Linear Algebra

Technical Report

Anming Gu
Boston University

July 12, 2023

Abstract

This is a short reference material on linear algebra following *Linear Algebra Done Right* by Axler.

Contents

1 Vector Spaces **2**

2 Finite-Dimensional Vector Spaces **2**
2.1 Span and Linear Independence . . . 2
2.2 Bases . . . 2
2.3 Dimension . . . 3

3 Linear Maps **3**
3.1 The Vector Space of Linear Maps . . . 3
3.2 Null Spaces and Ranges . . . 3
3.3 Matrices . . . 3
3.4 Invertability and Isomorphic Vector Spaces . . . 4
3.5 Products and Quotients of Vector Spaces . . . 4
3.6 Duality . . . 5

4 Eigenvalues, Eigenvectors, and Invariant Subspaces **6**
4.1 Invariant Subspaces . . . 6
4.2 Eigenvectors and Upper-Triangular Matrices . . . 6
4.3 Eigenspaces and Diagonal Matrices . . . 6

5 Inner Product Spaces **7**
5.1 Inner Products and Norms . . . 7
5.2 Orthonormal Bases . . . 7
5.3 Orthogonal Complements and Minimization Problems . . . 8

6 Operators on Inner Product Spaces **9**
6.1 Self-Adjoint and Normal Operators . . . 9
6.2 The Spectral Theorem . . . 10
6.3 Positive Operators and Isometries . . . 10
6.4 Polar Decomposition and Singular Value Decomposition . . . 11

7 Operators on Complex Vector Spaces **11**
7.1 Decomposition of an Operator . . . 11
7.2 Characteristic and Minimal Polynomials . . . 12
7.3 Jordan Form . . . 13

8 Operators on Real Vector Spaces **13**
8.1 Complexification . . . 13

8.2 Operators on Real Inner Product Spaces 14

9 Trace and Determinant **14**
9.1 Trace 14
9.2 Determinant 15

1 Vector Spaces

A *vector space* is a set V along with an addition on V and a scalar multiplication on V s.t. the following properties hold:

- commutativity: $u + v = v + u \ \forall u, v \in V$,
- associativity: $(u + v) + w = u + (v + w)$ and $(ab)v = a(bv) \ \forall u, v, w \in V, a, b \in \mathbb{F}$,
- additive identity: $\exists! 0 \in V$ s.t. $v + 0 = v \ \forall v \in V$,
- additive inverse: $\forall v \in \text{Var}, \exists! w \in V$ s.t. $v + w = 0$,
- multiplicative identity: $1v = v \ \forall v \in V$,
- distributive properties: $a(u + v) = au + av$ and $(a + b)v = av + bv \ \forall a, b \in \mathbb{F}, u, v \in V$.

If S is a set, then $\mathbb{F}^S$ denotes the set of functions from S to $\mathbb{F}$. Note that $\mathbb{F}^S$ is a vector space.

A subset $U \subset V$ is called a *subspace* of V if U is also a vector space (using the same addition and multiplication as on V).

A subset $U \subset V$ is a subspace of V iff $0 \in U$, U is closed under addition, and U is closed under scalar multiplications. Suppose $U_1, \dots, U_m \subset V$. The *sum* of $U_1, \dots, U_m$ is defined as $\sum U_i = \{\sum u_i \mid u_1 \in U_1, \dots, u_m \in U_m\}$.

Lemma 1.1. *$\sum U_i$ is the smallest subspace of V that contains $U_1, \dots, U_m$.*

Additionally, the sum is called a *direct sum* if each element of $\sum U_i$ can be written in only one way as a sum $\sum u_i$, where each $u_i \in U_i$. If $\sum U_i$ is a direct sum, we use the notation $\bigoplus U_i$.

Lemma 1.2. *Suppose $U_1, \dots U_m \subset V$ are subspaces. Then $\sum U_i$ is a direct sum iff the only way to write 0 as a sum $\sum u_i$, where each $u_i \in U_i$, is by letting $u_i = 0 \ \forall i$.*

Lemma 1.3. *Suppose $U, W \subset V$ are subspaces. Then $U + W$ is a direct sum iff $U \cap W = \{0\}$.*

2 Finite-Dimensional Vector Spaces

2.1 Span and Linear Independence

A *linear combination* $v_1, \dots, v_m$ of vectors in V is a vector of the form $\sum a_i v_i$, where $a_1, \dots, a_m \in \mathbb{F}$. The set of all linear combinations of a list of vectors $v_1, \dots, v_m$ is called the *span* of $v_1, \dots, v_m$, denoted $\text{span}\{v_1, \dots, v_m\}$.

Lemma 2.1. *The span of a list of vectors in V is the smallest subspace of V containing all the vectors in the list.*

A list $v_1, \dots, v_m$ of vectors in V is called *linearly independent* if the only choice of $a_1, \dots a_m \in \mathbb{F}$ that makes $\sum a_i = 0$ is $a_1 = \cdots = a_m = 0$. We also define the empty list to be linearly independent. A list of vectors in V is called *linearly dependent* if it is not linearly independent.

Lemma 2.2 (Linear dependence lemma)**.** *Suppose $v_1, \dots, v_m$ is linearly dependent in V. Then $\exists i \in [m]$ s.t. the following hold:*

(a) $v_i \in \text{span}\{v_1, \dots, v_{i-1}\}$,

(b) if the ith term is removed from $v_1, \dots, v_m$, the span of the remaining vectors equals $\text{span}\{v_1, \dots, v_m\}$.

2.2 Bases

A *basis* of V is a list of vectors in V that is linearly independent and spans V.

Lemma 2.3. *A list $v_1, \dots, v_n$ of vectors in V is a basis of V iff every $v \in V$ can be uniquely written in the form $v = \sum a_i v_i$, where $a_1, \dots, a_n \in \mathbb{F}$.*

Lemma 2.4. *Every spanning list in a vector space can be reduced to a basis of the vector space.*

Corollary 2.5. *Every finite-dimensional vector space has a basis.*

Lemma 2.6. *Every linearly independent list of vectors in a finite-dimensional vector space can be extended to a basis of the vector space.*

Lemma 2.7. *Suppose V is finite-dimensional and $U \subset V$ is a subspace. Then there exists a subspace $W \subset V$ s.t. $V = U \oplus W$.*

2.3 Dimension

We note that any two bases of a finite-dimensional vector space have the same length. The *dimension* $\dim V$ of a finite-dimensional vector space V is the length of any basis of the vector space.

Lemma 2.8. *If U_1, U_2 are subspaces of a finite-dimensional vector space, then*

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2).$$

3 Linear Maps

3.1 The Vector Space of Linear Maps

Suppose V, W are vector spaces. A *linear map* (or *transformation*) from V to W is a function $T : V \to W$ with the following properties:

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

We define the set of all linear maps from V to W by $\mathcal{L}(V, W)$.

Proposition 3.1. *Sjppose $v_1, \ldots, v_n$ is a basis of V and $w_1, \ldots, w_n \in W$. Then $\exists! T : V \to W$ linear s.t. $Tv_i = w_i\ \forall i \in [n]$.*

Theorem 3.2. *$\mathcal{L}(V, W)$ is a vector space.*

If $T \in \mathcal{L}(U, V), S \in \mathcal{L}(V, W)$, then the *product* $ST \in \mathcal{L}(U, W)$ is defined by $(ST)u = S(Tu)\ \forall u \in U$.

3.2 Null Spaces and Ranges

For $T \in \mathcal{L}(V, W)$, the *null space* of T is the subset of V consisting of those vectors that T maps to 0:

$$\ker T = \{v \in V \mid Tv = 0\}.$$

Proposition 3.3. *Suppose $T \in \mathcal{L}(V, W)$. Then $\ker T$ is a subspace of V.*

A function $T : V \to W$ is *injective* if $Tu = Tv$ implies $u = v$.

Theorem 3.4. *Let $T \in \mathcal{L}(V, W)$. Then T is injective iff $\ker T = \{0\}$.*

The *range* of T is the subset of W consisting of those vectors that are of the form Tv for some $v \in V$:

$$\operatorname{ran} T = \{Tv \mid v \in V\}.$$

If $T \in \mathcal{L}(V, W)$, then $\operatorname{ran} T$ is a subspace of W. A function $T : V \to W$ is *surjective* if its range equals W.

Theorem 3.5 (Fundamental theorem of linear maps)**.** *Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Then $\operatorname{ran} T$ is finite-dimesional and*

$$\dim V = \dim \ker T + \dim \operatorname{ran} T.$$

Corollary 3.6. *Suppose V, W are finite-dimensional vector spaces. If $\dim V > \dim W$, then no linear map $V \to W$ is injective. Conversely, if $\dim V < \dim W$, then no linear map $V \to W$ is surjective.*

3.3 Matrices

Let m, n be positive integers. An m-by-n *matrix* A is a rectangular array of elements $\mathbb{F}$ with m rows and n columns:

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

The notation $A_{j,k}$ denotes the entry in row j, column k of A. We denote the set of all m-by-n matrices with entries in $\mathbb{F}$ as $\mathbb{F}^{m,n}$.

3.4 Invertability and Isomorphic Vector Spaces

A linear map $T \in \mathcal{L}(V, W)$ is called *invertible* if $\exists S \in \mathcal{L}(W, V)$ s.t. $ST = I_V$ and $TS = I_W$. Such S is called an *inverse* of T. If T is invertible, then its inverse is denoted by T^{-1}.

Lemma 3.7. *An invertible linear map has a unique inverse.*

Lemma 3.8. *A linear map is invertible iff it is injetive and surjective.*

An *isomorphism* is an invertible linear map. Two vector spaces are *isomorphic* if there exists an isomorphism from one vector space onto the other one.

Lemma 3.9. *Two finite-dimensional vector spaces over $\mathbb{F}$ are isomorphic iff they have the same dimension.*

Lemma 3.10. *Suppose V, W are finite-dimensional. Then $\mathcal{L}(V, W)$ is finite-dimensional and*

$$\dim \mathcal{L}(V, W) = \dim V \cdot \dim W.$$

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

3.5 Products and Quotients of Vector Spaces

Suppose $V_1, \ldots, V_m$ are vector spaces over $\mathbb{F}$. The *product* $V_1 \times \cdots \times V_m$ is defined by

$$V_1 \times \cdots \times V_m = \{(v_1, \ldots, v_m) \mid v_1 \in V_1, \ldots, v_m \in V_m\}.$$

Example 3.11. $\mathbb{R}^2 \times \mathbb{R}^3$ is not equal to $\mathbb{R}^5$, but the two vector spaces are isomorphic.

Lemma 3.12. *Suppose $V_1, \ldots, V_m$ are finite-dimensional. Then $V_1 \times \cdots \times V_m$ is finite-dimensional and*

$$\dim(V_1 \times \cdots \times V_m) = \dim V_1 + \cdots + \dim V_m.$$

Suppose $v \in V$ and U is a subspace of V. Then $v + U$ is the subset of V defined by

$$v + U = \{v + u \mid u \in U\}.$$

An *affine subset* of V is a subset of V of the form $v + U$ for some $v \in V$ and some subspace U of V. The affine subset $v + U$ is said to be *parallel* to U. Then the *quotient space* V/U is the set of all affine subsets of V parallel to U, *i.e.*,

$$V/U = \{v + U \mid v \in V\}.$$

Define *addition* and *scalar multiplication* by the following:

$$(v + U) + (w + U) = (v + w) + U$$
$$\lambda(v + U) = (\lambda v) + U$$

for $v, w \in V, \lambda \in \mathbb{F}$. Then V/U is a vector space.

Suppose U is a subspace of V. The *quotient map* π is the linear map $\pi : V \to V/U$ defined by

$$\pi(v) = v + U$$

for $v \in V$.

Lemma 3.13. *Suppose V is finite-dimensional and U is a subspace of V. Then*

$$\dim V/U = \dim V - \dim U.$$

Suppose $T \in \mathcal{L}(V, W)$. Define $\tilde{T} : V/(\ker T) \to W$ by

$$\tilde{T}(v + \ker T) = Tv.$$

Lemma 3.14. *Suppose $T \in \mathcal{L}(V, W)$. Then*

- *(a) $\tilde{T}$ is a linear map from $V/(\ker T)$ to W,*
- *(b) $\tilde{T}$ is injective,*
- *(c) $\operatorname{ran} \tilde{T} = \operatorname{ran} T$,*
- *(d) $V/(\ker T)$ is isomorphic to $\operatorname{ran} T$.*

3.6 Duality

A *linear functional* on V is a linear map from V to $\mathbb{F}$. In other words, a linear functional is an element of $\mathcal{L}(V,\mathbb{F})$. The *dual space* of V, denoted V' is the vector space of all linear functionals on V. In other words, $V' = \mathcal{L}(V,\mathbb{F})$.

Lemma 3.15. *Suppose V is finite-dimensional. Then V' is also finite-dimensional and* $\dim V' = \dim V$.

If $v_1, \ldots, v_n$ is a basis of V, then the *dual basis* of $v_1, \ldots, v_n$ is the list $\varphi_1, \ldots, \varphi_n$ of elements of V', where each φ_j is the linear functional on V such that

$$\varphi_j(v_k) = \begin{cases} 1 & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases}$$

Lemma 3.16. *Suppose V is finite-dimensional. Then the dual basis of a basis of V is a basis of V'.*

If $T \in \mathcal{L}(V,W)$, then the *dual map* of T is the linear map $T' \in \mathcal{L}(W',V')$ defined by $T'(\varphi) = \varphi \circ T$ for $\varphi \in W'$.

Lemma 3.17. *Dual maps have the following properties:*

(a) $(S+T)' = S' + T'$ $\forall S,T \in \mathcal{L}(V,W)$.

(b) $(\lambda T)' = \lambda T'$ $\forall \lambda \in \mathbb{F}, T \in \mathcal{L}(V,W)$.

(c) $(ST)' = T'S'$ $\forall T \in \mathcal{L}(U,V), S \in \mathcal{L}(V,W)$.

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

$$U^0 = \{\varphi \in V' \mid \varphi(u) = 0 \ \forall u \in U\}.$$

Lemma 3.18. *Suppose $U \subset V$. Then U^0 is a subspace of V'.*

Lemma 3.19. *Suppose V is finite-dimensional and U is a subspace of V. Then*

$$\dim U + \dim U^0 = \dim V.$$

Lemma 3.20. *Suppose V,W are finite-dimensional. Then*

(a) $\ker T' = (\operatorname{ran} T)^0$,

(b) $\dim \ker T' = \dim \ker T + \dim W - \dim V$.

Lemma 3.21. *Suppose V,W are finite-dimensional and $T \in \mathcal{L}(V,W)$. Then T is surjective iff T' is injective.*

Lemma 3.22. *Suppose V,W are finite-dimensional and $T \in \mathcal{L}(V,W)$. Then*

1. $\dim \operatorname{ran} T' = \dim \operatorname{ran} T$,
2. $\operatorname{ran} T' = (\ker T)^0$.

Lemma 3.23. *Suppose V,W are finite-dimensional and $T \in \mathcal{L}(V,W)$. Then T is injective iff T' is surjective.*

The *transpose* of a matrix A, denoted $A^\top$, is the matrix obtained from A by interchanging the rows and columns. More specifically, $A^\top_{k,j} = A_{j,k}$.

Lemma 3.24. *If A is an $m \times n$ matrix and C is an $n \times p$ matrix, then*

$$(AC)^\top = C^\top A^\top.$$

Lemma 3.25. *Suppose $T \in \mathbf{L}(V,W)$. Then $\mathcal{M}(T') = (\mathcal{M}(T))^\top$.*

Suppose A is an $m \times n$ matrix with entries in $\mathbb{F}$. The *row rank* of A is the dimension of the span of the rows of A in $\mathbb{F}^{1,n}$ and the *column rank* of A is the dimension of the span of the columns of A in $\mathbb{F}^{m,1}$. The *rank* of a matrix $A \in \mathbb{F}^{m,n}$ is the column rank of A.

4 Eigenvalues, Eigenvectors, and Invariant Subspaces

4.1 Invariant Subspaces

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

Suppose $T \in \mathcal{L}(V)$. A number $\lambda \in \mathbb{F}$ is called an *eigenvalue* of T if $\exists v \in V$ s.t. $v \neq 0$ and $Tv = \lambda v$. Such v is called an *eigenvector* of T corresponding to λ.

Lemma 4.1. *Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T and $v_1, \dots, v_m$ are corresponding eigenvectors. Then $v_1, \dots, v_m$ is linearly independent.*

Corollary 4.2. *Suppose V is finite-dimensional. Then each operator on V has at most $\dim V$ distinct eigenvalues.*

Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V invariant under T. The *restriction operator* $T_{|U} \in \mathcal{L}(U)$ is defined by $T_{|U}(u) = Tu$ for $u \in U$. The *quotient operator* $T/U \in \mathcal{L}(V/U)$ is defined by $(T/U)(v + U) = Tv + U$ for $v \in V$.

4.2 Eigenvectors and Upper-Triangular Matrices

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

$$T^m = \underbrace{T \cdots T}_{m \text{ times}}.$$

T^0 is defined to be I_V. If T is invertible with inverse T^{-1}, then T^{-m} is defined by $T^{-m} = (T^{-1})^m$. Suppose $p \in \mathcal{P}(\mathbb{F})$ is a polynomial given by $p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$ for $z \in \mathbb{F}$. Then $p(T)$ is the operator defined by $p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m$.

Theorem 4.3. *Every operator on a finite-dimensional, nonzero, complex vector space has an eigenvalue.*

The *diagonal* of a square matrix consists of the entries along the line from the upper left corner to the bottom right corner. A matrix is called *upper triangular* if all the entries below the diagonal equal 0.

Lemma 4.4. *Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T has an upper triangular matrix w.r.t some basis of V.*

Lemma 4.5. *Suppose $T \in \mathcal{L}(V)$ has an upper triangular matrix wr.t. some basis of V. Then T is invertible iff all the entries on the diagonal of that upper triangular matrix are nonzero.*

Lemma 4.6. *Suppose $T \in \mathcal{L}(V)$ has an upper triangular matrix w.r.t. . some basis of V. Then the eigenvalues of T are precisely the entries on the diagonal of that upper triangular matrix.*

4.3 Eigenspaces and Diagonal Matrices

A *diagonal matrix* is a square matrix that is 0 everywhere except possibly along the diagonal. Suppose $T \in \mathcal{L}(V), \lambda \in \mathbb{F}$. The *eigenspace* of T corresponding to λ, denoted $E(\lambda, T)$, is defined by $E(\lambda, T) = \ker(T - \lambda I)$.

Lemma 4.7. *Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Suppose also that $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T. Then*

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

is a direct sum. Furthermore,

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

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

Theorem 4.8. *Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ 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 exists 1*-dimensional subspaces* $U_1, \dots, U_n$ *of* V*, each invariant under* T*, s.t.* $V = \bigoplus U_i$*,*

(d) $V = \bigoplus E(\lambda_i, T)$*,*

(e) $\dim V = \sum \dim E(\lambda_i, T)$.

5 Inner Product Spaces

5.1 Inner Products and Norms

For $x, y \in \mathbb{R}^n$, the *dot product* of x and y, denoted $x \cdot y$, is defined by

$$x \cdot y = x_1 y_y + \cdots + x_n y_n,$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$.

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

- *positivity*: $\langle v, v \rangle \geq 0 \ \forall v \in V$,
- *definiteness*: $\langle v, v \rangle = 0 \iff v = 0$,
- *additivity in first slot*: $\langle u + v, \rangle = \langle u, w \rangle + \langle v, w \rangle \ \forall u, v, w \in V$,
- *homogeneity in first slot*: $\langle \lambda u, v \rangle = \lambda \langle u, v \rangle \ \forall \lambda \in \mathbb{F}, u, v \in V$,
- *conjugate symmetry*: $\langle u, v \rangle = \overline{\langle v, u \rangle} \ \forall u, v \in V$.

An *inner product space* is a vector space V equipped with an inner product. The following are basic properties of an inner product:

(a) $\forall u \in V$ fixed, the function $v \mapsto \langle v, u \rangle$ is a linear map from V to $\mathbb{F}$.

(b) $\langle 0, u \rangle = 0 \ \forall u \in V$.

(c) $\langle u, 0 \rangle = 0 \ \forall u \in V$.

(d) $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle \ \forall u, v, w \in V$.

(e) $\langle u, \lambda v \rangle = \overline{\lambda} \langle u, v \rangle \ \forall \lambda \in \mathbb{F}, u, v \in V$.

For $v \in V$, the *norm* of v, denoted $\|v\|$ is defined by $\|v\| = \sqrt{\langle v, v \rangle}$. The norm has the following basic properties:

(a) $\|v\| = 0 \iff v = 0$.

(b) $\|\lambda v\| = |\lambda| \|v\| \ \forall \lambda \in \mathbb{F}$.

Two vectors $u, v \in V$ are called *orthogonal* if $\langle u, v \rangle = 0$, denoted $u \perp v$.

Theorem 5.1 (Pythagorean theorem)**.** *Suppose $u \perp v$ in V. Then*

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

Lemma 5.2. *Suppose $u, v \in V$, with $v \neq 0$. Set $c = \frac{\langle u, v \rangle}{\|v\|^2}$ and $w = u - \frac{\langle u, v \rangle}{\|v\|^2}$. Then $\langle w, v \rangle = 0$ and $u = cv + w$.*

The above lemma is used in the proof of Cauchy-Schwartz inequality.

Theorem 5.3 (Cauchy-Schwartz inequality)**.** *Suppose $u, v \in V$. Then $|\langle u, v \rangle| \leq \|u\| \|v\|$. Furthermore, the equality holds iff one of u, v is a scalar multiple of the other.*

Theorem 5.4 (Triangle inequality)**.** *Suppose $u, v \in V$. Then $\|u + v\| \leq \|u\| + \|v\|$. Furthermore, the equality holds iff one of u, v is a nonnegative multiple of the other.*

Theorem 5.5 (Parallelogram equality)**.** *Suppose $u, v \in V$, then $\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$.*

5.2 Orthonormal Bases

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

Lemma 5.6. *If $e_1, \ldots, e_m$ is an orthonormal list of vectors in V, then*

$$\|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

$\forall a_1, \ldots, a_m \in \mathbb{F}$.

An *orthonormal basis* of V is an orthonormal list of vectors in V that is also a basis of V.

Lemma 5.7. *Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $v \in V$. Then*

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n$$

and

$$\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2.$$

The following algorithm can be used to turn a linearly independent list of vectors into a orthonormal list with the same span as the original list:

Theorem 5.8 (Gram-Schmidt procedure)**.** *Suppose $v_1, \dots, v_m$ is a linearly independent list of vectors in V. Let $e_1 = v_1/\|v_1\|$. For $j = 2, \dots, m$, define e_j inductively by*

$$e_j = \frac{v_j - \langle v_j, e_1 \rangle e_1 - \cdots - \langle v_j, e_{j-1} \rangle e_{j-1}}{\|v_j - \langle v_j, e_1 \rangle e_1 - \cdots - \langle v_j, e_{j-1} \rangle e_{j-1}\|}.$$

Then $e_1, \dots, e_m$ is an orthonormal list of vectors in V s.t. $\operatorname{span}\{v_1, \dots, v_j\} = \operatorname{span}\{e_1, \dots, e_j\}$ *for $j = 1, \dots, m$.*

Theorem 5.9. *Every finite-dimensional inner product space has an orthonormal basis.*

Lemma 5.10. *Suppose V is finite-dimensional. Then every orthonormal list of vectors in V can be extended to an orthonormal basis of V.*

Theorem 5.11 (Schur's theorem)**.** *Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix w.r.t. some orthonormal basis of V.*

A *linear functional* on V is a linear map from V to $\mathbb{F}$. In other words, a linear functional is an element of $\mathcal{L}(V, \mathbb{F})$.

Theorem 5.12 (Riesz representation theorem)**.** *Suppose V is finite-dimensional and φ is a linear functional on V. Then $\exists! u \in V$ s.t. $\varphi(v) = \langle v, u \rangle \ \forall v \in V$.*

5.3 Orthogonal Complements and Minimization Problems

If $U \subset V$, 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:

$$U^\perp = \{v \in V \mid \langle v, u \rangle = 0 \ \forall u \in U\}.$$

The following are properties of the orthogonal complement:

(a) If $U \subset V$, then $U^\perp$ is a subspace of V.

(b) $\{0\}^\perp = V$.

(c) $V^\perp = \{0\}$.

(d) If $U \subset V$, then $U \cap U^\perp \subset \{0\}$.

(e) If $U, W \subset V$ and $U \subset W$, then $W^\perp \subset U^\perp$.

Theorem 5.13. *Suppose U is a finite-dimensional subspace of V. Then $V = U \oplus U^\perp$.*

Corollary 5.14. *Suppose V is finite-dimensional and U is a subspace of V, then* $\dim U^\perp = \dim V - \dim U$.

Theorem 5.15. *Suppose U is a finite-dimensional subspace of V. Then $U = (U^\perp)^\perp$.*

Suppose U is a finite-dimensional subspace of V. The *orthogonal projection* of V onto U is the operator $P_U \in \mathcal{L}(V)$ defined as follows: for $v \in V$, write $v = u + w$, where $u \in U, w \in U^\perp$. Then $P_U v = u$. The projection operator has the following properties:

(a) $P_U \in \mathcal{L}(V)$,

(b) $P_U u = u \ \forall u \in U$,

(c) $P_U w = 0 \ \forall w \in U^\perp$,

(d) $\operatorname{ran} P_U = U$,

(e) $\ker P_U = U^\perp$,

(f) $v - P_U v \in U^\perp$,

(g) $P_U^2 = P_U$,

(h) $\|P_U v\| \leq \|v\|$,

(i) for every orthonormal basis $e_1, \ldots, e_m$ of U,

$$P_U v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m.$$

Lemma 5.16. *Suppose U is a finite-dimensional subspace of V, $v \in V$, and $u \in U$. Then $\|v - P_U v\| \leq \|v - u\|$. Furthermore, the inequality is an equality iff $u = P_U v$.*

6 Operators on Inner Product Spaces

6.1 Self-Adjoint and Normal Operators

Suppose $T \in \mathcal{L}(V, W)$. The *adjoint* of T is the function $T^* : W \to V$ s.t. $\langle Tv, w\rangle = \langle v, T^* w\rangle\ \forall v \in V, w \in W$.

Lemma 6.1. *If $T \in \mathcal{L}(V, W)$, then $T^* \in \mathcal{L}(W, V)$.*

Adjoint operators have the following properties:

(a) $(S + T)^* = S^* + T^*\ \forall S, T \in \mathcal{L}(V, W)$,

(b) $(\lambda T)^* = \overline{\lambda} T^*\ \forall \lambda \in \mathbb{F}, T \in \mathcal{L}(V, W)$,

(c) $(T^*)^* = T\ \forall T \in \mathcal{L}(V, W)$,

(d) $I^* = I$,

(e) $(ST)^* = T^* S^*\ \forall T \in \mathcal{L}(V, W), S \in \mathcal{L}(W, U)$.

Theorem 6.2. *Suppose $T \in \mathcal{L}(V, W)$. Then*

(a) $\ker T^* = (\operatorname{ran} T)^\perp$,

(b) $\operatorname{ran} T^* = (\ker T)^\perp$,

(c) $\ker T = (\operatorname{ran} T^*)^\perp$,

(d) $\operatorname{ran} T = (\ker T^*)^\perp$.

An operator $T \in \mathcal{L}(V)$ is called *self-adjoint* if $T = T^*$. In otherwords, $T \in \mathcal{L}(V)$ is self-adjoint iff $\langle Tv, w\rangle = \langle v, Tw\rangle\ \forall v, w \in V$.

Lemma 6.3. *Every eigenvalue of a self-adjoint operator is real.*

Note that the following theorem is false for real inner product spaces.

Theorem 6.4. *Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Suppose $\langle Tv, v\rangle = 0 \forall v \in V$. Then $T = 0$.*

Lemma 6.5. *Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Then T is self-adjoint iff $\langle Tv, v\rangle \in \mathbb{R}\ \forall v \in V$.*

Lemma 6.6. *Suppose T is self-adjoint on V s.t. $\langle Tv, v\rangle = 0\ \forall v \in V$. Then $T = 0$.*

An operator on an inner product space is called *normal* if it commutes with its adjoint. In other words, $T \in \mathcal{L}(V)$ is normal if $TT^* = T^*T$. Obviously every self-adjoint operator is normal.

Lemma 6.7. *An operator $T \in \mathcal{L}(V)$ is normal iff $\|Tv\| = \|T^* v\|\ \forall v \in V$.*

Lemma 6.8. *Suppose $T \in \mathcal{L}(V)$ is normal and $v \in V$ is an eigenvector of T with eigenvalue λ. .Then v is also an eigenvector of T^* with eigenvalue $\overline{\lambda}$.*

Lemma 6.9. *Suppose $T \in \mathcal{L}(V)$ is normal. Then eigenvectors of T corresponding to distinct eigenvalues are orthogonal.*

6.2 The Spectral Theorem

The spectral theorem is probably the most useful tool in the study of operators on inner product spaces.

Theorem 6.10 (Complex spectral theorem)**.** *Let $T \in \mathcal{L}(V)$. 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 w.r.t. some orthonormal basis of V.

To prove the real spectral theorem, we need some preliminary results:

Lemma 6.11. *Suppose $T \in \mathcal{L}(V)$ is self-adjoint and $b, c \in \mathbb{R}$ s.t. $b^2 < 4c$. Then $T^2 + bT + cI$ is invertible.*

Lemma 6.12. *Suppose $V \neq \{0\}$ and $T \in \mathcal{L}(V)$ is self-adjoint. Then T has an eigenvalue.*

Lemma 6.13. *Suppose $T \in \mathcal{L}(V)$ is self-adjoint and U is a subspace of V that is invariant under T. Then*

(a) $U^\perp$ is invariant under T,

(b) $T_{|U} \in \mathcal{L}(U)$ is self-adjoint,

(c) $T_{|U^\perp} \in \mathcal{L}(U^\perp)$ is self-adjoint.

Theorem 6.14 (Real spectral theorem)**.** *Let $T \in \mathcal{L}(V)$. 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 w.r.t. some orthonormal basis of V.

6.3 Positive Operators and Isometries

An operator $T \in \mathcal{L}(V)$ is called *positive* if it is self-adjoint and $\langle Tv, v\rangle \geq 0\ \forall v \in V$. An operator R is called a *square root* of an operator T if $R^2 = T$.

Theorem 6.15. *Let $T \in \mathcal{L}(V)$. Then the following are equivalent:*

(a) T is positive,

(b) T is self-adjoint and all the eigenvalues of T are nonnegative,

(c) T has a positive square root,

(d) T has a self-adjoint square root,

(e) $\exists R \in \mathcal{L}(V)$ s.t. $T = R^*R$.

Lemma 6.16. *Every positive operator on V has a unique positive square root.*

We use $\sqrt{\cdot}$ to denote the square root of a positive operator. An operator $S \in \mathcal{L}(V)$ is called an *isometry* if $\|Sv\| = \|v\|\ \forall v \in V$. In other words, an operator is an isometry if it preserves norms.

Theorem 6.17. *Suppose $S \in \mathcal{L}(V)$. Then the following are equivalent:*

(a) S is an isometry,

(b) $\langle Su, Sv\rangle = \langle u, v\rangle\ \forall u, v \in V$,

(c) $Se_1, \ldots, Se_n$ is orthonormal for every orthonormal list of vectors $e_1, \ldots, e_n$ in V,

(d) there exists an orthonormal basis $e_1, \ldots, e_n$ of V s.t. $Se_1, \ldots, Se_n$ is orthonormal,

(e) $S^*S = I$,

(f) $SS^* = I$,

(g) S^* is an isometry,

(h) S is invertible and $S^{-1} = S^*$.

Theorem 6.18. *Suppose V is a complex inner product space and $S \in \mathcal{L}(V)$. Then the following are equivalent:*

(a) S is an isometry.

(b) There is an orthonormal basis of V consisting of eigenvectors of S whose corresponding eigenvalues all have absolute value 1.

6.4 Polar Decomposition and Singular Value Decomposition

Theorem 6.19 (Polar decomposition theorem)**.** *Suppose $T \in \mathcal{L}(V)$. Then there exists an isometry $S \in \mathcal{L}(V)$ s.t. $T = S\sqrt{T^*T}$.*

The polar decomposition theorem states that each operator on V is the product of an isometry and a positive operator.

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 6.20 (Singular value decomposition)**.** *Suppose $T \in \mathcal{L}(V)$ has singular values $s_1, \ldots, s_n$. Then there exists orthonormal bases $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ of V s.t.*

$$Tv = s_1\langle ev, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n$$

$\forall v \in V$.

Lemma 6.21. *Suppose $T \in \mathcal{L}(V)$. Then the singular values of T are the nonnegative square roots of the eigenvalues of T^*T, with each eigenvalue λ repeated $\dim E(\lambda, T^*T)$ times.*

7 Operators on Complex Vector Spaces

Lemma 7.1. *Suppose $T \in \mathcal{L}(V)$. Then*

$$\{0\} = \ker T^0 \subset \ker T^1 \subset \cdots \subset \ker T^k \subset \ker T^{k+1} \subset \cdots.$$

Furthermore, let $n = \dim V$. Then

$$\ker T^n = \ker T^{n+1} = \ker T^{n+2} = \cdots.$$

Theorem 7.2. *Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then*

$$V = \ker T^n \oplus \operatorname{ran} T^n.$$

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 $(T - \lambda I)^j v = 0$ for some $j \in \mathbb{N}$. In fact, we will see that every generalized eigenvector satisfies this equation with $j = \dim V$. 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. Note that $E(\lambda, T) \subset G(\lambda, T)$.

Lemma 7.3. *Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. Then $G(\lambda, T) = \ker(T - \lambda I)^{\dim V}$.*

Theorem 7.4. *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 generalized eigenvectors. Then $v_1, \ldots, v_m$ is linearly independent.*

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

Lemma 7.5. *Suppose $N \in \mathcal{L}(V)$ is nilpotent. Then $N^{\dim V} = 0$.*

7.1 Decomposition of an Operator

In this section, we will see that every operator on a finite-dimensional complex vector space has enough generalized eigenvectors to provide a decomposition.

Lemma 7.6. *Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbb{F})$. Then $\ker p(T)$ and $\operatorname{ran} p(T)$ are invariant under T.*

Theorem 7.7. *Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. Then*

(a) $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$,

(b) each $G(\lambda_j, T)$ is invariant under T,

(c) each $(T - \lambda_j I_{|G(\lambda_j, T)}$ is nilpotent.

Theorem 7.8. *Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Then there is a basis of V consisting of generalized eigenvectors of T.*

Suppose $T \in \mathcal{L}(V)$. The *multiplicity* of an eigenvalue λ of T is defined to be the dimension of the corresponding generalized eigenspace $G(\lambda, T)$. In other words, the multiplicity of an eigenvalue λ of T equals $\dim \ker(T - \lambda I)^{\dim V}$.

Lemma 7.9. *Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Then the sum of the multiplicites of all the eigenvalues of T equals $\dim V$.*

A *block diagonal matrix* is a square matrix of the form

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

where $A_1, \ldots, A_m$ are square matrices lying along the diagonal and all the other entries of the matrix equal 0.

Theorem 7.10. *Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T, with multiplicities $d_1, \ldots, d_m$. Then there is a basis of V w.r.t. which T has a block diagonal matrix of the form*

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

where each A_j is a $d_j \times d_j$ upper-triangular matrix of the form

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

Lemma 7.11. *Suppose $N \in \mathcal{L}(V)$ is nilpotent. Then $I + N$ has a square root.*

Lemma 7.12. *Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ is invertible. Then T has a square root.*

7.2 Characteristic and Minimal Polynomials

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

$$(z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}$$

is called the *characteristic polynomial* of T.

Lemma 7.13. *Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Then*

(a) the characteristic polynomial of T has degree $\dim V$,

(b) the zeros of the characteristic polynomial of T are the eigenvalues of T.

Theorem 7.14 (Cayley-Hamilton theorem)**.** *Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Le q denote the characteristic polynomial of T. Then $q(T) = 0$.*

A *monic polynomial* is a polynomial whose highest-degree coefficient equals 1.

Lemma 7.15. *Suppose $T \in \mathcal{L}(V)$. Then there is a unique monic polynomial p of smallest degree s.t. $p(T) = 0$.*

The above lemma justifies the following definition. Suppose $T \in \mathcal{L}(V)$. Then the *minimal polynomial* of T is the unique monic polynomial p of smallest degree s.t. $p(T) = 0$.

Lemma 7.16. *Suppose $T \in \mathcal{L}(V)$. Then the characteristic polynomial of T is a polynomial multiple of the minimal polynomial of T.*

Lemma 7.17. *Let $T \in \mathcal{L}(V)$. Then the zeros of the minimal polynomial are precisely the eigenvalues of T.*

7.3 Jordan Form

Theorem 7.18. *Suppose $N \in \mathcal{L}(V)$ is nilpotent. Then there exists vectors $v_1, \dots, v_n \in V$ and nonnegative integers $m_1, \dots, m_n$ s.t.*

1. *$N^{m_1}v_1, \dots, Nv_1, v_1, \dots, N^{m_n}v_n, \dots, Nv_n, v_n$ is a basis of V,*
2. *$N^{m_1+1}v_1 = \cdots = N^{m_n+1}v_n = 0$.*

Suppose $T \in \mathcal{L}(V)$. A basis of V is called a *Jordan basis* for T if w.r.t. 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

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

Theorem 7.19. *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 for T.*

8 Operators on Real Vector Spaces

8.1 Complexification

Suppose V is a real vector space. The *complexification* of V, denoted $V_{\mathbb{C}}$ equals $V \times V$. An element of $V_{\mathbb{C}}$ is an ordered pair (u, v), where $u, v \in V$, but we will write this as $u + iv$.

Lemma 8.1. *Suppose V is a real vector space.*

(a) If $v_1, \dots, v_n$ is a basis of V, then $v_1, \dots, v_n$ is a basis of $V_{\mathbb{C}}$.

(b) $\dim V_{\mathbb{C}} = \dim V$.

Suppose V is a real vector space and $T \in \mathcal{L}(V)$. The *complexification* of T, denoted $T_{\mathbb{C}}$, is the operator $T_{\mathbb{C}} \in \mathcal{L}(V_{\mathbb{C}})$ defined by

$$T_{\mathbb{C}}(u + iv) = Tu + iTv$$

for $u, v \in V$.

Suppose V is a real vector space with basis $v_1, \dots, v_n$ and $T \in \mathcal{L}(V)$. Then $\mathcal{M}(T) = \mathcal{M}(T_{\mathbb{C}})$, where both matrices are w.r.t. the basis $v_1, \dots, v_n$.

Theorem 8.2. *Every operator on a nonzero finite-dimensional vector space has an invariant subspace of dimension* 1 *or* 2*.*

Lemma 8.3. *Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Then the minimal polynomial of $T_{\mathbb{C}}$ equals the minimal polynomial of T.*

Lemma 8.4. *Suppose V is a real vector space, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{R}$. Then λ is an eigenvalue of $T_{\mathbb{C}}$ iff λ is an eigenvalue of T.*

Lemma 8.5. *Suppose V is a real vector space, $T \in \mathcal{L}(V)$, $\lambda \in \mathbb{C}$, $j \in \mathbb{Z}_{\geq 0}$, and $u, v \in V$. Then $(T_{\mathbb{C}} - \lambda I)^j(u + iv) = 0$ iff $(T_{\mathbb{C}}) - \overline{\lambda} I)^j(u - iv) = 0$.*

Lemma 8.6. *Suppose V is a real vector space, $T \in \mathcal{L}(V)$, $\lambda \in \mathbb{C}$, $j \in \mathbb{Z}_{\geq 0}$, and $u, v \in V$. Then $(T_{\mathbb{C}} - \lambda I)^j(u + iv) = 0$ iff $(T_{\mathbb{C}} - \overline{\lambda} I)^j(u - iv) = 0$.*

Corollary 8.7. *Suppose V is a real vector space, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{C}$. Then λ is an eigenvalue of $T_{\mathbb{C}}$ iff $\overline{\lambda}$ is an eigenvalue of $T_{\mathbb{C}}$.*

Lemma 8.8. *Suppose V is a real vector space, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{C}$ is an eigenvalue of $T_{\mathbb{C}}$. Then the multiplicity of λ as an eigenvalue of $T_{\mathbb{C}}$ equals the multiplicity of $\overline{\lambda}$ as an eigenvalue of $T_{\mathbb{C}}$.*

Corollary 8.9. *Every operator on an odd-dimensional real vector space has an eigenvalue.*

Lemma 8.10. *Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Then the coefficients of the characteristic polynomial of $T_{\mathbb{C}}$ are all real.*

Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Then the *characteristic polynomial* of T is defined to be the characteristic polynomial of $T_{\mathbb{C}}$.

Theorem 8.11. *Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Then*

(a) the coefficients of the characterstic polynomial of T are all real,

(b) the characteristic polynomial of T has degree $\dim V$,

(c) the eigenvalues of T are precisely the real zeros of the characteristic polynomial of T.

Theorem 8.12 (Cayley-Hamilton theorem)**.** *Suppose $T \in \mathcal{L}(V)$. Let q denote the characteristic polynomial of T. Then $q(T) = 0$.*

Theorem 8.13. *Suppose $T \in \mathcal{L}(V)$. Then*

(a) the degree of the minimal polynomial of T is at most $\dim V$,

(b) the characteristic polynomial of T is a polynomial multiple of the minimal polynomial of T.

8.2 Operators on Real Inner Product Spaces

Theorem 8.14. *Suppose V is a real inner product space, $T \in \mathcal{L}(V)$ is normal, and U is a subspace of V that is invariant under T. Then*

(a) $U^\perp$ is invariant under T,

(b) U is invariant under T^*,

(c) $(T_{|U})^* = (T^*)_{|U}$,

(d) $T_{|U} \in \mathcal{L}(U)$ and $T_{|U^\perp} \in \mathcal{L}(U^\perp)$ are normal operators.

Theorem 8.15. *Suppose V is a real inner product space and $T \in \mathcal{L}(V)$. Then the following are equivalent:*

(a) T is normal.

(b) There is an orthonormal basis of V w.r.t. which T has a block diagonal matrix s.t. each block is a 1×1 matrix or a 2×2 matrix of the form

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

with $b > 0$.

Theorem 8.16. *Suppose V is a real inner product space and $S \in \mathcal{L}(V)$. Then the following are equivalent:*

(a) S is an isometry.

(b) There is an orthonormal basis of V w.r.t. which S has a block diagonal matrix s.t. each block on the diagonal is a 1×1 matrix containing 1 or -1 or is a 2×2 matrix of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

with $\theta \in (0, \pi)$.

9 Trace and Determinant

9.1 Trace

Recall that a square matrix A is called *invertible* if there is a square matrix B of the same size s.t. $AB = BA = I$, which we call B the *inverse* of A and denote it by A^{-1}.

Theorem 9.1 (Change of basis formula)**.** *Suppose $T \in \mathcal{L}(V)$. Let $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ be bases of V. Let $A = \mathcal{M}(I, (u_1, \ldots, u_n), (v_1, \ldots, v_n))$. Then*

$$\mathcal{M}(T, (u_1, \ldots, u_n)) = A^{-1}\mathcal{M}(T, (v_1, \ldots, v_n))A.$$

Suppose $T \in \mathcal{L}(V)$. If $\mathbb{F} = \mathbb{C}$, then the *trace* of T is the sum of the elements of T, with each eigenvalue repeated according to its multiplicity. If $\mathbb{F} = \mathbb{R}$, then the trace of T is the same result as above but with the eigenvalues of $T_{\mathbb{C}}$. The trace of T is denoted by $\operatorname{trace} T$.

Lemma 9.2. *Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then $\operatorname{trace} T$ equals the negative of the coefficient of z^{n-1} in the characteristic polynomial of T.*

The *trace* of a square matrix A is defined to be the sum of the diagonal entries of A.

Lemma 9.3. *If A, B are square matrices of the same size, then $\operatorname{trace} AB = \operatorname{trace} BA$.*

Theorem 9.4. *Let $T \in \mathcal{L}(V)$. Suppose $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Then*

$$\operatorname{trace} \mathcal{M}(T, (u_1, \dots, u_n)) = \operatorname{trace} \mathcal{M}(T, (v_1, \dots, v_n)).$$

Corollary 9.5. *Suppose $T \in \mathcal{L}(V)$. Then $\operatorname{trace} T = \operatorname{trace} \mathcal{M}(T)$.*

Lemma 9.6. *Suppose $S, T \in \mathcal{L}(V)$. Then $\operatorname{trace}(S + T) = \operatorname{trace} S + \operatorname{trace} T$.*

The proof of the following theorem uses traces.

Theorem 9.7. *$\nexists S, T \in \mathcal{L}(V)$ s.t. $ST - TS = I$.*

9.2 Determinant

Suppose $T \in \mathcal{L}(V)$. If $\mathbb{F} = \mathbb{C}$, then the *determinant* of T is the product of the eigenvalues of T, with each eigenvalue repeated according to its multiplicity. If $\mathbb{F} = \mathbb{R}$, then the determinant of T is the same as above but using $T_{\mathbb{C}}$. The determinant of T is denoted by $\det T$.

Lemma 9.8. *Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then $\det T$ equals $(-1)^n$ times the constant term of the characteristic polynomial of T.*

Theorem 9.9. *An operator on V is invertible iff its determinant is nonzero.*

Theorem 9.10. *Suppose $T \in \mathcal{L}(V)$. Then the characteristic polynomial of T equals $\det(zI - T)$.*

A *permutation* of $[n]$ is a list that contains each of the numbers exactly once. Denote the set of all permutations of $[n]$ as $\operatorname{perm} n$. The *sign* of a permutation is defined to be 1 if the number of pairs of integers (j, k) with $1 \le j < k \le n$ s.t. j appears after k in the list is even and -1 if the number of such pairs is odd. Note that interchanging two entries in a permutation multiplies the sign of the permutation by -1.

Suppose A is an $n \times n$ matrix. The *determinant* of A, denoted $\det A$, is defined by

$$\det A = \sum_{\sigma \in \operatorname{perm} n} \operatorname{sgn}(\sigma) \cdot A_{m_1,1} \cdots A_{m_n,n}.$$

Lemma 9.11. *Suppose A is a square matrix and B is the matrix obtained from A by interchanging two columns. Then $\det A = -\det B$.*

Corollary 9.12. *If A is a square matrix that has two equal columns, then $\det A = 0$.*

Lemma 9.13. *Suppose A, B are square matrices of the same size. Then $\det AB = \det BA = \det A \cdot \det B$.*

Theorem 9.14. *Let $T \in \mathcal{L}(V)$. Suppose $u_1, \dots u_n$ and $v_1, \dots, v_n$ are bases of V. Then*

$$\det \mathcal{M}(T, (u_1, \dots, u_n)) = \det \mathcal{M}(T, (v_1, \dots, v_n)).$$

Theorem 9.15. *Suppose $T \in \mathcal{L}(V)$. Then $\det T = \det \mathcal{M}(T)$.*

Theorem 9.16. *Suppose $S, T \in \mathcal{L}(V)$. Then*

$$\det ST = \det TS = \det S \cdot \det T.$$

Theorem 9.17. *Suppose V is an inner product space and $S \in \mathcal{L}(V)$ is an isometry. Then $|\det S| = 1$.*

Theorem 9.18. *Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Then*

$$|\det T| = \det \sqrt{T^* T}.$$

For T a function defined on a set Ω, define $T(\Omega)$ by $T(\Omega) = \{Tx \mid x \in \Omega\}$.

Lemma 9.19. *Suppose $S \in \mathcal{L}(\mathbb{R}^n)$ is an isometry and $\Omega \subset \mathbb{R}^n$. Then $\operatorname{vol} S(\Omega) = \operatorname{vol} \Omega$.*

Theorem 9.20. *Suppose $T \in \mathcal{L}(\mathbb{R}^n)$ and $\Omega \subset \mathbb{R}^n$. Then $\operatorname{vol} T(\Omega) = |\det T| \cdot \operatorname{vol} \Omega$.*