalues.





## Restriction and Quotient Operators

If $T \in \mathcal{L}(V)$, and $U$ is a subspace of $V$ invariant under $T$, then $U$ determines two other operators:

**Definition (Restriction Operator $T|_U$; and Quotient Operator $T/U$)**

Suppose $T \in \mathcal{L}(V)$, and $U$ is a subspace of $V$ invariant under $T$

- The **restriction operator** $T|_U \in \mathcal{L}(U)$ is defined by
$$T|_U(u) = T(u), \ \ u \in U$$
- The **quotient operator** $T/U \in \mathcal{L}(V/U)$ is defined by
$$(T/U)(v+U) = T(v) + U, \ \ v \in V$$





## Restriction and Quotient Operators :: Example 1/2

**Example**

Define $T \in \mathcal{L}(\mathbb{F}^2)$ by $T(x,y) = (y,0)$. Let $U = \{(x,0) : x \in \mathbb{F}\}$

- $U$ is invariant under $T$ and $T|_U$ is the 0-operator on $U$:
  - $T(x,0) = (0,0) \in U$.

    So $U$ is invariant under $T$ and $T|_U$ is the 0-operator on $U$.
- $\nexists$ a subspace $W$ of $\mathbb{F}^2$ that is invariant under $T$, and $U \oplus W = \mathbb{F}^2$.
  - Since $\dim(\mathbb{F}^2) = 2$, $\dim(U) = 1$, we must have $\dim(W) = 1$. If $W$ is invariant under $T$, then all $w \in W$ are eigenvectors. However, the only eigenvalue is $\lambda = 0$, and $U$ contains the corresponding eigenvectors.

    Thus $W$ cannot be invariant under $T$.

## Restriction and Quotient Operators :: Example



**Example**

Define $T \in \mathcal{L}(\mathbb{F}^2)$ by $T(x,y) = (y,0)$. Let $U = \{(x,0) : x \in \mathbb{F}\}$

- $T/U$ is the 0-operator on $\mathbb{F}^2/U$:
  - $(x,y) \in \mathbb{F}^2$

$$\begin{aligned} (T/U)((x,y)+U) &= T(x,y)+U \\ &= (y,0)+U \\ &= 0+U \end{aligned}$$

the last equality holds because $(y,0) \in U$.

This example shows that sometimes the restriction and quotient operators do not provide (enough) information about $T$. Here, both are the 0-operators on their respective spaces, even though $T$ is not.





$\langle\langle\langle$ Live Math $\rangle\rangle\rangle$

e.g. 5A-$\{$**12**$\}$





## Live Math :: Covid-19 Version



**5A-12:** Define $T \in \mathcal{L}(\mathcal{P}_4(\mathbb{R}))$ by $(Tp)(x) = xp'(x)\ \forall x \in \mathbb{R}$. Find all eigenvalue and eigenvectors of $T$.

**Solution**

We use the eigenvalue/eigenvector characterization $T(p) = \lambda p$ — here $xp'(x) = \lambda p(x)$. We can write any $p \in \mathcal{P}_4(\mathbb{R})$ in the form $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$, which gives us

$$xp'(x) = a_1x + 2a_2x^2 + 3a_3x^3 + 4a_4x^4 = \lambda(a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4) = \lambda p(x)$$

In order for the equality to hold, the coeffcients for each power must be equal in the left and right expressions. Collecting those relations give us...





## Live Math :: Covid-19 Version



$$\begin{cases} 0a_0 &= \lambda a_0 \\ 1a_1 &= \lambda a_1 \\ 2a_2 &= \lambda a_2 \\ 3a_3 &= \lambda a_3 \\ 4a_4 &= \lambda a_4 \end{cases}$$

For any $j \in \{0, 1, 2, 3, 4\}$: a solution is given by

$$\{\ a_j \neq 0,\ \lambda = j,\ a_{k \neq j} = 0\ \},$$

which allows us to identify 5 eigenvalue-eigenvector pairs:

$$\{(0,1),\ (1,x),\ (2,x^2),\ (3,x^3),\ (4,x^4)\}$$

Technically, any non-zero scaling of the eigenvectors $\{1, x, x^2, x^3, x^4\}$ is also an eigenvector.

## Polynomials Applied to Operators

The main reason that a richer theory exists for operators (which map a vector space into itself) than for more general linear maps is that operators can be raised to powers — composed with themselves / applied multiple times:

**Definition ($T^m$)**

Suppose $T \in \mathcal{L}(V)$ and $m$ is a positive integer

- $T^m$ is defined by $T^m = \underbrace{T \circ \cdots \circ T}_{m \text{ times}}$
- $T^0$ is defined to be the identity operator on $V$
- If $T$ is invertible, with inverse $T^{-1}$, then $T^{-m} = (T^{-1})^m$





## The Operator $p(T)$

**Definition (The Operator $p(T)$)**

Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbb{F})$ is a polynomial given by

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m, \quad z \in \mathbb{F}$$

Then $p(T)$ is the operator defined by

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m$$

**Example ("The Gateway to Differential Equations.")**

Suppose $D \in \mathcal{L}(\mathcal{P}(\mathbb{R}))$ is the differentiation operator defined by $Dq = q'$, with $p$ being the polynomial defined by $p(x) = x^2 + k$, then $p(D) = D^2 + k$, and

$$p(D)q = q'' + kq, \ \forall q \in \mathcal{P}(\mathbb{R})$$

$p(D)q = 0$ is the Helmholtz Equation (in 1D).

https://en.wikipedia.org/wiki/Helmholtz_equation





## Product of Polynomials

**Definition (Product of Polynomials)**

If $p, q \in \mathcal{P}(\mathbb{F})$, then $pq \in \mathcal{P}(\mathbb{F})$ is the polynomial defined by

$$(pq)(z) = p(z)q(z), \quad z \in \mathbb{F}$$

**Theorem (Multiplicative Properties)**

*Suppose $p, q \in \mathcal{P}(\mathbb{F})$ and $T \in \mathcal{L}(V)$, then*

- $(pq)(T) = p(T)q(T)$
- $p(T)q(T) = q(T)p(T)$

The proof is purely "mechanical" (distributive property + bookkeeping)





## Existence of Eigenvalues

**Theorem (Existence of Eigenvalues)**

(!) *Every operator on a finite-dimensional, nonzero, complex vector space has an eigenvalue.* (!)

**Proof (Existence of Eigenvalues)**

Suppose $V$ is a complex vector space with dimension $n > 0$ and $T \in \mathcal{L}(V)$. Let $v \neq 0 \in V$, then

$$v, T(v), T^2(v), \ldots, T^n(v)$$

is not linearly independent, because $V$ has dimension $n$ and we have $(n+1)$ vectors. Thus there exist complex numbers $a_0, a_1, \ldots, a_n$, such that

$$0 = a_0 v + a_1 T(v) + a_2 T^2(v) + \cdots + a_n T^n(v).$$

Not all $a_1, \ldots, a_n$ can be zero, since that would force $a_0 = 0$ (and this would make the $(n+1)$ vectors linearly independent)...

## Existence of Eigenvalues

**Proof (Existence of Eigenvalues)**

Now, let the $a$'s be the coefficients of a polynomial; which by the [FUNDAMENTAL THEOREM OF ALGEBRA] has a factorization

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n = c(z - \lambda_1)\cdots(z - \lambda_m)$$

where $c$ is a nonzero complex number, $\lambda_j \in \mathbb{C}$, and the equation holds $\forall z \in \mathbb{C}$ (here $m$ is not necessarily equal to $n$, because $a_n$ may equal 0). We then have

$$\begin{aligned} 0 &= a_0 v + a_1 T(v) + a_2 T^2(v) + \cdots + a_n T^n(v) \\ &= (a_0 I + a_1 T + a_2 T^2 + \cdots + a_n T^n)(v) \\ &= c(T - \lambda_1 I)\cdots(T - \lambda_m I)v \end{aligned}$$

Thus $(T - \lambda_j I)$ is not injective for at least one $j$. $\Leftrightarrow$ $T$ has an eigenvalue.





## Upper-Triangular Matrices

**Definition (Matrix of an Operator, $\mathcal{M}(T)$)**

Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of $V$. The **matrix of $T$** with respect to this basis is the $(n \times n)$ matrix

$$\mathcal{M}(T) = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix}$$

whose entries $a_{j,k}$ are defined by

$$T(v_k) = a_{1,k} v_1 + \cdots + a_{n,k} v_n$$

If the basis is not "obvious from context," then we use the notation $\mathcal{M}(T, (v_1, \dots, v_n))$.





## Upper-Triangular Matrices :: Comments

Note that matrices of operators are square, rather than the more general rectangular case which we considered earlier for linear maps.

If $T$ is an operator on $\mathbb{F}^n$ and no basis is specified, assume that the basis in question is the standard basis. The $j$th column of $\mathcal{M}(T)$ is then $T$ applied to the $j$th basis vector.

A central ~~goal~~ milestone of linear algebra is to show that given an operator $T \in \mathcal{L}(V)$, there exists a basis of $V$ with respect to which $T$ has a reasonably simple matrix.

For instance, we might try to choose a basis of $V$ such that $\mathcal{M}(T)$ has many 0's.





## Upper-Triangular Matrices

If $V$ is a finite-dimensional complex vector space, there is a basis of $V$ with respect to which the matrix of $T$ looks like

| | $v$ | $w_1$ | $\cdots$ | $w_{n-1}$ |
|---|---|---|---|---|
| $v$ | $\lambda$ | $*$ | $*$ | $*$ |
| $w_1$ | 0 | $*$ | $*$ | $*$ |
| $\vdots$ | $\vdots$ | $*$ | $*$ | $*$ |
| $w_{n-1}$ | 0 | $*$ | $*$ | $*$ |

Let $\lambda$ be an eigenvalue of $T$ (existence is guaranteed); and let $v$ be the corresponding eigenvector. Extend $v$ to a basis of $V$: $v, w_1, \dots, w_{n-1}$ [LINEARLY INDEPENDENT LIST EXTENDS TO A BASIS (NOTES#2)]. Then the matrix of $T$ with respect to this basis has the form given.

## Upper-Triangular Matrices

**Definition (Diagonal of a Matrix)**

The diagonal of a square matrix consists of the entries along the line from the upper left corner to the bottom right corner.

— The $a_{i,i}$-entries.

**Definition (Upper-Triangular Matrix)**

A matrix is called upper triangular if all the entries below the diagonal equal 0.

— $a_{i,j} = 0 \; \forall i > j$. "The strictly lower-triangular part is filled with zeros."





## Conditions for Upper-Triangular Matrix

**Theorem (Conditions for Upper-Triangular Matrix)**

*Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of $V$. Then the following are equivalent:*

(a) *the matrix of $T$ with respect to $v_1, \dots, v_n$ is upper triangular*

(b) $T(v_k) \in \operatorname{span}(v_1, \dots, v_k), \; \forall k$

(c) $U_k = \operatorname{span}(v_1, \dots, v_k)$ *is invariant under $T$* $\forall k$

**Proof (Conditions for Upper-Triangular Matrix)**

(a) $\Leftrightarrow$ (b) follows from the definition, and (c) $\Rightarrow$ (b). The only part that requires work is (b) $\Rightarrow$ (c).

Suppose (b) holds. Fix $k \in \{1, \dots, n\}$ From (b) we know $T(v_i) \in \operatorname{span}(v_1, \dots, v_i) \subset \operatorname{span}(v_1, \dots, v_k)$, $i \in \{1, \dots, k\}$. Thus if $v = a_1 v_1 + \cdots + a_k v_k$, then $T(v) \in \operatorname{span}(v_1, \dots, v_k)$, which shows that $\operatorname{span}(v_1, \dots, v_k)$ is invariant under $T$.





## Over $\mathbb{C}$, Every Operator has an Upper-Triangular Matrix

**Theorem (Over $\mathbb{C}$, Every Operator has an Upper-Triangular Matrix)**

*Suppose $V$ is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then $T$ has an upper-triangular matrix with respect to some basis of $V$.*

**Comment**

The result does **not** hold on real vector spaces, because the first vector in a basis with respect to which an operator has an upper-triangular matrix is an eigenvector of the operator. Thus if an operator on a real vector space has no eigenvalues, then there is no basis with respect to which the operator has an upper-triangular matrix.

We skip the proof... but fear not, Axler provides 2 proofs in the book (pp.149–150).





## Determination of Invertibility from Upper-Triangular Matrix

The following two theorems indicate why we have gone through so much trouble to (isomorphically) link operators on abstract vector spaces to operators on $T \in \mathcal{L}(\mathbb{F}^n)$...

**Theorem (Determination of Invertibility from Upper-Triangular Matrix)**

*Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of $V$. Then $T$ is invertible if and only if all the entries on the diagonal of that upper-triangular matrix are nonzero.*

**Theorem (Determination of Eigenvalues from Upper-Triangular Matrix)**

*Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of $V$. Then the eigenvalues of $T$ are precisely the entries on the diagonal of that upper-triangular matrix.*

Unfortunately, identifying bases which reveal eigenvalues is non-trivial.

⟨⟨⟨ Live Math ⟩⟩⟩

e.g. 5B-$\{\mathbf{4}, 5, 6, 10\}$





## Live Math :: Covid-19 Version 5B-4

**5B-4:** Suppose $P \in \mathcal{L}(V)$ and $P^2 = P$.
Prove that $V = \text{null}(P) \oplus \text{range}(P)$.

⁜ (i) $\text{null}(P) \cap \text{range}(P) = \{0\}$ ⁜

Let $u \in \text{null}(P) \cap \text{range}(P)$. Then $P(u) = 0$, and $\exists w \in W : u = P(w)$. Applying $P$ to $u = P(w)$ gives

$$0 = \mathbf{P(u)} = \mathbf{P^2(w)} = P(w) = u,$$

hence the only vector in $\text{null}(P) \cap \text{range}(P)$ is $u = 0$.





## Live Math :: Covid-19 Version 5B-4

⁜ (ii) $V = \text{null}(P) + \text{range}(P)$ ⁜

Next, let $v \in V$, then

$$v = v + 0 = v + (P(v) - P(v)) = (v - P(v)) + P(v),$$

where

$$P(v-P(v)) = P(v)-P^2(v) = P(v)-P(v) = 0 \;\Rightarrow\; (v-P(v)) \in \text{null}(P),$$

and, by definition

$$P(v) \in \text{range}(P).$$

Thus

$$v = u + w, \quad u \in \text{null}(P), \; w \in \text{range}(P)$$

Since $v \in V$ was arbitrary, $V = \text{null}(P) + \text{range}(P)$.

⁜ (i) + (ii) $\Rightarrow V = \text{null}(P) \oplus \text{range}(P)$ ⁜

[Direct Sum of Two Subspaces (Notes#1)]





## Eigenspaces and Diagonal Matrices

**Definition (Diagonal Matrix)**

A diagonal matrix is a square matrix that is 0 everywhere except possibly along the diagonal.

**Note:** zeros ARE allowed on the diagonal.

If an operator has a diagonal matrix with respect to some basis, then the entries along the diagonal are precisely the eigenvalues of the operator.

**Definition (Eigenspace, $E(\lambda, T)$)**

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. The **Eigenspace** of $T$ corresponding to $\lambda$ denoted $E(\lambda, T)$ is defined to be

$$E(\lambda, T) = \text{null}(T - \lambda I)$$

$E(\lambda, T)$ is the set of all eigenvectors of $T$ corresponding to $\lambda$, along with the 0 vector.

## The Operator Restricted to an Eigenspace

**The Operator Restricted to an Eigenspace:**
If $\lambda$ is an eigenvalue of $T \in \mathcal{L}(V)$, then

$$T|_{E(\lambda,T)}(v) = \lambda v, \ \forall v \in E(\lambda, T)$$

this indicates that eigenspaces are (non-trivial, and highly useful) **invariant subspaces**; and we get a very simple description (scalar multiplication) of the operator when restricted to such a subspace.

*The is where our notation and language is starting to pay dividends!*





## Sum of Eigenspaces is a Direct Sum

**Theorem (Sum of Eigenspaces is a Direct Sum)**

*Suppose $V$ is finite-dimensional and $T \in \mathcal{L}(V)$. Suppose also that $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of $T$. Then*

$$E(\lambda_1, T) + \cdots + E(\lambda_m, T)$$

*is a direct sum:*

$$E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$$

*Furthermore,*

$$\dim\left(E(\lambda_1, T)\right) + \cdots + \dim\left(E(\lambda_m, T)\right) \leq \dim(V)$$

Eigenvalues and eigenspaces give us excellent understanding of operator behavior, so this is worth showing... However, we need a help-result...





## Filling in a Gap... Dimension of a Direct Sum of Finite Dimensional Subspaces

**Theorem (Dimension of a Direct Sum of Finite Dimensional Subspaces)**

*Suppose $U_1, \ldots, U_m$ are finite-dimensional subspaces of $V$ such that $U_1 + \cdots + U_m$ is a direct sum. Then $U_1 \oplus \cdots \oplus U_m$ is finite-dimensional and*

$$\dim\left(U_1 \oplus \cdots \oplus U_m\right) = \dim\left(U_1\right) + \cdots + \dim\left(U_m\right)$$

**Proof (Dimension of a Direct Sum of Finite Dimensional Subspaces)**

Let $W, U_1, \ldots, U_m$ be subspaces of $V$ such that

$$W = U_1 \oplus \cdots \oplus U_m$$

Choose a (finite) basis $\{u_{j,k}\}$ for each $U_j$. Concatenate the bases into a single list $\{w_\ell\}$. By construction the (finite) list $\{w_\ell\}$ spans $W = U_1 + \cdots + U_m$. Thus $W$ is finite dimensional... $\rightsquigarrow$





## Filling in a Gap... Dimension of a Direct Sum of Finite Dimensional Subspaces

**Proof (Dimension of a Direct Sum of Finite Dimensional Subspaces)**

(!) We need to show that the vectors in $\{w_\ell\}$ are linearly independent (and thus a basis for $W$), so that the dimension of $W$ equals the number of vectors in $\{w_\ell\}$, and therefore

$$\dim(W) = \dim\left(U_1\right) + \cdots + \dim\left(U_m\right).$$

**Assume:** $a_1 w_1 + \cdots + a_N w_N = 0$; we can group the sum in $m$ groups (depending on what space $U_j$ was the original source of the basis vector); and thus write this sum as $\hat{u}_1 + \cdots + \hat{u}_m = 0$, where each $\hat{u}_j \in U_j$.

Since $W = U_1 \oplus \cdots \oplus U_m$, this forces $0 \equiv \hat{u}_j \in U_j \ \forall j$; but since each such vector is formed by a linear combination of the basis vectors $\{u_{j,k}\}$, all the coefficients in each of those linear combinations must be 0; translated back to $a_1 w_1 + \cdots + a_N w_N = 0$, $a_i \equiv 0$, which makes $\{w_\ell\}$ linearly independent. $\surd$

## Sum of Eigenspaces is a Direct Sum

With the help-result in hand, we can show the theorem:

**Proof (Sum of Eigenspaces is a Direct Sum)**

To show that $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ is a direct sum, let

$$u_1 + \cdots + u_m = 0$$

where $u_j \in E(\lambda_j, T)$. Since the eigenvectors corresponding to distinct eigenvalues are linearly independent [Linearly Independent Eigenvectors], we get $u_j \equiv 0$. Now, using [Condition for a Direct Sum (Notes#1)] this implies that $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ is a direct sum. Using the [help-result], we now get

$$\begin{aligned} \dim\left(E(\lambda_1, T) + \cdots + E(\lambda_m, T)\right) &= \\ \dim\left(E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)\right) &= \\ \dim\left(E(\lambda_1, T)\right) + \cdots + \dim\left(E(\lambda_m, T)\right) &\leq \dim(V) \end{aligned}$$





## Diagonalizable Operators

**Definition (Diagonalizable)**

An operator $T \in \mathcal{L}(V)$ is called **diagonalizable** if the operator has a diagonal matrix with respect to some basis of $V$.

**Theorem (Conditions Equivalent to Diagonalizability)**

*Suppose $V$ is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of $T$. Then the following are equivalent:*

(a) *$T$ is diagonalizable.*

(b) *$V$ has a basis consisting of eigenvectors of $T$* "Eigenbasis"

(c) *$\exists$ 1-D subspaces $U_1, \ldots, U_n$ of $V$, each invariant under $T$, such that $V = U_1 \oplus \cdots \oplus U_n$*

(d) $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$

(e) $\dim(V) = \dim(E(\lambda_1, T)) + \cdots + \dim(E(\lambda_m, T))$





## Conditions Equivalent to Diagonalizability

**Proof (Conditions Equivalent to Diagonalizability)**

(a)⇔(b) $T \in \mathcal{L}(V)$ has a diagonal matrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ if and only if $T(v_k) = \lambda v_k$ for each $k$. $\surd$

(b)⇒(c) Suppose (b) holds; *i.e.* $V$ has a basis $v_1, \ldots, v_n$ consisting of eigenvectors of $T$. For each $k$, let $U_k = \mathrm{span}\,(v_k)$. By construction each $U_k$ is a 1-D subspace of $V$ that is invariant under $T$. Because $v_1, \ldots, v_n$ is a basis of $V$, each vector in $V$ can be written uniquely as a linear combination of $v_1, \ldots, v_n$. That is, each vector in $V$ can be written uniquely as a sum $u_1 + \cdots + u_n$, where $u_k \in U_k$. Thus $V = U_1 \oplus \cdots \oplus U_n$. $\surd$





## Conditions Equivalent to Diagonalizability

**Proof (Conditions Equivalent to Diagonalizability)**

(c)⇒(b) Suppose (c) holds; thus there are 1-dimensional subspaces $U_1, \ldots, U_n$ of $V$, each invariant under $T$, such that $V = U_1 \oplus \cdots \oplus U_n$. $\forall k$, let $v_k \neq 0 \in U_k$. Then each $v_k$ is an eigenvector of $T$. Because each vector in $V$ can be written uniquely as a sum $u_1 + \cdots + u_n$, where $u_k = \alpha_k v_k \in U_k$, we see that $v_1, \ldots, v_n$ is a basis of $V$. $\surd$

(b)⇒(d) Suppose (b) holds; thus $V$ has a basis consisting of eigenvectors of $T$. Hence every vector in $V$ is a linear combination of eigenvectors of $T$, which implies that

$$V = E(\lambda_1, T) + \cdots + E(\lambda_m, T)$$

Now [Sum of eigenspaces is a direct sum] shows that (d) holds:

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T). \qquad \surd$$

## Conditions Equivalent to Diagonalizability

**Proof (Conditions Equivalent to Diagonalizability)**

(d)⇒(e) [HELP-RESULT] √

(e)⇒(b) Suppose (e) holds, *i.e.* $\dim(V) = \dim(E(\lambda_1, T)) + \cdots + \dim(E(\lambda_m, T))$. Select a basis for each $E(\lambda_j, T)$; concatenate the basis into a list $v_1, \ldots, v_n$ of eigenvectors of $T$ ($n = \dim(V)$, by (e)). For linear independence, suppose $a_1 v_1 + \cdots + a_n v_n = 0$; let $u_j$ denote the sum of the group of vectors from $E(\lambda_j, T)$; and we get $u_1 + \cdots + u_m = 0$. Now, [LINEARLY INDEPENDENT EIGENVECTORS] forces $u_j = 0$, which in turn forces $a_i = 0$, which makes $v_1, \ldots, v_n$ linearly independent, and a basis for $V$ by [LINEARLY INDEPENDENT LIST OF THE RIGHT LENGTH IS A BASIS (NOTES#2)]. √

We now have
(a) ⇔ (b) ⇔ (c)
(b) ⇒ (d) ⇒ (e) ⇒ (b)

which means (a) ⇔ (b) ⇔ (c) ⇔ (d) ⇔ (e) √





## Diagonalizability is not Guaranteed

Unfortunately not every operator is diagonalizable. This can happen even on complex vector spaces, as was shown in one of our previous examples:

**Rewind**

The $T \in \mathcal{L}(\mathbb{F}^2)$ defined by $T(x, y) = (y, 0)$ has a single eigenvalue $\lambda = 0$, and $E(\lambda = 0, T) = \{(x, 0) : x \in \mathbb{F}\}$.
Since $\dim(\mathbb{F}^2) = 2$, and $\dim(E(0, T)) = 1$; we're out of luck

At some point (soon) we have to "do something" about non-diagonalizable operators.





## Enough Eigenvalues Implies Diagonalizability

**Theorem (Enough Eigenvalues Implies Diagonalizability)**

*If $T \in \mathcal{L}(V)$ has $n = \dim(V)$ distinct eigenvalues, then $T$ is diagonalizable.*

**Proof (Enough Eigenvalues Implies Diagonalizability)**

Let $T \in \mathcal{L}(V)$ have $n = \dim(V)$ distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. $\forall k$, let $v_k \in V$ be an eigenvector corresponding to $\lambda_k$. By [LINEARLY INDEPENDENT EIGENVECTORS] $v_1, \ldots, v_n$ is linearly independent, and by [LINEARLY INDEPENDENT LIST OF THE RIGHT LENGTH IS A BASIS (NOTES#2)] therefore a basis. With respect to this basis consisting of eigenvectors, $T$ has a diagonal matrix.

**Note:** this is a one-way result ⇒, not an if-and-only-if ⇔.





⟨⟨⟨ Live Math ⟩⟩⟩

e.g. 5C-{**8**}

## Live Math :: Covid-19 Version — 5C-8

**5C-8:** Suppose $T \in \mathcal{L}(\mathbb{F}^5)$ and $\dim(E(8,T)) = 4$. Prove that $(T-2I)$ or $(T-6I)$ is invertible.

**Solution**

We remind ourselves that

- $E(\lambda, T) = \text{null}(T - \lambda I)$ are the eigenspaces, and
- the sum of eigenspaces is a direct sum

Therefore

$$\underbrace{\dim(E(8,T))}_{4} + \underbrace{\dim(E(2,T))}_{\in Z^+} + \underbrace{\dim(E(6,T))}_{\in Z^+} \leq \dim(\mathbb{F}^5) = 5$$

which means that at least one (possibly both) of $\dim(E(2,T))$ and $\dim(E(6,T))$ must be zero. $\rightsquigarrow \exists \widehat{\lambda} \in \{2,6\}$ so that $E(\widehat{\lambda}, T) = \{0\}$, making $\widehat{\lambda}$ not an eigenvalue, and $(T - \widehat{\lambda} I)$ invertible.





## Suggested Problems

**5.A**—1, 2, 3, 4, 8, 9, 10, 12

**5.B**—1(a), 4‡, 5*, 7, 10, 14, 15

**5.C**—1, 2, 8

‡ This problem has "something" to do with Orthogonal Projections. (Familiar on $\mathbb{R}^n$, but here expressed on general vector spaces); also a matrix for which $P^2 = P$ holds is called *idempotent*. This problem appears in slightly different form in [Math 543]

* This problem relates to the eigenvalue structure, and the reason we (in Math 254) look for diagonalizing similarity transformations





## Assigned Homework — HW#5, Due Date in Canvas/Gradescope

**5.A**—4, 8, 10(a)

**5.B**—1(a), 7, 14, 15

**5.C**—1, 2

**Note:** Assignment problems are not official and subject to change until the first lecture on the chapter has been delivered (or virtually "scheduled.")

Upload homework to www.Gradescope.com





## Eigenvectors Corresponding to Distinct Eigenvalues are Linearly Independent

**Proof ((Alternative) Linearly Independent Eigenvectors)**

Let $\in \mathcal{L}(V)$. Suppose $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of $T$, and $v_1, \ldots, v_m$ are the corresponding eigenvectors. Consider $c_1 v_1 + \cdots + c_m v_m = 0$, where $c_j \in \mathbb{F}$. Let $k$ be the largest index so that $v_1, \ldots, v_k$ is linearly independent, but $v_1, \ldots, v_{k+1}$ is not and let (*) $c_1 v_1 + \cdots + c_k v_k = v_{k+1}$. Now:

$$\begin{aligned} \lambda_{k+1} v_{k+1} = T(v_{k+1}) &= T(c_1 v_1 + \cdots + c_k v_k) = c_1 T(v_1) + \cdots + c_k T(v_k) \\ &= c_1 \lambda_1 v_1 + \cdots + c_k \lambda_k v_k \quad (**) \end{aligned}$$

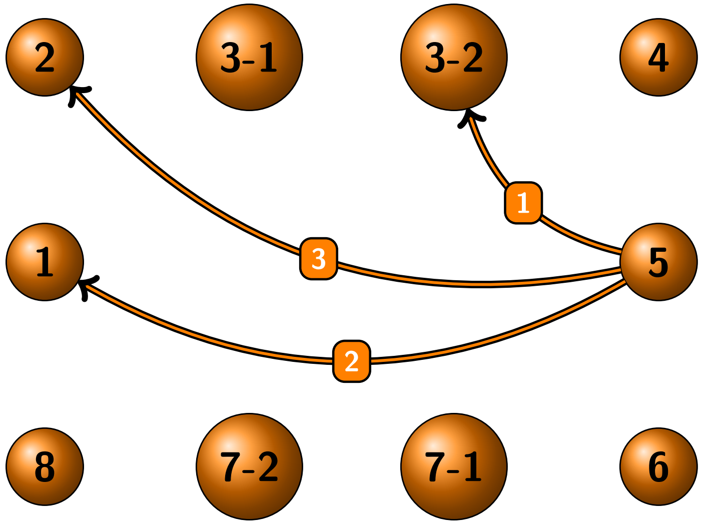
Multiplying (*) by $\lambda_{k+1}$ and subtract from (**):

$$c_1 \underbrace{(\lambda_1 - \lambda_{k+1})}_{\neq 0} v_1 + \cdots + c_k \underbrace{(\lambda_k - \lambda_{k+1})}_{\neq 0} v_k = 0$$

Since $v_1, \ldots, v_k$ are linearly independent, $c_1 = \cdots = c_k = 0$, which makes $v_{k+1} = 0$, but eigenvectors cannot be the zero vector; hence there is no such value $k$, and the vectors are linearly independent.

(It's the same proof, dressed up in a Halloween costume!)

# Explicit References to Previous Theorems or Definitions (with count)








# Explicit References to Previous Theorems or Definitions