Linear Algebra II

Course Notes

Lance Li

Published by Lance Li

www.math.toronto.edu/nhoell/MAT224

First edition, March 2022

Contents

I Notes

Notes

0. Introduction

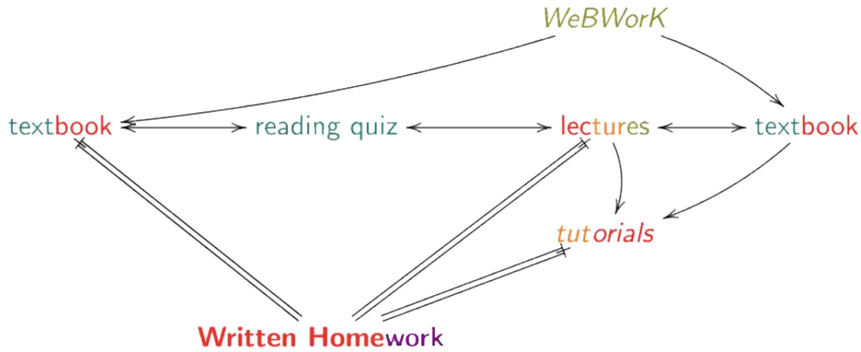
0.1 Introduction

Fields, complex numbers, vector spaces over a field, linear transformations, matrix of a linear transformation, kernel, range, dimension theorem, isomorphisms, change of basis, eigenvalues, eigenvectors, diagonalizability, real and complex inner products, spectral theorem, adjoint/self-adjoint/normal linear operators, triangular form, nilpotent mappings, Jordan canonical form.

0.1.1 Learning Outcome

- Read and understand new mathematical ideas and concepts on your own
- Communicate mathematical ideas of linear algebra clearly in words and in writing (proofs)
- Connect abstract knowledge to examples
- Approach challenging problems independently
- Use linear algebra as a computational tool

0.1.2 Course Structure



Our course has multiple components. Each component is built carefully to support several aspects of the learning outcomes of our course as listed in the syllabus. This page explains how different components of the course interact with one another and helps you navigate the course

Reading Quizzes

A typical week in our course starts by doing the reading assigned for that week and ends by revisiting the reading assignment for that week. Assigned readings are mainly from our primary textbook [1] followed by a Quercus quiz on your assigned readings for the week and that of the previous week. You should start your reading as soon as they are posted.

You should read all the assigned readings for the week and submit your quiz on Quercus. The weekly quiz is due at 10 am every Monday. The best strategy is to spread out the reading throughout the week. You are not expected to understand everything when you read the textbook before the class. Your learning will happen gradually during lectures, tutorials, and mostly when you do the suggested problems from the textbook and the written homework. You will need to review the textbook again after lectures. Reading Quizzes quizzes makeup 5% of your final grade.

Lectures

What to expect from lectures?

During lectures, your instructor will guide you through important concepts and invite you to participate in lecture activities. Lectures complement your reading, tutorials, and homework. Lectures are not meant to replace textbook reading. There will be concepts and theorems that are only discussed in the textbook or only during your lectures, or only in homework. You are encouraged to attend all lectures. While in your lecture you are expected to actively participate in all activities prompted by your instructor. While the delivery of the course is online, your instructor may record the lectures and make it available for you to watch asynchronously. Note that watching the lectures do not replace attending them as you will miss the in-class activities. You are asked to do a lecture reflection quiz at the end of your lecture week. The reflection due date is a day after your last lecture of the week. Reflections are a tool for your instructor to emphasize important concepts in the lectures and to collect feedback from you, and an opportunity for you to reflect on your understanding of the lecture material. Reflections are part of your participation grade as per our syllabus.

Tutorials

You must be enrolled in a tutorial. You will have weekly tutorial sessions starting on Jan 24th. During tutorials, you will work with your classmates in small groups on tutorial worksheets. Tutorial worksheets are carefully designed to be discussed in groups. Tutorial worksheets are roughly one week behind the lectures.

Ideally, you want to read, understand, and think about all the questions before your tutorial. This involves checking all the definitions and statements of the results you need to know to tackle questions. Set aside 15 minutes before your tutorial to go over questions. Your tutorial sessions are facilitated by your TAs. During your tutorial, your TA will ask you to sit with your tutorial group. You will work with your groupmates on selected problems from the worksheet. Each group works on a shared document that will be checked by the TA during the tutorial. Do not expect your TA to work out the questions for you or to teach the concepts. Your TA will give you feedback on what you already put down in your shared document. That is the more you work among your group the more your TA can help you. Your TA might choose to go over some of the questions for the entire class or answer your questions within your groups. You get the most out of your tutorials if you work on the problems ahead of time and stay active and engaged throughout the tutorial. After all, you can only get an answer to those questions you ask, either from your peers or the TA. You will get solutions to all tutorial questions at the end of the tutorial week. Tutorials are one of the most important components of the course because they facilitate your communication with your peers. Explain concepts to your group mates and ask them to do the same for you. At the end of each tutorial, you will submit your shared document as a group.

This submission is marked holistically, and not for correctness. You and all your group members will receive the same mark. Your active participation in your tutorial is measured by your group submission marks. Your active participation in tutorials is part of your participation grade.

Homework

You will have four types of homework. We already talked about two of them. That is reading assignments, due weekly, as part of your Reading Quiz, and Tutorial worksheets that you work on with help of your peers during your tutorial and submit with your group. The other two are

WeBWork. WeBWork is an online assignment system. You have weekly WeBWork due every Wednesday 11:59 pm. These questions are straightforward and cover what you learned in the same week and the past week. To access the homework, you should follow the link in the assignment posted under the Homework module. You will be automatically be directed to WeBWork environment. WARNING: You should check your grade for WeBWork inside the WebWork environment and not in Quercus. You may occasionally see some grades regarding this homework appear on Quercus. Those numbers are not accurate and will disappear! WeBWork makes 5% of your final grade.

Written Homework. There will be five written homework sets. These homework sets are rather long and you have about two weeks to do them. You should submit these homework sets individually. You will receive instructions on how to submit your written homework. Only selected questions from each set are graded. The lowest grade will be dropped. Written homework sets make 10% of your final grade.

0.1.3 Lecture Rules

- Come prepared (read the textbook)
- Be fully present
  - No distraction
  - Ready to engage
  - Ready to participate in activities
- Reflect
  - What did you learn?
  - Engage with the textbook
  - Write down your questions
  - Follow up on piazza

1. Fields, Complex Numbers and Vector Spaces

1.1 Fields and Complex Numbers

1.1.1 Fields

Definition 5.1.4 — Field. A *field* is a set[a] $\mathbb{F}$ with two operations, defined on ordered pairs of elements of $\mathbb{F}$, called *addition* and *multiplication*. Addition assigns to the pair x and $y \in \mathbb{F}$ their *sum*, which is denoted by $x + y$ and multiplication assigns to the pair x and y their *product*, which is denoted by $x \cdot y$ or xy. These two operations must satisfy the following properties for all x, y and $z \in \mathbb{F}$:

(i) Commutativity of addition $x + y = y + z$.
(ii) Associativity of addition: $(x + y) + z = x + (y + z)$.
(iii) Existence of an additive identity: There is an element $0 \in \mathbb{F}$, called zero, such that $x + 0 = a$.
(iv) Existence of additive inverses: For each x there is an element $-x \in \mathbb{F}$ such that $x + (-x) = 0$.
(v) Commutativity of multiplication: $xy = yx$.
(vi) Associativity of multiplication: $(xy)z = x(yz)$.
(vii) Distributivity: $(x + y)z = xz + yz$ and $x(y + z) = xy + xz$.
(viii) Existence of a multiplicative identity: There is an element $1 \in \mathbb{F}$, called 1, such that $x \cdot 1 = x$.
(ix) Existence of multiplicative inverses: If $x \neq 0$, then there is an element $x^{-1} \in \mathbb{F}$ such that $x \cdot x^{-1} = 1$.

[a] Note that such set must be **non-empty**.

1.1.2 Complex Numbers

Goal: build a field containing $\mathbb{R}$ such that all polynomials (such as $x^2 + 1$) have their roots.

Definition 5.1.2 The set of *complex numbers*, denoted $\mathbb{C}$, is the set of ordered pairs of real numbers(a, b) with the operations of addition and multiplication defined by

For all (a,b) and $(c,d) \in \mathbb{C}$, the *sum* of (a,b) and (c,d) is the complex number defined by $(a,b)+(c,d) = (a+c, b+d)$

and the *product* of (a,b) and (c,d) is the complex number defined by $(a,b)(c,d) = (ac-bd, ad+cb)$

The subset of $\mathbb{C}$ consisting of those elements with second coordinate zero, $\{(a,0)|a \in \mathbb{R}\}$, will be identified with the real numbers in the obvious way, $a \in \mathbb{R}$ is identified with $(a,0) \in \mathbb{C}$. If we apply our rules of addition and multiplication to the subset $\mathbb{R} \subset \mathbb{C}$, we obtain

$$(a,0)+(c,0) = (a+c,0)$$

and

$$(a,0)(c,0) = (ac - 0 \cdot 0)(a \cdot 0 + c \cdot 0) = (ac,0)$$

Proposition 5.1.5 The set of complex numbers is a field with the operations of addition and scalar multiplication as defined previously.

Proof. WTS $\mathbb{C}$ is a field[1].
(i), (ii), (v), (vi), and (vii) follow immediately.

(iii) The additive identity is $0 = 0 + 0i$ since
$(0+0i)+(a+bi) = (0+a)+(0+b)i = a+bi$

(iv) The additive inverse of $a+bi$ is $(-a)+(-b)i$.
$(a+bi)+((-a)+(-b)i) = (a+(-a))+(b+(-b))i = 0+0i = 0.$

(viii) The multiplicative identity is $1 = 1 + 0 \cdot i$ since
$(1+0 \cdot i)(a+bi) = (1 \cdot a - 0 \cdot b) + (1 \cdot b + 0 \cdot a)i = a+bi.$

(ix) Note first that if $a+bi \neq 0$, then either $a \neq 0$ or $b \neq 0$ and $a^2+b^2 \neq 0$. Further, note that $(a+bi)(a+(-b)i) = a^2+b^2$.

Therefore $(a+bi)\dfrac{a-ib}{a^2+b^2} = 1.$

Thus, $(a+ib)^{-1} = (a-ib)/(a^2+b^2)$. ■

Exercise 1.1 Compute the following.

1. $(3-5i)^{-1}$

$$\begin{aligned}(3-5i)^{-1} &= \frac{(3+5i)}{3^2+5^2}\\ &= \frac{3}{34} + \frac{5}{34}i\end{aligned}$$

[1]To show $\mathbb{F}$ is a field, we need to check commutativity, associativity, existence of additive identity and additive inverse, multiplicative identity and multiplicative inverse, and the distributivity between addition and multiplication.

2. $\dfrac{4-2i}{3-5i} = (4-2i)(3-5i)^{-1}$

$$\begin{aligned}\frac{4-2i}{3-5i} &= (4-2i)(3-5i)^{-1}\\ &= (4-2i)\left(\frac{1}{34}(3+5i)\right)\\ &= \frac{1}{34}(4-2i)(3+5i)\\ &= \frac{1}{34}(12+10+20i-6i)\\ &= \frac{1}{34}(22+14i)\end{aligned}$$

■

Example 1.1 Let $\mathbb{F}_2 = \{0, 1\}$.

$+$	0	0
0	0	1
1	1	0

$\times$	0	0
0	0	0
1	0	1

Discussion
- Is $+$ commutative? How to see it visually?
  - Yes.
  - The diagonals align up (this tabel is symmetric).
- Is $\times$ commutatiove? How to see it visually?
  - Yes.
  - The diagonals align up (this tabel is symmetric).
- Does $(\mathbb{F}_2, +, \times)$ have additive and multiplicative identities? How to see them visually?
  - Yes, they all have their identities.
  - 0 is the additive identity.
    The corresponding rows and columns of 0 are copies of the index row/column.
  - 1 is the multiplicative identity.
    The corresponding rows and columns of 1 are copies of the index row/column.

■

1.3 Vector Spaces

1.3.1 Vector Space

Let $\mathbb{F}$ be a field[2].

Definition 5.2.1 — Vector Space over a Field. A *vector space over $\mathbb{F}$* is a set V (whose elements are called *vectors*) together with two operations:
- A binary operation called **vector addition**, which for each pair of vectors $\vec{v}, \vec{w} \in V$ produces a vector denoted $\vec{v} + \vec{w} \in V$, and
- an operation called **multiplication by a scalar**[a] (a field element), which for each

[2]The differences between a field and a vector space:
- Over *fields*, we have two binary operations;
- Over *vector spaces*, we have one binary operation and one scalar multiplication.

vector $\vec{v} \in V$, and each scalar $c \in \mathbb{F}$ produces a vector denoted $c\vec{v} \in V$.

[a] This is also called a **scalar multiplication**

$$+: \begin{array}{ccc} V \times V & \to & V \\ (\vec{v}, \vec{w}) & \mapsto & \vec{v} + \vec{w} \end{array} \qquad \times: \begin{array}{ccc} \mathbb{F} \times V & \to & V \\ (c, \vec{v}) & \mapsto & c\vec{v} \end{array}$$

Furthermore, the two operations must satisfy the following axioms:

(1) For all vectors $\vec{u}$, $\vec{v}$, and $\vec{w} \in V$, $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$. (addition is *associative*)
(2) For all vectors $\vec{v}$ and $\vec{w} \in V$, $\vec{v} + \vec{w} = \vec{w} + \vec{v}$. (addition is *commutative*)
(3) There exists a vector $\vec{0} \in V$ with the property that $\vec{v} + \vec{0} = \vec{v}$ for all vectors $\vec{v} \in V$. ($\exists$ an *additive identity*)
(4) For each vector $\vec{v} \in V$, there exists a vector denoted $-\vec{v}$ with the property that $\vec{v} + -\vec{v} = \vec{0}$. ($\exists$ *additive inverse*)
(5) For all vectors $\vec{v}$ and $\vec{w} \in V$ and all scalars $c \in \mathbb{F}$, $c(\vec{v} + \vec{w}) = c\vec{v} + c\vec{w}$. (*distributive* property 1)
(6) For all vectors $\vec{v} \in V$, and all scalars c and $d \in \mathbb{F}$, $(c + d)\vec{v} = c\vec{v} + d\vec{v}$. (*distributive* property 2)
(7) For all vectors $\vec{v} \in V$, and all scalars c and $d \in \mathbb{F}$, $(cd)\vec{v} = c(d\vec{v})$. (multiplication is *associative*)
(8) For all vectors $\vec{v} \in V$, $1\vec{v} = \vec{v}$. ($\exists$ an *multiplicative identity*)

■ Example 1.2

Define $\mathbb{R}^n = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, a_i \in \mathbb{R} \right\}$, $\mathbb{C}^n = \left\{ \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}, a_i \in \mathbb{C} \right\}$.

Take $\vec{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$, $\vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$, then $\vec{z} + \vec{w} = \begin{pmatrix} z_1 + w_1 \\ \vdots \\ z_n + w_n \end{pmatrix}$[3].

In general, $\mathbb{F}^n := \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, a_i \in \mathbb{F} \right\}$ is a vector space over the field $\mathbb{F}$ [4].

- Additive identity: $\vec{0} := \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$[5]
- Take $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{F}^n$, then $-\vec{v} = \begin{pmatrix} -v_1 \\ \vdots \\ -v_n \end{pmatrix}$[6].

■

■ Example 1.3

Define $\mathcal{P}_n(\mathbb{F}) = \{a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \mid a_i \in \mathbb{F}\}$.

- $\mathcal{P}_n(\mathbb{F})$ is an $\mathbb{F}$_V.S.
  Take $p(x) = a_0 + a_1 x + \cdots + a_n x^n$, $q(x) = b_0 + b_1 x + \cdots + b_n x^n \in P_n(\mathbb{F})$.
  $p(x) + q(x) := (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n b_n)x^n$.

[3] Note that the two $+$'s ($\vec{z} + \vec{w}$ and $z_i + w_i$) here are different. The former is vector addition, while the later is addition over $\mathbb{C}$.

[4] "A vector space over the field $\mathbb{F}$" can be shortened to "$\mathbb{F}$ vector space", or "$\mathbb{F}$_V.S.".

[5] Note that $\vec{0}$ is the additive identity of $\mathbb{F}^n$, while the 0's are the additive identity over the field $\mathbb{F}$.

[6] Note that the two $-$'s ($-\vec{v}$ and $-v_i$) here are different. The former is the symbol for the additive inverse inside $\mathbb{F}^n$, while the later is additive inverse over the field $\mathbb{F}$.

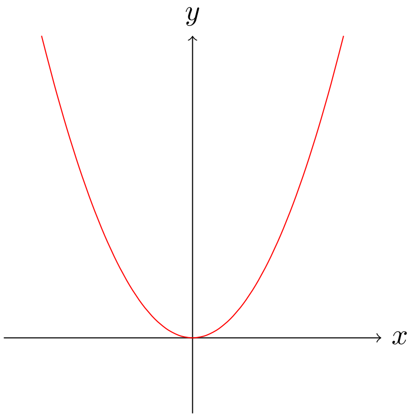
$$cp(x) := ca_0 + \overbrace{ca_1}^{c\,\times\,a_1 \text{ on } \mathbb{F}} x + \cdots + ca_n x^n.$$

■

■ Example 1.4

$F(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ is a function}\}$ is a real vector space.
Consider $f(x) = x^2 \in F(\mathbb{R})$.

$$\begin{aligned} f :\ \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^2 \end{aligned}$$



- Define vector addition and scalar multiplication
  Take $f, g \in F(\mathbb{R})$.
  - $(f+g)(x) := f(x) + g(x)$ for all $x \in \mathbb{R}$.
  - $(rf)(x) = r\,(f(x))$ for all $x \in \mathbb{R}$ and all $x \in \mathbb{R}$.
- Additive identity of $F(\mathbb{R})$:

$$\begin{aligned} \underset{F(\mathbb{R})}{0}\quad \mathbb{R} &\to \mathbb{R} \\ x &\mapsto 0_{\mathbb{R}} \end{aligned}$$

  - Prove that $\underset{F(\mathbb{R})}{0}$ is the additive identity of $F(\mathbb{R})$.

    Proof. WTS $\forall f \in F(\mathbb{R})$, $f + \underset{F(\mathbb{R})}{0} = f$.
    Pick arbitrary $f \in F(\mathbb{R})$, $f : \mathbb{R} \to \mathbb{R}$.
    WTS $\forall x \in \mathbb{R}$, $(f + \underset{F(\mathbb{R})}{0})(x) = f(x)$.

$$\begin{aligned} \text{Pick } x \in \mathbb{R}, (f + \underset{F(\mathbb{R})}{0})(x) &= f(x) + \underset{F(\mathbb{R})}{0}(x) && \text{by the definition of } + \text{ in } F(\mathbb{R}) \\ &= f(x) + 0_{\mathbb{R}} && \text{by the definition of } \underset{F(\mathbb{R})}{0} \\ &= f(x) && \text{since } 0_{\mathbb{R}} \text{ is the additive identity in } \mathbb{R} \end{aligned}$$

    We have shown that $\forall x \in \mathbb{R}$, $(f + \underset{F(\mathbb{R})}{0})(x) = f(x)$. ■

  - Prove that $\forall f \in F(\mathbb{R})$, $\exists -f \in F(\mathbb{R})$ s.t. $f + (-f) = \underset{F(\mathbb{R})}{0}$.

    Proof. WTS $\forall f \in F(\mathbb{R})$, $\exists -f \in F(\mathbb{R})$ s.t. $f + (-f) = \underset{F(\mathbb{R})}{0}$.
    Pick arbitrary $f \in F(\mathbb{R})$.

Let $-f$ be the function

$$\begin{array}{rccc} -f: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & -f(x) \end{array}$$

WTS $f+(-f)=\underset{F(\mathbb{R})}{0}$.
WTS $\forall x\in\mathbb{R}$, $(f+(-f))(x)=\underset{F(\mathbb{R})}{0}(x)$.

Pick arbitrary $x\in\mathbb{R}$.

$$\begin{aligned}
(f+(-f))(x) &= f(x)+(-f)(x) && \text{by the definition of} + \text{in } \mathbb{R}\\
&= f(x)+(-f(x)) && \text{by the definition of} -f\\
&= \underset{\mathbb{R}}{0} && \text{since } \underset{\mathbb{R}}{0} \text{ is the additive identity in } \mathbb{R}\\
&= \underset{F(\mathbb{R})}{0} && \text{since} \forall x\in\mathbb{R},\ \underset{F(\mathbb{R})}{0}(x)=\underset{\mathbb{R}}{0}
\end{aligned}$$

Thus, $\forall f\in F(\mathbb{R})$, $\exists -f\in F(\mathbb{R})$ s.t. $f+(-f)=\underset{F(\mathbb{R})}{0}$. ■

■

■ **Example 1.5** Prove $\underset{2\times3}{M}(\mathbb{R})\,\{[a_{ij}]\mid a_{ij}\in\mathbb{R}\}$ is a vector space.

