AS.110.212 Honors Linear Algebra

Textbook

Linear Algebra Done Right by Sheldon Axler

Linear Algebra by Hoffman and Kunze (Additional Resource)

Contents

1 Introduction to Proofs and Systems of Linear Equations

1.1 Proofs

Implication Symbols

1. $\Rightarrow$ (Therefore), also denoted by $\therefore$, is used when the preceding statement implies the succeeding statement

A	T	T	F	F
B	T	F	T	F
A $\Rightarrow$ B	T	F	T	T

2. $\Leftarrow$ (Because), also denoted by $\because$, is used when the succeeding statement implies the preceding statement

3. $\Leftrightarrow$ (Equivalently) stands for $\Rightarrow$ and $\Leftarrow$, means the preceding and the succeeding statement both imply each other

Solution Solution to a linear system or inequality is the set of all n-tuples $(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ which satisfy the system or inequality.

Mathematical Induction Given statements $P(n)$ one for each $n \in \mathbb{N}$. To prove $P(n)$ for all $n \in \mathbb{N}$, it is suffice to prove:

1. Base: $P(0)$ is true, and
2. Inductive Step: $P(n) \Rightarrow P(n+1)\ \forall\ n \in \mathbb{N}$. In other words, assume $P(n)$ is true (called induction hypothesis), then $P(n+1)$ is true.

1.2 Definitions

Cartesian Product Given sets A and B, their Cartesian product, or product, is a set, denoted $A \times B$, whose elements are all ordered pairs (a, b) with a $a \in A$ and $b \in B$, or

$$A \times B := \{(a, b) : a \in A, b \in B\}$$

Function Given sets A and B, a function from A to B is a subset $f \subset A \times B$ with the property that

$$\forall\ a \in A,\ \exists!\ b \in B : (a, b) \in f$$

Injective, Surjective, and Bijective

A function $f : A \to B$ is said to be one-to-one, or injective if

$$\forall\, x, y \in A,\ \ f(x) = f(y) \Rightarrow x = y$$

It is onto, or surjective if

$$\forall\, b \in B,\ \exists\, a \in A : f(a) = b$$

And it is bijective if it is both injective and surjective.

1.3 Systems of Linear Equations

Linear Equations n linear equations in n unknowns

$$\begin{cases} a_{11}x_1 & + a_{12}x_2 & + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 & + a_{22}x_2 & + \cdots + a_{2n}x_n & = b_1 \\ & \vdots & & = \vdots \\ a_{m1}x_1 & + a_{m2}x_2 & + \cdots + a_{mn}x_n & = b_m \end{cases}$$

The system is **homogeneous** if $b_1 = b_2 = \cdots = b_m = 0$.

Linear Combination Selecting m scalars $c1, \cdots, c_m$, multiply the j-th equation by c_j and then add, we obtain the **linear combination** of the equations:

$$c_1(A_{11}x_1 + \cdots + A_{1n}x_n) + \cdots + c_m(A_{m1}x_1 + \cdots + A_{mn}x_n) = c_1y_1 + \cdots + c_my_m$$

Equivalent System Two systems of linear equations are **equivalent** if each equation in each system is a linear combination of the equations in the other system.

Solutions between equivalent linear systems

Equivalent systems of linear equations have exactly the same solutions.

1.4 Matrices and Elementary Row Operations

We can abbreviate the system by $AX = Y$ where

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \text{and} \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

A is called the **matrix of coefficients**, and the entries of the matrix A are the scalars $A(i, j) = A_{ij}$.

Elementary Row Operations

1. $R_i \leftrightarrow R_j$

2. $R_i \to cR_i$ where $c \in \mathbb{R}^*$
3. $R_i \to R_i + cR_j$ where $j \neq i$ and $c \in \mathbb{R}$

Inverse operation of elementary Row operations

To each elementary row operation e there corresponds an elementary row operation e_1, of the same type as e, such that $e_1(e(A)) = e(e_1(A)) = A$ for each A. In other words, the inverse operation (function) of an elementary row operation exists and is an elementary row operation of the same type.

Row-Equivalent If A and B are $m \times n$ matrices over the field F, we say that B is row-equivalent to A if B can be obtained from A by a finite sequence of elementary row operations.

Solutions between systems formed by row-equivalent matrix of coefficients

If A and B are row-equivalent $m \times n$ matrices, the homogeneous systems of linear equations $AX = 0$ and $BX = 0$ have exactly the same solutions.

Row-Reduced Matrix An $m \times n$ matrix R is called row-reduced if:

(a) The first non-zero entry in each non-zero row of R is equal to 1;

(b) Each column of R which contains the leading non-zero entry of some row has all its other entries 0.

Matrix is row-equivalent to a row-reduced matrix

Every $m \times n$ matrix over the field F is row-equivalent to a row-reduced matrix.

1.5 Row-Reduced Echelon Matrices

Row-Reduced Echelon Matrix An $m \times n$ matrix R is called a row-reduced echelon matrix if:

(a) R is row-reduced;

(b) Every row of R which has all its entries 0 occurs below every row which has a non-zero entry;

(c) If rows $1, \cdots, r$ are the non-zero rows of R, and if the leading nonzero entry of row i occurs in column k_i, $i = 1, \cdots, r$, then $k_1 < k_2 < \cdots < k_r$

Matrix is row-Equivalent to a row-Reduced echelon matrix

Every $m \times n$ matrix A is row-equivalent to a row-reduced echelon matrix.

Non-trivial solution for homogeneous system

If A is an $m \times n$ matrix and $m < n$, then the homogeneous system of linear equations $AX = 0$ has a non-trivial solution.

Identity coefficient matrix produces trivial solution

If A is an $n \times n$ (square) matrix, then A is row-equivalent to the $n \times n$ identity matrix if and only if the system of equations $AX = 0$ has only the trivial solution.

2 Fields and Polynomials

2.1 Fields

Notation of F **F** denote either the set of real numbers ($\mathbb{R}$) or the set of complex numbers ($\mathbb{C}$).

Field A field is a set F together with two binary operations:

1. Addition: $\forall\ a, b \in F, a + b \in F$
2. Multiplication: $\forall\ a, b \in F, ab \in F$

with the following properties:

1. F is a ***commutative group*** for addition:
   (a) Commutativity: $a + b = b + a\ \forall\ a, b \in F$
   (b) Associativity: $(a + b) + c = a + (b + c)\ \forall\ a, b, c \in F$
   (c) Identity/Neutral Element: $\exists\ 0 \in F : a + 0 = a\ \forall\ a \in F$
   (d) Inverse: $\forall\ a \in F,\ \exists\ b \in F : a + b = 0$ (Notation: $b = -a$)
2. $F^* = F \backslash \{0\}$ is a commutative group for multiplication:
   (a) Commutativity: $a \cdot b = b \cdot a\ \forall\ a, b \in F$
   (b) Associativity: $(a \cdot b) \cdot c = a \cdot (b \cdot c)\ \forall\ a, b, c \in F$
   (c) Identity/Neutral Element: $\exists\ 1 \in F : 1 \cdot a = a\ \forall\ a \in F$
   (d) Inverse: $\forall\ a \in F^*,\ \exists\ b \in F : a \cdot b = 1$ (Notation: $b = a^{-1} = 1/a$)
3. Distributive Property: $a \cdot (b + c) = a \cdot b + b\dot{c}\ \forall\ a, b, c \in F$

Subfield A subfield of the field C is a set F of complex numbers which is itself a field under the usual operations of addition and multiplication of complex numbers.

Characteristic If a finite number of unit 1 sum to 0 in F, then the least n such that the sum of $n1$'s is 0 is called the **characteristic** of the field F. If it does not happen in F, then F is called a field of **characteristic zero**.

2.2 Polynomials

2.2.1 Complex Conjugate and Absolute Value

Re z, Im z Suppose $z = a + bi$, where a and b are real numbers. The ***real part*** of z, denoted Re z, is defined by $\operatorname{Re} z = a$. The ***imaginary part*** of z, denoted Im z, is defined by $\operatorname{Im} z = b$.

Complex conjugate $\bar{\mathbf{z}}$, Absolute value $|\mathbf{z}|$ The complex conjugate of a complex number $z \in \mathbb{C}$, denoted $\bar{z}$, is defined by

$$\bar{z} = \operatorname{Re} z - (\operatorname{Im} z)i.$$

The absolute value of a complex number z, denoted $|z|$, is defined by

$$|z| = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2}.$$

2.2.2 Uniqueness of Coefficients for Polynomials

Polynomial, $\mathcal{P}(\mathbf{F})$ A function $p : \mathbf{F} \to \mathbf{F}$ is called polynomial with coefficients in $\mathbf{F}$ is there exist $a_0, a_1, \cdots a_m \in \mathbf{F}$ such that

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

for all $z \in \mathbf{F}$.

$\mathcal{P}(\mathbf{F})$ is the set of all polynomials with coefficients in F.

Degree of a Polynomial, deg p A polynomial $p \in \mathcal{P}(\mathbf{F})$ is said to have degree m if there exist scalars $a_0, \cdots a_m \in \mathbf{F}$ with $a_m \neq 0$ such that

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m$$

for all $z \in \mathbf{F}$. If p has degree m, we write $\deg p = m$.

The polynomial that is identically 0 is said to have degree $-\infty$.

$\mathcal{P}_{\mathbf{m}}(\mathbf{F})$ For m a non-negative integer, $\mathcal{P}_m(\mathbf{F})$ denotes the set of all polynomials with coefficients in $\mathbf{F}$ and degree at most m

If a polynomial is the zero function , then all coefficients are 0

Suppose $a_0, a_1, \cdots, a_m \in \mathbf{F}$. If

$$a_0 + a_1 x + \cdots + a_m z^m = 0$$

for every $z \in \mathbf{F}$, then $a_0 = a_1 = \cdots = a_m = 0$.

2.2.3 The Division Algorithm for Polynomials

Division Algorithm for Polynomials

Suppose that $p, s \in \mathcal{P}(\mathbf{F})$, with $s \neq 0$. Then there exist unique polynomials $q, r \in \mathcal{P}(\mathbf{F})$ such that

$$p = sq + r$$

and $\deg r < \deg s$.

2.2.4 Zeros of Polynomials

Zero of Polynomial A number $\lambda \in \mathbf{F}$ is called a ***zero*** (or ***root***) of a polynomial $p \in \mathcal{P}(\mathbf{F})$ if

$$p(\lambda) = 0$$

Factor A polynomial $s \in \mathcal{P}(\mathbf{F})$ is called a factor of $p \in \mathcal{P}(\mathbf{F})$ if there exists a polynomial $q \in \mathcal{P}(\mathbf{F})$ such that $p = sq$.

Each zero of a polynomial corresponds to a degree-1 factor

Suppose $p \in \mathcal{P}(\mathbf{F})$ and $\lambda \in \mathbf{F}$. Then $p(\lambda) = 0$ if and only if there is a polynomial $q \in \mathcal{P}(\mathbf{F})$ such that

$$p(z) = (z - \lambda)q(z)$$

for every $z \in \mathbf{F}$.

A polynomial has at most as many zeros as its degree

Suppose $p \in \mathcal{P}(\mathbf{F})$ is a polynomial with degree $m \geq 0$. Then p has at most m distinct zeros in $\mathbf{F}$.

2.2.5 Factorization of Polynomials over $\mathbb{C}$

Fundamental Theorem of Algebra

Every non-constant polynomial with complex coefficients has a zero.

Factorization of a polynomial over $\mathbb{C}$

If $p \in \mathcal{P}(\mathbb{C})$ is a non-constant polynomial, then p has a unique factorization (except for the order of the factors) of the form

$$p(z) = c(z - \lambda_1) \cdots (z - \lambda_m)$$

where $c, c_1, \cdots c_m \in \mathbb{C}$.

2.2.6 Factorization of Polynomials over $\mathbb{R}$

Polynomials with real coefficients have zeros in pairs

Suppose $p \in \mathcal{P}(\mathbb{C})$ is a polynomial with real coefficients. If $\lambda \in \mathbb{C}$ is a zero of p, then so is $\bar{\lambda}$.

Factorization of a quadratic polynomial

Suppose $b, c \in \mathbb{R}$. Then there is a polynomial factorization of the form

$$x^2 + bx + x = (x - \lambda_1)(x - \lambda_2)$$

with $\lambda_1, \lambda_2 \in \mathbb{R}$ if and only if $b^2 < 4c$.

Factorization of a polynomial over $\mathbb{R}$

Suppose $p \in \mathcal{P}$ is a non-constant polynomial. Then p has a unique factorization (except for the order of the factors) of the form

$$p(x) = c(x - \lambda_1) \cdots (x - \lambda_m)(x^2 + b_1 x + c_1) \cdots (x^2 + b_M x + x_M)$$

where $x, \lambda_1, \cdots, \lambda_m, b_1, \cdots, b_M, c_1, \cdots c_M \in \mathbb{R}$ with $b_j^2 < 4c_j$ for each j

3 Matrix Multiplication and Vector Spaces

3.1 Matrix Multiplication and Invertible Matrix

3.1.1 Matrix Multiplication

Matrix Multiplication Suppose A is an $m \times n$ matrix and C is an $n \times p$ matrix. Then AC is defined to be the m-by-p matrix whose entry in row j , column k, is given by

$$(AC)_{jk} = \sum_{r=1}^{n} A_{jr} C_{rk}$$

Matrix is defined if and only if the number of columns in the first matrix coincides with the number of rows in the second matrix, and matrix multiplication is not commutative.

Multiplicative **associativity** holds for matrix multiplication.

Elementary Matrix An $m \times n$ matrix is said to be an elementary matrix if it can be obtained from the $m \times m$ identity matrix by means of a single elementary row operation.

Elementary Row Operation is Equivalent to Multiplication with Elementary Matrix

Let e be an elementary row operation and let E be the $m \times m$ elementary matrix $E = e(I)$. Then, for every $m \times n$ matrix A,

$$e(A) = EA$$

3.1.2 Invertible Matrix

Invertible Let A be an $n \times n$ (square) matrix over the field F.

- An $n \times n$ matrix B such that $BA = I$ is called a ***left inverse*** of A;
- An $n \times n$ matrix B such that $AB = I$ is called a ***right inverse*** of A.
- If $AB = BA = I$, then B is called a ***two-sided inverse*** of A. A is said to be ***invertible***, and the inverse is denoted as A^{-1}.

Elementary Matrices are Invertible

Elementary matrices are invertible.

The Inverse and The Product of Invertible Matrices are Invertible

Let A and B be $n \times n$ matrices over F.

1. If A is invertible, so is A^{-1} and $(A^{-1})^{-1} = A$.
2. If both A and B are invertible, so is AB, and $(AB)^{-1} = B^{-1}A^{-1}$.

Properties of Invertible Matrix

If A is an $n \times n$ matrix, the following are equivalent.

1. A is invertible.
2. A is row-equivalent to the $n \times n$ identity matrix.
3. A is a product of elementary matrices.
4. The homogeneous system $AX = 0$ has only the trivial solution $X = 0$.
5. The system of equations $AX = Y$ has a solution X for each $n \times 1$ matrix Y.

3.2 $\mathbb{R}^{\mathbf{n}}$ and $\mathbb{C}^{\mathbf{n}}$

F $\mathbf{F}$ stands for either $\mathbb{R}$ or $\mathbb{C}$.

List A ***list*** of length $n : n \in \mathbb{Z}$ is an ordered collection of n elements separated by commas and surrounded by parentheses.

Two lists are equal if and only if they have the same length and the same elements in the same order.

F$^{\mathbf{n}}$ $\mathbf{F}^n$ is the set of all lists of length n of elements of $\mathbf{F}$:

$$\mathbf{F}^n = \{(x_1, \cdots, x_n) : x_j \in \mathbf{F} \text{ for } j = 1, \cdots, n\}$$

For $(x_1, \cdots, x_n) \in \mathbf{F}^n$ and $j \in \{1, 2, \cdots, n\}$, x_j denotes j-th ***coordinate*** of $(x_1, \cdots, x_n)$.

Addition in F$^{\mathbf{n}}$ Addition in $\mathbf{F}^n$ is defined by adding corresponding coordinates:

$$(x_1, \cdots, x_n) + (y_1, \cdots, y_n) = (x_1 + y_1, \cdots, x_n + y_n)$$

Commutativity of Addition in F^n

If $x, y \in \mathbf{F}^n$, then $x + y = y + x$.

0 List 0 denote the list of length n whose coordinates are all 0:

$$0 = (0, \cdots, 0)$$

Additive Inverse in $\mathbf{F^n}$ For $x \in \mathbf{F}^n$, the ***additive inverse*** of x, denoted $-x$, is the vector $-x \in \mathbf{F}^n$ such that

$$x + (-x) = 0$$

In other words, if $x = (x_1, \cdots, x_n)$, then $-x = (-x_1, \cdots, -x_n)$.

Scalar Multiplication in $\mathbf{F^n}$ The product of a scalar λ and a vector in $\mathbf{F}^n$ is computed by multiplying each coordinate of the vector by λ:

$$\lambda(x_1, \cdots, x_n) = (\lambda x_1, \cdots, \lambda x_n)$$

here $\lambda \in \mathbf{F}$ and $(x_1, \cdots, x_n) \in \mathbf{F}^n$

3.3 Definition of Vector Spaces

Vector Space and Subspace Vector Space over a field F is a set V along with an addition and a scalar multiplication on V (closed under addition and scalar multiplication) such that the following properties hold:

1. $(V, +)$ is a commutative group (Commutativity, Associativity, Identity, and Inverse)
2. Multiplicative Identity
3. Distributive Properties

$\mathbf{F^S}$ If S is a set, then $\mathbf{F}^S$ denotes the set of functions from S to $\mathbf{F}$. For $f, g \in \mathbf{F}^S,\ \lambda \in \mathbf{F}$ the addition $f + g \in \mathbf{F}^S$ and scalar multiplication $\lambda f \in \mathbf{F}^S$ are defined by

$$(f + g)(x) = f(x) + g(x)$$

$$(\lambda f)(x) = \lambda f(x)$$

for all $x \in S$.

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

Conditions for a Subspace

A subset U of V is a subspace of V if and only if U satisfies the following three conditions:

1. Additive Identity: $0 \in U$
2. Closed Under Addition: $\forall\ u, v \in U,\ u + v \in U$
3. Closed Under Scalar Multiplication: $\forall\ u \in U$ and $a \in \mathbf{F},\ au \in U$

Sum of Subsets Suppose $U_1, \cdots, U_m$ are subsets of V. The ***sum*** of $U_1, \cdots, U_m$, denoted $U_1 + \cdots U_m$, is the set of all possible sums of elements of $U_1, \cdots, U_m$,

$$U_1 + \cdots U_m = \{u_1 + \cdots + u_m : u_1 \in U_1, \cdots, u_m \in U_m\}$$

Sum of Subspaces is the Smallest Containing Subspace

(Theorem 1.39) Suppose $U_1, \cdots, U_m$ are subspaces of V. Then $U_1 + \cdots + U_m$ is the smallest subspace of V containing $U_1, \cdots, U_m$.

Direct Sum Suppose $U_1, \cdots, U_m$ are subspaces of V. The sum $U_1 + \cdots + U_m$ is called a ***direct sum***, denoted as $U_1 \oplus \cdots \oplus U_m$, if each element of the sum can be written in only one way as a sum $u_1 + \cdots + u_m$, where each $u_j \in U_j$.

Proving Direct Sum To prove $U \oplus W = V$, we need to prove $U + W = V$ and $U + W$ is a direct sum (either by definition, theorem 1.44, or theorem 1.45).

Condition for a Direct Sum

(Theorem 1.44) Suppose $U_1, \cdots, U_m$ are subspaces of V. Then $U_1 + \cdots + U_m$ is a direct sum if and only if the only way to write 0 as a sum $u_1 + \cdots + u_m$, where each $u_j \in U_j$, is by taking each u_j equal to 0.

Direct Sum of Two Subspaces

(Theorem 1.45) Suppose U, W are subspaces of V. Then $U + W$ is a direct sum if and only if $U \cap W = \{0\}$

4 Finite-Dimensional Vector Spaces

4.1 Span and Linear Independence

Linear Combination A ***linear combination*** of a list $v_1, \cdots, v_m$ of vectors in V is a vector of the form

$$a_1 v_1 + \cdots + a_m v_m$$

where $a_1, \cdots, a_m \in \mathbf{F}$

Span The set of all linear combinations of a list of vectors $v_1, \cdots, v_m$ in V is called the ***span*** of $v_1, \cdots, v_m$, denoted $\text{span}(v_1, \cdots, v_m)$. In other words,

$$\text{span}(v_1, \cdots, v_m) = \{a_1 v_1 + \cdots + a_m v_m : a_1, \cdots, \in \mathbf{F}\}$$

The span of the empty list () is defined to be $\{0\}$.

If $\text{span}(v_1, \cdots, v_m)$ equals V, we say that $v_1, \cdots, v_m$ ***spans*** V.

Span is the Smallest Containing Subspace

(Theorem 2.7) The span of a list of vectors in V is the smallest subspace of V containing all the vectors in the list.

Finite-Dimensional and Infinite-Dimensional Vector Space A vector space is called ***finite-dimensional*** if some list of vectors in it spans the space. If it is not finite-dimensional, it is called ***infinite-dimensional***.

Linearly Independent and Dependent A list $v_1, \cdots, v_m$ of vectors in V is called ***linearly independent*** if the only choice of $a_1, \cdots, a_m \in \mathbf{F}$ that makes $a_1 v_1 + \cdots + a_m v_m = 0$ is $a_1 = \cdots = a_m = 0$.

The empty list () is also declared to be linearly independent.

A list of vectors in V is called ***linearly dependent*** if it is not linearly independent.

Linear Dependence Lemma

(Theorem 2.21) Suppose $v_1, \cdots, v_m$ is a linearly dependent list in V. Then there exists $j \in \{1, 2, \cdots, m\}$ such that the following hold:

1. $v_j \in \text{span}(v_1, \cdots, v_{j-1})$
2. If the j-th term is removed from $v_1, \cdots, v_m$, the span of the remaining list equals $\text{span}(v_1, \cdots, v_m)$.

Length of Linearly Independent List $\leq$ Length of Spanning list

(Theorem 2.23) In a finite-dimensional vector space, the length of every linearly independent list of vectors is less than or equal to the length of every spanning list of vectors.

Theorem 2.23 Sketch of Proof Suppose $V = \text{span}(v_1, \cdots, v_n)$ and $W = (u_1, \cdots, u_m)$ is linearly independent. By using linear dependence lemma (theorem 2.21), we can replace the elements in the spanning list by elements of linearly independent list W without changing the span. If $m > n$, $u_1, \cdots, u_n$ spans V and $u_{n+1} \in V$, so the list $u_1, \cdots, u_n, u_{n+1}$ cannot be linearly independent. Hence, the linearly independent list in V has at most n elements.

Finite-Dimensional Subspaces

(Theorem 2.26) Every subspace of a finite-dimensional vector space is finite-dimensional.

4.2 Bases

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

Criterion for Basis

(Theorem 2.29) A list $v_1, \cdots, v_n$ of vectors in V is a basis of V if and only if every $v \in V$ can be written uniquely in the form

$$v = a_1 v_1 + \cdots + a_n v_n$$

Spanning List Contains a Basis

(Theorem 2.31) Every spanning list in a vector space can be reduced to a basis of the vector space.

Basis of Finite-Dimensional Vector Space

(Theorem 2.32) Every finite-dimensional vector space has a basis.

Linearly Independent List Extends to a Basis

(Theorem 2.33) Every linearly independent list of vectors in a finite-dimensional vector space can be extended to a basis of the vector space.

Every Subspace of V is Part of a Direct Sum Equal to V

(Theorem 2.34) Suppose V is finite-dimensional and U is a subspace of V. Then there is a subspace W of V such that $V \oplus W = V$.

4.3 Dimensions

Basis Length Does Not Depend on Basis

(Theorem 2.35) Any two bases of a finite-dimensional vector space have the same length.

Dimensions, dim V The dimension of a finite-dimensional vector space is the length of any basis of the vector space. The dimension of V (if V is finite-dimensional) is denoted by $\dim V$.

Dimension of a Subspace

(Theorem 2.38) If V is finite-dimensional and U is a subspace of V, then $\dim U \leq \dim V$.

Linearly Independent List of the Right Length is a Basis

(Theorem 2.39) Suppose V is finite-dimensional. Then every linearly independent list of vectors in V with length $\dim V$ is a basis of V.

Spanning List of The Right Length is a Basis

(Theorem 2.42) Suppose V is finite-dimensional. Then every spanning list of vectors in V with length $\dim V$ is a basis of V.

Dimension of a Sum

If U_1 and U_2 are subspaces of a finite-dimensional vector space, then

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

5 Linear Maps

5.1 The Vector Space of Linear Maps

5.1.1 Definition of Linear Maps

Linear Map A linear map from V to W is a function $T : V \to W$ with the following properties:

1. Additivity: $T(u + v) = Tu + Tv \ \forall\ u, v \in V$
2. Homogeneity: $T(\lambda v) = \lambda(Tv) \ \forall\ \lambda \in \mathbf{F}, v \in V$

A map $T : V \to W$ is well-defined if and only if

1. $\forall\ v \in V$, Tv is defined.
2. $\forall\ v \in V$, $Tv \in W$.
3. $\forall\ v \in V$, $\exists!\ w \in W : Tv = w$ (namely, T maps v to an unique element).

$\mathcal{L}(\mathbf{V}, \mathbf{W})$ The set of all linear maps from V to W is denoted $\mathcal{L}(V, W)$.

Linear Maps and Basis of Domain

(theorem 3.5) Suppose $v_1, \cdots, v_n$ is a basis of V and $w_1, \cdots, w_n \in W$. Then there exists a unique linear map $T : V \to W$ such that

$$Tv_j = w_j$$

for each $j = 1, \cdots, n$.

5.1.2 Algebraic Operations on $\mathcal{L}(V, W)$

Addition and Scalar Multiplication on $\mathcal{L}(\mathbf{V}, \mathbf{W})$ Suppose $S, T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbf{F}$. The ***sum*** $S + T$ and the ***product*** λT are the linear maps from V to W defined by

$$(S + T)(v) = Sv + Tv \quad \text{and} \quad (\lambda T)(v) = \lambda(Tv)$$

for all $v \in V$.

$\mathcal{L}(\mathbf{V}, \mathbf{W})$ is a Vector Space

(Theorem 3.7) With the operations of addition and scalar multiplication as defined above, $\mathcal{L}(V, W)$ is a vector space.

Product of Linear Maps If $T, S \in \mathcal{L}(V, W)$, then the ***product*** $ST \in \mathcal{L}(V, W)$ is defined by

$$(ST)(u) = S(Tu)$$

for $u \in U$.

Algebraic Properties of Products of Linear Maps

(Theorem 3.9)

1. Associativity: $(T_1T_2)T_3 = T_1(T_2T_3)$
2. Identity: $TI = IT = T$
3. Distributive Properties: $(S_1 + S_2)T = S_1T = S_2T$ and $S(T_1 + T_2) = ST_1 + ST_2$

Note that the multiplication of linear maps is not commutative. In other words, it is not necessarily true that $ST = TS$.

Linear Maps Take 0 to 0

(theorem 3.11) Suppose T is a linear map from V to W. Then $T(0) = 0$.

5.2 Null Space and Ranges

5.2.1 Null Space and Injectivity

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

$$\text{null } T = \{v \in V : Tv = 0\}$$

The Null Space is a Subspace

(Suppose 3.14) Suppose $T \in \mathcal{L}(V, W)$. Then null T is a subspace of V.

Injective A function $T : V \to V$ is called ***injective*** (one-to-one) if $Tu = Tv$ implies $u = v$.

Injectivity is Equivalent to Null Space Equals $\{0\}$

(theorem 3.16) Let $T \in \mathcal{L}(V, W)$. Then T is injective if and only if null $T = \{0\}$.

5.2.2 Range and Surjectivity

Range For $T \in \mathcal{L}(V, W)$, the range of T is the subset of W consisting of those vectors that are of the form Tv for some $v \in V$:

$$\text{range } T = \{T : v \in V\}$$

The Range is a Subspace

(Suppose 3.19) Suppose $T \in \mathcal{L}(V, W)$. Then range T is a subspace of W.

Surjectivity A function $T : V \to W$ is called ***surjective*** (onto) if its range equals W.

5.2.3 Fundamental Theorem of Linear Maps

Rank-Nullity Theorem (Fundamental Theorem of Linear Maps)

(Theorem 3.22) Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Then range T is finite-dimensional and

$$\dim V = \dim \text{null } T + \dim \text{range } T$$

This theorem is called Rank-Nullity Theorem because

$$\text{rank } T := \dim(\text{range } T)$$
$$\text{nullity } T := \dim(\text{null } T)$$

A Map to a Smaller Dimensional Space is Not Injective

(Theorem 3.23) Suppose V and W are finite-dimensional vector spaces such that $\dim V > \dim W$. Then no linear map from V to W is injective.

A Map to a Larger Dimensional Space is Not Surjective

(Theorem 3.24) Suppose V and W are finite-dimensional vector spaces such that $\dim V < \dim W$. Then no linear map from V to W is surjective.

Homogeneous System of Linear Equations

(Theorem 3.26) A homogeneous system of linear equations with more variables than equations has nonzero solutions.

Inhomogeneous System of Linear Equations

(Theorem 3.29) An inhomogeneous system of linear equations with more equations than variables has no solution for some choice of the constant terms.

5.3 Matrices

5.3.1 Representing a Linear Map by a Matrix

Matrix, $\mathbf{A_{j,k}}$, $\mathbf{A_{\cdot,k}}$, $\mathbf{A_{j,\cdot}}$ Let m and n denote positive integers. An m-by-n matrix A is a rectangular array of elements of F with m rows and n columns:

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

- $A_{j,k}$ denotes the entry in row j , column k of A.
- $Aj,\cdot$ denotes the 1-by-n matrix consisting of row j of A, where $1 \leq j \leq m$.
- $A\cdot,k$ denotes the m-by-1 matrix consisting of column k of A, where $1 \leq j \leq n$.

Matrix of a Linear Map, $\mathcal{M}(\mathbf{T})$ Suppose $T \in \mathcal{L}(V, W)$ and $v_1, \cdots, v_n$ is a basis of V and $w_1, \cdots, w_m$ is a basis of W. The matrix of T with respect to these bases is the m-by-n matrix, $\mathcal{M}(T)$, whose entries $A_{j,k}$ are defined by

$$Tv_k = A_{1,k}w_1 + \cdots A_{m,k}w_m$$

If the bases are not clear from the context, then the notation $\mathcal{M}(T, (v_1, \cdots, v_n), (w_1, \cdots, w_m))$ is used.

Suppose $\mathfrak{B} = (v_1, \cdots, v_n)$ and $\mathfrak{C} = (w_1, \cdots, w_m)$. By the definition of the matrix of the linear map

$$\mathcal{M}(T)_{\cdot,k} = (Tv_k)_{\mathfrak{C}}$$

In other words, the k-th column of $\mathcal{M}(T)$ is the coordinate of Tv_k of the basis of W.

The linear map can be represented by the matrix multiplication

$$[T(v)]_{\mathfrak{C}} = c_1[T(v_1)]_{\mathfrak{C}} + \cdots + c_n[T(v_n)]_{\mathfrak{C}}$$

$$[T(v)]_{\mathfrak{C}} = \mathcal{M}(T)(v)_{\mathfrak{B}}$$

$\mathbf{F^{m,n}}$ For m and n positive integers, the set of all m-by-n matrices with entries in F is denoted by $\mathbf{F}^{m,n}$.

$\mathbf{dim\ F^{m,n} = mn}$

(Theorem 3.40) Suppose m and n are positive integers. With addition and scalar multiplication defined as above, $\mathbf{F}^{m,n}$ is a vector space with dimension mn.

Matrix Addition and Scalar Multiplication

– The sum of two matrices of the same size is the matrix obtained by adding corresponding entries in the matrices: $(A+C)_{j,k} = A_{j,k} + C_{j,k}$.

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

– The product of a scalar and a matrix is the matrix obtained by multiplying each entry in the matrix by the scalar: $(\lambda A)_{j,k} = \lambda A_{j,k}$.

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

The Matrix of the Sum of Linear Maps

(Theorem 3.36) Suppose $S, T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(S+T) = \mathcal{M}(S) + \mathcal{M}(T)$.

The Matrix of a Scalar Times a Linear Map

(Theorem 3.38) Suppose $\lambda \in \mathbf{F}$ and $T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

5.3.2 Matrix Multiplication

Matrix Multiplication Suppose A is an m-by-n matrix and C is an n-by-p matrix. Then AC is defined to be the m-by-p matrix whose entry in row j , column k, is given by the following equation:

$$(AC)^{j,k} = \sum_{r=1}^{n} A_{j,r} C_{r,k}$$

In other words, the entry in row j , column k, of AC is computed by taking row j of A and column k of C, multiplying together corresponding entries, and then summing.

Entry of Matrix Product Equals Row Times Column

(Theorem 3.47) Suppose A is an m-by-n matrix and C is an n-by-p matrix. Then

$$(AC)_{j,k} = A_{j,\cdot} C_{\cdot,k}$$

for $1 \leq j \leq m$ and $1 \leq k \leq p$.

Column of Matrix Product Equals Matrix Times Column

(Theorem 3.49) Suppose A is an m-by-n matrix and C is an n-by-p matrix. Then

$$(AC)_{\cdot,k} = AC_{\cdot,k}$$

for $1 \leq k \leq p$.

Linear Combination of Columns

(Theorem 3.52) Suppose A is an m-by-n matrix and $c = \text{col}(c_1, \cdots, c_n)$. Then

$$Ac = c_1 A_{\cdot,1} + \cdots + c_n A_{\cdot,n}$$

In other words, Ac is a linear combination of the columns of A, with the scalars that multiply the columns coming from c.

5.4 Invertibility and Isomorphic Vector Spaces

5.4.1 Invertible Linear Maps

Invertible, Inverse A linear map $T \in \mathcal{L}(V, W)$ is called ***invertible*** if there exists a linear map $S \in \mathcal{L}(W, V)$ such that ST equals the identity map on V and V equals the identity map on W.

A linear map $S \in \mathcal{L}(W, V)$ satisfying $ST = I$ and $TS = I$ is called an ***inverse*** of T (note that the first I is the identity map on V and the second I is the identity map on W).

Note that the definition of invertible is not equivalent to $ST = TS = I$ because $TS = I_W$ where $ST = I_V$.

Inverse is Unique

(Theorem 3.54) An invertible linear map has a unique inverse.

$\mathbf{T^{-1}}$ If T is invertible, then its inverse is denoted by T^{-1}. In other words, if $T \in \mathcal{L}(V, W)$ is invertible, then T^{-1} is the unique element of $\mathcal{L}(W, V)$ such that $T^{-1}T = I$ and $TT^{-1} = I$.

Invertibility is Equivalent to Injectivity and Surjectivity

(Theorem 3.56) A linear map is invertible if and only if it is injective and surjective.

5.4.2 Isomorphic Vector Spaces

Isomorphism, Isomorphic An ***isomorphism*** is an invertible linear map.

Two vector spaces are called ***isomorphic*** if there is an isomorphism from one vector space onto the other one.

Dimension Shows Whether Vector Spaces are Isomorphic

(Theorem 3.59) Two finite-dimensional vector spaces over $\mathbf{F}$ are isomorphic if and only if they have the same dimension.

$\mathcal{L}(\mathbf{V}, \mathbf{W})$ and $\mathbf{F}^{\mathbf{m},\mathbf{n}}$ are Isomorphic

(Theorem 3.60) Suppose $v_1, \cdots, v_n$ is a basis of V and $w_1, \cdots, w_m$ is a basis of W. Then $\mathcal{M}$ is an isomorphism between $\mathcal{L}(V, W)$ and $\mathbf{F}^{m,n}$.

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

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

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

5.4.3 Linear Maps Thought of as Matrix Multiplication

Matrix of a Vector, $\mathcal{M}(\mathbf{v})$ Suppose $v \in V$ and $v_1, \cdots, v_n$ is a basis of V. The ***matrix*** of v with respect to this basis is the n-by-1 matrix

$$\mathcal{M}(v) = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix},$$

where $c_1, \cdots, c_n$ are the scalars such that

$$v = c_1 v_1 + \cdots + c_n v_n$$

Note the matrix of a vector is the coordinate of the vector.

$\mathcal{M}(\mathbf{T})_{\cdot,\mathbf{k}} = \mathcal{M}(\mathbf{v_k})$

(Theorem 3.64) Suppose $T \in \mathcal{L}(V, W)$ and $v_1, \cdots, v_n$ is a basis of V and $w_1, \cdots, w_n$ is a basis of W. Let $1 \leq k \leq n$. Then the k^{th} column of $\mathcal{M}(T)$, which is denoted by $\mathcal{M}(T)_{\cdot,k}$, equals $\mathcal{M}(v_k)$.

Linear Maps Act Like matrix multiplication

(Theorem 3.65) Suppose $T \in \mathcal{L}(V, W)$ and $v \in V$. Suppose $v_1, \cdots, v_n$ is a basis of V and $w_1, \cdots, w_n$ is a basis of W. Then

$$\mathcal{M}(Tv) = \mathcal{M}(T)\mathcal{M}(v)$$

5.4.4 Operators

Operator, $\mathcal{L}(\mathbf{V})$ A linear map from a vector space to itself is called an ***operator***.

The notation $\mathcal{L}(V)$ denotes the set of all operators on V. In other words, $\mathcal{L}(V) = \mathcal{L}(V, V)$.

Injectivity is Equivalent to Surjectivity in Finite Dimensions

(Theorem 3.69) Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then the following are equivalent:

1. T is invertible;
2. T is injective;
3. T is surjective.

Theorem 3.69 Sketch of Proof Prove by Rank-Nullity Theorem.

5.5 Products and Quotients of Vector Spaces

5.5.1 Products of Vectors Spaces

Product of Vector Spaces Suppose $V_1, \cdots, V_m$ are vector spaces over $\mathbf{F}$, the ***product*** $V_1 \times \cdots \times V_m$ is defined by

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

Addition and scalar multiplication on $V_1 \times \cdots \times V_m$ is defined by

$$(u_1, \cdots, u_m) + (v_1, \cdots, v_m) = (u_1 + v_1, \cdots, u_m + v_m)$$
$$\lambda(v_1, \cdots, c_m) = (\lambda v_1, \cdots, \lambda v_m)$$

Product of vector spaces is a Vector space

(Theorem 3.73) Suppose $V_1, \cdots, V_m$ are vector spaces over $\mathbf{F}$. Then $V_1 \times \cdots \times V_m$ is a vector space over $\mathbf{F}$.

Dimension of a Product is the Sum of Dimensions

(Theorem 3.76) Suppose $V_1, \cdots, V_m$ are finite-dimensional vector spaces. Then $V_1 \times \cdots \times V_m$ is finite-dimensional and

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

Theorem 3.76 Sketch of Proof Construct a basis for each V_j, consider the element of the product equals the basis vector in the j^{th} slot and 0 in other slots. The list of all such vectors is the basis of the product.

5.5.2 Products and Direct Sums

Products and Direct Sums

(Theorem 3.77) Suppose that $U_1, \cdots, U_m$ are subspaces of V. Define a linear map $\Gamma : U_1 \times \cdots \times U_m \to V_1 + \cdots + V_m$ by

$$\Gamma(u_1, \cdots, u_m) = u_1 + \cdots + u_m$$

Then $U_1 + \cdots U_m$ is a direct sum if and only if Γ is injective.

Theorem 3.77 Sketch of Proof Γ is injective if and only null $\Gamma = \{0\}$, i.e., there is a unique way to write 0 as the sum, thus the sum is a direct sum (Theorem 1.44).

A Sum is a Direct Sum If and Only If Dimensions Add Up

(Theorem 3.78) Suppose V is finite-dimensional and $U_1, \cdots, U_m$ are subspaces of V. Then $U_1 + \cdots + U_m$ is a direct sum if and only if

$$\dim(U_1 + \cdots + U_m) = \dim U_1 + \cdots + \dim U_m$$

Theorem 3.77 Sketch of Proof Γ is surjective. By the Rank-Nullity Theorem, Γ is injective if and only if $\dim(U_1 \times \cdots \times U_m) = \dim(U_1 + \cdots + U_m)$, and by Theorem 3.76, we can have the equation as desired.

5.5.3 Quotients of vectors Spaces

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

$$v + U = \{v + U : u \in U\}$$

Affine Subset, Parallel

- 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.
- For $v \in V$ and U a subspace of V, the affine subset $v + U$ is said to be ***parallel*** to U.

Quotient Space, V/U Suppose U is a subspace of V. Then the ***quotient space*** V/U is the set of all affine subsets of V parallel to U. In other words,

$$V/U = \{v + U : v \in V\}$$

Two Affine Subsets Parallel to U are Equal or Disjoint

(Theorem 3.85) Suppose U is a subspace of V and $v, w \in V$. Then the following are equivalent:

(a) $v - w \in U$

(b) $v + U = w + U$

(c) $(v + U) \cap (w + U) \neq \varnothing$

Addition and Scalar Multiplication on V/U Suppose U is a subspace of V. Then addition and scalar multiplication are defined on V/U by

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

for $v, w \in V$ and $\lambda \in \mathbf{F}$.

Quotient Space is a Vector Space

(Theorem 3.87) Suppose U is a subspace of V. Then V/U, with the operations of addition and scalar multiplication as defined above, is a vector space.

Theorem 3.87 Note To show that the quotient space is a vector space, we need to show the operations are well-defined and make sense.

Quotient Map 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 $v \in V$.

Dimension of a Quotient Space

(Theorem 3.89) Suppose V is finite-dimensional and U is a subspace of V. Then

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

Theorem 3.89 Sketch of Proof By the definition of π, π is surjective, and null $\pi = U$ by Theorem 3.85. The desired equation can be derived by the Rank-Nullity Theorem.

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

$$\tilde{T}(v + \text{null } T) = Tv$$

Null Space and Range of $\tilde{\text{T}}$

(Theorem 3.91) Suppose $T \in \mathcal{L}(V, W)$. Then

(a) $\tilde{T}$ is a linear map from $V/(\text{null } T)$ to W

(b) $\tilde{T}$ is injective

(c) range $\tilde{T}$ = range T

(d) $V/(\text{null } T)$ is isomorphic to range W

5.6 Duality

5.6.1 The Dual Space and the Dual Map

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

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

dim V′ = dim V

(Theorem 3.95) Suppose V is finite-dimensional. Then V' is also finite-dimensional and dim V' = dim V.

Theorem 3.95 Sketch of Proof Directly follows from 3.61.

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

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

Equivalently,

$$\varphi_j(c_1 v_1 + \cdots + c_n v_n) = c_j$$

Dual Basis is a Basis of the Dual space

(Theorem 3.98) Suppose V is finite-dimensional. Then the dual basis of a basis of V is a basis of V'.

Dual Map, T′ 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 $\varphi \in W'$.

Algebraic Properties of Dual Maps

(Theorem 3.101)

- $(S+T)' = S' + T'$ for all $S, T \in \mathcal{L}(V, W)$
- $(\lambda T)' = \lambda T'$ for all $\lambda \in \mathbf{F}$ and $T \in \mathcal{L}(V, W)$
- $(ST)' = T'S'$ for all $T \in \mathcal{L}(U, V)$ and all $S \in \mathcal{L}(V < W)$

5.6.2 The Null Space and Range of the Dual of a Linear Map

Annihilator, $\mathbf{U^0}$ For $U \subset V$, the ***annihilator*** of U, denoted U^0, is defined by

$$U^0 = \{\phi \in V' : \varphi(u) = 0 \text{ for all } u \in U\}$$

The Annihilator is a Subspace

(Theorem 3.105) Suppose $U \subset V$. Then U^0 is a subspace of V'.

Dimension of the Annihilator

(Theorem 3.106) Suppose V is finite-dimensional and U is a subspace of V. Then

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

Theorem 3.106 Sketch of Proof Suppose W is the complement of U on V (i.e., $U \oplus W = V$). Construct the dual basis of U, $\varphi_1, \cdots, \phi_n$, and the dual basis of W, $\psi_1, \cdots, \psi_m$. Given the basis is linearly independent, we can prove that $U^0 = \text{span}(\psi_1, \cdots, \psi_m)$, namely $U^0 = W'$. By the dimension of the direct sum, $\dim U + \dim U^0 = V' = V$.

The Null Space of $\mathbf{T'}$

(Theorem 3.107) Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then

(a) $\text{null } T' = (\text{range } T)^0$

(b) $\dim \text{null } T' = \dim \text{null } T + \dim W - \dim V$

T Surjective is Equivalent to $\mathbf{T'}$ Injective

(Theorem 3.108) Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then T is surjective if and only if T' is injective.

The Range of T′

(Theorem 3.109) Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then

(a) dim range $T' =$ dim range T

(b) range $T' = (\text{null } T)^0$

T Injective is Equivalent to T′ Surjective

(Theorem 3.110) Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then T is injective if and only if T' is surjective.

5.6.3 The Matrix of the Dual of a Linear Map

Transpose, $\mathbf{A^t}$ The transpose of a matrix A, denoted A^t, is the matrix obtained from A by interchanging the rows and columns. More specifically, if A is an m-by-n matrix, then A^t is the n-by-m matrix whose entries are given by the equation

$$(A^t)_{k,j} = A_{j,k}$$

The Transpose of the Product of Matrices

(Theorem 3.113) If A is an m-by-n matrix and C is an n-by-p matrix, then

$$(AC)^t = C^t A^t$$

The Transpose of the Product of Matrices

(Theorem 3.113) Suppose $T \in \mathcal{L}(V, W)$, then $\mathcal{M}(T') = (\mathcal{M}(T))^t$.

5.6.4 The Rank of a Matrix

Row Rank, Column Rank, Rank Suppose A is an m-by-n matrix with entries in $\mathbf{F}$.

- The ***row rank*** of A is the dimension of the span of the rows of A in $\mathbf{F}^{1,n}$
- The ***column rank*** of A is the dimension of the span of the columns of A in $\mathbf{F}^{m,1}$
- The ***rank*** of A is the row/column rank of A

Dimension of Range T Equals Column Rank of $\mathcal{M}$(T)

(Theorem 3.117) Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then dim range T equals the column rank of $\mathcal{M}(T)$.

Row Rank Equals Column Rank

(Theorem 3.118) Suppose $A \in \mathbf{F}^{m,n}$, then the row rank of A equals the column rank of A.

6 Eigenvalues, Eigenvectors, and Invariant Subspaces

6.1 Invariant Subspaces

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

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

Equivalent Conditions to be an Eigenvalue

(Theorem 5.6) Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{F}$. Then the following are equivalent

(a) λ is an eigenvalue of T.

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

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

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

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

Linear Independent Eigenvectors

(Theorem 5.10) Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \cdots, \lambda_m$ are distinct eigenvalues of T and $v_1, \cdots, v_m$ are corresponding eigenvectors. Then $v_1, \cdots, v_m$ is linearly independent.

Number of Eigenvalues

(Theorem 5.13) Suppose V is finite-dimensional. Then each operator on V has at most $\dim V$ distinct eigenvalues.

Restriction Operator $\mathbf{T|_U}$ and Quotient Operator $\mathbf{T/U}$ Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V invariant under T. The ***restriction operator*** $T|_U \in \mathcal{L}(U)$ is defined by

$$T|_U(u) = Tu$$

for $u \in U$. The ***quotient operator*** $T/U \in \mathcal{L}(V/U)$ is defined by

$$(T/U)(v + U) = Tv + U$$

for $v \in V$.

6.2 Eigenvectors and Upper-Triangular Matrices

6.2.1 Polynomials Applied to Operators

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

- T^m is defined by $T^m = T \cdots T$ (m times).
- T^0 is defined to be the identity operator I in V.
- If T is invertible withe inverse T^{-1}, then T^{-m} is defined by $T^{-m} = (T^{-1})^m$.

$\mathbf{p(T)}$ Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(F)$ is a polynomial given by $p(z) = a_0 + a_1 z + \cdots + a_m z^m$ for $z \in \mathbf{F}$. Then $p(T)$ is the operator defined by

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

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

$$(pq)(z) = p(z)q(z)$$

for $z \in \mathbf{F}$.

Multiplicative Properties

(Theorem 5.20) Suppose $p, q \in \mathcal{P}(\mathbf{F})$ and $T \in \mathcal{L}(V)$. Then

(a) $(pq)(T) = p(T)q(T)$;

(b) $p(T)q(T) = q(T)p(T)$.

6.2.2 Existence of Eigenvalues

Operators on Complex Vector Spaces Have an Eigenvalue

(Theorem 5.21) Every operator on a finite-dimensional, nonzero, complex vector space has an eigenvalue.

Theorem 5.21 Sketch of Proof Construct a linear-independent list $v, