Linear Algebra Done Right Solutions

Michael You

Contents

Forward

These solutions are for the 3rd edition *Linear Algebra Done Right* textbook by Sheldon Axler.

I will do my best to provide some guidance along the way in this book as I read it. I will highlight important theorems for each chapter.

Please note that I'm mostly going through this book for review[1], so I will likely skip over more introductory topics. I might also be lazy with proofs, so definitely feel free to submit feedback about problems that I could've done better. I still hope that whenever I say "intuition" or give sketches of solutions, that it helps.

[1] my third time doing linear algebra, although first two times were both introductory level

Chapter 1

Vector Spaces

1.1 $\mathbf{R}^n$ and $\mathbf{C}^n$

Exercise 1.1.1
Multiply both sides by $\frac{a-bi}{a-bi}$.

This is known as *rationalizing the denominator*.

Exercise 1.1.2
I think this is meant to be an arithmetic problem, but we can notice that

$$\frac{-1+\sqrt{3}i}{2} = \operatorname{cis}\left(\frac{2\pi}{3}\right) \implies \left(\frac{-1+\sqrt{3}i}{2}\right)^3 = \operatorname{cis}(2\pi) = 1$$

which is just 1 because we are at the rightmost part of the unit circle.

Exercise 1.1.3
$i = \operatorname{cis}\pi/2 + 2\pi k$ so if we take the square root we get solutions in the form of

$$\operatorname{cis}\pi/4 + \pi k$$

Exercise 1.1.4
Apply commutativity to real and imaginary parts.

Exercise 1.1.5
Separate real and imginary parts and use field properties.

Exercise 1.1.6
Separate real and imginary parts and use field properties.

Exercise 1.1.7
skip

Exercise 1.1.8
skip

Exercise 1.1.9
skip

Exercise 1.1.10
$x = (1/2, 6, -7/2, 1/2)$.

Exercise 1.1.11
We can equate two pairs of tuple values and show that no such λ exists after simplifications.

Skip the rest

1.2 Definition of Vector Space

The vector space definition in this chapter is very important!

Definition 1.2.1
A **vector space** is a set V along with an addition on V and a scalar multiplication on V such that the following properties hold:

- commutativity
$$u + v = v + v \quad \forall u, v \in V$$
- associativity
$$\begin{aligned} &\forall u, v, w \in V, a, b, \in \mathbf{F} \\ &(u + v) + w = u + (v + w) \text{ and } (ab)v = a(bv) \end{aligned}$$
- additive identity
$$\exists 0 \in V \quad \forall v\,, v + 0 = v$$
- additive inverse
$$\forall v \in V, \exists w \in V \quad v + w = 0$$
- multiplicative identity
$$\forall v \in V, 1v = v$$
- distributive properties
$$\begin{aligned} &\forall a, b \in \mathbf{F} \text{ and } \forall u, v \in V \\ &a(u + v) = au + av \text{ and } (a + b)v = av + bv \end{aligned}$$

Exercise 1.2.1
The additive inverse of v is $-v$. The additive inverse of $-v$ is $-(-v)$. We have

$$\begin{aligned} v + -v = 0 &= -v + -(-v) \\ v &= -(-v) \end{aligned}$$

Exercise 1.2.2
Just do it by cases.

I think we have to prove it is not possible for $a \neq 0$ and $v \neq 0$ though... In this case you'd do an AFSOC, but showing $av \neq 0$ I'm not sure how to do without using information e.g. that two nonzero elements multiplied together is not zero (this is not trivial btw, e.g. you can have two non zero matrices that multiply to be 0).

Exercise 1.2.3
Because $x = \frac{1}{3}(w - v)$

Exercise 1.2.4
Since the empty set doesn't have elements, any statement with $\forall$ will be vacuously true. The only statement that doesn't look like that is the additive identity, which we can confirm is not true because $0 \notin \emptyset$.

Exercise 1.2.5
If we can replace the additive identity, then we should be able to derive it from our existing properties.

So let's start with some $v \in V$, we want to find some $w \in V$ such that $v + w = 0$.

I think we can choose $w = 0v - v$?

Exercise 1.2.6
skip

1.3 Subspaces

Subspace property here is very important

Definition 1.3.1
$U \subset V$ is a **subspace** of V if U is also a vector space.

The conditions for a subspace $U \subset V$ are:

1. additive identity
$$0 \in U$$
2. closed under addition
$$u, w \in U \implies u + w \in U$$
3. closed under scalar multiplication
$$a \in \mathbf{F} \text{ and } u \in U \implies au \in U$$

Exercise 1.3.1

(a) yes

(b) no scalar fails

(c) Not closed under addition

(d) Yes

Exercise 1.3.2
We have to confirm, I'll just point out the ones that don't work.

(a) if $b \neq 0$ then scaling doesn't work

(b) yes

(c) yes

(d) not close dunder addition if $b \neq 0$.

(e) Limit 0 is key, another limit point and this won't work.

Exercise 1.3.15
$U + U = U$, since $\forall u \in U, 2u \in U$

Exercise 1.3.16
Commutativity should hold.

Exercise 1.3.17
Yes should be associative via field properties.

Exercise 1.3.18
I think they all should, since 0 is in the subspace and scalar is closed, so we have $v, -v$ for any $v \in V$.

Exercise 1.3.19
False, $U_1 = (u, 0), U_2 = (0, u), W = (x, y)$

Exercise 1.3.20
$W = (0, w_1, 0, w_2)$

Exercise 1.3.21
Can we not just choose a trivial W like $W = \mathbf{F}^5$?

Otherwise, we notice that x, y are the only independent variables in the tuple, so we need 3 more tuples, so we just choose $W = (0, 0, x, y, z)$.

Exercise 1.3.22
As we saw in 1.3.21, we have 3 degrees of freedom, so we can choose

$$W_1 = (0, 0, w_1, 0, 0)$$
$$W_2 = (0, 0, 0, w_2, 0)$$
$$W_3 = (0, 0, 0, 0, w_3)$$

Exercise 1.3.23
False.

The intuition here is that the part W is "missing" from V is constant, but the trick is that we can have one of the U_1 "over contribute" and the other "under contribute".

So we can choose like $W = (x, y)$ and then $U_1 = (x, 0)$ and then $U_2 = (x, y)$ and the result is both sums are $\mathbb{R}^2$.

Exercise 1.3.24
skip

Chapter 2

Finite-Dimensional Vector Spaces

2.1 Span and Linear independence

Linear combination, span and linear independence (it's in the title duh) are the most definitions in this chapter.

Definition 2.1.1
A **linear combination** of a list $v_1, \ldots, v_m$ of vectors in V is a vector of the form

$$\sum_{i=1}^{m} a_i v_i,$$

where $a_1, \ldots, a_m \in \mathbf{F}$.

Definition 2.1.2
The set of all linear combinations of a list of vectors $v_1, \ldots, v_m$ in V is called the **span** of the list, denoted

$$\operatorname{span}(v_1, \ldots, v_m) = \left\{ \sum_{i=1}^{m} a_i v_i \,:\, a_1, \ldots, a_m \in \mathbf{F}. \right\}$$

The span of the empty list () is defined to be [0].

Definition 2.1.3

- A list $v_1, \ldots, v_m$ of vectors in V is called **linearly independent** if the only choice $a_1, \ldots, a_m \in \mathbf{F}$ that makes
$$\sum_{i=1}^{m} a_i v_i = 0$$
is $a_1 = \cdots = a_m = 0$.
- The empty list () is also declared to be linearly independent.

Exercise 2.1.1
We have

$$\begin{aligned} \operatorname{span}(v_4, v_3 - v_4) &= \operatorname{span}(v_4, v_3) \\ \operatorname{span}(v_4, v_3 - v_4, v_2 - v_3) &= \operatorname{span}(v_4, v_3, v_2 - v_3) = \operatorname{span}(v_4, v_3, v_2) \\ \operatorname{span}(v_4, v_3 - v_4, v_2 - v_3, v_1 - v_2) &= \operatorname{span}(v_4, v_3, v_2, v_1 - v_2) = \operatorname{span}(v_4, v_3, v_2, v_1) \end{aligned}$$

Exercise 2.1.2
(a) $v = 0$ does not work because we have $a_1 \neq 0, a_1 v = 0$. Otherwise, $v \neq 0 \implies av \neq 0$ for $a \neq 0$.

(b) If one is a scalar multiple of another, then you can write $v_1 = kv_2 \implies v_1 - kv_2 = 0$.

(c) Yes, tuple values are in separate coordinates.

(d) Yes.

Exercise 2.1.3
We want $(5, 9, t)$ to be a linear combination of the first two vectors, i.e.

$$x(3, 1, 4) + y(2, -3, 5) = (5, 9, t)$$

So we have $t = 4x + 5y$, and if we just need to solve

$$\begin{aligned} 3x + 2y &= 5 \\ x - 3y &= 9 \end{aligned}$$

I believe this is $y = -2, x = 3$.

Exercise 2.1.4
This is basically the same as the previous problem.

Exercise 2.1.5
(a) $\forall k \in \boldsymbol{R}, k(1 + i) \neq 1 - i$

(b) $-i(1 + i) = 1 - i$

Exercise 2.1.6
This is very similar to 2.1.1.

Exercise 2.1.7
True.

Exercise 2.1.17
If $\forall j, p_j(2) = 0$, then either $\exists p_j = 0$, in which case these polynomials are not linearly independent.

The other case is that $\forall p_j \neq 0$. In this case, we know they all have a zero at 2, which means we can divide all of them by $(x - 2)$.

Now we have $m + 1$ polynomials who degree is at most $m - 1$. We know that $\mathcal{P}_{m-1}(\mathbf{F})$ has rank m, so $m + 1$ elements in this space will definitely be linear dependent.

2.2 Bases

This chapter should be pretty familiar now that you already understand what span and linear independence are.

Exercise 2.2.1
These vector spaces are a scalar multiple of a single element.

Exercise 2.2.2
too lazy

Exercise 2.2.3

(a) $(3, 1, 0, 0, 0), (0, 0, 7, 1, 0), (0, 0, 0, 0, 1)$.

(b) add $(1, 0, 0, 0, 0), (0, 0, 1, 0, 0)$

(c) Take what we added above

Exercise 2.2.4
lazy

Exercise 2.2.5
False...I think.

Exercise 2.2.6
Just show fun span stuff.

2.3 Dimension

Dimension should be nothing new, it's just the number of vectors in a basis.

too lazy to do these exercises rip

Chapter 3

Linear Maps

3.1 The Vector Space of Linear Maps

The property of linearity is very crucial.

Definition 3.1.1
A **linear map** from $V \to W$ is a function $T : V \to W$ with the following properties:

- additivity
$$\forall u, v \in V \quad T(u+v) = Tu + Tv$$
- homogeneity
$$\forall \lambda \in \mathbf{F}, v \in V \quad T(\lambda v) = \lambda(Tv)$$

Exercise 3.1.1
You know $T(0) = 0$ so use that fact for b. For c just try any nonzero x, y, z.

Exercise 3.1.2
skip

Exercise 3.1.3
The general idea here is that we use the span of $x_1, \ldots, x_n$ to create each coordinate of the element in $\boldsymbol{F}^m$.

Exercise 3.1.4
Suppose $v_1, \ldots, v_m$ is linearly dependent, then for some v_k we can write it as a linear combination of the other v_i's, so we have

$$\begin{aligned} v_k &= \sum_{i \neq k} a_i v_i \qquad (a_i \text{ not all } 0) \\ T(v_k) &= T\left(\sum_{i \neq k} a_i v_i\right) \\ T(v_k) &= \sum_{i \neq k} a_i T(v_i) \\ \sum_{i \neq k} a_i T(v_i) - T(v_k) &= 0 \end{aligned}$$

the last part is a contradiction because we know Tv_i are linearly dependent.

Exercise 3.1.5
Tedious, just verify properties of a vector space in 1.2.1.

Exercise 3.1.6
skip

Exercise 3.1.7
If $\dim V = 1$, then $V = \operatorname{span}\{v\}$ for some $v \in V$. That means every element of V can be written as λv.

Exercise 3.1.8

$$\phi([v_1, v_2]) = \sqrt{v_1^2 + v_2^2}$$

Exercise 3.1.9
I'm stuck, I tried $\operatorname{cis}, \arg$ and no luck.

Exercise 3.1.10
AFSOC T is a linear map.

Suppose $0 \in U$. Then we have $T(0) = 0 = S(0)$ which is not true by $s \neq 0$ assumption.

So suppose $0 \in V \setminus U$. But then U is not a subspace since $0 \notin U$.

Exercise 3.1.11
I think we can use the construction from before and do $Tv = Sv$ if $v \in U$ and 0 otherwise. If there's no $S \neq 0$ restriction I think the construction works.

Exercise 3.1.12
We can trivially map the first $\dim V$ tuple elements of v to the first $\dim V$ tuple elements of w, and then pad with a zero, and then pad with two zeros and so on, for an infinite number of maps.

Exercise 3.1.13
We just have to choose m linearly independent vectors in W, which wouldn't allow $Tv_k = w_k$, otherwise we have $v_1, \ldots, v_m$ linearly independent.

Exercise 3.1.14
skip

3.2 Null Spaces and Ranges

Definition 3.2.1
The **null space** is the subset of inputs that gets mapped to zero.
More formally, $T \in \mathcal{L}(V, W)$,

$$\operatorname{null} T = \{v \in V \; : \; Tv = 0\} \tag{3.1}$$

We start learning about relationships between dimension, linearity, null spaces, ranges and domains.

Theorem 1 *Fundamental Theorem of Linear Maps. Suppose V is finite dimensional and $T \in \mathcal{L}(V, W)$. Then* range T *is finite-dimensional and*

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T \tag{3.2}$$

Exercise 3.2.1
$T(v_1, v_2, v_3, v_4, v_5) = (v_1, v_2, 0, 0, 0)$.

Exercise 3.2.2

$$\begin{aligned}(ST)^2 v &= (ST)(ST)v \\ &= (ST)w && (w \in \operatorname{null} T) \\ &= S(Tw) \\ &= S(0) \\ &= 0\end{aligned}$$

Exercise 3.2.3
(a) $\dim \operatorname{null} T = 0$

(b) $z_i \neq 0$ for some set of $\forall i v_i$ are linearly independent

Exercise 3.2.4
$\dim \boldsymbol{R}^5 = 5 \neq \dim T + \dim \operatorname{range} T = 2 + 2$.

Exercise 3.2.5
Let

$$T(w, x, y, z) = (0, 0, w, x)$$

Then

$$\begin{aligned}\operatorname{range} T &= (0, 0, v_1, v_2) \\ \operatorname{null} T &= (0, 0, v_1, v_2)\end{aligned}$$

Exercise 3.2.6
$\dim \operatorname{range} T \neq \dim \operatorname{null} T$ since they have to add up to 5, so they cannot be equal.

gonna power through rest of the book

3.3 Matrices

Matrices are just a tool to represent linear maps, but they are no different or special.

Exercise 3.3.1
We can imagine $\dim \operatorname{range} T$ as the number of linearly independent columns in the matrix, which has to be at least the dimension, and those are nonzero.

Exercise 3.3.2
I think you just want to shift the constant term, so

$$\begin{matrix} a_1 \\ a_2 \\ a_3 \\ a_0 \end{matrix}$$

We didn't really learn much in this chapter except for matrix operations, so going to skip this.

3.4 Invertibility and Isomorphic Vector Spaces

The author says the definition of an operator is extremely important.

Definition 3.4.1
We define **operator** $\mathcal{L}(V)$ to be

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

$$\mathcal{L}(V) = \mathcal{L}(V, V) \tag{3.3}$$

I'm going to be honest that I don't really understand this definition (as in how it works and why it's important), so I'm getting rekt in the problems already.

These problems are getting harder for sure.

Exercise 3.4.1
We can verify

$$(ST)^{-1}(ST) = T^{-1}S^{-1}ST = T^{-1}IT = T^{-1}T = I$$
$$(ST)(ST)^{-1} = STT^{-1}S^{-1} = SIS^{-1} = SS^{-1} = I$$

Exercise 3.4.2
The operators are noninvertible, so they cannot be surjective or injective.

A little stuck here but I think by using that fact you can show that they may not be closed under addition or something of that sort.

Exercise 3.4.3
If T invertible exists, then S must be injective, otherwise if $S(x) = S(x') = v$ then $T(v)$ cannot be defined.

If S is injective, then we know for every $v \in V$, there is at most one $u \in U$ such that $S(u) = v$, so we just define T to invert these v that have a corresponding u.

Exercise 3.4.4
If $T_1 = ST_2$, then suppose $v \in \operatorname{null} T_1$. Then we have

$$\begin{aligned} T_1 v = 0 &= ST_2 v \\ S^{-1}(0) &= S^{-1}ST_2 v && \text{(Since } S \text{ invertible.)} \\ 0 &= T_2 v \end{aligned}$$

so $v \in \operatorname{null} T_2$ as well

If $\operatorname{null} T_1 = \operatorname{null} T_2$...**TODO** My intuition is that since T_1, T_2 span the same space (I think...bc of the null spaces being the same), their basis vectors are similar enough where you can make them the same by using some linear map.

Exercise 3.4.5
I think the $T_1 = T_2 S$ direction is pretty much the same as 3.4.4.

The other direction...also **TODO**

Exercise 3.4.6
TODO

Exercise 3.4.7

(a) $0 \in E$, $(T_1 + T_2)v = T_1 v + T_2 v = 0 + 0 = 0$, so $T_1 + T_2 \in E$, and $\lambda T v = T(\lambda v) = 0$.

(b) I think the dimension is 1??? I want to say something about $\operatorname{null} V$, but we are only taking a single vector.

Exercise 3.4.8
skip

Exercise 3.4.9
We know $(ST)^{-1} = T^{-1}S^{-1}$, so if ST is invertible, both T, S need to be invertible[1].
Alternatively, if T, S are invertible, then we can define $(ST)^{-1} = T^{-1}S^{-1}$.

Exercise 3.4.10
Forward direction,

$$\begin{aligned} ST &= I \\ STS &= IS = S \\ S^{-1}STS &= TS = S^{-1}S = I \end{aligned}$$

Backwards direction,

$$\begin{aligned} TS &= I \\ STS &= SI = S \\ STSS^{-1}TS &= SS^{-1} = I \end{aligned}$$

Exercise 3.4.11

$$\begin{aligned} STU &= I \\ T &= S^{-1}U^{-1} \\ T^{-}1 &= US \end{aligned}$$

Sorry did a lackluster job, on we go.

[1] sorry this is probably not an actual proof

3.5 Products and Quotients of Vector Spaces

Quotients are something we don't usually get in our intro linear algebra class, so that's cool.

Exercise 3.5.1
If T is a linear map, then we get that the graph of T is a subspace pretty easily with linear properties.

If the graph of T is a subspace, then we have that

$$(v_1, Tv_1) + (v_2, Tv_2) = (v_1 + v_2, Tv_1 + Tv_2) \in \text{graph of } T$$

here we see that $Tv_1 + Tv_2 = Tv'$ for some $v' \in V$. Then we must have (v', Tv') in the graph as well, which means

$$(v_1 + v_2, Tv_1 + Tv_2) = (v', Tv')$$

and we have $v' = v_1 + v_2$, so we conclude that $Tv_1 + Tv_2 = T(v_1 + v_2)$.

The proof for scalar is pretty similar.

Exercise 3.5.2
If $\exists V_j$ infinite-dimensional, then the cross product could be infinite dimensional.

Exercise 3.5.3
Let $U_1, U_2 = (u_1, u_2, u_3, \ldots)$, then

$$U_1 \times U_2 = \boldsymbol{F}^\infty = U_1 + U_2,$$

but we clearly have multiple choices for $u \in U_1, u' \in U_2$ such that their sum is the same.

Exercise 3.5.4
Just intuition here, one is mapping an m tuple's elements each to W, the other is a m-tuple of $V_i \to W$ mappings.

Exercise 3.5.5
Very similar intuition to Exercise 3.5.4.

Exercise 3.5.6
Let's take advantage of the last two exercises, so we have

$$\begin{aligned}\mathcal{L}(\boldsymbol{F}^n, V) &= \prod_{i=1}^{n} \mathcal{L}(\boldsymbol{F}, V) \\ &= \mathcal{L}(\boldsymbol{F}, V^n) = V^n\end{aligned}$$

Exercise 3.5.7
AFSOC $U \neq W$, then WLOG $\exists w \in W, \notin U$. Then we have that

$$v \in V, \forall u \in U \quad v + w \neq v + u$$

which violates our assumption.

Exercise 3.5.14

(a) We can check a few things here

- $0 \in U$
- Closed by addition since finite $x_j \neq 0$ will still result in that property
- Closed by scalar since the finite $x_j \neq 0$ still holds

Thus we can conclude the subspace

(b) Consider the mapping $f : U \to \boldsymbol{F}^\infty$ where for $u \in U$, say j is the largest $x_j \neq 0$, then

$$f(u) = (\ldots, x_j, x_j, x_j, \ldots)$$

This $f(u)$ has infinitely many $x_i \neq 0$, so $f(u) \notin u$, but $f(u) \in \boldsymbol{F}^\infty$.

Since there are infinite number of coordinates that are nonzero, we must have infinite dimension.

3.6 Duality

Duality has been hard for me to understand, but it seems to be some corresponding linear operator that can be constructed from any basis of a vector space.

Definition 3.6.1
The **dual space** of V, denoted V', is the vector space of all linear functionals on V. In other words, $V' = \mathcal{L}(V, \boldsymbol{F})$.

Definition 3.6.2
For $U \subset V$, the **annihilator** of U, denoted U^0, is defined by

$$U^0 = \{\phi \in V' \,:\, \phi(u) = 0, \forall u \in U\} \tag{3.4}$$

Yeah I tried reading through the section and 80% made sense, but I would still myself rekt by not really understanding what a dual it and how to use it.

At least the matrix definitions are pretty trivial.

If you want to prove the row and column rank are the same, you can use Gaussian elimination instead, although we haven't learned that in this book yet. But most intro courses would definitely introduce Gaussian elimination, since it's an easy way to solve linear equations.

Too lazy to do exercises, hope I can get by this course without them...

Chapter 4

Polynomials

Short chapter let's go...

Not a huge fan of using the Complex Analysis Liouville's Theorem to prove the Fundamental Theorem of Algebra, but it works.

Okay yeah this chapter was truly short and I would say anyone who has taken Algebra II can probably skip this chapter.

I'm going to omit the exercises since there wasn't any new content introduced in this chapter.

Chapter 5

Eigenvalues, Eigenvectors, and Invariant Subspaces

I think the first time learning about Eigenvalues was super lit. Just the name sounds legendary, and when you see the results from it, you not only get cool results, but applicable properties that you find in so many fields, especially in Machine Learning and Stats (e.g. PCA).

5.1 Invariant Subspaces

Definition 5.1.1
Suppose $T \in \mathcal{L}(V)$. A subspace U of V is called **invariant** under T if $u \in U$ implies $Tu \in U$.

I think having this definition of an invariant subspace is cool, as usually we think of eigenvalues and eigenvectors as magical $Tv = \lambda v$ property, but not as the invariant subspace of $\operatorname{span} v$.

Definition 5.1.2
Suppose $T \in \mathcal{L}(V)$. A number $\lambda \in \boldsymbol{F}$ is called an **eigenvalue** of T if there exists $v \in V$ such that $v \neq 0$ and $Tv = \lambda v$. This v is called an **eigenvector** of T corresponding to λ.

I think the cool part of linear algebra is that the linear independence, span, dimension and other properties we learned earlier all apply to linearly-behaving entities. See the next Theorem.

Theorem 2 *Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T and $v_1, \ldots, v_m$ are corresponding eigenvectors. Then $v_1, \ldots, v_m$ is linearly independent.*

I think this chapter is lacking practical techniques to solve for eigenvalues and eigenvectors, but it is a second course on Linear Algebra after all...and it's also more pure math anyway.

Exercise 5.1.1

(a) $u \in U$ means $Tu = 0 \in U$.

(b) $u \in U$ means $Tu \in \operatorname{range} T \implies Tu \in U$.

Exercise 5.1.2
Choose $v \in \operatorname{null} S$,

$$\begin{aligned} Sv &= 0 \\ TSv &= T(0) = 0 \\ STv &= 0 && \text{(assumption of problem)} \\ S(Tv) &= 0 \end{aligned}$$

Therefore $Tv \in \operatorname{null} S$.

Exercise 5.1.3
I think this is essentially a repeat of 5.1.2.

Exercise 5.1.4
We can use the distributive property of T since it is linear, and case on each U_i.

Exercise 5.1.5
If we let this collection be U_i, then if we have some $u \in \bigcap_i U_i$, then $Tu \in U_i, \forall i \implies \in \bigcap_i U_i$.

Exercise 5.1.6
Seems true.

Exercise 5.1.7
Our system of equations is

$$\begin{aligned} -3y &= \lambda x \\ x &= \lambda y \end{aligned}$$

so we have

$$-3y = \lambda^2 y \implies \lambda^2 = -3$$

and we find $\lambda = \pm\sqrt{3}i$.

Exercise 5.1.8
Our system of equations is

$$\begin{aligned} w &= \lambda z \\ z &= \lambda w \end{aligned}$$

so we have

$$w = \lambda^2 w \implies \lambda^2 = 1$$

and we find $\lambda = \pm 1$.

Exercise 5.1.9
Our system of equations is

$$\begin{aligned} 2z_2 &= \lambda z_1 \\ 0 &= \lambda z_2 \\ 5z_3 &= \lambda z_3 \end{aligned}$$

I think here we have to conclude that $z_2 = 0$, otherwise if $\lambda = 0$ then all $z_i = 0$, and that wouldn't be an eigenvector.

Then we have $0 = \lambda z_1$, and we conclude $z_1 = 0$. We are just left with $\lambda = 5$, and our eigenvector is $(0, 0, 1)$.

Exercise 5.1.10

(a) Eigenvectors are the standard basis, with $\lambda = i$ for e_i.

(b) Invariant subspaces are just the spans of e_i individually.

5.2 Eigenvectors and Upper-Triangular Matrices

Exercise 5.2.1

(a) Not actually sure how to prove this is invertible...**TODO**

(b) It's like adding all the parts of T back in.

Exercise 5.2.2
Fundamental Theorem of Algebra.

Exercise 5.2.3
$T^2 - I = (T - I)(T + I) = 0$. If $\lambda \neq -1$, then $T \neq -I$, so $T = I$ in this case.

Exercise 5.2.4
$P(P - I) = 0$ means that $\operatorname{null} P = \operatorname{range}(P - I)$. Stuck here...

Exercise 5.2.5
The key here is that $(STS^{-1})^n$ expands in a way that cancels out all S, S^{-1} except for the outer ones. E.g.

$$\left(STS^{-1}\right)^3 = (STS^{-1})(STS^{-1})(STS^{-1}) = ST\left(S^{-1}S\right)T\left(S^{-1}S\right)TS^{-1} = ST^3S^{-1}$$

Exercise 5.2.6
Applying T^n to U will still be invariant in U.

Exercise 5.2.7
$T^2 = 9 \implies (T - 3)(T + 3) = 0$.

Exercise 5.2.8
Not sure if this exists.

Exercise 5.2.9
By FTA we can write $p(T) = \prod_i (T - \lambda_i I)$ So for any $(T - \lambda_i I)v = 0$, this implies λ_i is an eigenvalue, which is a zero of p.

Exercise 5.2.10

$$\begin{aligned} p(T)v &= \left(\sum_{i=0}^{n} a_i T^i\right) v \\ &= \sum_{i=0}^{n} a_i T^i v \\ &= \sum_{i=0}^{n} a_i \lambda_i^i v \\ &= \left(\sum_{i=0}^{n} a_i \lambda_i^i\right) v \\ &= p(\lambda) v \end{aligned}$$

5.3 Eigenspaces and Diagonal Matrices

Theorem 3 *Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T. Then the following are equivalent:*

(a) *T is diagonalizable*

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

(c) *There exist 1-dimensional subspaces $U_1, \ldots, U_n$ of V, each invariant under T, such that*

$$V = \bigoplus_{i=1}^{n} U_i \tag{5.1}$$

(d) $V = \bigoplus_{i=1}^{m} E(\lambda_i, T)$

(e) $\dim V = \sum_{i=1}^{m} \dim E(\lambda_i, T)$

Definition 5.3.1
Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \boldsymbol{F}$. The **eigenspace** of T corresponding to λ, denoted $E(\lambda, T)$, is defined by

$$E(\lambda, T) = \operatorname{null}(T - \lambda I). \tag{5.2}$$

So $E(\lambda, T)$ is the span of all eigenvectors corresponding to λ plus the 0 vector.

Exercise 5.3.1
$V = \operatorname{range} T$ since T is diagonalizable, and $\operatorname{null} T = \{0\}$.

Exercise 5.3.2
I don't think it is always true, we just need a $\operatorname{null} T$ that is nontrivial.

Exercise 5.3.3
If we have (a) then (b) follows directly. I think (b) also gives (a) pretty easily because we know $\operatorname{null} T, \operatorname{range} T$ cannot have common elements besides 0.

I think (b) gives us (c) directly from the discussion earlier. I think we still need to show (c) $\implies$ (a) or (b)...

Exercise 5.3.4
TODO

I should do more of these problems :)

Chapter 6

Inner Product Spaces

6.1 Inner Products and Norms

Working a lot with $\boldsymbol{R}^2$ in this chapter, so there is geometry, which is easy to visualize.

Definition 6.1.1
An **inner product** on V is a function that takes each ordered pair (u, v) of elements of V to a number $\langle u, v\rangle \in \boldsymbol{F}$ and has the following properties:

- positivity
$$\langle v, v\rangle \geq 0, \forall v \in V \tag{6.1}$$
- definiteness
$$\langle v, v\rangle = 0 \text{ iff } v = 0 \tag{6.2}$$
- additivity in first slot
$$\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle, \forall u, v, w \in V \tag{6.3}$$
- homogeneity in first slot
$$\langle \lambda u, v\rangle = \lambda\langle u, v\rangle, \forall \lambda \in \boldsymbol{F} \text{ and all } u, v, w \in V \tag{6.4}$$
- conjugate symmetry
$$\langle u, v\rangle = \overline{\langle v, u\rangle}, \forall u, v \in V \tag{6.5}$$

Theorem 4 *Suppose $u, v \in V$. Then*
$$|\langle u, v\rangle| \leq \|u\|\|v\| > \tag{6.6}$$
The inequality is an equality iff one of u, v is a scalar multiple of the other.

The Cauchy-Schwarz inequality is **very commonly used** in many parts of math, so I would pay attention to this and remember it!

Exercise 6.1.1
I don't think additivity in first slot works

Exercise 6.1.2
definiteness does not work, since we aren't considering the second coordinate.

Exercise 6.1.3
not sure how to do this **TODO**

Exercise 6.1.4

(a) We have

$$\begin{aligned}\langle u+v, u-v\rangle &= \langle u, u-v\rangle + \langle v, u-v\rangle \\ &= \langle u,u\rangle - \langle u,v\rangle + \langle v,u\rangle - \langle v,v\rangle \\ &= \|u\|^2 - \|v\|^2 && \text{(Since } u, v \text{ real, } \langle v,u\rangle = \langle u,v\rangle\text{)}\end{aligned}$$

(b) From above, then the inner product is 0.

(c) The diagonals of a rhombus are $u+v, u-v$, and all sides are equal, so $\|u\| = \|v\|$.

Exercise 6.1.5

If $\|Tv\| \leq \|v\|$, then it must be the case that if there is an eigenvalue, then $\lambda \leq 1 < \sqrt{2}$, so thus we know $\sqrt{2}$ is not an eigenvalue and $T - \sqrt{2}I$ is invertible.

Exercise 6.1.6

Forward direction is trivial, we just see $0 \leq X$ where X is some inner product, which must be nonnegative.

Backward direction we have

$$\begin{aligned}\|u\|^2 &\leq \langle u+av, u+av\rangle \\ &\leq \langle u, u+av\rangle + \langle av, u+av\rangle \\ &\leq \overline{\langle u,u\rangle + \langle av,u\rangle + \langle u,av\rangle + \langle av,av\rangle} \\ 0 &\leq \overline{a}\langle u,v\rangle + a\langle v,u\rangle + \|a\|^2\|v\|^2 \\ &\leq 2\,\mathrm{Re}(a\langle v,u\rangle) + \|a\|^2\|v\|^2\end{aligned}$$

If we want this to hold $\forall a$, I think we need the real part to be zero, so we need $\langle v,u\rangle = 0$? But this is not $\langle u,v\rangle$.

Exercise 6.1.7

$$\begin{aligned}\|au+bv\|^2 &= \langle au+bv, au+bv\rangle \\ &= a\langle u, au+bv\rangle + b\langle v, au+bv\rangle \\ &= a\overline{\langle au+bv, u\rangle} + b\overline{\langle au+bv, v\rangle} \\ &= a\big(\overline{a}\langle u,u\rangle + \overline{b}\langle u,v\rangle\big) + b\big(\overline{a}\langle v,u\rangle + \overline{b}\langle v,v\rangle\big) \\ &= a\overline{a}\|u\|^2 + a\overline{b}\langle u,v\rangle + b\overline{a}\langle v,u\rangle + b\overline{b}\|v\|^2\end{aligned}$$

$$\begin{aligned}\|bu+av\|^2 &= \langle bu+av, bu+av\rangle \\ &= b\langle u, bu+av\rangle + a\langle v, bu+av\rangle \\ &= b\big(\overline{b}\langle u,u\rangle + \overline{a}\langle u,v\rangle\big) + a\big(\overline{b}\langle v,u\rangle + \overline{a}\langle v,v\rangle\big) \\ &= a\overline{a}\|v\|^2 + a\overline{b}\langle v,u\rangle + b\overline{a}\langle u,v\rangle + b\overline{b}\|u\|^2\end{aligned}$$

If we have $\|u\| = \|v\|$, we just have to show

$$a\overline{b}\langle u,v\rangle + b\overline{a}\langle v,u\rangle = a\overline{b}\langle v,u\rangle + b\overline{a}\langle u,v\rangle$$

I might've missed something...don't think this should be this hard **TODO**

Exercise 6.1.8

$$|\langle u,v\rangle| = 1$$

means by the Cauchy-Schwarz inequality that $u = kv$. If we have that $\|u\| = \|v\| = 1$, then it must be the case that $k = 1$, so we have $u = v$ as desired.

Exercise 6.1.9
This is plug and chug Cauchy-Schwarz.

Exercise 6.1.10
I think the system of equations to solve is

$$\begin{aligned} c &= \frac{\langle u, (1,2)\rangle}{5} \\ v &= u - \frac{\langle u, (1,2)\rangle}{5}(1,2) \\ u &= c(1,2) + v \end{aligned}$$

The hardest part of the problem is re-assigning the variables from earlier (u, v, w) to the ones corresponding to this problem.

Exercise 6.1.11
Let

$$\begin{aligned} u &= \left(\sqrt{a}, \sqrt{b}, \sqrt{c}, \sqrt{d}\right) \\ v &= \left(1/\sqrt{a}, 1/\sqrt{b}, 1/\sqrt{c}, 1/\sqrt{d}\right) \end{aligned}$$

and use the Cauchy-Schwarz inequality,

$$\begin{aligned} |\langle u, v\rangle| = \left|1^2 + 1^2 + 1^2 + 1^2\right| = 4 &\leq \sqrt{a+b+c+d}\sqrt{1/a+1/b+1/c+1/d} = \|u\|\|v\| \\ 16 &\leq (a+b+c+d)(1/a+1/b+1/c+1/d) = \|u\|\|v\| \quad \text{(square both sides)} \end{aligned}$$

Exercise 6.1.12
This is just Cauchy-Schwarz with

$$\begin{aligned} u &= (x_1, x_2, \ldots, x_n) \\ v &= (1, 1, \ldots, 1) && (n \text{ 1s}) \end{aligned}$$

Exercise 6.1.13
Let's do some basic geometry

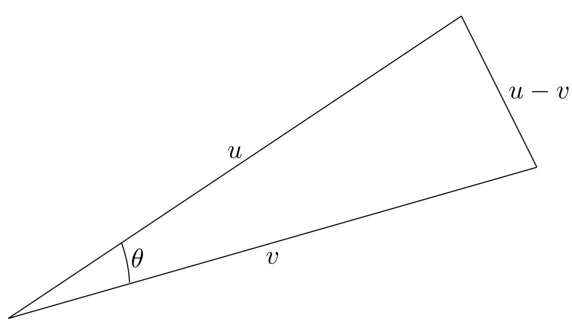

Figure 6.1: Showing how to apply law of cosines.

By the law of cosines, this gives us

$$\begin{aligned} \|u-v\|^2 &= \|u\|^2 + \|v\|^2 - 2\|u\|\|v\|\cos\theta \\ \langle u-v, u-v\rangle = \langle u, u\rangle + \langle v, v\rangle - 2\langle u, v\rangle &= \|u\|^2 + \|v\|^2 - 2\|u\|\|v\|\cos\theta \\ \langle u, v\rangle &= \|u\|\|v\|\cos\theta \end{aligned}$$

Exercise 6.1.14
arccos is only defined with inputs $\in [-1, 1]$, so the Cauchy-Schwarz Inequality gives us this guarantee.

Exercise 6.1.15
Pretty sure this is just Cauchy-Schwarz again. Whenever you see sums and inequalities, try Cauchy-Schwarz.

Exercise 6.1.16
Let's draw this out,

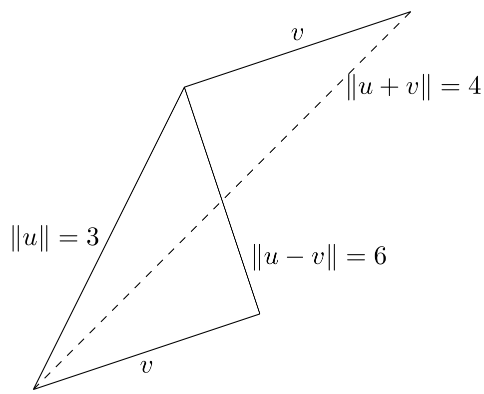

Figure 6.2: In geometry problems, always draw a good diagram.

Cool, so we have a parallelogram, so we can use 6.22 the parallelogram equality to give

$$4^2 + 6^2 = 2(3^2 + \|v\|^2) \implies \|v\| = \sqrt{17}.$$

6.2 Orthonormal Bases

Calculating orthonormal bases can be a pain, as you probably remember from Gram-Schmidt.

Definition 6.2.1
A list of vectors is called **orthonormal** if each vector in the list has norm 1 and is orthogonal to all other vectors in the list.

E.g. a list $e_1, \ldots, e_m$ of vectors in V is orthonormal if

$$\langle e_j, e_k \rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases} \tag{6.7}$$

Theorem 5 *Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V and $v \in V$. Then*

$$v = \sum_{i=1}^{n} \langle v, e_i \rangle e_i \tag{6.8}$$

and

$$\|v\|^2 = \sum_{i=1}^{n} |\langle v, e_i \rangle|^2 \tag{6.9}$$

The following theorem is pretty surprising,

Theorem 6 *Suppose V is finite-dimensional and ϕ is a linear functional on V. Then there is a unique vector $u \in V$ such that*

$$\phi(v) = \langle v, u \rangle \tag{6.10}$$

for every $v \in V$.

Specifically, the vector u is

$$u = \sum_{i=1}^{n} \overline{\phi(e_i)} e_i \tag{6.11}$$

Exercise 6.2.1

(a) Multiply them together, and use $\sin^2 \theta + \cos^2 \theta = 1$ for the norm.

(b) Unsure how to do this tbh **TODO**

Exercise 6.2.2
Decompose v into the orthogonal parts, and then take the inner product with itself to find the norm.

I'm too lazy to do G-S and the rest of the problems, rip, but I think the results in this chapter were cool, esp with Riesz.

6.3 Orthogonal Complements and Minimization Problems

Definition 6.3.1
If U is a subset of V, then the **orthogonal complement** of U, denoted $U^\perp$, is the set of all vectors in V that are orthogonal to every vector in U:

$$U^\perp = \{v \in V : \langle v, u\rangle = 0, \forall u \in U\} \tag{6.12}$$

The following theorem is a good optimization technique used in many fields,

Theorem 7 *Suppose U is a finite-dimensional subspace of $V, v \in V$, and $u \in U$. Then*

$$\|v - P_U v\| \leq \|v - u\|. \tag{6.13}$$

Furthermore, the inequality is an equality iff $u = P_U v$.

Exercise 6.3.1
We can show this by using $V = U \oplus U^\perp$.

Exercise 6.3.2
We can show this by using $V = U \oplus U^\perp$.

Exercise 6.3.3
Our earlier exercises tell us the u_i span U and w_1 span $U^\perp$.

Now if we apply Gram-Schmidt, the corresponding orthonormal vectors will still span the same spaces.

Exercise 6.3.4
Too lazy to do G-S.

Chapter 7

Operators on Inner Product Spaces

7.1 Self-Adjoint and Normal Operators

Adjoint operators are used a lot in higher-level physics classes, especially in quantum.

Definition 7.1.1
Suppose $T \in \mathcal{L}(V, W)$. The **adjoint** of T is the function $T^* : W \to V$ such that

$$\langle Tv, w \rangle = \langle v, T^* w \rangle \tag{7.1}$$

for every $v \in V$ an every $w \in W$.

Theorem 8 *If $T \in \mathcal{L}(V, W)$, then $T^* \in \mathcal{L}(W, V)$.*

Definition 7.1.2
An operator on an inner product space is called **normal** if it commutes with its adjoint.
So $T \in \mathcal{L}(V)$ is normal if

$$TT^* = T^*T. \tag{7.2}$$

Exercise 7.1.1

$$\begin{aligned}
\langle (z_1, \dots, z_n), T^*(x_1, \dots, x_n) \rangle &= \langle T(z_1, \dots, z_n), (x_1, \dots, x_n) \rangle \\
&= \langle (0, z_1, \dots, z_{n-1}), (x_1, \dots, x_n) \rangle \\
&= \sum_{i=2}^{n} x_i z_{i-1} \\
&= \langle (z_1, \dots, z_{n-1}, z_n), (x_2, \dots, x_n, 0) \rangle
\end{aligned}$$

So we see that

$$T^*(x_1, \dots, x_n) = (x_2, \dots, x_n, 0). \tag{7.3}$$

Exercise 7.1.2

$$\begin{aligned}
\langle \lambda v, v \rangle &= \langle Tv, v \rangle \\
&= \langle v, T^* v \rangle \\
&= \left\langle v, \overline{\lambda} v \right\rangle
\end{aligned}$$

Exercise 7.1.3
Let $u \in U, u' \in U^{\perp}$,

$$\langle Tu, u' \rangle = 0 \iff \langle u, T^* u' \rangle = 0 \tag{7.4}$$

Exercise 7.1.4
skip

Exercise 7.1.5
skip

Exercise 7.1.6

(a) We see that

$$\begin{aligned}\langle Tp, q\rangle &= p_1 x \\ \langle p, Tq\rangle &= q_1 x\end{aligned}$$

and these expressions are not equal.

(b) Our basis is not orthonormal.

Exercise 7.1.7
If ST is self-adjoint, then

$$\begin{aligned}ST &= (ST)^* && \text{(self-adjoint)} \\ &= T^*S^* && \text{(adjoint properties)} \\ &= TS && (S, T \text{ self-adjoint})\end{aligned}$$

If $ST = TS$, then we have

$$\begin{aligned}(ST)^* &= T^*S^* \\ &= TS && (S, T \text{ self-adjoint}) \\ &= ST. && \text{(by assumption)}\end{aligned}$$

Exercise 7.1.8

(i) 0 is self-adjoint.

(ii) Suppose S, T are self-adjoint. Then

$$(S+T)^* = S^* + T^* = S + T$$

(iii) Suppose S is self-adjoint. Then

$$\begin{aligned}\langle \lambda Sv, w\rangle &= \lambda\langle Sv, w\rangle \\ &= \lambda\langle v, S^*w\rangle \\ &= \langle v, \overline{\lambda} Sw\rangle\end{aligned}$$

I think $\lambda \in \boldsymbol{F}$ so we're good here, so $\overline{\lambda} = \lambda$?

Exercise 7.1.9
Additivity fails because if we try

$$\begin{aligned}(T+S)(S+T)^* &= (T+S)(S^* + T^*) \\ &= TS^* + SS^* + ST^* + TT^*,\end{aligned}$$

we're quickly stuck because T and S have no guarantees between them about commutativity.

7.2 The Spectral Theorem

One of the most lit theorems in Linear Algebra, listen up.

Theorem 9 *Suppose $\boldsymbol{F} = \boldsymbol{C}$ and $T \in \mathcal{L}(V)$. Then the following are equivalent:*

(a) T is normal.

(b) V has an orthonormal basis consisting of eigenvectors of T.

(c) T has a diagonal matrix with respect to some orthonormal basis of V.

Theorem 10 *Suppose $\boldsymbol{F} = \boldsymbol{R}$ and $T \in \mathcal{L}(V)$. Then the following are equivalent:*

(a) T is self-adjoint.

(b) V has an orthonormal basis consisting of eigenvectors of T.

(c) T has a diagonal matrix with respect to some orthonormal basis of V.

Exercise 7.2.1
Seems like this is true, the spectral theorem is about orthonormal bases and diagonal matrices. In general we can have T with eigenvectors that span $\boldsymbol{R}^3$.

Exercise 7.2.2
Do we just factor this to $(T - 2I)(T - 3I) = 0$?

Exercise 7.2.3
If we choose

$$T = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$

Then we have the expression equal to

$$T^2 - 5T + 6I = \begin{pmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Exercise 7.2.4
If T is normal, by the Complex Spectral Theorem (9), we know these statements are true.

Exercise 7.2.5
This is the Real Spectral Theorem (10).

Exercise 7.2.6
TODO

7.3 Positive Operators and Isometrics

Definition 7.3.1
An operator $T \in \mathcal{L}(V)$ is called **positive** if T is self-adjoint and

$$\forall v \in V \quad \langle Tv, v\rangle \geq 0 \tag{7.5}$$

Definition 7.3.2
An operator $S \in \mathcal{L}(V)$ is called an **isometry** if

$$\|Sv\| = \|v\|, \forall v \in V. \tag{7.6}$$

An operator is an isometry if it preserves norms.

Exercise 7.3.1
Consider some $v \in V$. We can represent v as part of the orthonormal basis

$$v = \sum_{i=1}^{n} a_i e_i$$

and now if we consider the positive operator definition

$$\begin{aligned}\langle Tv, v\rangle &= \sum_{i=1}^{n} \langle T(a_i e_i), v\rangle \\ &= \sum_{i=1}^{n} \langle T(a_i e_i), a_i e_i\rangle && \text{(Since } \langle e_i, e_j\rangle = 0, i \neq j) \\ &= \sum_{i=1}^{n} a_i^2 \langle T(e_i), e_i\rangle \geq 0.\end{aligned}$$

So we can conclude that T is a positive operator.

Exercise 7.3.3
If U is invariant under $T|_U$, then $T|_U u \in U$.
Now, we have

$$\langle T|_U u, u\rangle \geq 0$$

since $u \in U \subset V \implies u \in V$, and T is a positive operator on V.

This solution seems kinda fishy since I didn't really use the invariant property. Perhaps it's just to say that T only operates within U? We already know that any element $v \in V$ satisfies the positive operator condition, and every element of U is in V...

Exercise 7.3.4
For $v \in V$,

$$\langle T^*Tv, v\rangle = \langle Tv, Tv\rangle \geq 0$$

For $w \in W$,

$$\langle TT^*w, w\rangle = \langle T^*v, T^*v\rangle \geq 0$$

Exercise 7.3.5
Suppose we have S, T positive operators on V. Then for $v \in V$,

$$\begin{aligned}\langle (S+T)v, v\rangle &= \langle Sv + Tv, v\rangle \\ &= \langle Sv, v\rangle + \langle Tv, v\rangle \geq 0\end{aligned}$$

Exercise 7.3.6
We proceed by induction.

Base case $k = 1$ is easy since that's just the definition.

Suppose we have $\left\langle T^k v, v \right\rangle \geq 0$ for some $k \geq 1$. Now, consider $\left\langle T^{k+1} v, v \right\rangle$. We know that T has an orthonormal basis with nonnegative eigenvalues λ_i. Then if we consider

$$\left\langle T^k v, v \right\rangle = \sum_{i=1}^{n} \langle a_i e_i, e_i \rangle$$

then if we multiply by T, we have

$$\left\langle T^{k+1} v, v \right\rangle = \sum_{i=1}^{n} \langle \lambda_i a_i e_i, e_i \rangle = \geq 0$$

We have to show that $a_i \geq 0$, but this we can prove with a lemma with the positive eigenvalues.

7.4 Polar Decomposition and Singular Value Decomposition

SVD is used extensively in Computer Vision.

Theorem 11 *Polar Decomposition. Suppose $T \in \mathcal{L}(V)$. Then $\exists$ isometry $S \in \mathcal{L}(V)$ such that*

$$T = S\sqrt{T^*T}. \tag{7.7}$$

Note that in the proof, we used the fact that $\sqrt{T^*T}$ is a self-adjoint square root, so we were able to pull off

$$\left\langle \sqrt{T^*T}\sqrt{T^*T}v, v \right\rangle = \left\langle \sqrt{T^*T}v, \sqrt{T^*T}v \right\rangle$$

Definition 7.4.1
Suppose $T \in \mathcal{L}(V)$. The **singular values** of T are the eigenvalues of $\sqrt{T^*T}$, with each eigenvalue λ repeated $\dim E(\lambda, \sqrt{T^*T})$ times.

Theorem 12 *Singular Value Decomposition. Suppose $T \in \mathcal{L}(V)$ has singular values $s_1, \ldots, s_n$. Then $\exists$ orthonormal bases $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ of V such that*

$$Tv = \sum_{i=1}^{n} s_i \langle v, e_i \rangle f_i, \forall v \in V. \tag{7.8}$$

Too lazy to do these exercises :(

Chapter 8

Operators on Complex Vector Spaces

I think most intro linear algebra courses would end at Chapter 7, but hey, here we are, ready to learn more math. Give yourself a pat on the back.

8.1 Generalized Eigenvectors and Nilpotent Operators

Definition 8.1.1
Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. A vector $v \in V$ is called a **generalized eigenvector** of T corresponding to λ if $v \neq 0$ and

$$(T - \lambda I)^j v = 0 \tag{8.1}$$

for some $j \in \boldsymbol{Z}^+$.

Definition 8.1.2
Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \boldsymbol{F}$. The **generalized eigenspace** of T corresponding to λ, denoted $G(\lambda, T)$, is defined to be the set of all generalized eigenvectors of T corresponding to λ, along with the 0 vector.

Definition 8.1.3
An operator is called **nilpotent** if some power of it equals 0.

Exercise 8.1.1
Guessing eigenvalues will give us 0. We can also solve the system

$$\begin{aligned} \lambda w &= z \\ \lambda z &= 0. \end{aligned}$$

Now to find the generalized spaces, we raise T minus the eigenvalue to 2 since that's the dimension of the space,

$$T^2(x, y) = (0, 0) \implies (x, y) \text{ in null space}$$

so we have $G(0, T) = \{(x, y) : x, y \in \boldsymbol{C}\}$.

We can check our answer that this sole eigenspace spans $\boldsymbol{C}^2$.

Exercise 8.1.2
To find the eigenvalues,

$$\begin{aligned} \lambda w &= -z \\ \lambda z &= w \\ \implies \lambda w &= -\frac{w}{\lambda} \\ \implies \lambda^2 w &= -1 \end{aligned}$$

which leads us to see that $\lambda = \pm i$.

We can see that

$$T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

So if we calculate

$$(T - (i)I)^{\dim(\boldsymbol{C}^2)} = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix}^2 = \begin{pmatrix} -2 & 2i \\ -2i & -2 \end{pmatrix}$$
$$(T - (-i)I)^{\dim(\boldsymbol{C}^2)} = \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix}^2 = \begin{pmatrix} -2 & -2i \\ 2i & -2 \end{pmatrix}$$

The null spaces are

$$G(i, T) = \operatorname{span}\left\{ \begin{pmatrix} i \\ 1 \end{pmatrix} \right\}$$
$$G(-i, T) = \operatorname{span}\left\{ \begin{pmatrix} -i \\ 1 \end{pmatrix} \right\}$$

Exercise 8.1.3

We now need to show that the $G\left(\frac{1}{\lambda}, T^{-1}\right) = G(\lambda, T)$ is the same. We can show this by

$$\begin{aligned}
\operatorname{null}(T - \lambda I)^{\dim V} v &= 0 \\
\operatorname{null}(I - \lambda T^{-1} I)^{\dim V} v &= 0 && \text{(Multiply by } \left(T^{-1}\right)^{\dim V}) \\
\operatorname{null}\left(\frac{1}{\lambda} I - T^{-1}\right)^{\dim V} v &= 0 && \text{(Multiply by } (1/\lambda)^{\dim V}) \\
\operatorname{null}\left(T^{-1} - \frac{1}{\lambda} I\right)^{\dim V} v &= 0 && \text{(Notice negative doesn't matter, RHS= 0)}
\end{aligned}$$

Exercise 8.1.4

AFSOC $\exists v \neq 0$ such that $v \in G(\alpha, T) \cap G(\beta, T)$.

Then v, v must be linearly independent by 8.13, since generalized eigenvectors corresponding to distinct eigenvalues must be linearly independent. But that's a contradiction, since v, v is clearly linearly dependent.

Exercise 8.1.5

I think first, we have to show that $T^k v \neq 0$ for $k < m$. This is pretty easy to show, because if $T^j v = 0$ for some $j < m$, then $T^k v = 0$ for $k \in [j, m]$, which is a contradiction since we know $T^{m-1} v \neq 0$.

Now for the main proof, we will proceed by induction.

$m = 0$ case is trivial, since we only have one vector.

Suppose we have $m > 0$, so we have $\left\{T^k v, k < m\right\}$ are linearly independent.

$$T^m v = 0$$

AFSOC $T^m v$ is linearly dependent on the previous $\left\{T^k v, k < m\right\}$ vectors. Then we have

$$T^m v = \sum_{i=0}^{m-1} a_i T^i v$$
$$0 = \sum_{i=0}^{m-1} a_i T^i v$$

which implies $a_i = 0, \forall i$, which is a contradiction, since that means $T^m v$ is not linearly dependent.

Therefore we conclude that $\left\{T^k v, 0 \leq k \leq m\right\}$ are linearly independent.

8.2 Decomposition of an Operator

Exercise 8.2.1
By the definition of the eigenvalue, we know $\exists j$ such that for $v \in V$,

$$(N - \lambda I)^j v = N^j v = 0$$

so N is nilpotent.

Exercise 8.2.2
Just from definitions we have $(T - 0I) = T$ should be nilpotent, but that's over a complex space.

Why does this not work in the real space? I'm struggling...I think for $\boldsymbol{R}^2$ I was not able to find anything...

Exercise 8.2.3

$$\begin{aligned} Tv &= \lambda v \\ STvS^{-1} &= S(\lambda v)S^{-1} \\ STS^{-1}v &= SS^{-1}\lambda v \\ STS^{-1}v &= \lambda v \end{aligned}$$

Exercise 8.2.4
no idea **TODO**

Exercise 8.2.5
If V has a basis of eigenvectors of T, then every generalized eigenvector is an eigenvector of T by definition.

If every generalized eigenvector of T is an eigenvector as well, then since we know $V = \bigoplus_i G(\lambda_i, T)$, then it must be the case that V has a basis consisting of eigenvectors of T.

Exercise 8.2.6
Too lazy, but just use the proof outline in the book **TODO**

Exercise 8.2.7
We can use the same proof in the textbook for square roots, except use the Taylor expansion of $(1 + x)^{1/3}$ for inspiration.

Exercise 8.2.8
rip didn't understand this either, related to problem 4 I believe

Exercise 8.2.9
This is just straight bookkeeping and computation.

8.3 Characteristic and Minimal Polynomials

Definition 8.3.1
Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T, with multiplicities $d_1, \ldots, d_m$. The polynomial

$$\prod_{i=1}^{m} (z - \lambda_i)^{d_i} \tag{8.2}$$

is called the **characteristic polynomial** of T.

Theorem 13 *Cayley-Hamilton Theorem. Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let q denote the characteristic polynomial of T. Then $q(T) = 0$.*

NOTE: I'm guessing for the most part for some of the operators since I'm lazy.

Exercise 8.3.1
This follows directly from the Cayley-Hamilton Theorem 14, after factoring out the extra multiplicity.

Exercise 8.3.2
This follows directly from the Cayley-Hamilton Theorem 14, after factoring out the extra multiplicity. It's pretty much just another special case compared to Exercise 8.4.1.

Exercise 8.3.3
We just need $\lambda = 7, 8$, both multiplicity of 2. So the following should do

$$T = \begin{pmatrix} 7 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}$$

Exercise 8.3.4
I'm just guessing, but

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

My guess is that having not just an entirely diagonal matrix should help us with the minimum polynomial only being able to being reduced to $(z - 5)^2$.

Exercise 8.3.5

$$T = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

Exercise 8.3.6

$$T = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

Exercise 8.3.7
We can refactor to

$$P(P-I)=0,$$

and if we consider $P(P-I)v=0$, then we get our desired result, $Pv=0$ for our null space, and $(P-I)v=0$ will give our complementary space, which will be $\dim \operatorname{range} P$[1].

Exercise 8.3.8
Not sure if there are tricks here, but I think you factor, and then take $1/\lambda$ for all the roots λ in the characteristic polynomial of T.

Exercise 8.3.9
We know what the characteristic polynomial of T^{-1} looks like, so let's manipulate it. Call the char. poly. of T^{-1} $q(z)$, then

$$\begin{aligned}
q(z) &= \prod_{i=1}^{\dim V}\left(z-\frac{1}{\lambda_i}\right)\\
&= z^{\dim V}\prod_{i=1}^{\dim V}\left(1-\frac{1}{z\lambda_i}\right)\\
&= \left(\prod_{i=1}^{\dim V}-\frac{1}{\lambda_i}\right)z^{\dim V}\prod_{i=1}^{\dim V}\left(1-\frac{1}{z}\right)\\
&= \frac{1}{p(0)}z^{\dim V}p\left(\frac{1}{z}\right)
\end{aligned}$$

The last result comes from the fact that

$$\begin{aligned}
p(0) &= \prod_{i=1}^{\dim V}(0-\lambda_i)\\
&= \prod_{i=1}^{\dim V}(-\lambda_i).
\end{aligned}$$

Exercise 8.3.10
Just use Exercise 8.4.9.

[1] this last part I'm not actually sure about, and I didn't formally prove it

8.4 Characteristic and Minimal Polynomials

Definition 8.4.1
Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T, with multiplicities $d_1, \ldots, d_m$. The polynomial

$$\prod_{i=1}^{m}(z-\lambda_i)^{d_i} \tag{8.3}$$

is called the **characteristic polynomial** of T.

Theorem 14 *Cayley-Hamilton Theorem. Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let q denote the characteristic polynomial of T. Then $q(T) = 0$.*

NOTE: I'm guessing for the most part for some of the operators since I'm lazy.

Exercise 8.4.1
This follows directly from the Cayley-Hamilton Theorem 14, after factoring out the extra multiplicity.

Exercise 8.4.2
This follows directly from the Cayley-Hamilton Theorem 14, after factoring out the extra multiplicity. It's pretty much just another special case compared to Exercise 8.4.1.

Exercise 8.4.3
We just need $\lambda = 7, 8$, both multiplicity of 2. So the following should do

$$T = \begin{pmatrix} 7 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}$$

Exercise 8.4.4
I'm just guessing, but

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

My guess is that having not just an entirely diagonal matrix should help us with the minimum polynomial only being able to being reduced to $(z-5)^2$.

Exercise 8.4.5

$$T = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

Exercise 8.4.6

$$T = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

Exercise 8.4.7
We can refactor to

$$P(P-I)=0,$$

and if we consider $P(P-I)v=0$, then we get our desired result, $Pv=0$ for our null space, and $(P-I)v=0$ will give our complementary space, which will be $\dim \operatorname{range} P$[2].

Exercise 8.4.8
Not sure if there are tricks here, but I think you factor, and then take $1/\lambda$ for all the roots λ in the characteristic polynomial of T.

Exercise 8.4.9
We know what the characteristic polynomial of T^{-1} looks like, so let's manipulate it. Call the char. poly. of T^{-1} $q(z)$, then

$$\begin{aligned}
q(z) &= \prod_{i=1}^{\dim V}\left(z-\frac{1}{\lambda_i}\right)\\
&= z^{\dim V}\prod_{i=1}^{\dim V}\left(1-\frac{1}{z\lambda_i}\right)\\
&= \left(\prod_{i=1}^{\dim V}-\frac{1}{\lambda_i}\right)z^{\dim V}\prod_{i=1}^{\dim V}\left(1-\frac{1}{z}\right)\\
&= \frac{1}{p(0)}z^{\dim V}p\left(\frac{1}{z}\right)
\end{aligned}$$

The last result comes from the fact that

$$\begin{aligned}
p(0) &= \prod_{i=1}^{\dim V}(0-\lambda_i)\\
&= \prod_{i=1}^{\dim V}(-\lambda_i).
\end{aligned}$$

Exercise 8.4.10
Just use Exercise 8.4.9.

[2]this last part I'm not actually sure about, and I didn't formally prove it

Chapter 9

Operators on Real Vector Spaces

9.1 Complexification

Definition 9.1.1
Suppose V is a real vector space.

- The **complexification** of V, denoted $V_{\boldsymbol{C}}$, equals $V \times V$. An element of $V_{\boldsymbol{C}}$ is an ordered pair (u, v), where $u, v \in V$, but we will write this as $u + iv$.
- Addition on $V_{\boldsymbol{C}}$ is defined by

$$(u_1 + iv_1) + (u_2 + iv_2) = (u_1 + u_2) + i(v_1 + v_2), u_1, u_2, v_1, v_2 \in V \tag{9.1}$$

- Complex scalar multiplication on $V_{\boldsymbol{C}}$ is defined by

$$(a + bi)(u + iv) = (au - bv) + i(av + bu), a, b \in \boldsymbol{R} \text{ and } u, v \in V \tag{9.2}$$

This chapter was mostly formalities, felt logistics-like, just proving the same results from complex into the real space.

Don't really feel like doing these exercises...

9.2 Operators on Real Inner Product Spaces

Don't feel like doing these either.

Chapter 10

Trace and Determinant

10.1 Trace

Ah...we're finally dealing with concrete matrices and numbers. This is stuff you should be familiar with from Linear Algebra I.

Definition 10.1.1
Suppose $T \in \mathcal{L}(V)$.

- If $\boldsymbol{F} = \boldsymbol{C}$, then the **trace** of T is the sum of eigenvalues of T, with each eigenvalue repeated according to its multiplicity.
- If $\boldsymbol{F} = \boldsymbol{R}$, then the **trace** of T is the sum of eigenvalues of $T_{\boldsymbol{C}}$, with each eigenvalue repeated according to its multiplicity.

The trace of T is denoted $\operatorname{trace} T$.

We also later see that the **trace** of a square matrix A, denoted $\operatorname{trace} A$, is defined to be the sum of the diagonal entries of A.

Exercise 10.1.1
Not sure how the author wants this to be proved, but every T has a corresponding matrix $\mathcal{M}$, so if T is invertible, then its matrix for the inverse must exist as well.

Exercise 10.1.2
Inverses are unique so we know $B = A^{-1}$.

Exercise 10.1.3
This means for the matrix $\mathcal{M}$ of T, we have

$$\mathcal{M} = A^{-1}\mathcal{M}A$$

for any change of basis. This isn't formal, but only $\mathcal{M} = I$ satisfies this equation it seems $\forall A$.

Exercise 10.1.4
T acts like the identity. We have to formalize this, but I think you can just use the linear combination fact, so first choose some $w \in V$, then

$$w = \sum a_i v_i$$
$$w = \sum b_i u_i$$

And now we can show $Tw = w$, by two steps, first

$$Tw = \sum T(a_i v_i) \text{ and } Tw = \sum T(b_i u_i)$$
$$\implies \sum (a_i - b_i) v_i = 0$$

so we know that $a_i = b_i$.

Now, we can show that

$$\begin{aligned} Tw &= \sum T(a_i v_i) \\ &= \sum a_i u_i \\ &= \sum b_i u_i (a_i = b_i) \\ &= w \end{aligned}$$

so we have shown that $T = I$.

Exercise 10.1.5
Change the basis to the eigenvectors, or any basis with an upper triangular matrix.

Exercise 10.1.6
I think we want complex eigenvalues like i twice, so T^2 will have -2 as the trace.

Exercise 10.1.7
We have the eigenvector basis, and we know that the trace of T^2 is just $\sum |\lambda|^2$.

I think...I'm assuming our matrix is diagonal, although I think it's actually only upper triangular.

Alright, too lazy to do the rest and there's one more chapter left.

10.2 Determinant

Definition 10.2.1
Suppose $T \in \mathcal{L}(V)$.

- If $\boldsymbol{F} = \boldsymbol{C}$, then the **determinant** of T is the product of the eigenvalues of T, with each eigenvalue repeated according to its multiplicity.

- If $\boldsymbol{F} = \boldsymbol{R}$, then the **determinant** of T is the product of the eigenvalues of $T_{\boldsymbol{C}}$, with each eigenvalue repeated according to its multiplicity.

The determinant of T is denoted by $\det T$.

We later find out to calculate the determinant, for

$$A = \begin{pmatrix} A_{1,1} & \cdots A_{1,n} & \\ \vdots & & \vdots \\ A_{n,1} & \cdots A_{n,n} & \end{pmatrix}, \tag{10.1}$$

we calculate

$$\det A = \sum_{(m_1,\ldots,m_n)\in \text{perm } n} (\text{sign}(m_1, \ldots, m_n))A_{m_1,1} \cdots A_{m_n,n}. \tag{10.2}$$

Most of the proofs in this chapter are pretty bash-y, meaning we just have to explicitly calculate sums and products of matrix entries.

The ending stuff about geometries and polar coordinates and spherical coordinates felt pretty rushed, and was a weird addition to a mostly pure-math book. I think those topics suit better in a course with more physics or geometry applications.

Sorry...not going to do these exercises.

Was overall a fun read though :)

Appendix A

Useful Proof Techniques

A.1 Showing Something is Linear

You need to show additivity and scalar multiplication.

A.2 AFSOC

Reminder that AFSOC is still a great proof technique! Assume a contradiction and then look for something wrong.

A.3 Linear Independence

Remember the definition of linear independence,

Definition A.3.1

- A list $v_1, \ldots, v_m$ of vectors in V is called **linearly independent** if the only choice $a_1, \ldots, a_m \in \mathbf{F}$ that makes
$$\sum_{i=1}^{m} a_i v_i = 0$$
is $a_1 = \cdots = a_m = 0$.
- The empty list () is also declared to be linearly independent.

A.4 Decomposition

If we can decompose an operator (matrix) into an orthonormal basis, or into eigenvalues, it can often be helpful in the proofs, as we know a lot of these particular bases.

Appendix B

Computation Skills

Linear algebra often has a lot of computation skills. We list how to do certain common operations here.

B.1 Null Space

You are trying to find for $T \in \mathcal{L}(V), v \in V$, which

$$Tv = 0.$$

In practice, the best way to do this is to reduce T into row-echelon form, and then write each pivot in terms of other variables, which will be the "free variables."

Example 1 Find the null space for the following operator,

$$\begin{pmatrix} 1 & -2 & 0 & 3 & 0 & 0 & 2 \\ 0 & 0 & 3 & 7 & -8 & 0 & 9 \\ 0 & 0 & 0 & 0 & 0 & 1 & -4 \end{pmatrix} v = 0$$

We can see for $v = (x_1, x_2, x_3, x_4, x_5, x_6, x_7)$ that we can find the equations

$$\begin{aligned} x_1 &= 2x_2 - 3x_4 - 2x_7 \\ x_3 &= -\frac{7}{3}x_4 + \frac{8}{3}x_5 - 3x_7 \\ x_6 &= 4x_7 \end{aligned}$$

so our null space is

$$\operatorname{null} T = x_2 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -3 \\ 0 \\ -\frac{7}{3} \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} 0 \\ 0 \\ \frac{8}{3} \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_7 \begin{pmatrix} -2 \\ 0 \\ -3 \\ 0 \\ 0 \\ 0 \\ 4 \\ 1 \end{pmatrix} = \operatorname{span} \left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -3 \\ 0 \\ -\frac{7}{3} \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \frac{8}{3} \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -2 \\ 0 \\ -3 \\ 0 \\ 0 \\ 0 \\ 4 \\ 1 \end{pmatrix} \right\}$$

The idea here is to write our null space vectors in terms of the "free variables," which together determine the pivot variables.

B.2 Eigenvectors and Eigenvalues

B.2.1 Generalized Eigenvectors and Eigenvalues

Appendix C

Some Thoughts

I thought this book was a very readable book, especially for someone who has taken an intro linear algebra course. Therefore, difficulty probably is somewhere around a medium (maybe 5/10).

I think the results in this book generalize nicely above a more application-based linear algebra course, which will hopefully transition well to more pure math after this book.

Some improvements on the book

- There isn't much practical advice given on how to find eigenvalues and eigenvectors. There is an assumption that Gaussian elimination, finding null space and all the fun skills that you learn in an intro class are known by the reader, but it would be good to employ it in some cases, finding null space and all the fun skills that you learn in an intro class are known by the reader, but it would be good to employ it in some cases.
- A lot of the examples are pretty trivial, which doesn't really help the reader's problem solving skills.