Example: $\begin{bmatrix}1&2&0\\0&0&0\end{bmatrix}\in\underset{2\times3}{M}(\mathbb{R})$.

- $[a_{ij}]+[b_{ij}]=[a_{ij}+b_{ij}]$
- $r\,[a_{ij}]=[ra_{ij}]$
- Additive identity in $\underset{2\times3}{M}(\mathbb{R})$ is $\underset{2\times3}{0}=\begin{bmatrix}0&0&0\\0&0&0\end{bmatrix}$.
- Given $A=[a_{ij}]\in\underset{2\times3}{M}(\mathbb{R})$, $-A=[-a_{ij}]$ is the additive inverse of A.

■

Proposition 1.1.6 Let V be a vector space. Then

a) The zero vector $\vec{0}$ (additive identity in V) is unique.

Proof. WTS Additive identity in V is unique[7].
Suppose $\vec{0}$ and $\vec{0}'$ are additive identity in V.
Since $\vec{0}$ is additive identity, $\forall\vec{v}\in V$, $\vec{v}+\vec{0}=\vec{v}$. In particular, $\vec{0}'+\vec{0}=\vec{0}'$.
Since $\vec{0}'$ is additive identity, $\forall\vec{v}\in V$, $\vec{v}+\vec{0}'=\vec{v}$. In particular, $\vec{0}+\vec{0}'=\vec{0}$.
So $\vec{0}'=\vec{0}+\vec{0}'=\vec{0}$, so there is only one additive identity in V. ■

b) For all $\vec{v}\in V$, $\underset{\mathbb{F}}{0}\vec{v}=\vec{0}_V$.

Proof. WTS $\forall\vec{v}\in V$, $\underset{\mathbb{F}}{0}\vec{v}=\vec{0}_V$.

[7] To prove that something is unique, a common technique is to assume we have two examples of the object in question, then show that those two examples must in fact be equal.

Pick $\vec{v} \in V$.

$$\begin{aligned}
\underset{\mathbb{F}}{0}\vec{v} &= (\underset{\mathbb{F}}{0} + \underset{\mathbb{F}}{0})\vec{v} \\
&= \underset{\mathbb{F}}{0}\vec{v} + \underset{\mathbb{F}}{0}\vec{v} \\
(-\underset{\mathbb{F}}{0}\vec{v}) + \underset{\mathbb{F}}{0}\vec{v} &= (-\underset{\mathbb{F}}{0}\vec{v}) + (\underset{\mathbb{F}}{0}\vec{v} + \underset{\mathbb{F}}{0}\vec{v}) \\
-\underset{\mathbb{F}}{0}\vec{v} + \underset{\mathbb{F}}{0}\vec{v} &= (-\underset{\mathbb{F}}{0}\vec{v} + \underset{\mathbb{F}}{0}\vec{v}) + \underset{\mathbb{F}}{0}\vec{v} \\
\vec{0}_V &= \vec{0}_V + \underset{\mathbb{F}}{0}\vec{v} \\
&= \underset{\mathbb{F}}{0}\vec{v}
\end{aligned}$$

Thus, $\forall \vec{v} \in V$, $\underset{\mathbb{F}}{0}\vec{v} = \vec{0}_V$. ■

c) For each $\vec{v} \in V$, the additive inverse $-\vec{v}$ is unique.

Proof. We use the same idea as in the proof of part a. Given $\vec{v} \in V$, if $-\vec{v}$ and $(-\vec{v})'$ are two additive inverses of $\vec{v}$, then on one hand we have $\vec{v} + -\vec{v} + (-\vec{v})' = (\vec{v} + -\vec{v}) + (-\vec{v})' = \vec{0} + (-\vec{v})' = (-\vec{v})'$, by axioms 1, 4, and 3. On the other hand, if we use axiom 2 first before associating, we have $\vec{v} + -\vec{v} + (-\vec{v})' = \vec{v} + (-\vec{v})' + -\vec{v} = (\vec{v} + (-\vec{v})') + -\vec{v} = \vec{0} + -\vec{v} = -\vec{v}$. Hence $-\vec{v} = (-\vec{v})'$, and the additive inverse of $\vec{v}$ is unique. ■

d) For all $\vec{v} \in V$, and all $c \in \mathbb{R}$, $(-c)\vec{v} = -(c\vec{v})$.

Proof. We have $c\vec{v} + (-c)\vec{v} = (c + -c)\vec{v} = 0\vec{v} = \vec{0}$ by axiom 6 and part b. Hence $(-c)\vec{v}$ also serves as an additive inverse for the vector $c\vec{v}$. By part c, therefore, we must have $(-c)\vec{v} = -(c\vec{v})$. ■

1.3.2 Subspace

■ **Example 1.6** Let $W \subseteq \mathbb{R}^3$.
Let $\vec{w}_1, \vec{w}_2 \in W$.
Then, $\vec{w}_1 + \vec{w}_2 \in W$.

$$+: \begin{array}{ccc} W \times W & \to & W \\ (\vec{w}_1, \vec{w}_2) & \mapsto & \vec{w}_1 + \vec{w}_2 \end{array} \qquad \cdot: \begin{array}{ccc} \mathbb{R} \times W & \to & W \\ (r, \vec{w}) & \mapsto & r\vec{w} \end{array}$$

We can conclude that W is a subspace of $\mathbb{R}^3$ [8]. ■

Definition 1.2.6 — Subspace. Let V be a vector space and let $W \subseteq V$ be a *non-empty* subset. Then W is a (vector) *subspace* of V if W is a vector space itself under the operations of vector sum and scalar multiplication from V.

Theorem 1.2.8 — Subset Test. Let V be a vector space, and let W be a non-empty subset of V. Then W is a subspace of V if and only if for all $\vec{v}, \vec{w} \in W$, and all $c \in \mathbb{R}$. we have $c\vec{v} + \vec{w} \in W$ [a].

[a] This is equivalent to "W is closed under vector addition and under scalar multiplication of V".

[8] A short hand of "*W is a subspace of V*" is "$W \underset{S.S.}{\subseteq} V$".

■ **Example 1.7** Is $\mathbb{R}^2$ a subspace of $\mathbb{R}^3$?

Note that $\mathbb{R}^2 = \left\{ \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \mid a_1, a_2 \in \mathbb{R} \right\}$ but $\mathbb{R}^3 = \left\{ \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \mid a_1, a_2, a_3 \in \mathbb{R} \right\}$. Clearly, $\mathbb{R}^2 \not\subseteq \mathbb{R}^3$. ■

■ **Example 1.8** Let $\mathbb{F}$ be a field.

Define $\mathbb{F}^3 = \left\{ \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \mid a_i \in \mathbb{F} \right\}$.

Let $W = \left\{ \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \in \mathbb{F}^3 \mid a_1 + a_2 + a_3 = \underset{\mathbb{F}}{0} \right\}$.

Proof. WTS W is a subspace of $\mathbb{F}^3$.

- Show that W is non-empty.
  $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \in W$, so W is non-empty.
- Show that $W \subseteq \mathbb{F}^3$.
  Since $\forall \vec{v} \in W$, $\vec{v} \in \mathbb{F}^3$, we have $W \subseteq \mathbb{F}^3$.
- Show that W is closed under vector addition.
  Let $\vec{w}_1 = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \vec{w}_2 = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \in W$, $a_i, b_i \in \mathbb{F}$.
  Then, $\vec{w}_1 + \vec{w}_2 = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{bmatrix}$.
  Note that $(a_1 + b_1) + (a_2 + b_2) + (a_3 + b_3) = (a_1 + a_2 + a_3) + (b_1 + b_2 + b_3) = \underset{\mathbb{F}}{0} + \underset{\mathbb{F}}{0} = \underset{\mathbb{F}}{0}$.
  Thus, $\vec{w}_1 + \vec{w}_2 \in W$.
  That is, W is closed under vector addition.
- Show that W is closed under scalar multiplication.
  Pick $r \in \mathbb{F}$, pick $\vec{w} \in W$.
  $r\vec{w} = r \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} ra_1 \\ ra_2 \\ ra_3 \end{bmatrix}$.
  $ra_1 + ra_2 + ra_3 = r(a_1 + a_2 + a_3) = r(\underset{\mathbb{F}}{0}) = \underset{\mathbb{F}}{0}$.
  Thus, $r\vec{w} \in W$.
  Thus is, W is closed under scalar multiplication.

Thus, $W \underset{S.S.}{\subseteq} \mathbb{F}^3$. ■

■

1.4 Linear Combinations

Definition 1.3.1 Let S be a subset of a vector space V.

- A *linear combination* of vectors in S is any sum $a_x\vec{v}_1 + \cdots + a_n\vec{v}_n$, where the $a_i \in \mathbb{R}$, and the $\vec{v}_i \in S$.
- If $5 \neq \emptyset$ (the empty subset of V), the set of all linear combinations of vectors in S is called the (linear) *span* of S, and denoted $\text{Span}(S)$. If $S = \emptyset$, we define

Span$(S) = \{\vec{0}\}$.
- If $W = \text{Span}(S)$, we say S *spans* (or generates) W.
- We think of the span of a set S as the set of all vectors that can be "built up" from the vectors in S by forming linear combinations.

How do we describe Span(S) in set builder notation?
- If S is finite or infinite

$$\text{Span}(S) = \{a_1\vec{x}_1 + a_2\vec{x}_2 + \cdots + a_n\vec{x}_n \mid a_i \in \mathbb{F}, n \in \mathbb{N}, \vec{x}_i \in S\}$$

- If S is finite

$$S = \{\vec{w}_1, \dots, \vec{w}_n\}$$

$$\text{Span}(S) = \{a_1\vec{w}_1 + \cdots + a_k\vec{w}_k \mid a_i \in \mathbb{F}, \vec{w}_i \in S\}$$

■ **Example 1.9**

Let $S = \{\sin^2 x, \cos^2 x\} \subseteq F(\mathbb{R})$.
- Describe Span(S).
  Span$(S) = \{a_1 \sin^2 x + a_2 \cos^2 x \mid a_i \in \mathbb{R}\}$
- True or False: Span(S) contains all constant functions
  True.
  Let $f(x) = c$, $c \in \mathbb{R}$.

$$\begin{aligned}\text{Then, } f(x) &= c \cdot 1\\ &= c(\sin^2 x + \cos^2 x)\\ &= c\sin^2 x + c\cos^2 x \in \text{Span}(S)\end{aligned}$$

Let $S' = \left\{\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}\right\} \subseteq \underset{2\times 2}{M}(\mathbb{R})$.
- Describe Span(S).
  Span(S) = $\left\{\begin{bmatrix}a_1 & a_2\\0 & 0\end{bmatrix} \mid a_1, a_2 \in \mathbb{R}\right\}$
- True or False: Span(S') contains no symmetric matrix.
  False.
  Consider $0_{2\times 2} = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$.

■

■ **Example 1.10** Let $\mathcal{P}(\mathbb{R})$:= the set of all polynomials with real coefficients.
- Give s spanning set for $\mathcal{P}(\mathbb{R})$.

$$S = \{x^i \mid i \in \mathbb{N} \cup \{0\}\}$$

- Can we find a finite spanning set for ?
  No.

■

Theorem 1.4.4 Let V be a vector space and let S be any subset of V. Then Span(S) is a subspace of V.

Proof. V is an $\mathbb{F}$_V.S. $S \subseteq V$. WTS Span$(S) \underset{S.S.}{\subseteq} V$.
- Show Span$(S) \subseteq V$.
  Pick $\vec{v} \in \text{Span}(S) = \{a_1\vec{v}_1 + \cdots + a_m\vec{v}_m \mid a_i \in \mathbb{F}, \vec{v}_i \in S, m \in \mathbb{N}\}$.
  Then, $\vec{v} = a_1\vec{v}_1 + \cdots + a_m\vec{v}_m$ for some $\vec{v}_1, \dots, \vec{v}_m \in S$ and $a_1, \dots a_m \in \mathbb{F}$.
  Since $\vec{v}_i \in S \subseteq V$, $\vec{v}$ is a linear combination of vectors in V, and V is a vector space.
  Thus, $\vec{v} \in V$.
- Show Span$(S) \neq \emptyset$
- Show Span(S) is closed under vector addition.

- Show Span(S) is closed under scalar multiplication.

■

■ **Example 1.11**
Let $S_1 = \{1, 2\sin^2 x, 3\cos^2 x\}$, $S_2 = \{\sin^2 x, \cos^2 x\}$.
$\text{Span}(S_1) \stackrel{?}{=} \text{Span}(S_2)$?

- $\text{Span}(S_1) \stackrel{?}{\subseteq} \text{Span}(S_2)$?
  Yes.
  $S_1 \subseteq \text{Span}(S_2)$.
  $\vec{v} = a \underbrace{\sin^2 x}_{\in \text{Span}(S_2)} + b \underbrace{\cos^2 x}_{\in \text{Span}(S_2)} \in \text{Span}(S_2)$ since $\text{Span}(S_2) \underset{S.S.}{\subseteq} V$.
- $\text{Span}(S_2) \stackrel{?}{\subseteq} \text{Span}(S_1)$?
  Yes.
  $S_2 \subseteq \text{Span}(S_1)$.
  $\vec{v} = a \cdot \underbrace{1}_{\in \text{Span}(S_1)} + b \underbrace{\sin^2 x}_{\in \text{Span}(S_1)} + c \underbrace{\cos^2 x}_{\in \text{Span}(S_1)} \in \text{Span}(S_1)$ since $\text{Span}(S_1) \underset{S.S.}{\subseteq} V$.

■

1.5 Linear Dependence and Linear Independence

Definition 1.4.2 — Linear Dependence. Let V be a vector space, and let S be a subset of V.

a) A *linear dependence*[a] among the vectors of S is an equation

$$a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = 0$$

where the $\vec{v}_i \in S$ and the $a_i \in \mathbb{R}$ are not all zero (i.e., at least one of the $a_i \neq 0$).

b) the set S is said to be *linearly dependent* if there exists a linear dependence among the vectors in S.

[a] A linear dependence is often called a **non**-*trivial relation*.

■ **Example 1.12**
Is $S = \left\{ \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \right\} \subseteq \underset{2\times 2}{M}(\mathbb{R})$ linearly independent?
Yes.

$$0 \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + 0 \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix} + 1 \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \vec{0}_{2\times 2}$$

■

■ **Example 1.13** Let V be a vector space over the field $\mathbb{F}$.
Is $S' = \{\vec{0}, \vec{v}_1, \ldots, \vec{v}_n\} \subseteq V$ linearly independent?
Yes.

$$1_{\mathbb{F}}\vec{0} + 0_{\mathbb{F}}\vec{v}_1 + \cdots + 0_{\mathbb{F}}\vec{v}_n = \vec{0}$$

■

Definition 1.4.4 — Linearly Independent. A subset S of a vector space V is *linearly independent* if whenever we have $a_i \in \mathbb{R}$ and $\vec{v}_i \in S$ such that $a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = 0$, then $a_i = 0$ for all i.

■ **Example 1.14** Is $S = \{e^x, e^{2x}, e^{3x}\} \subseteq \mathcal{C}(\mathbb{R})$ linearly independent?
No.

Assume $\forall x \in \mathbb{R}, a_1e^x + a_2e^{2x} + a_3e^{3x} = \vec{0}$.
In particular when $x = 0$, $a_1 + a_2 + a_3 = 0$.

Taking the derivative on both sides. $\forall x \in \mathbb{R}$, $a_1e^x + 2a_2e^{2x} + 3a_3e^{3x} = \vec{0}$.
Again, when $x = 0$, $a_1 + 2a_2 + 3a_3 = 0$.

Take the derivative again, $\forall x \in \mathbb{R}$, $a_1e^x + 4a_2e^{2x} + 9a_3e^{3x} = \vec{0}$.
When $x = 0$, $a_1 + 4a_2 + 9a_3 = 0$.

Thus, $a_1 = a_2 = a_3 = 0$.
So S is linearly independent. ■

Proposition 1.4.7

- Let S be a linearly dependent subset of a vector space V, and let S' be another subset of V that **contains** S. Then S' is also linearly dependent.
- Let S be a linearly independent subset of a vector space V and let S' be another subset of V that is **contained in** S. Then S' is also linearly independent.

■ **Example 1.15** Let S_1 and S_2 be linearly independent subsets of a vector space V.

a) Is $S_1 \cup S_2$ always linearly independent? Why or why not?
No.
$\{1\} \in \mathbb{R}$, $\{2\} \in \mathbb{R}$ and they are linearly independent.
$\{1\} \cup \{2\} = \{1, 2\} \in \mathbb{R}$ and it is linearly dependent.

b) Is always linearly independent? Why or why not?
Yes.
$S_1 \cap S_2 \subseteq S_1$ and $S_1 \cap S_2 \subseteq S_2$. Then by the proposition $S_1 \cap S_2$ is linearly independent.

c) Is always linearly independent? Why or why not?
Yes.
$S_1 \setminus S_2 \subseteq S_1$. Then by the proposition $S_1 \setminus S_2$ is linearly independent.

■

Lemma 1.6.8 Let S be a linearly independent subset of V and let $\vec{v} \in V$, but $\vec{v} \notin S$. Then $S \cup \{\vec{v}\}$ is linearly independent if and only if $\vec{v} \notin \text{Span}\,(S)$.

Proof. WTS $S \cup \{\vec{v}\}$ is linearly independent if and only if $\vec{v} \notin \text{Span}((S)$.

- $\rightarrow$: Assume $S \cup \{\vec{v}\}$ is linearly independent. WTS $\vec{v} \notin \text{Span}\,(S)$.

  Proof by contradiction. Assume $\vec{v} \in \text{Span}(S)$.
  Then, $\exists r_1, \ldots, r_m \in \mathbb{F}$, $\vec{s}_1, \ldots, \vec{s}_m \in S$ s.t. $\vec{v} + r_1\vec{s}_1 + \cdots + r_m\vec{s}_m$.
  i.e $\vec{v} - r_1\vec{s}_1 - \cdots - r_m\vec{s}_m = \vec{0}$. This is a non-trivial relation among vectors in $S \cup \{\vec{v}\}$, $S \cup \{\vec{v}\}$ is linearly dependent.
  Contradiction with the assumption.

  Thus, $\vec{v} \notin \text{Span}\,(S)$.
- $\leftarrow$: Assume $\vec{v} \notin \text{Span}(S)$. WTS $S \cup \{\vec{v}\}$ is linearly independent.

  (Prove the contrapositive $\neg(S \cup \{\vec{v}\}$ is linearly independent) $\implies \neg(\vec{v} \notin \text{Span}(S))$, that is, we prove that "$S \cup \{\vec{v}\}$ is linearly dependent $\implies \vec{v} \in \text{Span}(S)$").

  There exists a dependency relation among vectors in $S \cup \{\vec{v}\}$.
  This relation should contain $\vec{v}$.

$\exists r_1, \dots, r_m, a \in \mathbb{F}$ s.t. $r_1\vec{s}_1 + \cdots + r_m\vec{s}_m + a\vec{v} = \vec{0}$.

Thus, $a\vec{v} = -r\vec{s}_1 - \cdots - r_m\vec{s}_m$

$$\vec{v} = -\frac{r_1}{a}\vec{s}_1 - \cdots - \frac{r_m}{a}\vec{s}_m$$

That is, $x \in \operatorname{Span}(S)$.

Thus, $S \cup \{\vec{v}\}$ is linearly independent if and only if $\vec{v} \notin \operatorname{Span}((S)$. ■

1.6 Bases and Dimension

Definition 1.6.1 — Basis. A subset S of a vector space V is called a *basis* of V if $V = \operatorname{Span}(S)$ and S is *linearly independent*.

Theorem 1.6.3 Let V be a vector space, and let S be a nonempty subset of V. Then S is a basis of V if and only if every vector $\vec{v} \in V$ may be written uniquely as a linear combination of the vectors in S.

Proof. WTS ■

■ Example 1.16

$$\mathbb{F}^n = \left\{ \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \mid a_i \in \mathbb{F} \right\}$$

Find a basis $\mathcal{B}$ of $\mathbb{F}^n$.

Note that $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = a_1 \begin{bmatrix} 1_{\mathbb{F}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1_{\mathbb{F}} \\ \vdots \\ 0 \end{bmatrix} + \cdots + a_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1_{\mathbb{F}} \end{bmatrix}$.

$\vec{e}_i := \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1_{\mathbb{F}} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ with the i-th element $1_{\mathbb{F}}$ and the rest 0.

$\{\vec{e}_1, \dots, \vec{e}_n\}$ is linearly independent, and $\operatorname{Span}(\vec{e}_1, \dots, \vec{e}_n) = \mathbb{F}^n$.
Then $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ is a basis for $\mathbb{F}^n$. ■

■ Example 1.17

Consider $\mathbb{C}^2 = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \mid z_i \in \mathbb{C} \right\}$.

- Suppose $\mathbb{C}^2$ is a $\mathbb{C}$_V.S., what is $\dim \mathbb{C}^2$?
  - $w \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} wz_1 \\ wz_2 \end{pmatrix}, w \in \mathbb{C}$.
  - $\mathcal{E} = \left\{ \vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$
- Suppose $\mathbb{C}^2$ is a $\mathbb{R}$_V.S., what is $\dim \mathbb{C}^2$?

- $r\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} rz_1 \\ rz_2 \end{pmatrix}, r \in \mathbb{R}$.
- $\mathcal{V} = \left\{ \vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \vec{v}_3 = \begin{pmatrix} i \\ 0 \end{pmatrix}, \vec{v}_4 = \begin{pmatrix} 0 \\ i \end{pmatrix} \right\}$
- $z_1 = a_1 + b_1 i$, $z_2 = a_2 + b_2 i$, then $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{C}^2 = a_1\vec{v}_1 + a_2\vec{v}_2 + b_1\vec{v}_3 + b_2\vec{v}_4$

■

Theorem 1.6.10 Let V be a *finitely generate* vector space and let S be a spanning set for V, which has m elements. Then no linearly independent set in V can have more than m elements.

Corollary 1.6.11 Let V be a vector space and let S and S' be two bases of V, with m and m' elements, respectively. Then $m = m'$.

Proof. WTS if S and S' are two bases of V, with m and m' elements, respectively, then $m = m'$.

S is a basis for $V \overset{\text{by definition of basis}}{\Longrightarrow} S$ is linearly independent.
S' is a basis for $V \overset{\text{by definition of basis}}{\Longrightarrow} \operatorname{Span}(S') = V$.
Applying the theorem above, $m = |S| \leq |S'| = m'$.

S is a basis for $V \overset{\text{by definition of basis}}{\Longrightarrow} \operatorname{Span}(S) = V$.
S' is a basis for $V \overset{\text{by definition of basis}}{\Longrightarrow} S'$ is linearly independent.
Applying the theorem above, $m' = |S'| \leq |S| = m$.

So $m \leq m'$ and $m' \leq m$. That is, $m = m'$. ■

2. Linear Transformations

2.1 Linear Transformations

Definition 2.1.1 — Linear Transformation. A function $T : V \to W$ is called a *linear mapping* or a *linear transformation* if it satisfies

i) $T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2)$ for all $\vec{v}_1$ and $\vec{v}_2 \in V$.

ii) $T(r\vec{v}) = rT(\vec{v})$ for all $r \in \mathbb{R}$ and $\vec{v} \in V$.

V is called the *domain* of T and W is called the *target* of T.

■ **Example 2.1** Consider the following examples.

$$\begin{array}{rccc} \mathrm{id}: & V & \to & V \\ & \vec{x} & \mapsto & \vec{x} \end{array} \qquad \begin{array}{rccc} R_\theta: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & \vec{x} & \mapsto & \vec{x} \text{ rotated through } \theta \end{array}$$

$$\begin{array}{rccc} P_l: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & \vec{x} & \mapsto & \mathrm{proj}_l \vec{x} \end{array} \qquad \begin{array}{rccc} T_M: & \mathbb{R}^3 & \to & \mathbb{R} \\ & \vec{x} & \mapsto & \begin{bmatrix} 1 & 2 & 2 \end{bmatrix} \vec{x} \end{array}$$

These are all commonly seen linear transformations. ■

Given a $m \times n$ matrix A, $T_A : \begin{array}{ccc} \mathbb{R}^m & \to & \mathbb{R}^n \\ \vec{x} & \mapsto & A\vec{x} \end{array}$ is a linear transformation.

Conversely for every linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$, $\exists n \times n$ matrix M_T s.t. $T(\vec{v}) = M_T\vec{v}$.

Given V, W be $\mathbb{F}$_V.S., and a linear transformation $T : V \to W$.

Can we describe T using matrices?

■ **Example 2.2** Which function is a linear transformation?

a) Let $\mathcal{C}^\infty(\mathbb{R})$ be the space of smooth functions

$$\begin{array}{rccc} D: & \mathcal{C}^\infty(\mathbb{R}) & \to & \mathcal{C}^\infty(\mathbb{R}) \\ & f & \mapsto & f' \end{array}$$

b) Let $\mathcal{C}[0,1]$ be the space of continuous functions on $[0,1] \to \mathbb{R}$.

$$\begin{array}{rccc} I: & \mathcal{C}[0,1] & \to & \mathbb{R} \\ & f & \mapsto & \int_0^1 f(t)\,dt \end{array}$$

- a) is a linear transformation.
  Let $f, g \in \mathcal{C}^\infty(\mathbb{R})$, $r \in \mathbb{R}$.
  $D(f+g) = (f+g)' = f' + g' = D(f) + D(g)$
  $D(rf) = (rf)' = rf' = rD(f)$.
- b is also a linear transformation.
  Let $f, g \in \mathcal{C}[0,1]$, $r \in \mathbb{R}$.
  $I(f+g) = \int_0^1 (f+g)(t)\,dt = \int_0^1 f(t) + g(x)\,dt = \int_0^1 f(t)\,dt + \int_0^1 g(t)\,dt = I(f) + I(g)$.
  $I(rf) = \int_0^1 (rf)(t)\,dt = r\int_0^1 f(t)\,dt = rI(f)$.

∎

Properties of a Linear Transformation

- Linear transformations $\vec{0}$ to $\vec{0}$.
- Linear transformations preserve linear combinations

Lemma 2.0.1 Let $T: V \to W$ be a linear transformation. Then $T(\vec{0}_V) = \vec{0}_W$.

Proof. WTS $T(\vec{0}_V) = \vec{0}_W$.

$$\begin{aligned} T(\vec{0}_V) &= T(\vec{0}_V + \vec{0}_V) && \text{since } \vec{0}_V \text{ is the additive identity in } V \\ &= T(\vec{0}_V) + T(\vec{0}_V) && \text{by linearity of } T \\ T(\vec{0}_V) + (-T(\vec{0}_V)) &= (T(\vec{0}_V) + T(\vec{0}_V)) + (-T(\vec{0}_V)) && \text{since the additive inverse is unique} \\ \vec{0}_W &= T(\vec{0}_V) + (T(\vec{0}_V) + (-T(\vec{0}_V))) && \text{since addition is associative} \\ &= T(\vec{0}_V) + \vec{0}_W \\ &= T(\vec{0}_V) && \text{since } \vec{0}_W \text{ is the additive identity in } W \end{aligned}$$

Thus, $T(\vec{0}_V) = \vec{0}_W$. ∎

Proposition 2.1.2 A function $T: V \to W$ is a linear transformation if and only if for all a and $b \in \mathbb{F}$ and all $\vec{u}$ and $\vec{v} \in V$

$$T(a\vec{u} + b\vec{v}) = aT(\vec{u}) + bT(\vec{v})$$

Proof. WTS a function $T: V \to W$ is a linear transformation if and only if for all a and $b \in \mathbb{F}$ and all $\vec{u}$ and $\vec{v} \in V$, $T(a\vec{u} + b\vec{v}) = aT(\vec{u}) + bT(\vec{v})$.

- $\to$: WTS $T: V \to W$ is a linear transformation $\implies T(a\vec{u} + b\vec{v}) = aT(\vec{u}) + bT(\vec{v})$.

$$\begin{aligned} T(a\vec{u} + b\vec{v}) &= T(a\vec{u}) + T(b\vec{v}) && \text{by (1) of the definition of L.T.} \\ &= aT(\vec{u}) + bT(\vec{v}) && \text{by (2) of the definition of L.T.} \end{aligned}$$

- $\leftarrow$: WTS $T(a\vec{u} + b\vec{v}) = aT(\vec{u}) + bT(\vec{v}) \implies T: V \to W$ is a linear transformation.

∎

Corollary 2.1.3 A function $T: V \to W$ is a linear transformation if and only if for all

$a_1, \dots a_k \in \mathbb{F}$ and for all $\vec{v}_1, \dots, \vec{v}_k \in V$:

$$T(\sum_{i=1}^{k} a_i\vec{v}_i) = \sum_{i=1}^{k} a_i T(\vec{v}_i)$$

Proof. WTS $T : V \to W$ is a linear transformation if and only if for all $a_1, \dots a_k \in \mathbb{F}$ and for all $\vec{v}_1, \dots, \vec{v}_k \in V$, $T(\sum_{i=1}^{k} a_i\vec{v}_i) = \sum_{i=1}^{k} a_i T(\vec{v}_i)$.

- $\to$: WTS $T : V \to W$ is a linear transformation $\implies T(\sum_{i=1}^{k} a_i\vec{v}_i) = \sum_{i=1}^{k} a_i T(\vec{v}_i)$.

  Prove by induction.
  - **Base case:** WTS $T(a_1\vec{v}_1) = a_1T(\vec{v}_1)$
    This follows immediately from (2) of the definition of linear transformation.
  - **Induction hypothesis:** $T(\sum_{i=1}^{k} a_i\vec{v}_1) = \sum_{i=1}^{k} a_i T(\vec{v}_i)$.

    WTS $T(\sum_{i=1}^{k+1} a_i\vec{v}_1) = \sum_{i=1}^{k+1} a_i T(\vec{v}_i)$.

    By (1) of the definition of linear transformation, $T(\sum_{i=1}^{k+1} a_i\vec{v}_i) = T(\sum_{i=1}^{k} a_i\vec{v}_i) + T(a_{k+1}\vec{v}_{k+1})$.
    $T(\sum_{i=1}^{k} a_i\vec{v}_i) = \sum_{i=1}^{k} a_i T(\vec{v}_i)$ by the induction hypothesis.
    $T(a_{k+1}\vec{v}_{k+1}) = a_{k+1}T(\vec{v}_{k+1})$ by (2) of the definition of linear transformation.

    So
    $$\begin{aligned} T(\sum_{i=1}^{k+1} a_i\vec{v}_i) &= T(\sum_{i=1}^{k} a_i\vec{v}_i) + T(a_{k+1}\vec{v}_{k+1}) \\ &= \sum_{i=1}^{k} a_i T(\vec{v}_i) + a_{k+1}T(\vec{v}_{k+1}) \\ &= \sum_{i=1}^{k} a_i T(\vec{v}_i) \end{aligned}$$

    Thus, $T(\sum_{i=1}^{k} a_i\vec{v}_i) = \sum_{i=1}^{k} a_i T(\vec{v}_i)$.
- $\leftarrow$: WTS $T(\sum_{i=1}^{k} a_i\vec{v}_i) = \sum_{i=1}^{k} a_i T(\vec{v}_i) \implies T : V \to W$ is a linear transformation.

■

Proposition 2.1.14 If $T : V \to W$ is a linear transformation and V is finite-dimensional, then T is uniquely determined by its value on the members of a basis of V.

Definition 2.0.2 Let A, B be sets.
Let $f : A \to B$ be a function.
- f is *injective* (one to one) if $\forall a_1, a_2 \in A,\ f(a_1) = f(a_2) \implies a_1 = a_2$.

- f is *surjective* (onto) if $\forall b \in B$, $\exists a \in A$ s.t. $f(a) = b$.
- f is *bijective* if f is both injective and surjective.

2.2 Kernel and Image

Definition 2.3.1 — Kernel. The *kernel*[a] of T, denoted $\ker(T)$. is the subset of V consisting of all vectors $\vec{v} \in V$ such that $T(\vec{v}) = 0$.

$$\ker T = \{\vec{v} \in V \mid T(\vec{v}) = \vec{0}\}$$

[a]Kernel is equivalent to the null space.

Definition 2.3.10 — Image. The subset of W consisting of all vectors $\vec{w} \in W$ for which there exists a $\vec{v} \in V$ such that $T(\vec{v}) = \vec{w}$ is called the *image* of T and is denoted by $\mathrm{Im}(T)$.

$$\mathrm{Im}\ T = \{\vec{w} \in W \mid T(\vec{v}) = \vec{w} \text{ for some } \vec{v} \in V\}$$

Proposition 2.0.3 Let T be a linear transformation. T is injective if and only if $\ker T = \{\vec{0}\}$.

Proof. WTS T is injective iff $\ker T = \{\vec{0}\}$.

- $\rightarrow$: Assume T is injective. WTS $\ker T = \{\vec{0}\}$.
  Pick $\vec{v} \in T$.
  Suppose $T(\vec{v}) = \vec{0}_W$.

  Then, $T(\vec{v}) = \vec{0}_W = T(\vec{0})$.
  Since T is injective, $\vec{v} = 0$.

  Thus, $\forall \vec{v} \in \ker T, \vec{v} = \vec{0}$.
  That is, $\ker T = \vec{0}$.
- $\leftarrow$: Assume $\ker T = \{\vec{0}\}$. WTS T is injective.
  WTS $T(\vec{v}_1) = T(\vec{v}_2) \implies \vec{v}_1 = \vec{v}_2$.
  Pick $\vec{v}_1, \vec{v}_2 \in V$.
  Suppose $T(\vec{v}_1) = T(\vec{v}_2)$.

  Then, $T(\vec{v}_1 - \vec{v}_2) = \vec{0}_W$.
  By linearity of T, $\vec{0}_W = T(\vec{v}_1 - \vec{v}_2) = T(\vec{v}_1) - T(\vec{v}_2)$.

  Then, $\vec{v}_1 - \vec{v}_2 \in \ker T = \{\vec{0}_V\}$.
  Thus, $\vec{v}_1 - \vec{v}_2 = 0$.
  That is, $\vec{v}_1 = \vec{v}_2$.

Thus, T is injective iff $\ker T = \{\vec{0}\}$. ■

Proposition 2.0.4 Let T be a linear transformation. T is surjective if and only if $\mathrm{Im}\ T = W$.

■ **Example 2.3** Let $\begin{array}{rccc} D: & \mathcal{C}^\infty(\mathbb{R}) & \to & \mathcal{C}^\infty(\mathbb{R}) \\ & f & \mapsto & f' \end{array}$ be a linear transformation.

- Describe:
  - $\ker D$.

$$\begin{aligned} \ker D &= \{f \in \mathcal{C}^\infty(\mathbb{R}) \mid D(f) = \vec{0}\} \\ &= \{f \in \mathcal{C}^\infty(\mathbb{R}) \mid f' = \vec{0}\} \\ &= \{f \in \mathcal{C}^\infty(\mathbb{R}) \mid f(x) = c \text{ for some } c \in \mathbb{R}\} \end{aligned}$$

- Im D.

$$\begin{aligned}\text{Im } D &= \{g \in \mathcal{C}^\infty(\mathbb{R}) \mid g = D(f) \text{ for some } f \in \mathcal{C}^\infty(\mathbb{R})\} \\ &= \{g \in \mathcal{C}^\infty(\mathbb{R}) \mid g = f' \text{ for some } f \in \mathcal{C}^\infty(\mathbb{R})\} \\ &= \mathcal{C}^\infty(\mathbb{R})\end{aligned}$$

- Is D:
  - injective?
    No.
    The kernel is NOT trivial.
  - surjective?
    Yes.
    Im $g = \mathcal{C}^\infty(\mathbb{R})$(the image is equal to the codomain).

■

■ Example 2.4 Let $T: \begin{array}{ccc} P_2(\mathbb{R}) & \to & P_3(\mathbb{R}) \\ p(x) & \mapsto & xp(x) \end{array}$ be a linear transformation.

- Describe:
  - $\ker D$.

$$\begin{aligned}\text{Ker } T &= \{p(x) \in P_2(x) \mid xp(x) = 0\} \\ &= \{\vec{0}\}\end{aligned}$$

  - Im D.

$$\text{Im } T = \{xp(x) \mid p(x) \in P_2(\mathbb{R})\}$$

- Is D:
  - injective?
    Yes.
    The kernel of T is trivial.
  - surjective?
    No.
    Take $p(x) = 2 \in P_3(\mathbb{R})$. $2 \notin \text{Im } T$.

■

Proposition 2.3.1 Let $T: V \to W$ be a linear transformation. $\ker(T)$ is a subspace of V.

Proof. We will use subspace test.

- Show that $\ker T \subseteq V$.
  The follows immediately form the definition of kernel.
- Show that $\ker T \neq \emptyset$.
  Since $T(\vec{0}) = \vec{0}$, $\vec{0} \in \ker T$. So $\ker T \neq \emptyset$.
- Show that $\ker T$ is closed under addition.
  Pick $\vec{v}_1, \vec{v}_2 \in \ker T$.

$$\begin{aligned}T(\vec{v}_1 + \vec{v}_2) &= T(\vec{v}_1) + T(\vec{v}_2) && \text{by linearity of } T \\ &= \vec{0}_W + \vec{0}_W && \text{since } \vec{v}_1, \vec{v}_2 \in \ker T \\ &= \vec{0}_W\end{aligned}$$

- Show that $\ker T$ is closed under scalar multiplication.

Pick $\vec{v} \in V$, $r \in \mathbb{F}$.

$$\begin{aligned} T(r\vec{v}) &= rT(\vec{v})) && \text{by linearity of } T \\ &= r\vec{0}_W && \text{since } \vec{v} \in \ker T \\ &= \vec{0}_W \end{aligned}$$

Thus, $\ker T \underset{S.S.}{\subseteq} V$. ■

■ **Example 2.5** True or false.

- Let T be a linear transformation and $\dim \ker T = 0$. Then T is injective.
  True.
- Let T be a linear transformation and $\dim \text{Im } T = \dim W$. Then is surjective.
  True
  - $\text{Im } T \underset{S.S.}{\subseteq} W \overset{?}{\Longrightarrow} \text{Im } T = W$
  - Assume $\dim(\text{Im } T) = \dim W$.
  - Let $\mathcal{B} = \left(\vec{b}_1, \ldots, \vec{b}_k\right)$ be a basis for $\text{Im } T \implies \mathcal{B}$ is linearly independent in W.
  - $\mathcal{B}$ is a basis for W, and $\mathcal{B}$ has k elements.

■

■ **Example 2.6**

Let $A = \begin{bmatrix} 1 & 4 & 1 & 2 \\ 3 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$.

$$A \sim \underbrace{\begin{bmatrix} 1 & 4 & 1 & 2 \\ 0 & -8 & -3 & -5 \\ 0 & 0 & 0 & 4 \end{bmatrix}}_{REF} \sim \underbrace{\begin{bmatrix} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & \frac{3}{8} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{RREF}.$$

$$\begin{aligned} T_A : \mathbb{R}^4 &\to \mathbb{R}^3 \\ \vec{x} &\mapsto A\vec{x} \end{aligned}$$

1. Find a basis for $\underbrace{\ker T_A}_{\text{Nul}(A)}$ and $\dim(\ker T_A)$.

   $\ker T_A = \text{Nul } A = \{\vec{x} \in \mathbb{R}^4 \mid A\vec{x} = \vec{0}\}$.
   Solution to $A\vec{x} = \vec{0}$.
   From RREF of : x_3 is a free variable.
   - $x_1 = \frac{1}{2}x_3$
   - $x_2 = -\frac{3}{8}x_3$
   - $x_4 = 0$

   $\vec{x} = x_3 \begin{pmatrix} \frac{1}{2} \\ -\frac{3}{2} \\ 1 \\ 0 \end{pmatrix}$, $x_3 \in \mathbb{R}$.

   $\ker T_A = \text{Nul } A = \text{Span}\left(\begin{pmatrix} \frac{1}{2} \\ -\frac{3}{8} \\ 1 \\ 0 \end{pmatrix}\right)$.

That is, a basis for $\ker T_A = \left\{ \begin{pmatrix} \frac{1}{2} \\ -\frac{3}{8} \\ 1 \\ 0 \end{pmatrix} \right\}$, $\dim(\ker T_A) = 1$.

2. Find a basis for $\underbrace{\text{Im } T_A}_{\text{Col}(A)}$ and $\dim(\text{Im } T_A)$.

$$\begin{aligned} \text{Im } T_A &= \text{Cal}(A) = \{A\vec{x} \mid \vec{x} \in \mathbb{R}^4\} \\ &= \text{Span}\left(\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix} \right) \end{aligned}$$

Take $x_3 = 1$:

$$A \begin{pmatrix} \frac{1}{2} \\ -\frac{3}{8} \\ 1 \\ 0 \end{pmatrix} = \vec{0}$$

$$\vec{a}_3 = -\frac{1}{2}\vec{a}_1 + \frac{3}{8}\vec{a}_2$$

Thus, a basis for $\text{Im } T_A = \left\{ \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix} \right\}$. $\dim(\text{Im } T_A) = 3$.

3. Find the relationship between the rank and nullity of T_A.

$$\dim \text{dom } A = \dim \text{Im } T_A + \dim \ker T_A$$

■

Theorem 2.3.17 — Rank-Nullity Theorem. If V is a **finite-dimensional** vector space and $T : V \to W$ is a linear transformation, then

$$\dim V = \dim \text{Im } T + \dim \ker T$$

Proof. Proof: WTS $\dim V = \dim \text{Im } T + \dim \ker T$.

Consider a basis $\vec{u}_1, \ldots, \vec{u}_m$ for $\ker T$.
$\vec{u}_1, \ldots, \vec{u}_m$ is linearly independent in V, thus it can be extended to a basis for V.
Suppose $\vec{u}_1, \ldots, \vec{u}_m, \vec{v}_1, \ldots, \vec{v}_n$ is a basis for V we get.
Note that $\dim \ker T = m$, and $\dim V = m + n$.

WTS $\dim \text{Im } T = n$.
Take $\vec{w} \in \text{Im } T$.
Then, $\exists \vec{v} \in V$ s.t. $T(\vec{v}) = \vec{w}$.

$$\begin{aligned} \vec{w} &= T(\vec{v}) \\ &= T(a_1\vec{u}_1 + \cdots + a_m\vec{u}_m + b_1\vec{v}_1 + \cdots + b_n\vec{v}_n) \text{ for some } a_i, b_i \in \mathbb{F} \\ &= a_1T(\vec{u}_1) + \cdots + a_mT(\vec{u}_m) + b_1T(\vec{v}_1) + \cdots + b_nT(\vec{v}_n) \end{aligned}$$

Since $\vec{v}_1, \dots, \vec{v}_n$ is a basis for $\ker T$, $a_i T(\vec{u}_i) = \vec{0}$.

$$\vec{w} = b_1 T(\vec{v}_1) + \cdots + b_n T(\vec{v}_n)$$

Thus, $\text{Im } T = \text{Span}\,(T(\vec{v}_1), \dots, T(\vec{v}_n))$.

Show that $T(\vec{v}_1), \dots, T(\vec{v}_n)$ is linearly independent.
$c_1 T(\vec{v}_1), \dots, c_n T(\vec{v}_n) = \vec{0}$, WTS $c_i = 0$.

$$T(\underbrace{c_1\vec{v}_1, \dots, c_n\vec{v}_n}_{\in \ \ker T}) = \vec{0}$$
$$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = r_1\vec{u}_1 + \cdots + r_m\vec{u}_m \text{ for some } r_i \in \mathbb{F}$$
$$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n - r_1\vec{u}_1 - \cdots - r_m\vec{u}_m = \vec{0}$$

$\{\vec{v}_i, \vec{u}_j \mid i \leq u \leq m, 1 \leq j \leq n\}$ is a basis for V.
Thus, $c_1 = \cdots = c_n = r_1 = \cdots = r_n = 0$.
So $T(\vec{v}_1), \dots, T(\vec{v}_n)$ is linearly independent, and so it is a basis for V.
$\dim \text{Im } T = n$. ■

2.3 The Inverse of a Linear Transformation

2.4 Change of Basis

3. The Spectral Theorem in $\mathbb{R}^n$

3.1 Diagonalizability

3.1.1 Eigenvectors, Eigenvalues and Eigenspaces

Definition 4.1.2 — Eigenvector and Eigenvalue. Let V be a vector space over the field $\mathbb{F}$. Let $T : V \to V$ be a linear mapping.

a) A vector $\vec{v} \in V$ is called an *eigenvector of T* if $\vec{v} \neq \vec{0}$ and there exists a scalar $\lambda \in \mathbb{R}$ such that $T(\vec{v}) = \lambda\vec{v}$.

b) If $\vec{v}$ is an eigenvector of T and $T(\vec{v}) = \lambda\vec{v}$, the scalar λ is called the *eigenvalue of T corresponding to $\vec{v}$*.

Definition 4.1.6 — Eigenspace. $\forall\lambda \in \mathbb{F}$, the *$\lambda$-eigenspace of T*, denoted E_λ, is the set

$$E_\lambda = \{\vec{v} \in V \mid T(\vec{v}) = \lambda\vec{v}\} = \{\text{all eigenvectors of } \lambda\} \cup \{\vec{0}\}$$

That is, E_k is the set containing all the eigenvectors of T with eigenvalue λ, together with the vector $\vec{0}$. (If λ is not an eigenvalue of T, then we have $E_\lambda = \vec{0}$.)

■ **Example 3.1**

Consider the linear transformation $D: \begin{array}{ccc} \mathcal{C}^\infty(\mathbb{R}) & \to & \mathcal{C}^\infty(\mathbb{R}). \\ f & \mapsto & f' \end{array}$

- 1 is a eigenvector with a eigenvalue of 0.
- e^x is an eigenvector is with a eigenvalue of 1.
- e^{Mx} is an eigenvector with the eigenvalue of M.

■

■ **Example 3.2** Define the plane $W : x + y + z = 0$.

Consider the linear transformation $R: \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R}^3 \\ \vec{x} & \mapsto & \vec{x} \text{ reflected w.r.t. } W \end{array}$

$\forall\vec{x} \in W,\ R(\vec{v}) = 1\vec{v}$, so $\lambda = 1$ is an eigenvalue, and $E_1 = W$.

$R(\vec{w}) = -\vec{w}$ for all $\vec{w} \perp W$, so $\lambda = -1$ is an eigenvalue, and $E_{-1} = \mathrm{Sp}(\vec{w})$. ■

Proposition 4.1.7 E_λ is a subspace of V for all λ.

Proof.
- $T(\vec{0}) = \vec{0} = \lambda\vec{0} \implies \vec{x} \in E_\lambda$
- Pick $\vec{v}, \vec{w} \in E_\lambda$.
  WTS $\vec{v} + \vec{w} \in E_\lambda$.
  $T(\vec{v}+\vec{w}) = T(\vec{v}) + T(\vec{w}) = \lambda\vec{v} + \lambda\vec{w} = \lambda(\vec{v}+\vec{w}) \in E_\lambda$
- Pick $\vec{v} \in E_\lambda$, $r \in \mathbb{F}$.
  WTS $T(r\vec{v}) = rT(\vec{v})$.
  $T(r\vec{v}) = \lambda(r\vec{v}) = r(\lambda\vec{x}) = rT(\vec{v}) \in E_\lambda$.

■

Eigenbasis

Consider the same linear transformation R discussed in example 2.

$R(\vec{x}) = A\vec{x}$

$$A = [R]_{\mathcal{E}} = \begin{bmatrix} | & | & | \\ R(\vec{e}_1) & R(\vec{e}_2) & R(\vec{e}_3) \\ | & | & | \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{bmatrix}$$

Eigenbasis: linearly independent and spans $\mathbb{R}^3$.

$$\text{For example: } \mathcal{B} = \left\{ \underbrace{\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}}_{\vec{v}_1}, \underbrace{\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}}_{\vec{v}_2}, \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_{\vec{w}} \right\}$$

$$\text{Then, } [R]_{\mathcal{B}} = \begin{bmatrix} | & | & | \\ [R(\vec{v}_1)]_{\mathcal{B}} & [R(\vec{v}_2)]_{\mathcal{B}} & [R(\vec{w})]_{\mathcal{B}} \\ | & | & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$$[R(\vec{v}_1)]_{\mathcal{B}} = [\vec{v}_1]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$[R(\vec{v}_2)]_{\mathcal{B}} = [\vec{v}_2]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$[R(\vec{w})]_{\mathcal{B}} = [\vec{w}]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

Let $T : V \to V$ be a linear transformation.

$\vec{v} \in V$ is an eigenvector if and only if $T(\vec{v}) = \lambda\vec{v}$, by the definition.

$$\begin{aligned} T(\vec{v}) - \lambda\vec{v} &= \vec{0} \\ T(\vec{v}) - \lambda\ id(\vec{v}) &= \vec{0} \\ (T - \lambda\ id)\vec{v} &= 0 && \text{by sum of linear transformations} \\ \vec{v} &\in \ker(T - \lambda \text{ id}) && \text{for some } \lambda \in \mathbb{F} \end{aligned}$$

Proposition 4.1.5 A vector $\vec{v}$ is an eigenvector of T with eigenvalue λ if and only if $\vec{v} \neq \vec{0}$ and $\vec{v} \in \ker(T - \lambda\text{id})$.

■ Example 3.3

Let $\dim V = n$, and fix a basis $\mathcal{B}$ for V.

$$\begin{aligned} \text{Then, } \mathcal{L}(V,V) &\overset{\cong}{\to} \underset{n\times n}{M}(\mathbb{F}) \\ T &\mapsto [T]_{\mathcal{B}} = B \end{aligned}$$

Suppose $\vec{v}$ is an eigenvector of T.

$T(\vec{v}) - \lambda\vec{v}$ means that $[T]_{\mathcal{B}}[\vec{v}]_{\mathcal{B}} = [\lambda\vec{v}]_{\mathcal{B}} = \lambda[\vec{v}]_{\mathcal{B}}$.

$$\begin{aligned}[T - \lambda\ id]_{\mathcal{B}} &= [T]_{\mathcal{B}} - \lambda[id]_{\mathcal{B}} \\ &= B - \lambda I_{n\times n}\end{aligned}$$

λ is an eigenvalue for T iff $B - \lambda I$ is not invertible.
λ is an eigenvalue for T iff $\det(B - \lambda I) = 0$.
This is called that **characteristic polynomial** of the linear transformation T. ■

Definition 4.1.11 — Characteristic Polynomial. Let $A \in M_{n\times n}(\mathbb{R})$. The polynomial $\det(A = \lambda I)$ is called the *characteristic polynomial* of A.

■ **Example 3.4**

$$\begin{array}{rccc} T: & \mathcal{P}_3(\mathbb{R}) & \to & \mathcal{P}_3(\mathbb{R}) \\ & p & \mapsto & p' + 2p \end{array}$$

1. Find the characteristic polynomial of T.

   Pick the basis $\mathcal{B} = (1, x, x^2, x^3)$. Then, $B = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 2 \end{bmatrix}$.

   $$\operatorname{char}(T) = \det(B - \lambda I) = (\lambda - 2)^4 \quad [1]$$

   Note: $\operatorname{char}(T)$ is well defined (does NOT depend on the choice of basis).
2. Find all eigenvalues and the corresponding eigenspace of T.

   $$E_2 = \operatorname{nul}(B - \lambda I) = \operatorname{sp}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

   $\dim(E_2) = 1$ is the *geometric multiplicity* of $\lambda = 2$.

■

3.2 Diagonalizability

Definition 4.2.1 — Diagonalizable. Let V be a finite-dimensional vector space, and let $T : V \to V$ be a linear mapping. T is said to be *diagonalizable* if there exists a basis of V, all of whose vectors are eigenvectors of T.

Proposition 4.2.2 $T : V \to V$ is diagonalizable if and only if, for any basis α of V, the matrix $[T]_\alpha^\alpha$ is similar to a diagonal matrix.

Proof.

- $\to$: WTS T is diagonalizable $\implies$ $\exists$ a basis of $\mathcal{B}$ s.t. is diagonal.
  T has an eigenbasis $\mathcal{B} = (\vec{b}_1, \ldots, \vec{b}_n)$.
  Note that $T(\vec{b}_i) = \lambda_i\vec{b}_i$ for some $\lambda_i \in \mathbb{F}$.

$$[T]_{\mathcal{B}} = \begin{bmatrix} | & & | \\ [T(\vec{b}_1)]_{\mathcal{B}} & \ldots & [T(\vec{b}_n)]_{\mathcal{B}} \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ [\lambda_1\vec{b}_1]_{\mathcal{B}} & \ldots & [\lambda_n\vec{b}_n]_{\mathcal{B}} \\ | & & | \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & \vdots & & 0 \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

[1] Note that the power of 4 means the *algebraic multiplicity* of $\lambda = 2$ is 4.

- $\Leftarrow$: WTS $\exists$ a basis $\mathcal{B}$ of T s.t. $[T]_\mathcal{B}$ is diagonal $\implies$ T is diagonalizable.

  $\exists$ a basis $\mathcal{B} = (\vec{b}_1, \dots, \vec{b}_n)$ s.t. $[T]_\mathcal{B} = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix}$.

  $$[T]_\mathcal{B} = \begin{bmatrix} | & & | \\ [T(\vec{b}_1)]_\mathcal{B} & \dots & [T(\vec{b}_n)]_\mathcal{B} \\ | & & | \end{bmatrix} = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix}$$

  $$\forall i, [T(\vec{b}_i)]_\mathcal{B} = a_i \vec{e}_i$$
  $$[a_i \vec{b}_i]_\mathcal{B} = a_i \vec{e}_i$$

  So $T(\vec{b}_i) = a_i \vec{b}_i$, since $\gamma_\mathcal{B}$ is injective.

■

Proposition 3.0.1 — Diagonalizable. $T : V \to V$ is *diagonalizable* if and only if $\exists$ a basis $\mathcal{B}$ of V s.t. $[T]_\mathcal{B}$ is diagonal

Theorem 4.2.7 (Computational Criteria)
Let V be a vector space, $\dim V = n$.
Let $T : V \to V$ be a linear transformation.
Let $\lambda_1, \dots, , \lambda_n$ be **distinct** eigenvalues of T.

Let m_i be the algebraic multiplicity of λ_i.
Let M_i be the geometric multiplicity of λ_i.

T is diagonalizable if and only if

1. $\displaystyle\sum_{i=1}^{k} m_i = n$
2. $\forall i,\ M_i = m_i$.

- $\text{char}(T) = (x - \lambda_i)^{m_i} g(x)$
- $\dim(E_{\lambda_i}) = M_i$

■ **Example 3.5** Let $T : \mathcal{P}_2(\mathbb{R}) \to \mathcal{P}_2(\mathbb{R})$ be a linear transformation.
Let $\mathcal{A} = (x^2, x, 1)$ be a basis for $\mathcal{P}_2(\mathbb{R})$.

$$A = [T]_\mathcal{A} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\text{char}(T) = \det(A - \lambda I) = (2 - \lambda^2)(1 - \lambda)$$

$$E_2 = \text{Span}\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right)$$

$$E_1 = \text{Span}\left(\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right)$$

■

$$\mathrm{Nul}(A-2I) = \mathrm{Span}\left(\begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\1\end{pmatrix}\right)$$

- For $\lambda = 2$, the algebraic multiplicity is 2.
- For $\lambda = 1$, the algebraic multiplicity is 1.

$$2+1=3$$

The geometric multiplicity of $\lambda = 2$ is $\dim\ker(T-2\mathrm{id}) = \dim E_2$
$\dim(\mathrm{Nul}(A-2I)) = 2$

- The geometric multiplicity of $\lambda = 2$ is 2.
- The geometric multiplicity of $\lambda = 1$ is 1.

$$\mathrm{Nul}(A-2I) = \mathrm{Span}\left(\begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\1\end{pmatrix}\right) \underset{S.S.}{\subseteq} \mathbb{R}^3$$

$\ker(T-2\mathrm{id}) = E_2$

$$\begin{array}{ccc} E_2 & \underset{S.S.}{\subseteq} & \mathcal{P}_2(\mathbb{R}) \\ \cong & & \downarrow \lambda_{\mathcal{A}} \\ \mathrm{Nul}(A-2I) & \underset{S.S.}{\subseteq} & \mathbb{R}^3 \end{array}$$

$\dim \mathrm{Nul}(A-2I) = 2$
$\mathrm{Nul}(A-2I) \cong E_2$
$\dim E_2 = \dim \mathrm{Nul}(A-2I)$
Find an eigenbasis $\mathcal{B}$ for T.
Note: We need the eigenbasis for $\mathcal{P}_{\in}(\mathbb{R})$, not $\mathbb{R}^3$!
$\{x, -x^2+1, -x+1\}$ is an eigenbasis for T

$$\begin{bmatrix}2&0&0\\0&2&0\\0&0&1\end{bmatrix}$$

- $E_{\lambda_1} = \{x_1, \ldots, x_{l_1}\}$
- $E_{\lambda_2} = \{x_1, \ldots, x_{l_2}\}$
- $E_{\lambda_k} = \{x_1, \ldots, x_{l_1}\}$

$E_{\lambda_1} \cup \cdots \cup E_{\lambda_k}$ is an eigenbasis for T.

Proposition 4.2.6 Let V be finite-dimensional, and let $T : V \to V$ be linear.

Let A be an eigenvalue of T, and assume that A is an m-fold root of the characteristic polynomial of T. Then we have

$$1 \leq \dim(E_\lambda) \leq m$$

Proposition 4.2.4 Let $\vec{v}_1, \ldots, \vec{v}_n$ be eigenvectors corresponding to distinct eigenvalues of $T : V \to V$. Then $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is linearly independent.

Proof. We know that $T(\vec{v}_i) = \lambda_i \vec{v}_i$, $1 \leq i \leq k$ and $\lambda_i \neq \lambda_j$ for $i \neq j$.

WTS $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is linearly independent.

Rough Work

$k = 1$, $0 \neq \vec{v}_1$ is linearly independent.

$k = 2$, $\vec{v}_1, \vec{v}_2$, $T(\vec{v}_1) = T(\alpha\vec{v}_2) = \alpha T(\vec{v}_1) = \alpha\lambda_2\vec{v}_2 = \lambda_2\alpha\vec{v}_2 = \lambda_2\vec{v}_1$. Contradiction, because the λ's should be distinct.

Induction on k

- **Base case:** $k = 1$
  $0 \neq \vec{v}_1$ is linearly independent.
- **Induction hypothesis:** suppose $\vec{v}_1, \dots, \vec{v}_k$ is linearly independent.
  Prove that $\vec{v}_1, \dots, \vec{v}_k, \vec{v}_{k+1}$ is linearly independent.
  Suppose $a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k + a_{k+1}\vec{v}_{k+1} = \vec{0}$. (*)
  WTS $a_1 = \cdots = a_k = a_{k+1} = 0$.
  Apply T to (*):

$$a_1T(\vec{v}_1) + a_kT(\vec{v}_k) + a_{k+1}T(\vec{v}_{k+1}) = \vec{0}$$
$$a_1\lambda_1\vec{v}_1 + a_k\lambda_k\vec{v}_k + a_{k+1}\lambda_{k+1}\vec{v}_{k+1} \tag{1}$$

  Multiply (*) by λ_{k+1} (if $\lambda_{k_i} = 0$, switch with λ_k):

$$a_1\lambda_{k+1}\vec{v}_1 + \cdots + a_k\lambda_{k+1}\vec{v}_k + a_{k+1}\lambda_{k+1}\vec{v}_{k+1} \tag{2}$$

  $1 - 2$: $a_1(\lambda_1 - \lambda_{k+1})\vec{v}_1 + \cdots + a_k(\lambda_k - \lambda_{k+1})\vec{v}_k = \vec{0}$
  We have $a_1 = \cdots = a_k = 0$.

  Plug back into (*):
  $a_{k+1} = 0$. So $\vec{v}_1, \dots, \vec{v}_{k+1}$ are linearly independent.

■

3.3 Geonetry in $\mathbb{R}^n$

Let $\mathbb{R}^n = \begin{pmatrix} r_1 \\ r_1 \\ \vdots \\ r_n \end{pmatrix}$.

We know that $V \overset{\gamma\mathcal{B}}{\cong} \mathbb{F}^n$, $v \mapsto [v]_\mathcal{B}$ and $T \mapsto [T]_\mathcal{B}$
In $\mathbb{R}^3$ and $\mathbb{R}^3$, we have these concepts intuitively

- Norm
- Distance
- Angle

In $\mathbb{R}^n$,

- $\|\vec{v}\| = \sqrt{{v_1}^2 + \cdots + {v_n}^2} = \sqrt{\vec{v} \cdot \vec{v}}$
- $d(\vec{v}, \vec{w}) = \|\vec{v} - \vec{w}\|$
  - $\vec{v} \perp \vec{w} \iff \vec{v} \cdot \vec{w} = 0$
- $\theta = \cos^{-1} \frac{\vec{v} \cdot \vec{w}}{\|\vec{v}\|^2 \|\vec{w}\|^2}$

Definition 4.3.1 The *standard inner product* (or *dot product*) on $\mathbb{R}^n$ is the function $\langle , \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by the following rule. If $\vec{x} = (x_1, \dots, x_n)$ and $\vec{y} = (y_1, \dots, y_n)$ in standard coordinates, then $\langle \vec{x}, \vec{y} \rangle$ is the scalar: $\langle \vec{x}, \vec{y} \rangle = x_1y_1 + \cdots + x_ny_n$.

Definition 3.0.2 — Inner Product.

$$\begin{aligned} \langle\ ,\ \rangle :\ \ V \times V &\to \mathbb{F} \\ (\vec{v}, \vec{w}) &\mapsto \langle \vec{v}, \vec{w} \rangle \end{aligned}$$

An *inner product* on V is a function that takes each ordered pair of (u, v) of elements of V to a number $\langle u, v \rangle \in \mathbb{F}$ and has the following properties:

a) **Positivity**
$\langle v, v\rangle \geq 0$ for all $v \in V$;
b) **Definiteness**
$\langle v, v\rangle = 0$ if and only if $v = 0$
c) **Additivity in the first slot**
$\langle u+v, w\rangle = \langle u+w\rangle + \langle v+w\rangle$ for all $u, v, w \in V$;
d) **Homogeneity in the first slot**
$\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$ for all $u, v \in V$
e) **Conjugate symmetry**
$\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$.

Proposition 4.3.2
a) For all $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$ and all $c \in \mathbb{R}$, $\langle c\vec{x}+\vec{y}, \vec{z}\rangle = c\langle \vec{x}, \vec{z}\rangle + \langle \vec{y}, \vec{z}\rangle$. (In other words, as a function of the first variable, the inner product is a *linear* mapping.)
b) For all $\vec{x}, \vec{y} \in \mathbb{R}^n$, $\langle \vec{x}, \vec{y}\rangle = \langle \vec{y}, \vec{x}\rangle$ (i.e., the inner product is "*symmetric*").
c) For all $\vec{x} \in \mathbb{R}^n$, $\langle \vec{x}, \vec{x}\rangle \geq 0$, and $\langle \vec{x}, \vec{x}\rangle = 0$ if and only if $\vec{x} = \vec{0}$. (We say the standard inner product is "*positive-definite*".)

Proof. Many of the simpler facts about the standard inner product may be derived by expanding out the expressions involved, using Definition (4.3.1).
a) Write $\vec{x} = (x_1, \ldots, x_n)$, $\vec{y} = (y_1, \ldots, y_n)$, and $\vec{z} = (z_1, \ldots, z_n)$. Then $\langle c\vec{x}+\vec{y}\rangle = (cx_1 + y_1, \ldots, x_n + y_n)$, so we have

$$\langle \vec{x}+\vec{y}, \vec{z}\rangle = (cx_1 + y_1)z_1 + \cdots + (cx_n + y_n)z_n$$

After multiplying this out and rearranging the terms, we obtain

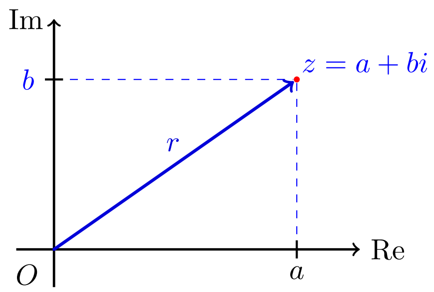
$$\langle \vec{x}+\vec{y}, \vec{z}\rangle = c(x_1z_1 + \cdots + x_nz_n) + \cdots + (y_1z_1 + \cdots + y_nz_n) = c\langle \vec{x}, \vec{z}\rangle + \langle \vec{y}, \vec{z}\rangle$$

b) We have:
$$\begin{aligned}\langle \vec{x}, \vec{y}\rangle &= x_1y_1 + \cdots + x_ny_n \\ &= y_1x_1 + \cdots + y_nz_n \\ &= \langle \vec{y}, \vec{x}\rangle.\end{aligned}$$

c) By the definition, $\langle \vec{x}, \vec{x}\rangle = x_1^2 + \cdots + x_n^2$. Since the components of $\vec{x}$ are all real numbers, each $x_i^2 \geq 0$, so $\langle \vec{x}, \vec{x}\rangle \geq 0$. Furthermore, if $\langle \vec{x}, \vec{x}\rangle = 0$, then it is clear that $x_i = 0$ for all i, or in other words, that $\vec{x} = \vec{0}$.

■

Note that $\mathbb{C} \cong \mathbb{R}^2$
Let $z \in \mathbb{C}$, $z = a + bi$



$$\|z\| = \sqrt{a^2+b^2}$$
$$\overline{z} = a - bi$$
$$a^2+b^2 = z\overline{z}$$
$$\|z\|^2 = a^2+b^2 = z\overline{z}$$
$$\|\vec{z}\|^2 = z_1\overline{z_1} + z_2\overline{z_2} + \cdots + z_n\overline{z_n} = \langle \vec{z}, \vec{z}\rangle_{\mathbb{C}}$$

■ **Example 3.6** The following are commonly seen examples of inner products.

- dot product in $\mathbb{R}^n$
  $\vec{x}\cdot\vec{y} = x_1y_1 + \cdots + x_ny_n$
- Euclidian inner product (or Hermitan inner product)
  $$\vec{z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$
  $$\langle \vec{z}, \vec{x}\rangle = z_1\overline{w_1} + \cdots + z_n\overline{w_n}$$

■

Definition 5.3.1 — Hermitian Inner Product. Let V be a complex vector space. A *Hermitian inner product* on V is a complex valued function on pairs of vectors in V, denoted by $\langle \vec{u}, \vec{v}\rangle \in \mathbb{V}$ for $\vec{u}, \vec{v} \in V$, which satisfies the following properties:

a) For all $\vec{u}, \vec{v}$, and $\vec{w} \in V$ and $a, b \in \mathbb{C}$, $\langle a\vec{u} + b\vec{v}, \vec{w}\rangle = a\langle \vec{u}, \vec{w}\rangle + b\langle \vec{v}, \vec{w}\rangle$,
b) For all $\vec{u}, \vec{v} \in V$, $\langle \vec{u}, \vec{v}\rangle = \overline{\langle \vec{v}, \vec{u}\rangle}$, and
c) For all $\vec{v} \in V$, $\langle \vec{v}, \vec{v}\rangle \geq 0$, and $langle\vec{v}, \vec{v}\rangle = 0$ implies $\vec{v} = \vec{0}$.

We call a complex vector space with a Hermitian inner product a *Hermitian inner product space*.

$$\langle z, w\rangle_{\mathbb{C}} \stackrel{?}{=} \overline{\langle w, z\rangle}$$ [2]

$$\langle \vec{z}, \vec{w}\rangle_{C} = \sum_{i=1}^{n} z_i\overline{w_i} = \sum_{i=1}^{n} \overline{w_i}z_i = \overline{\overline{\sum_{i=1}^{n} \overline{w_i}z_i}} = \overline{\sum_{i=1}^{n} \overline{\overline{w_i}z_i}} = \overline{\sum_{i=1}^{n} \overline{\overline{w_i}}\overline{z_i}} = \overline{\sum_{i=1}^{n} w_i\overline{z_i}} = \overline{\langle \vec{w}, \vec{z}\rangle_{\mathbb{C}}}$$

Definition 3.0.3 — Inner Product Space. An *inner product space* is a vector space V along with an inner product on V.

Definition 4.3.4 — Norm.

a) The *length* (or *norm*) of $\vec{v} \in \mathbb{R}^n$ is the scalar
$$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v}\rangle}$$
b) $\vec{v}$ is called a *unit vector* if $\|\vec{v}\| = 1$.

Definition 4.3.5 The angle, θ, between two nonzero vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$ is defined to be $\theta = \cos^{-1}\dfrac{\langle \vec{x}, \vec{y}\rangle}{\|x\|\cdot\|y\|}$.

Definition 4.3.7 — Orthogonal. Two vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$ are said to be *orthogonal* (or *perpendicular*) if $\langle \vec{x}, \vec{y}\rangle = 0$.

■ **Example 3.7** Define $V = \mathcal{C}^0[0,1] = \{f : [0,1] \to \mathbb{R} | f \text{ continuous}\}$

[2] $\overline{w+z} = \overline{w} + \overline{z}$.

$\langle f,g\rangle = \int_0^1 f(x)g(x)\,dx$

Show $\langle\ ,\ \rangle$ is an inner product on V.

- Show that $\langle\ ,\ \rangle$ has positivity
  $\langle f,f\rangle = \int_0^1 f^(x)\,dx \geq 0$. $\langle\ ,\ \rangle$ is positive
- Show that $\langle\ ,\ \rangle$ has definiteness
  $\int_0^1 f^2(x)\,dx = 0$ if and only if $f(x)=0$. $\langle\ ,\ \rangle$ is definite
- Show that $\langle\ ,\ \rangle$ has additivity and homogeneityin the first slot

$$\begin{aligned}
\langle rf+g,h\rangle &= \int_0^1 (rf+g)(x)h(x)\,dx \\
&= \int_0^1 (rf(x)+g(x))h(x)\,dx \\
&= \int_0^1 rf(x)h(x)+g(x)h(x)\,dx \\
&= r\int_0^1 f(x)h(x)\,dx + \int_0^1 g(x)h(x)\,dx \\
&= r\langle f,h\rangle + \langle g,h\rangle
\end{aligned}$$

- Show that $\langle\ ,\ \rangle$ has conjugate summery

$$\begin{aligned}
\langle g,h\rangle &= \int_0^1 g(x)h(x)\,dx \\
&= \int_0^1 \overline{h(x)g(x)}\,dx \\
&= \overline{\langle h,g\rangle}
\end{aligned}$$

■

Note that inner products have the following properties:

- $\langle\vec{0},\vec{u}\rangle = 0$ for every $\vec{u}\in V$.
- $\langle\vec{u},\vec{0}\rangle = 0$ for every $\vec{u}\in V$
- $\langle\vec{u},\vec{v}+\vec{w}\rangle = \langle\vec{x}+\vec{v}\rangle + \langle\vec{u}+\vec{w}\rangle$ for all $\vec{u},\vec{v},\vec{w}\in V$
- $\langle\vec{u},\lambda\vec{v}\rangle = \overline{\lambda}\langle\vec{u},\vec{v}\rangle$ for all $\lambda\in\mathbb{F}$ and $\vec{u},\vec{v}\in V$.

Let $\vec{v}\in V$

- $\|\vec{v}\| = 0$ if and only if $\vec{v}=\vec{0}$.
- $\|\lambda\vec{v}\| = |\lambda|\|\vec{v}\|$ for all $\lambda\in\mathbb{F}$.

Definition 3.0.4 — Trace.
Let $A\in M_{n\times n}(\mathbb{F})$. The *trace* of A, denoted $\mathrm{tr}(A)$, is defined by

$$\mathrm{tr}(A) = \sum_{i=1}^{\infty} a_{ij}, A_{n\times n} = a_{ij}$$

(the sum of the diagonal entries).

- $\mathrm{tr}(A+B) = \mathrm{tr}A + \mathrm{tr}B$
- $\mathrm{tr}(AB) = \mathrm{tr}(BA)$

■ **Example 3.8**

Let $V = \underset{2\times 2}{M}(\mathbb{R})$ be an inner product space.

Define $\langle\ ,\ \rangle_V : \begin{array}{ccc} VxV & \rangle & \mathbb{R} \\ (A,B) & \mapsto & \frac{1}{2}\mathrm{tr}(A^T B) \end{array}.$

Let $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$. What is $\|A\|$?

$$\begin{aligned} \|A\|^2 &= \langle A, A\rangle \\ &= \frac{1}{2}\mathrm{tr}(A^T A) \\ &= \frac{1}{2}(2) \\ &= 1 \end{aligned}$$

So $\|A\| = \sqrt{1} = 1$. ■

Definition 4.3.9 — Orthogonal (set) and Orthonormal.

a) A set of vectors $S \subset \mathbb{R}^n$ is said to be *orthogonal* if for every pair of vectors $\vec{x}, \vec{y} \in S$ with $\vec{x} \neq \vec{y}$, wc have $\langle \vec{x}, \vec{y}\rangle = 0$.

b) A set of vectors $S \subset \mathbb{R}^n$ is said to be *orthonormal* if S is orthogonal and, in addition, every vector in S is a unit vector.

Pythagorean Theorem

Suppose u and v are orthogonal vectors in V. Then

$$\|u + v\|^2 = \|u\|^2 + |v\|^2.$$

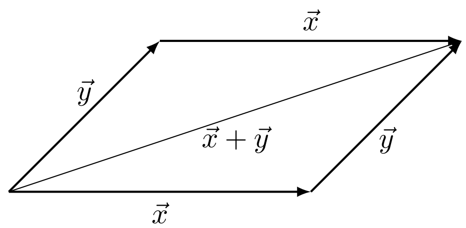
Proof. WTS $\vec{u}, \vec{v} \in V$ s.t. $\langle \vec{u}, \vec{v}\rangle = 0$. Then $\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$.

$$\begin{aligned} \|\vec{u} + \vec{v}\|^2 &= \langle \vec{u} + \vec{v}, \vec{u} + \vec{v}\rangle \\ &= \langle \vec{u}, \vec{u} + \vec{v}\rangle + \langle \vec{v} + \vec{u} + \vec{v}\rangle \\ &= \langle \vec{u}, \vec{u}\rangle + \underbrace{\langle \vec{u}, \vec{v}\rangle}_{0} + \underbrace{\langle \vec{v}, \vec{u}\rangle}_{0} + \langle \vec{v}, \vec{v}\rangle \\ &= \|\vec{u}\|^2 + \|\vec{v}\|^2 \end{aligned}$$

■

Proposition 4.3.4 — The Triangle Inequality and The Cauchy-Schwarz Inequality.

a) (The triangle inequality) For all vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$, $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$.
(Geometrically, in a triangle, the length of any side is less than or equal to the sum of the lengths of the other two sides.)



b) (The Cauchy-Schwarz inequality[3]) For all vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$,

$$|\langle \vec{x}, \vec{x} \rangle \leq \|\vec{x}\| \cdot \|\vec{y}\|.$$

$$\langle \vec{u}_i, \vec{u}_j \rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

■ **Example 3.9** Let $\mathbb{R}^2$ with the dot product be an inner product space.

$\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\}$ is a basis for $\mathbb{R}^2$.

$\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ is an orthogonal basis for $\mathbb{R}^2$.

$\left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \right\}$ is an orthonormal basis for $\mathbb{R}^2$. ■

■ **Example 3.10** Example of Inner Product Space

$W \underset{S.S.}{\subseteq} V, \ \dim W < \infty$

$$W^\perp = \{\vec{v} \in V, \langle \vec{v}, \vec{w} \rangle = 0 \ \forall \vec{w} \in V\}$$

$$V^\perp = ?$$

- $\langle \vec{0}, \vec{v} \rangle = 0$, so $\vec{0} \in V^\perp$
- $\vec{v} \in V^\perp$: $\langle \vec{v}, \vec{w} \rangle = 0 \ \forall \vec{w} \in V$
  $\langle \vec{v}, \vec{v} \rangle = 0 \implies \vec{v} = \vec{0}$

So $V^\perp = \{\vec{0}\}$

$$\{\vec{0}\}^\perp = ?$$

- $\{\vec{0}\}^\perp = \left\{ \vec{v} \in V \middle| \langle \vec{v}, \vec{0} \rangle = 0 \right\} = V$

■

Proposition 4.4.3 — Orthogonal Decomposition.

a) For every subspace W of $\mathbb{R}^n$, $W^\perp$ is also a subspace of $\mathbb{R}^n$.
b) We have $\dim(W) + \dim(W^\perp) = \dim(\mathbb{R}^n) = n$.
c) For all subspaces W of $\mathbb{R}^n$, $W \cap W^\perp = \{\vec{0}\}$.
d) Given a subspace W of $\mathbb{R}^n$, every vector $\vec{v} \in \mathbb{R}^n$ can be written **uniquely** as $\vec{v} = \vec{v}_1 + \vec{v}_2$, where $\vec{v}_1 \in W$ and $\vec{v}_2 \in W^\perp$. In other words, $\mathbb{R}^n = W \oplus W^\perp$.

3.4 Orthogonal Projection and the Gram-Schmidt Process

Definition 3.0.5 — Orthogonal Projection.

Let V be an inner product space. The orthogonal projection of the vector $\vec{v} \in V$ onto the subspace $W \subseteq V$, denoted P_W or $\text{proj}_W(\vec{v})$, is the mapping

$$P_W := proj_W : \begin{array}{ccc} V & \to & V \\ \underset{\vec{v}_W + \vec{v}_{W^\perp}}{\vec{v}} & \mapsto & \vec{v}_W \end{array}$$

where $\vec{v} = \vec{v}_W + \vec{v}_{W^\perp}$, with $\vec{v}_W \in W$ and $\vec{v}_{W^\perp} \in W^\perp$.

[3] This inequality if an equality if and only if one of u, v is a scalar multiple of the other.

Proposition 4.4.5

a) P_W is a linear mapping.
b) $\mathrm{Im}(P_W) = W$, and if $\vec{w} \in W$, then $P_W(\vec{w}) = \vec{w}$.
c) $\ker(P_W) = W^\perp$.

Proposition 4.4.6 Let $\{\vec{w}_1, \dots, \vec{w}_n\}$ be an orthonormal basis for the subspace $W \subseteq V$.

a) For each $w \in W$, we have

$$\vec{w} = \sum_{i=1}^{n} \langle \vec{w}, \vec{w}_i \rangle \vec{w}_i$$

b) For all $\vec{v} \in V$, we have

$$P_W(\vec{v}) = \sum_{i=1}^{n} \langle \vec{v}, \vec{w}_i \rangle \vec{w}_i$$

Proof. (idea) say $\{\vec{u}_1, \vec{u}_2\}$ $(\dim V = 2)$, $\vec{v} \in V$
$\vec{v} = r_1\vec{u}_1 + r_2\vec{u}_2,\ r_i \in \mathbb{F}$.

$$\begin{aligned}
\langle \vec{v}, \vec{u}_1 \rangle &= \langle r_1\vec{u}_1 + r_2\vec{u}_2, \vec{u}_1 \rangle \\
&= r_1\langle \vec{u}_1, \vec{u}_1 \rangle + r_2\langle \vec{u}_2, \vec{u}_1 \rangle \\
&= r_1 \cdot 1 + r_2 \cdot 0 \\
&= r_1
\end{aligned}$$

■

Lemma 3.0.6 Let $\mathcal{B} = (\vec{b}_1, \dots, \vec{b}_n)$ be a basis for W. $\vec{v} \in W^\perp \iff \forall \vec{b}_i \in \mathcal{B}, \langle \vec{v}, \vec{b}_i \rangle = 0$.

■ **Example 3.11**

P_W is a linear transformation.

- $\ker P_W = ?$
  $W^\perp$
  Let $\vec{w} \in W^\perp$, $P_W(\vec{w}) = P_W(\vec{0} + \vec{w}) = \vec{0}$, so $W^\perp \in \ker P_W$
  Let $\vec{w} \in \ker P_W$, so $P_W(\vec{w}) = \vec{0} \implies P_W(\vec{w}_W + \vec{w}_{W^\perp})$, so $\vec{w} = \vec{0} + \vec{w}_{W^\perp}$, so $\vec{w} = \vec{w}_{W^\perp}$, $\vec{w} \in W^\perp$.
- $\mathrm{Im}\ P_w = ?$
  W
  $$\vec{v} = \vec{v}_W + \vec{v}_{W^\perp}$$
  $$\mathrm{Im}(P_W) \subseteq W$$
  $\vec{w} \in W, P_W(\vec{w} + \vec{0}) = \vec{w}$
- Assume $\dim V < \infty$, what does rank-nullity say about P_w?
  $\dim V = \dim W + \dim W^\perp$

■

Proof. WTS $V = W + W^\perp$.
Suppose $\mathcal{U} = (\vec{u}_1, \dots, \vec{u}_m)$ is an orthonormal basis for W.
$\forall \vec{v} \in V$, $\vec{w} := \langle \vec{v}, \vec{u}_1 \rangle \vec{u}_1 + \langle \vec{v}, \vec{u}_2 \rangle \vec{u}_2 + \cdots + \langle \vec{v}, \vec{u}_m \rangle \vec{u}_m$
$\vec{v} = \underset{\in W}{\vec{w}} + \underset{\overset{?}{\in} W^\perp}{(\vec{v} - \vec{w})}$
WTS $\vec{v} - \vec{w}$ is in $W^\perp$.
WTS $\langle \vec{v} - \vec{w}, \vec{u} \rangle = 0 \forall \vec{u} \in W$.

This is true $\iff \langle \vec{v} - \vec{w}, \vec{u}_i \rangle = 0$ (by lemma).

$$\begin{aligned}
\langle \vec{v} - \vec{w}, \vec{u}_i \rangle &= \langle \vec{v}, \vec{u}_i \rangle - \langle \vec{w}, \vec{u}_i \rangle \\
&= \langle \vec{v}, \vec{u}_1 \rangle - \langle \langle \vec{v}, \vec{u}_1 \rangle \vec{u}_1 + \cdots + \langle \vec{v}, \vec{u}_i \rangle \vec{u}_i + \cdots + \langle \vec{v}, \vec{u}_m \rangle \vec{u}_m, \vec{u}_i \rangle \\
&= \langle \vec{v}, \vec{u}_i \rangle - \langle \langle \vec{v}, \vec{u}_i \rangle \vec{u}_i, \vec{u}_i \rangle \\
&= \langle \vec{v}, \vec{u}_i \rangle - \langle \vec{v}, \vec{u}_i \rangle \\
&= 0
\end{aligned}$$

$$\langle \vec{v}, \vec{u}_i \rangle \langle \vec{u}_i, \vec{u}_i \rangle = \begin{cases} \langle \vec{v}, \vec{u}_i \rangle \\ 0 \end{cases}$$

■

Definition 4.4.8 — Gram-Schmidt Procedure. By induction, suppose that we have already constructed the vectors $\{\vec{v}_1, \ldots, \vec{v}_k\}$. Note that the vectors $\left\{ \frac{\vec{v}_1}{\|\vec{v}_1\|}, \ldots, \frac{\vec{v}_j}{\|\vec{v}_j\|} \right\}$form an orthonormal basis for the subspace W_j. Then we have

$$P_{W_j}(\vec{v}) = \sum_{i=1}^{j} \frac{\langle \vec{v}_i, \vec{v} \rangle}{\langle \vec{v}_i, \vec{v}_i \rangle} \vec{v}_i$$

Substituting this expression into our formula for $\vec{v}_{j+1}$, w obtain

$$\begin{aligned}
\vec{v}_1 &= \vec{u}_1 \\
\vec{v}_2 &= \vec{u}_2 - \frac{\langle \vec{v}_1, \vec{u}_2 \rangle}{\langle \vec{v}_1, \vec{v}_1} \vec{v}_2 \\
&\vdots \\
\vec{v}_{j+1} &= \vec{u}_{j+1} - \sum_{i=1}^{j} \frac{\langle \vec{v}_i, \vec{v} \rangle}{\langle \vec{v}_i, \vec{v}_i \rangle} \vec{v}_i
\end{aligned}$$

Theorem 4.4.9 Every finitely generated inner product space has an orthonormal basis.

Proof. (idea)

$W \underset{S.S.}{\subseteq} V$, $\dim W < \infty$.

Let $\mathcal{B} = (\vec{b}_1, \ldots, \vec{b}_m)$ be a basis for W.

Apply the Gram–Schmidt process, we get $\mathcal{U} = (\vec{u}_1, \ldots, \vec{u}_m)$, an orthonormal basis.

- $\dim W = 1$
  $\mathcal{B} = (\vec{b}_1)$
  $\vec{u}_1 = \dfrac{\vec{b}}{\|\vec{v}\|}$
- $\dim W = 2$
  $\mathcal{B} = (\vec{b}_1, \vec{b}_2)$
  $\vec{u}_1 = \dfrac{\vec{b}_1}{\|\vec{b}_1\|}$
  Take $\vec{p}_2 = \vec{b}_2^{\perp} = \vec{b}_2 + \mathrm{proj}_{\mathrm{Sp}(\vec{b}_1)} \vec{b}_2$, $\vec{u}_2 = \dfrac{\vec{p}_2}{\|\vec{p}_2\|}$
- $\dim W = 3$
  $\mathcal{B} = (\vec{b}_1, \vec{b}_2, \vec{b}_3)$

$$\vec{u}_1 = \frac{\vec{b}_1}{\|\vec{b}_1\|} <$$
$$\vec{p}_2 = \vec{b}_2 - \mathrm{proj}_{\mathrm{Sp}(\vec{b}_1)}\vec{b}_2,\ \vec{u}_2 = \frac{\vec{p}_2}{\|\vec{p}_2\|}$$
$$\vec{p}_3 = \vec{b}_3 - \mathrm{proj}_{\mathrm{Sp}(\vec{b}_1,\vec{b}_2)}\vec{b}_3,\ \vec{u}_3 = \frac{\vec{p}_3}{\|\vec{p}_3\|}$$

- $\cdots$

If $\vec{u}_1, \ldots, \vec{u}_m$ is an orthonormal basis for $W \underset{S.S.}{\subseteq} V$, $\forall \vec{v} \in V$, $\mathrm{proj}_W \vec{v} - \langle \vec{v}, \vec{u}_1 \rangle \vec{u}_1 + \cdots + \langle \vec{v}, \vec{u}_m \rangle \vec{u}_m$.

$$\vec{v} = \underset{\in W}{\vec{w}} + \underset{\in W}{\vec{v} - \vec{w}}$$

■

■ **Example 3.12** $\mathcal{P}_2(\mathbb{R}) \in \mathcal{C}^0(\mathbb{R})$.

$\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$

$\mathcal{B} = (1 + x, x^2) \xrightarrow{G-S} (u_1, u_2, u_2)$ prthonormal basis for $\mathcal{P}_2(\mathbb{R})$.

$\|1\| = \sqrt{\int_0^1 1 \cdot 1\,dx}$

$\|x\| = 1$ ■

Theorem 3.0.7 $\forall \vec{v} \in V$, $\|\vec{v} - \mathrm{proj}_W \vec{v}\| \leq \|\vec{v} - \vec{w}\|\ \forall \vec{w} \in W$.

Proof.

$$\|\vec{v} - \mathrm{proj}_W \vec{v}\|^2 \leq \|\vec{v} - \mathrm{proj}_W \vec{v}\|^2 + \|\mathrm{proj}_W \vec{v} - \vec{w}\|^2$$
$$\|\vec{v} - \mathrm{proj}_W \vec{v}\|^2 \leq \|\vec{v} - \vec{w}\|^2$$

■

■ **Example 3.13**

Previously we have:

$\mathcal{P}_2(\mathbb{R}) \underset{S.S.}{\subseteq} \mathcal{C}^0(\mathbb{R})$.

$\mathcal{B} = (1, x, x^2) \xrightarrow{G-S} \mathcal{U} = (u_1, u_2, u_3)$

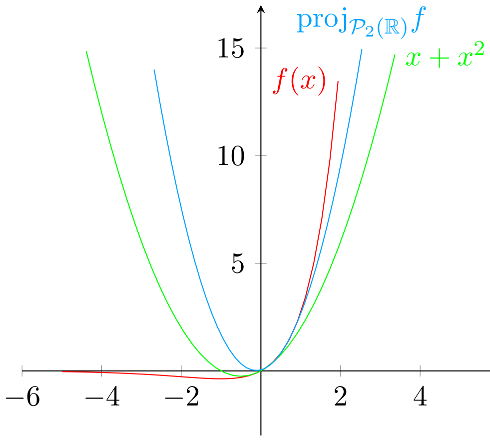
- $\vec{u}_1 = 1$
- $\vec{u}_2 = 2\sqrt{3}(x - \frac{1}{2})$
- $\vec{u}_3 = 6\sqrt{5}(x^2 - x + \frac{1}{6})$

Let $f(x) = xe^x \in \mathcal{C}^0(\mathbb{R})$.

Find the closest polynomial in $\mathcal{P}_2(\mathbb{R})$ to $f(x)$.

$$\mathrm{proj}_{\mathcal{P}_2(\mathbb{R})} f = \sum_{i=1}^{3} \frac{\langle f, \vec{u}_i \rangle}{\langle \vec{u}_i, \vec{u}_i \rangle} \vec{u}_i = \sum_{i=1}^{3} \langle f, \vec{u}_i \rangle \vec{u}_i \approx \frac{1}{500}(1\ 062x^2 + 246x + 23)$$

Tayler appension (for $x = 2$) would be $x + x^2$



■

Example 3.14

Let $T : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R}^3 \\ \vec{x} & \mapsto & A\vec{x} \end{array}$ be a linear transformation.

Let $A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}$.

$\text{char}(T) = (2 - 2\lambda)(1 - \lambda)$.

$E_2 = \text{Sp}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ (v1 and v2)

$E_1 = \text{Sp}\left\{ \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$ (v3)

$\mathcal{U} = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$.

$[T]_{\mathcal{U}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

$E_1 \not\perp E_2$.

$\mathbb{R}^3 \ni \vec{v} = c_1\vec{v}_1 + c_2\vec{v}_3 + c_3\vec{v}_3$ uniquely.

$\mathbb{R}^3 = E_1 \oplus E_2$. ■

Example 3.15

Let $S : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R}^3 \\ \vec{x} & \mapsto & D\vec{x} \end{array}$ be a linear transformation.

Let $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

$\mathcal{E} = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ is an eignebasis.

- $S(\vec{e}_1) = 2\vec{e}_1$
- $S(\vec{e}_2) = 2\vec{e}_2$
- $S(\vec{e}_3) = 1\vec{e}_3$

$\mathbb{R}^3 = E_2 \oplus E_1$. ■

Example 3.16

Let $L : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R}^3 \\ \vec{x} & \mapsto & B\vec{x} \end{array}$ be a linear transformation.

Let $B = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$.

$\operatorname{char}(T) = (\lambda - 1)(\lambda - \sqrt{3})(\lambda + \sqrt{3})$.

- $E_1 = \operatorname{Sp}\left\{\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right\}$
- $E_{\sqrt{3}} = \operatorname{Sp}\left\{\begin{pmatrix} \sqrt{3} \\ -1 \\ 1 \end{pmatrix}\right\}$
- $E_{-\sqrt{3}} = \operatorname{Sp}\left\{\begin{pmatrix} -\sqrt{3} \\ -1 \\ 1 \end{pmatrix}\right\}$

L is diagonalizable because it has 3 distinct eigenvalues.

$\mathbb{R}^3 = E_1 \oplus E_{\sqrt{3}} \oplus E_{-\sqrt{3}}$

$E_i \perp E_j \ \forall i \neq j \leq 3$.

Note that B is symmetric since $B = B^T$. ■

■ **Example 3.17** Let $T : \begin{array}{ccc} \mathbb{C}^2 & \to & \mathbb{C}^2 \\ \vec{x} & \mapsto & \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \vec{x} \end{array}$ be a linear transformation.

$\operatorname{char}(T) = \lambda(\lambda - 2)$.

- $E_0 = \operatorname{Sp}\left\{\begin{pmatrix} -i \\ 1 \end{pmatrix}\right\}$
- $E_1 = \operatorname{Sp}\left\{\begin{pmatrix} i \\ 1 \end{pmatrix}\right\}$

T is diagonalizable because it has 2 distinct eigenvalues.

$\mathbb{C}^2 = E_0 \oplus E_2$.

$\langle \begin{pmatrix} -i \\ i \end{pmatrix}, \begin{pmatrix} i \\ 1 \end{pmatrix} \rangle = (-i)\overline{(i)} + (1)\overline{(1)} = -1 + 1 = 0$, so $E_1 \perp E_2$.

Note that if $A = (a_{ij})$, then $A^T = (a_{ji})$ and $A^* = (\overline{a_{ji}})$ (conjugate transpose).

$\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}^* = \begin{bmatrix} \overline{1} & \overline{i} \\ \overline{-i} & \overline{1} \end{bmatrix} = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$ ■

3.5 Spectral Theorem

Theorem 5.3.12 Let $T : V \to V$ be a self-adjoint transformation of a complex vector space V with Hermitian inner product $\langle\ ,\ \rangle$. Then there is an orthonormal basis of V consisting of eigenvectors for T and, in particular, T is diagonalizable.

Theorem 4.6.1 Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a symmetric linear mapping. Then there is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of T. In particular, T is diagonalizable.

$T : V \to V$, V inner product space TFAE[4].

1. T is diagonalizable AND eigenspaces of T are orthogonal.
2. $V = E_{\lambda_1} \oplus \cdots + E_{\lambda_k}$ AND $E_{\lambda_i} \perp E_{\lambda_j}$ for $i \neq j$, $\lambda_1, \ldots, \lambda_k$ are eigenvalues.
3. T has an orthonormal eigenbasis.

[4] This means "The Following Are Equivalent".

- $2 \to 3$
  $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$, $E_{\lambda_i} \perp E_{\lambda_j}$
  $\overset{G-S}{\rightsquigarrow} \mathcal{U}_1 \cup \cdots \cup \mathcal{U}_k$ is an orthonormal eignebasis.
- $3 \to 2$
  $(\vec{v}_1, \ldots, \vec{v}_n)$ [5] is an orthonormal eigenbais, $\vec{v} \in V$.
  $\vec{v} = r_1\vec{v}_1 + r_2\vec{v}_2 + \cdots + r_n\vec{v}_n$
  $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_n}$

Definition 5.3.7 — Adjoint. Let V be a finite dimensional Hermitian inner product space and let α be an **orthonormal** basis for V. The *adjoint* of the linear transformation $T : V \to V$ is the linear transformation T^* whose matrix with respect to the orthonormal basis α is the matrix $([\overline{T}]_\alpha^\alpha)^T$; that is, $[T^*]_\alpha^\alpha = ([\overline{T}]_\alpha^\alpha)^T$.

Proposition 5.3.8 Let V be a finite dimensional Hermitian inner product space. The adjoint of $T : V \to V$ satisfies $\langle T(\vec{v}), \vec{w}\rangle = \langle \vec{v}, T(\vec{w})\rangle$ for all $\vec{v}$ and $\vec{w} \in V$.

Questions:

- Does T^* exist?
- Is T^* unique?

Definition 5.3.9 — Self-adjoint. $T : V \to V$ is called *Hermitian* or *self-adjoint* if $\langle T(\vec{u}), \vec{v}\rangle = \langle \vec{u}, T(\vec{v})\rangle$ for all $\vec{u}$ and $\vec{v} \in V$. Equivalently, T is Hermitian or self-adjoint if $T = T^*$ or $[\overline{T}]_\alpha^\alpha = [T]_\alpha^\alpha$ for an orthonormal basis α. An $n \times n$ complex matrix is called *Hermitian* or *self-adjoint* if $A = A^*$.

- $A \in M_n(\mathbb{R})$, $A = A^T \implies T_A$ is self-adjoint
- $B \in M_n(\mathbb{C})$, $B = B^* \implies T_B$ is self-adjoint.

Observation 1

Let $T : \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^n \\ \vec{x} & \mapsto & A\vec{x} \end{array}$ be a linear transformation.

If $A = A^T$, then T is self-adjoint.

$\forall \vec{v}, \vec{w} \in \mathbb{R}^n$.

$\langle T(\vec{v}), \vec{w}\rangle = \langle A\vec{v}, \vec{w}\rangle = (A\vec{v}) \cdot \vec{w}$.

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

That is, $\vec{x} \cdot \vec{y} = \vec{x}^T \vec{y}$

$(AB)^T = B^T A^T$

[5] $\vec{v}_i$'s are eigenvectors corresponding to the eigenvales.

$$\begin{aligned}
\langle T(\vec{v}), \vec{w}\rangle &= \langle A\vec{v}, \vec{w}\rangle \\
&= (A\vec{v}) \cdot \vec{w} \\
&= (A\vec{v})^T \vec{w} \\
&= \vec{v}^T A^T \vec{w} \\
&= \vec{v}^T A \vec{w} \\
&= \vec{v} \cdot \underbrace{(A\vec{w})}_{T(\vec{w})} \\
&= \langle \vec{v}, T(\vec{w})\rangle
\end{aligned}$$

Theorem 3.0.8 Let $T \in \mathcal{L}(V, V)$.
Let $\mathcal{U} = (\vec{u}_1, \dots, \vec{u}_n)$ be an orthonormal basis for V.
Let $A = [T]_{\mathcal{U}}$.
Then, $[T^*]_{\mathcal{U}} = A^*$ [a].

[a] For $\mathbb{F} = \mathbb{R}$, $M^* = M^T$,

Observation 2
Let $T: \begin{array}{ccc} \mathbb{C}^2 & \to & \mathbb{C}^2 \\ \vec{x} & \mapsto & B\vec{x} \end{array}$ be a linear transformation.
$B = B^* = \overline{B^T}$, B is Hermitian.
Then, T is self adjoint.
That is, $\langle T(\vec{v}), \vec{w}\rangle_{\mathbb{C}} = \langle \vec{v}, T(\vec{w})\rangle_{\mathbb{C}}$ [6].

$$\vec{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}, \vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{C}^n$$

$\langle \vec{z}, \vec{w}\rangle = z_1\overline{w_1} + \cdots + z_n\overline{w_n} = \vec{w}^*\vec{z}$

That is, $\langle \vec{z}, \vec{w}\rangle_{\mathbb{C}} = \vec{w}^*\vec{z}$.

$(BA)^* = A^*B^*$

$$\begin{aligned}
\langle \vec{v}, T(\vec{w})\rangle_{\mathbb{C}} &= \langle \vec{v}, B\vec{w}\rangle_{\mathbb{C}} \\
&= (B\vec{w})^*\vec{v} \\
&= \vec{w}^* B^* \vec{v} \\
&= \vec{w}^*(B\vec{v}) \\
&= \langle B\vec{v}, \vec{w}\rangle_{\mathbb{C}} \\
&= \langle T(\vec{v}), \vec{w}\rangle
\end{aligned}$$

Proof. Let $A = [T]_{\mathcal{U}} = \begin{bmatrix} | & | & & | \\ [T(\vec{u}_1)]_{\mathcal{U}} & [T(\vec{u}_2)]_{\mathcal{U}} & \cdots & [T(\vec{u}_n)]_{\mathcal{U}} \\ | & | & & | \end{bmatrix}$, where $\mathcal{U} = (\vec{u}_1, \dots, \vec{u}_n)$ is an orthonormal basis for V.

- The $i - j^{th}$ entry of $[T(\vec{u}_j)]_{\mathcal{U}}$:
  - $[T(\vec{u}_j)]_{\mathcal{U}} = \sum_{k=1}^{n} \langle T(\vec{u}_j), \vec{u}_k\rangle \vec{u}_k$

[6] Recall that $\langle u, v\rangle_{\mathbb{C}} = u^*v = \overline{u^T}v = u^T\overline{w}$.

- The $i-j^{th}$ entry of A is $\langle T(\vec{u}_j), \vec{u}_i\rangle$
- The $i-j^{th}$ entry of $[T^*(\vec{u}_j)]_{\mathcal{U}}$:
  - $[T^*(\vec{u}_j)]_{\mathcal{U}} = \sum_{k=1}^{n} \langle T^*((\vec{u}_j), \vec{u}_k\rangle \vec{u}_k$
  - The $i-j^{th}$ entry of $[T^*(\vec{u}_j)]_{\mathcal{U}}$ is $\langle T^*(\vec{u}_j), \vec{u}_i\rangle$
- The $i-j^{th}$ entry of A is $\langle T(\vec{u}_j), \vec{u}_i\rangle$
- The $i-j^{th}$ entry of A^T is $\langle T(\vec{u}_i), \vec{u}_j\rangle$
- The $i-j^{th}$ entry of $\overline{A}^T = A^*$ is $\overline{\langle T(\vec{u}_i), \vec{u}_j\rangle} = \overline{\langle \vec{u}_i, T^*(\vec{u}_j)\rangle} = \langle T^*(\vec{u}_j), \vec{u}_i\rangle = i-j^{th}$ column of $[T^*]_{\mathcal{U}}$

■

Corollary 3.0.9 $T \in \mathcal{L}(V)$ is self-adjoint if and only if $[T]_{\mathcal{U}} = ([T]_{\mathcal{U}})^*$.

Proposition 5.3.10 If λ is an eigenvalue of the self-adjoint linear transformation T, then $\lambda \in \mathbb{R}$.

Proof. Let λ be an eigenvalue.
$T(\vec{v}) = \lambda\vec{v}$ for some $\vec{v} \in V$.
$\langle \lambda\vec{v}, \vec{v}\rangle = \lambda\langle T(\vec{v}), \vec{v}\rangle = \langle \vec{v}, T(\vec{v})\rangle = \langle \vec{v}, \lambda\vec{v}\rangle = \overline{\lambda}\langle \vec{v}, \vec{v}\rangle$.
$(\lambda - \overline{\ })\langle \vec{v}, \vec{v}\rangle = 0$.
Since $\langle \vec{v}, \vec{v}\rangle \neq 0$, $\lambda = \overline{\lambda} \implies \lambda \in \mathbb{R}$. ■

■ **Example 3.18**

Is $\lambda = 3 - i$ an eigenvalue of $A = \begin{bmatrix} 1 & -2 & 3 \\ -2 & 5 & 2 \\ 3 & 2 & 7 \end{bmatrix}$?

No. A is a symmetric, so all of its eigenvectors must be in real. ■

Proposition 5.3.11 If $\vec{u}$ and $\vec{v}$ are eigenvectors, respectively, for the distinct eigenvalues λ and μ of $T : V \to V$, then $\vec{u}$ and $\vec{v}$ are orthogonal, $\langle \vec{u}, \vec{v}\rangle = 0$.

Proof. Let $\lambda_1 \neq \lambda_2$ be eigenvalues.
$T(\vec{v}_1) = \lambda_1\vec{v}_1$, and $T(\vec{v}_2) = \lambda_2\vec{v}_2$, $v_i \neq 0$, $\vec{v} \in V$.
$\langle \lambda_1\vec{v}_1, \vec{v}_2\rangle = \langle T(\vec{v}_1), \vec{v}_2\rangle = \langle \vec{v}_1, T(\vec{v}_2\rangle = \langle v_1, \lambda_2\vec{v}_2)\rangle = \overline{\lambda_2}\langle \vec{v}_1, \vec{v}_2\rangle$.
Then, $(\lambda_1 - \lambda_2)\langle \vec{v}_1, \vec{v}_2\rangle = 0$.
$\lambda_1 - \lambda_2 \neq 0 \to \langle \vec{v}_1, \vec{v}_2\rangle = 0$. ■

■ **Example 3.19**

Let $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$.

Let $\vec{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ be an eigenvector.

A must have another eigenvector of $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. ■

Proposition 3.0.10 Let $T \in \mathcal{L}(V)$ be self-adjoint. $U \underset{S.S.}{\subseteq} V$ s.t. $\underbrace{T(U) \in U}_{U \text{ is invariant}}$.

1. $T(U^{\perp}) \underset{S.S.}{\subseteq} U^{\perp}$

2. $T\Big|_U : U \to U$ s.t. $\vec{u} \mapsto T(\vec{u})$ [7] is self-adjoint.
3. $T\Big|_{U^\perp} : U^\perp \to_\perp U^\perp$ is self-adjoint.

Proof. 1. Let $\vec{v} \in U^\perp$. WTS $T(\vec{v}) \in U^\perp$.
WTS $\langle T(\vec{v}), \vec{u}\rangle = 0$ for all $\vec{u} \in U$.
$\langle T(\vec{v}), \vec{u}\rangle = \langle \underset{\in U^\perp}{\vec{v}}, \underset{\in U}{T(\vec{u})}\rangle = 0$

2. Exercise
3. Exercise

■

Theorem 3.0.11 — Spectral Theorem.

- Spectral Theorem ($\mathbb{F} = \mathbb{C}$)
  $T \in \mathcal{L}(V)$ is self-adjoint $\implies$ T admits an orthonormal eigenbasis.
- Spectral Theorem ($\mathbb{F} = \mathbb{R}$)
  $T \in \mathcal{L}(V)$ is self-adjoint $\iff$ T admits an orthonormal eigenbasis.

■ **Example 3.20**

$T_A : \mathbb{C}^3 \to \mathbb{C}^3$ defined by $\vec{x} \mapsto A\vec{x}$.

$$A = \begin{bmatrix} i & 1 & -1 \\ 0 & 2+i & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

- True of False:
  - $Sp(\vec{e}_1) \underset{S.S.}{\subseteq} \mathbb{C}^3$ is invariant.
    True
    $\vec{v} \in Sp(\vec{e}_1) \implies k\vec{e}_1$
    $$T(\vec{e}_1) = \begin{pmatrix} i \\ 0 \\ 0 \end{pmatrix} = i\vec{e}_1 \in Sp(\vec{e}_1)$$
    $T_A(\vec{v}) = T_A(k\vec{e}_1) = ki\vec{e}_1 \in Sp(\vec{e}_1)$
  - $Sp(\vec{e}_1, \vec{e}_2) \underset{S.S.}{\subseteq} \mathbb{C}^3$ is invariant.
    $\vec{v} \in Sp(\vec{e}_1, \vec{e}_2), T(\vec{v}) \overset{?}{\in} Sp(\vec{e}_1, \vec{e}_2)$
    $$T(\vec{e}_1) = \begin{pmatrix} i \\ 0 \\ 0 \end{pmatrix} \in Sp(\vec{e}_1) \subset Sp(\vec{e}_1, \vec{e}_2)$$
    $$T(\vec{e}_2 = \begin{pmatrix} 1 \\ 2+i \\ 0 \end{pmatrix}) \in Sp(\vec{e}_1, \vec{e}_2) <\mathrm{br}> T(\vec{v}) = T(r\vec{e}_1 + s\vec{e}_2) \in Sp(\vec{e}_1, \vec{e}_2).$$
  - $Sp(\vec{e}_1, \vec{e}_2, \vec{e}_3) = \mathbb{C}^3 \underset{S.S.}{\subseteq} ?$ is invariant.
    True
- Give an example of a non-invariant subspace of $\mathbb{C}^3$.
  $Sp(\vec{e}_1, \vec{e}_3) \underset{S.S.}{\subseteq} \mathbb{C}^3$

[7] $T\Big|_U$ means T restricted to U.

$$T(\vec{e}_3) = \begin{pmatrix} -i \\ 1 \\ 3 \end{pmatrix} = (-i)\vec{e}_1 + (1)\vec{e}_2 + 3\vec{e}_3 \notin Sp(\vec{e}_1, \vec{e}_2).$$

■

■ **Example 3.21** $\dim V < \infty$, $T \in \mathcal{L}(V)$. λ is an eigenvalue of T.

- True or False: $\mathrm{Im}(T - \lambda\mathrm{id}) \underset{S.S.}{\subsetneq} V$

  True

  $\exists \vec{v} \neq \vec{0} \in \ker(T - \lambda\mathrm{id})$

  $\dim\ker(T - \lambda\mathrm{id}) \geq 1 \implies \dim\mathrm{Im}(T - \lambda\mathrm{id}) < \dim V$, $\mathrm{Im}(T - \lambda\mathrm{id}) \subsetneq V$.
- True or False: $Im(T - \lambda id)$ is invariant.

  True

  Let $\vec{u} \in \mathrm{Im}(T - \lambda\mathrm{id})$, $T(\vec{u}) \overset{?}{\in} \mathrm{Im}(T - \lambda\mathrm{id})$

  $T(\vec{u}) = \underbrace{T(\vec{u}) - \lambda\vec{u}}_{(T-\lambda id)\vec{u} \in \mathrm{Im}(T-\lambda\mathrm{id})} + \underbrace{\lambda\vec{u}}_{\in \mathrm{Im}(T-\lambda\mathrm{id})} \in \mathrm{Im}(T - \lambda\mathrm{id}).$

■

4. Jordan Canonical Form

4.0 Moltivation

Let $T : V \to V$ be a linear transformation.

- Let $\mathcal{B}$ be an eigenbasis of V.
$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$
- Let $\mathcal{U}$ be an orthonormal eigenbasis of V.
$$[T]_{\mathcal{U}} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$
- $E_{\lambda_j} \perp E_{\lambda_j}$, $i \neq j$

We know that if T is self-adjoint then an orthonormal eigenbasis exists. But what if there is no eigenbasls? Then we need something called a Jordan basis.

Let $\mathcal{B}$ be a Jordan basis of V. Then, $[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 1 & & & & \\ 1 & \lambda_1 & & & & \\ & & \lambda_2 & 1 & 0 & \\ & & & \lambda_2 & 1 & \\ & & & & \lambda_2 & \\ & & & & & \lambda_3 \end{bmatrix}$

Theorem 4.0.1 ($\mathbb{F} = \mathbb{C}$)
$\dim V < \infty$.
Let $T \in \mathcal{L}(V)$. There exists a basis $\mathcal{B}$ s.t. $[T]_{\mathcal{B}}$ is upper triangular.

Proof. Prove by induction on $\dim V$

- $\dim V = 1$
  $\mathcal{B} = \{\vec{b}_1\}$
  $[T]_{\mathcal{B}} = [a]_{1\times 1}$ is upper triangular.
- Induction hypothesis: all $T \in \mathcal{L}(W)$, $\dim W < n$ have an upper triangular matrix representation.
  Let $\dim V = n$, $T \in \mathcal{L}(V)$.
  Let λ be an eigenvalue of T.
  $U = Im(T - \lambda id) \underset{S.S.}{\subsetneq} V$ is invariant.
  $T\Big|_U : U \to U$, $\dim U \lneq n$.
  By induction hypothesis, $\exists \mathcal{U} = \{\vec{u}_1, \ldots, \vec{u}_m\}$ be a basis for U s.t. $[T\Big|_U]_{\mathcal{U}}$ is a diagonal matrix.
  IDEA: extend $\mathcal{U}$ to a basis for V.
  $\mathcal{B} = (\vec{u}_1, \ldots, \vec{u}_m, \vec{v}_1, \ldots, \vec{v}_e)$, $m + e = m$
  Claim $[T]_{\mathcal{B}}$ is upper triangular.

■

4.1 Triangular Form

Definition 6.1.1 — Invariant. Let $T : V \to V$ be a linear mapping. A subspace $W \subset V$ is said to be *invariant* (or *stable*) under T if $T(W) \subset W$.

■ **Example 4.1**
$\mathrm{Im}(T - \lambda I) \subsetneq V$ invariant. ■

Proposition 6.1.4 — Triangularizable. We say that a linear mapping $T : V \to V$ on a finite-dimensional vector space V is *triangularizable* if there exists a basis β uch that $[T]_\beta^\beta$ is upper-triangular.

Let $\mathcal{B} = (\vec{v}_1, \cdots, \vec{v}_n)$ be basis for V.

Assume $[T]_{\mathcal{B}} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{bmatrix}$

Then, $[T(\vec{v}_1)]_{\mathcal{B}} = [T]_{\mathcal{B}}[\vec{v}_1]_{\mathcal{B}} = \begin{pmatrix} b_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

- $T(\vec{v}_1) = b_{11}\vec{v}_1 \in \mathrm{Sp}(\vec{v}_1)$
- $T(\vec{v}_2) = b_{12}\vec{v}_1 + b_{22}\vec{v}_2 \in \mathrm{Sp}(\vec{v}_1, \vec{v}_2)$.
- $\cdots$
- $T[\vec{v}_n] = b_{1n}\vec{v}_1 + \cdots + b_{nn}\vec{v}_n \in \mathrm{Sp}(\vec{v}_1, \cdots, \vec{v}_n)$

$\forall j$, $T(\vec{v}_j) \in \mathrm{Sp}(\vec{v}_1, \cdots, \vec{v}_j)$
$\mathrm{Sp}(\vec{v}_1, \cdots, \vec{v}_j) \underset{S.S.}{\subseteq} V$ is an invariant subspace of V.
$T(\vec{v}_i) \in \mathrm{Sp}(\vec{v}_1, \ldots, \vec{v}_i) \underset{S.S.}{\subseteq} \mathrm{Sp}(\vec{v}_1, \ldots, \vec{v}_j)$, $1 \le i \le j$

Proof. Prove Proposition 6.1.4.

Prove by induction on $\dim V$

- $\dim V = 1$
  This follows immediately.
- True for $T \in \mathcal{L}(W)$, $\dim W < n$.
  WTS True for $T \in \mathcal{V}$, $\dim V = n$.
  Consider $U = Im(T = \lambda I) \underset{S.S.}{\subsetneq} V$ invariant under T.

  $T\Big|_U : U \to U,\ u \mapsto T(u)$

  By induction hypothesis, $\exists$ a basis $\mathcal{U} = (\vec{u}_1, \cdots, \vec{u}_m)$ for U s.t. $[T\Big|_U]_{\mathcal{U}}$ is upper triangular.
  Extend $\mathcal{U}$ to a basis $\mathcal{B}$ for V: $\vec{u}_1, \cdots, \vec{u}_m, \vec{v}_1, \cdots, \vec{v}_l$, $l + m = n$.
  WTS $[T]_{\mathcal{B}}$ is upper triangular.
  $\underbrace{T\Big|_U (\vec{u}_i)}_{=T(\vec{u}_i)} \in Sp(\vec{u}_1, \cdots, \vec{u}_i)$

$$
\begin{aligned}
T(\vec{v}_i) &= T(\vec{v}_i) - \lambda\vec{v}_1 + \lambda\vec{v}_i \\
&= \underbrace{(T - \lambda I)(\vec{v}_i)}_{\in Im(T-\lambda I)=Sp(\vec{v}_1,\cdots,\vec{v}_m)} + \underbrace{\lambda\vec{v}_i}_{\in Sp(\vec{v}_i)} \\
&\in \mathrm{Sp}(\vec{u}_1, \cdots, \vec{u}_m, \vec{v}_1, \cdots, \vec{v}_i)
\end{aligned}
$$

■

Theorem 4.0.2 — Schur's Theorem. Suppose V is a finite-dimensional comples vector space and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some orthonormal basis of V.

Proof. Take a basis $\mathcal{B} = (\vec{b}_1, \cdots, \vec{b}_n)$.
After G-S:
$\mathrm{Sp}(\vec{u}_1) = \mathrm{Sp}(\vec{b}_1)$
$\mathrm{Sp}(\vec{u}_1, \cdots, \vec{u}_j) = \mathrm{Sp}(\vec{b}_1, \cdots, \vec{b}_j)$
$[T]_{\mathcal{U}}$ is triangular iff $\forall i$, $T(\vec{u}_i) \in Sp(\vec{u}_1, \cdots, \vec{u}_i) = Sp(\vec{b}_1, \cdots, \vec{b}_i)$ ■

Theorem 6.1.12 — Cayley-Hamilton Theorem. Let $T : V \to V$ be a linear mapping on a finite-dimensional vector space V, and let $p(t) = \det(r - tI)$ be its characteristic polynomial. Assume that $p(t)$ has $\dim(V)$ roots in the field $\mathbb{F}$ over which V is defined. Then $p(T) = 0$ (the zero mapping on V)[a].

[a] $q(x) = x^n + a_{n-1}x^n - 1 + \cdots + a_1 x = a_0$. $q(T) = T^n + a_{n-1}T^{n-1} + \cdots + a_1 T + a_0 = 0$.

Proof. Let $Q(T)$ be the characteristic polynomial of T. The $Q(T) = 0 \in \mathcal{L}(V)$.
Note $Q(t) = (t - \lambda_1)^{m_1}(y - \lambda_2)^{m_2} \cdots (T - \lambda_k)^{m_k}$, m_i's are the algebraic multiplicities of λ_i.
WTS $Q(T)\vec{v} = \vec{0} \in V$ for all $\vec{x} \in V$.
$Q(T) : V = G_{\lambda_1} \oplus \cdots \oplus G_{\lambda_k} \to G_{\lambda_1} \oplus \cdots \oplus G_{\lambda_k}$.
Note that $Q(T)|_{G_\lambda + i} : G_\lambda + i \to G_\lambda + i$ is G_{λ_i} invariant.
$\forall \vec{v} \in \vec{v}$, $\vec{v} = \vec{v}_1 + \vec{v}_2 + \cdots + \vec{v}_k$, $\vec{v}_i \in G_{\lambda_i}$.

WTS $Q(T)\big|_{G_{\lambda_i}}(\vec{v}_i) = \vec{0}$.
Take $\vec{v}_i \in G_{\lambda_i} = \ker(T - \lambda_i \mathrm{id})^{m_i}$.
$Q(T)\big|_{G_{\lambda_i}} = () \cdots (T - \lambda_i \mathrm{id})^{m_i} \vec{v}_i = \vec{0}$. ■

■ Example 4.2
Let $T : \begin{array}{ccc} \mathbb{C}^n & \to & \mathbb{C}^n \\ \vec{x} & \mapsto & A\vec{x} \end{array}$ be a linear transformation.
$\mathrm{char}(T) = \mathrm{char}(A) = \det(A - \lambda U)$
$\det(A - A) = \det(0) = 0 \in \mathbb{C}$. ■

4.2 Canonical Form for Nilpotent Mappings

Definition 4.0.3 — Niplotent. The linear transformation $T : V \to V$ is *nilpotent* if $\exists n \geq 1$ s.t $T^n = 0$. The smallest such n is called the *index* of T. Its signature is $(\dim\ker T, \cdots, \dim\ker T^k)$.

Note: If $T \in \mathcal{L}(V)$ is nilpotent, then λ is an eigenvalue of $T \iff \lambda = 0$

■ Example 4.3 True or False ($\mathbb{F} = \mathbb{C}$):
Suppose $T \in \mathcal{L}(V)$ is nilpotent with index k.
Then, $k \leq \dim V$.
True.
$p(x) = char(T) = x^{\dim V} \implies T^{\dim V} = 0$ (by Cayley-Hamilton). ■

Matrix of a nilpotent operator

Suppose N is a nilpotent operator on V. Then there is a basis V with respect to which the matrix of N has the form

$$\begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix};$$

here all the entries on and below the diagonals are 0's.

Definition 4.0.4 — Jordan Basis. Suppose $T \in \mathcal{L}(\mathrm{V})$. A basis of V is called a *Jordan basis* for T if with respect to this basis T has a block diagonal matrix

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix}$$

where each A_j is an upper triangular matrix of the form

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

Definition 6.2.1 — Cycle. Let N, $\vec{v} \neq 0$ and k be as before.

a) The set $\{N^{k-1}(\vec{v}), N^{k-2}(\vec{v}, \ldots, \vec{v}$ is called. the *cycle* generated by $\vec{v}$. $\vec{v}$ is called the *initial vector* of the cycle.

b) The subspace Span($\{N^{k-1}(\vec{v}), N^{k-2}(\vec{v}, \dots, \vec{v}\}$) is called thet *cycle subspace* generated by $\vec{v}$, and denoted $C(\vec{v})$.
c) The integer k is called the *length* of the cycle.

Definition 6.2.5 — Non-overlapping Cycle. We say that the cycles $\alpha_i = \{N^{k_i-1}(\vec{v}_i), \dots, \vec{v}_i\}$ are *non-overlapping* cycles if $\alpha_1 \cup \dots \cup \alpha_r$ is linearly independent.

Definition 6.2.7 — Canonical Basis. Let $N : V \to V$ be a nilpotent mapping on a finite-dimensional vector space V. We call a basis β for V a *canonical basis* (with respect to N) if β is the union of a collection of nonoverlapping cycles for N.

■ **Example 4.4** $N : \mathbb{C}^4 \to \mathbb{C}^4$ is nilpotent.

$$\exists \mathcal{B} = (\vec{b}_1, \cdots, \vec{b}_4) \text{ s.t. } [N]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$ ■

Theorem 4.2.1 — Canonical Form for Nilpotent Mappings. Suppose V is a complex vector space. If $T \in \mathcal{L}(V)$, then there is a basis of V that is a Jordan basis (canonical basis) for T.

4.3 Jordan Canonical Form

■ **Example 4.4 — count..**

- What is index of N?
  The index of N is 4.
- What is signature of N?
  The signature of N is $(1, 2, 3, 4)$
  - $\dim \ker N^1 = 1$
  - $\dim \ker N^2 = 2$
  - $\dim \ker N^3 = 3$
  - $\dim \ker N^4 = 4$
- What can we say about $\mathcal{B}$?
  $\vec{0} \xleftarrow{N} \vec{b}_1 \xleftarrow{N} \vec{b}_2 \xleftarrow{N} \vec{b}_3 \xleftarrow{N} \vec{b}_4$ is a cycle.
  - $\vec{b}_1$ is an eigenvector, so $\vec{b}_1 \in \ker(N - \lambda I) = \ker N$ since $\lambda = 0$.
  - $\vec{b}_2 \in \ker(N - \lambda I)^2 = \ker N^2$
  - $\vec{b}_3 \in \ker(N - \lambda I)^3 = \ker N^3$
  - $\vec{b}_4 \in \ker(N - \lambda I)^4 = \ker N^4$

■

Definition 6.3.2 Let $T : V \to V$ be a linear mapping on a finite-dimensional vector space V. Let λ be an eigenvalue of T with multiplicity m.
a) The *λ-generalized eigenspace*, denoted by K_λ (or G_λ), is the kernel of the mapping $(T - \lambda I)^m$ on V.
b) The nonzero elements of K_λ are called *generalized eigenvectors* of T.

In other words, suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. A vector $\vec{v} \in V$ is called a *generalized eigenvector* of T corresponding to λ if $v \neq 0$ and

$$(T - \lambda I)^j \vec{v} = 0$$

for some positive integer j. The *generalized eigenspace* of T corresponding to λ is defined to be the set of all generalized eigenvectors of T corresponding to λ, along with $\vec{0}$.

Note that $T(\vec{v}) = \lambda\vec{v}$, $\vec{v}$ is a generalized vector.

■ Example 4.5

Let $N : \begin{array}{ccc} \mathbb{C}^4 & \to & \mathbb{C}^4 \\ \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} & \mapsto & \begin{pmatrix} 0 \\ z_1 \\ z_2 \\ z_3 \end{pmatrix} \end{array}$ be a linear mapping.

Fina a basis $\mathcal{B}$ such that $[N]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix}$

$0 \leftarrow N^3(\vec{e}_1) \leftarrow N^2(\vec{e}_1) \leftarrow N(\vec{e}_1) \leftarrow \vec{e}_1$

$\mathcal{B} = (N^3(\vec{e}_1), N^2(\vec{e}_1), N(\vec{e}_1), \vec{e}_1)$. ■

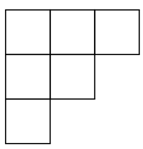
■ Example 4.6

Let $N : \begin{array}{ccc} \mathbb{C}^4 & \to & \mathbb{C}^4 \\ \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \end{pmatrix} & \mapsto & \begin{pmatrix} 0 \\ z_1 \\ z_2 \\ 0 \\ z_4 \\ 0 \end{pmatrix} \end{array}$ be a linear mapping.

$\exists \mathcal{B}$ s.t. $[N]_{\mathcal{B}} = \begin{bmatrix} A_1 & & \\ & A_2 & \\ & & A_3 \end{bmatrix}$ where $A_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $A_3 = \begin{bmatrix} 0 \end{bmatrix}$.

(That is, $[N]_{\mathcal{B}} = \begin{bmatrix} \begin{bmatrix} 0 & 1 & \\ & 0 & 1 \\ & & 0 \end{bmatrix} & & 0 \\ & \begin{bmatrix} 0 & 1 \\ & 0 \end{bmatrix} & \\ 0 & & \begin{bmatrix} 0 \end{bmatrix} \end{bmatrix}$)

- What is index of N?
  - $\vec{0} \leftarrow \vec{b}_1 \leftarrow \vec{b}_2 \leftarrow \vec{b}_3$
  - $\vec{0} \leftarrow \vec{v}_4 \leftarrow \vec{b}_5$
  - $\vec{0} \leftarrow \vec{b}_6$
- What is signature of N?
  $(3, 5, 6)$



This is called the "Young Tableau" or the "Cycle Tableau".[1]

[1] Let us assume that the cycles have been arranged so that $k_1 \geq k_2 \geq \cdots \geq k_r$. Then the cycle tableau consists of r rows of boxes, where the boxes in the i-th row represent the vectors in the i-th cycle. (Thus, there are k_i boxes in the i-th row of the tableau.) We will always arrange the tableau so that the leftmost boxes in each row are in the same column. In other words, cycle tableaux are always left-justified.

- What can we say about $\mathcal{B}$?
  $\vec{b}_i$'s are generalized eigenvectors.
- Find such a $\mathcal{B}$.
  - $0 \leftarrow N^2(\vec{e}_1) \leftarrow N(\vec{e}_1) \leftarrow \vec{e}_1$
  - $0 \leftarrow N(\vec{e}_4) \leftarrow \vec{e}_4$
  - $0 \leftarrow \vec{e}_6$

  $\mathcal{B} = (N^2(\vec{e}_1), N(\vec{e}_1), \vec{e}_1, N(\vec{e}_4), \vec{e}_4, \vec{e}_6)$
- Assume $\mathcal{B}' = (N(\vec{e}_4), \vec{e}_4, N^3(\vec{e}_1), N(\vec{e}_1), \vec{e}_1, \vec{e}_6)$. What is $[T]_{\mathcal{B}'}$?

  $$[T]_{\mathcal{B}'} = \begin{bmatrix} A_2 & & \\ & A_1 & \\ & & A_3 \end{bmatrix}$$

■

4.4 Computing Jordan Form

Definition 6.4.1 — Computation of Jordan Canonical Form.

a) Find all the eigenvalues of T and their multiplicities by factoring the characteristic polynomial completely. (We will usually work over an algebraically closed field such as $\mathbb{C}$ to ensure that this is always possible.)

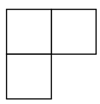
b) For each distinct eigenvalue λ_i, in turn, construct the cycle tableau for a canonical basis of K_{λ_i} with respect to the mapping $N_i = (T - \lambda_i I)\big|_{K_{\lambda_i}}$ using the method of Lemma (6.2.9), in the following modified form. For each j, the number of boxes in the j-th column of the tableau for λ_i will be

$$\dim\ker(T - \lambda_i I)^j - \dim\ker(T - \lambda_i I)^{j-1}$$

(computed on V). This is the same as $\dim\ker(N_i^j) - \dim\ker(N_i^{j-1})$, since $\ker(T - \lambda_i I)^j$ is contained in JK_{λ_i} for all $j \geq 1$. As a result it is never actually necessary to construct the matrix of the restricted mapping N_i.

c) Form the corresponding Jordan blocks and assemble the matrix of T.

■ **Example 4.7** $N : V \to V$ is a nilpotent map.
The Young tableau is

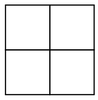


- What is the signature?
  Index is 2 (there are two rows in the cycle tableau).
  The signature is $(2, 3)$ (there are 2 vectors in the first column (N), and 1 vector in the second column N^2).
- What is Jordan form of N?
  - Since N is nilpotent, λ is an eigenvalue $\implies \lambda = 0$.
  - Since the cycle tableau has two rows, of N takes the Jordan form $\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$.
  - Since there are two vectors in the first row of the cycle tableau, $A_1 = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$.
  - Since there is one vectors in the first row of the cycle tableau, $A_1 = \begin{bmatrix} \lambda \end{bmatrix}$.

  Thus, the Jordan form of N is $\begin{bmatrix} 0 & 1 & \\ 0 & 0 & \\ & & 0 \end{bmatrix}$

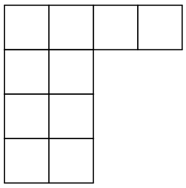
■

■ **Example 4.8** Let $T : \mathbb{C}^4 \to \mathbb{C}^4$ be a nilpotent linear mapping.
Known $\dim\ker T^2 = 2$, $\dim\ker T^3 = \dim\ker T^4 = 4$.



Then $[T]_{\mathcal{B}} = \begin{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & \\ & \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \end{bmatrix} = N_2 \oplus N_2$ ■

■ **Example 4.9** Let $T : \mathbb{C}^{10} \to \mathbb{C}^{10}$ be a nilpotent linear mapping.
Known $\dim\ker T = 4$, $\dim\ker T^2 = 8$, $\dim\ker T^3 = 9$ and $\dim\ker T^4 = 10$.



Then, $[T]_{\mathcal{B}} = N_4 \oplus N_1 \oplus N_1 \oplus N_1$ ■

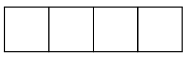
■ **Example 4.10 — Single Block.** Let $T : \mathbb{C}^4 \to \mathbb{C}^4$ be a nilpotent linear mapping.

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix}$$

- The eigenvalue is λ with an algebraic multiplicity of 4.
- The geometric multiplicity of λ is 1.

Thus, $T = \lambda\mathrm{id} : \mathbb{C}^4 \to \mathbb{C}^4 : [T - \lambda\mathrm{id}] = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}$

$T(\vec{b}_1) = \lambda\vec{b}_1$
$T(\vec{b}_i) = \lambda\vec{b}_i + \vec{b}_i$
$[T - \lambda\mathrm{id}]$ has a cycle of $\vec{b}_1 \leftarrow \vec{b}_2 \leftarrow \vec{b}_3 \leftarrow \vec{b}_4$.



■

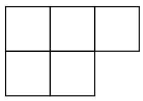
■ **Example 4.11 — Multiple Block.** Let $T : \mathbb{C}^8 \to \mathbb{C}^8$ be a nilpotent linear mapping.

$$\exists \text{ a basis } \mathcal{B} \text{ s.t. } [T]_{\mathcal{B}} = \begin{bmatrix} \begin{bmatrix} -i & 1 & \\ & -i & 1 \\ & & -i \end{bmatrix} & & & \\ & \begin{bmatrix} -i & i \\ & -i \end{bmatrix} & & \\ & & \begin{bmatrix} 3 & 1 \\ & 3 \end{bmatrix} & \\ & & & [2] \end{bmatrix}$$

- The eigenvalues are
  - $-i$ with an algebraic multiplicity of 5
  - 3 with an algebraic multiplicity of 2
  - 2 with an algebraic multiplicity of 1
- For $\lambda = -i$

$$[T-(-i)\mathrm{id}] = \begin{bmatrix} \begin{bmatrix} 0 & 1 & \\ & 0 & 1 \\ & & 0 \end{bmatrix} & & & \\ & \begin{bmatrix} 0 & 1 \\ & 0 \end{bmatrix} & & \\ & & \begin{bmatrix} * & 1 \\ & * \end{bmatrix} & \\ & & & \begin{bmatrix} * \end{bmatrix} \end{bmatrix}$$

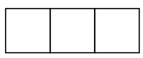
  - the geomrtric mutiplicity of $\lambda = -i$ is 2 ().
    $\mathcal{B} = (\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_8)$
    $\vec{0} \leftarrow \vec{b}_1 \leftarrow \vec{b}_2 \leftarrow \vec{b}_3$
    $\vec{0} \leftarrow \vec{b}_4 \leftarrow \vec{b}_5$



  - The generalized eigenspace[2] , $G_i = \mathrm{Sp}(\vec{b}_1, \vec{b}_2, \vec{b}_3, \vec{b}_4, \vec{b}_5)$
- For $\lambda = 3$

$$[T-3\mathrm{id}]_{\mathcal{B}} = \begin{bmatrix} \begin{bmatrix} * & 1 & \\ & * & 1 \\ & & * \end{bmatrix} & & & \\ & \begin{bmatrix} * & 1 \\ & * \end{bmatrix} & & \\ & & \begin{bmatrix} 0 & 1 \\ & 0 \end{bmatrix} & \\ & & & \begin{bmatrix} * \end{bmatrix} \end{bmatrix}$$

  - $\vec{0} \leftarrow \vec{b}_6 \leftarrow \vec{b}_7$
  - $G_3 = \mathrm{Sp}(\vec{b}_6, \vec{b}_7)$



- For $\lambda = 2$

$$[T-2\mathrm{id}]_{\mathcal{B}} = [T-3\mathrm{id}]_{\mathcal{B}} = \begin{bmatrix} \begin{bmatrix} * & 1 & \\ & * & 1 \\ & & * \end{bmatrix} & & & \\ & \begin{bmatrix} * & 1 \\ & * \end{bmatrix} & & \\ & & \begin{bmatrix} * & 1 \\ & * \end{bmatrix} & \\ & & & \begin{bmatrix} 0 \end{bmatrix} \end{bmatrix}$$

  - $G_2 = \mathrm{Sp}(\vec{b}_8) = E_2$.

■

■ Example 4.12

Let $T: \begin{array}{ccc} \mathbb{C}^4 & \to & \mathbb{C}^4 \\ \vec{x} & \mapsto & A\vec{x} \end{array}$ be a linear transformation.

---

[2] $G_\lambda = \ker(T-\lambda\mathrm{id})^l$, $l \geq 1$

Let $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{bmatrix}$.

Given the eigenvalue $\lambda_1 = 2$ with an algebraic multiplicity of 3, and the eigenvalue $\lambda_2 = 3$ with an algebraic multiplicity of 1.

Given $\ker(T - 1\mathrm{id}) = \mathrm{Sp}\left\{\begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}\right\}$, $\ker(T - 1\mathrm{id})^2 = \mathrm{Sp}\left\{\begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}\right\}$.

Given $\ker(T - 2\mathrm{id}) = \mathrm{Sp}\left\{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right\}$ [3], and $\ker(T - 2\mathrm{id})^2 = \left\{\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right\}$

- $\vec{0} \leftarrow \vec{b}_1 \leftarrow \vec{b}_2$
  Pick a $\vec{b}_2$, compute its image and that will be $\vec{b}_1$.
- $\vec{0} \leftarrow \vec{b}_3$

■

Lemma 4.4.1 Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. Then $G(\lambda, T) = \mathrm{null}(T - \lambda I)^{\dim V}$.

Recall: V is a $\mathbb{C}$_V.S., $\dim V = n$. $T \in \mathcal{L}(V)$. $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues.

Theorem 4.4.2 T admits a Jordan basis.
$\exists$ a basis $\mathcal{B} = (\vec{b}_1, \ldots, \vec{x}_n)$ for V s.t.

$$[T]_{\mathcal{B}} = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_l \end{bmatrix}$$

, with each $A_i = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$

Observations

- A_i's with the same λ are the matrix representation of $(T - \lambda I)\big|_{G(\lambda,T)}$.
  $T\big|_{G_\lambda} : G_\lambda \to G_\lambda$ is invariant. $T(G_\lambda) \underset{S.S.}{\subseteq} G_\lambda$.
- $\vec{b}_i$'s are generalized eigenvectors.
  $V = \bigoplus_{\lambda = \{\lambda_1, \ldots, \lambda_k\}} G_\lambda = G_{\lambda_1} \oplus G_{\lambda_2} \oplus \cdots \oplus G_{\lambda_k}$.
- $T - \lambda\mathrm{id}\big|_{G_\lambda}$ is nilpotent.

Theorem 4.4.3

a) G_λ is T invariant: $T(G_\lambda) = G_\lambda$

[3] This tells us that we have two Jordan blocks with eigenvalues of 2

b) $V = \bigoplus_{\lambda=\{\lambda_1,\ldots,\lambda_k\}} G_\lambda = G_{\lambda_1} \oplus G_{\lambda_2} \oplus \cdots \oplus G_{\lambda_k}$

c) $T - \lambda \mathrm{id}\big|_{G_\lambda}$ is nilpotent.

Proof.

a) $G_\lambda = \ker(T - \lambda \mathrm{id})^n$.
Let $\vec{v} \in G_\lambda$.
Then, $(T - \lambda \mathrm{id})^n \vec{v} = 0$.

WTS $T(\vec{x}) \in \ker(T - \lambda \mathrm{id})^n$.
$(T - \lambda \mathrm{id})^n T(\vec{v})$
Let $p(T) = (T - \lambda \mathrm{id})^n$.
$p(T)T(\vec{v}) = Tp(T)(\vec{v}) = T(\vec{0}) = 0$

OR
$T(\vec{v}) = (T - \lambda \mathrm{id})(\vec{v}) + \lambda\vec{v} = \vec{0} + \lambda\vec{v} = \lambda\vec{v}$.

b) Proof by induction on $\dim V$.
- **Base case:** $\dim V = 1$
  λ is an eigenvalues of T, so $E_\lambda = V = G_\lambda$.
- **Induction Hypothesis:** Assume $W = \bigoplus_{\lambda=\{\lambda_1,\ldots,\lambda_n\}} G_\lambda = G_{\lambda_1} \oplus \cdots \oplus G_{\lambda_k}$ holds for any $S \in \mathcal{L}(W)$, $\dim W < n$.
  Let $T \in \mathcal{L}(V)$m $\dim V = n$. $V = \underbrace{G_{\lambda_1}}_{=\ker(T-\lambda_1)\mathrm{id}^n} \oplus$ something.
  Consider $T - \lambda_1 \mathrm{id} : B \to V$ by :) $V = \underbrace{\ker(T - \lambda \mathrm{id})^n}_{G_{\lambda_1}} + \underbrace{\mathrm{Im}(T - \lambda_1 \mathrm{id})^n}_{U}$.
  Note U is invariant:

  Exercise 4.1 $S : V \to V$. Prove $S(\mathrm{Im}\ S) \underset{S.S.}{\subseteq} \mathrm{Im}\ S$.

  Proof. Let $\vec{w} \in \mathrm{Im}\ S$. $S(\vec{w}) \in \mathrm{Im}\ S$. ∎

  ■

  $T\big|_U : U \to U$, $\dim U < n$.
  By induction hypothesis, $U = G_{\lambda_2} \oplus \cdots \oplus G_{\lambda_k}$

  Note that $G_{\lambda_1} = G(\lambda_1, T)$, but $G_{\lambda_2} = G(\lambda_2, T\big|_U)$.
  WTS $G(\lambda_i, T\big|_U) = G(\lambda_i, T)$.
  - $G(\lambda_i, T\big|_U) \underset{S.S.}{\subseteq} G(\lambda_i, T)$
    Let $\vec{v} \in G(\lambda_i, T\big|_U)$.
    $(T\big|_U - \lambda_i \mathrm{id})^n \vec{v} = \vec{0}$.
    $\underbrace{(T - \lambda_i \mathrm{id})^n}_{p(T)} \vec{v} = (T\big|_U - \lambda_i \mathrm{id})^n \vec{v} = \vec{0}$.
  - $G(\lambda_i, T) \underset{S.S.}{\subseteq} G(\lambda_i, T\big|_U)$
    Let $\vec{v} \in G(\lambda_i, T)$.
    Show that $\vec{v} \in U$.
    $\vec{v} = \vec{v}_1 + \vec{u}_2 + \cdots + \vec{u}_k =$ a component in G_{λ_1} + a component in U.

c) WTS $T - \lambda \mathrm{id}\big|_{G_\lambda}$ is nilpotent.
Let $\vec{v} \in G_\lambda$, $T\big|_{G_\lambda} - \lambda \mathrm{id}\big|_{G_\lambda}(\vec{x}) = T(\vec{v}) - \lambda \mathrm{id}(\vec{v}) \in G_\lambda$.

$\forall \vec{v} \in G_\lambda$, $(T - \lambda \mathrm{id})^n = \vec{0}$ by the description of a generalized eigenspace.

■

■ **Example 4.13** Provide counter examples of $T : \ker T \oplus \mathrm{Im}\ T$.

$$\begin{aligned} T: \quad \mathbb{R} &\to \quad \mathbb{R} \\ \vec{x} &\mapsto \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \vec{v} \end{aligned}$$

■

■ **Example 4.14** Prove $V = \ker T^n \oplus \mathrm{Im}\ T^n$:)

Proof. Let $\vec{v} \in \ker T^n \cap \mathrm{Im}\ T^n$. WTS $\vec{v} = 0$.

Since $\vec{v} \in \ker T^n$, $T^n(\vec{v}) = \vec{0}$.

Since $\vec{v} \in \mathrm{Im}\ T^n$, $\exists \vec{w} \in V$ s.t. $T(\vec{w}) = \vec{x}$.

$T^n(T^n(\vec{w})) = T^n(\vec{v}) = \vec{0}$.

$T^{2n}(\vec{x}) = \vec{0}$.

From HW2: $\ker T \subsetneq \ker T^2 \subsetneq \cdots \subsetneq \ker T^n = \ker T^{n+1} = \cdots$.

So $\ker T^{2n} = \ker T^n$, so $T^{2n}(\vec{w}) = 0 \implies T^n(\vec{w}) = \vec{0} = \vec{v}$.

WTS $\ker T^n \oplus \mathrm{Im}\ T^n \underset{S.S.}{\subseteq} V \implies \ker T^n \oplus \mathrm{Im} T^n = V$ by Rank-Nullity Theorem. ■

■

■ **Example 4.15**

Let $T : \mathbb{C}^4 \to \mathbb{C}^4$ be a linear transformation defined by $\vec{x} \mapsto A\vec{x}$.

$$\mathrm{Nul}(A - 2I) = \mathrm{Sp}\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

$$\mathrm{Nul}(A - t3I) = \mathrm{Sp}\left\{ \begin{pmatrix} 1 \\ -1 \\ -1 \\ 3 \end{pmatrix} \right\}$$

$$\mathrm{Nul}(A - 3I) = \mathrm{Sp}\left\{ \begin{pmatrix} 2 \\ 10 \\ 1 \\ 0 \end{pmatrix} \right\}$$

- Fina a Jordan Canonical form for A

$$\begin{pmatrix} [2] & & \\ & [2] & \\ & & \begin{bmatrix} 3 & 1 \\ & 3 \end{bmatrix} \end{pmatrix}$$

- Find the corresponding Jordan basis
  - $\mathcal{B} = (\vec{b}_1, \vec{b}_2, \vec{b}_3, \vec{b}_4)$.
  - $[T(\vec{b}_1)]_{\mathcal{B}} = [2\vec{b}_1]_{\mathcal{B}} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

- Take $\vec{b}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $\vec{b}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$.
- $\vec{0} \leftarrow \vec{b}_3 \leftarrow \vec{b}_4$ is a cycle.
- Take $\vec{b}_4 = \begin{pmatrix} 2 \\ 10 \\ 1 \\ 0 \end{pmatrix}$ s.t. $\vec{b}_4 \in \mathrm{Nul}(A - 3I)^2$ and $\vec{x}_4 \notin \mathrm{Nul}(A - 3I)$.
- $\vec{b}_3 = (A - 3I)(\vec{b}_4) = \begin{pmatrix} 3 \\ -3 \\ -3 \\ 9 \end{pmatrix} \in \mathrm{Nul}(A - 3I)$.

So $\mathcal{B} = \left(\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -3 \\ -3 \\ 9 \end{pmatrix}, \begin{pmatrix} 2 \\ 10 \\ 1 \\ 0 \end{pmatrix} \right)$ is a Jordan basis for T.

■

Part Two

Bibliography

Books

[DL11] David B. Diamiano and John B. Little. *A Course in Linear Algebra*. Dover edition. 2011 (cited on page 8).

Index

K

L

M

N

O

P

R

S

T

U

V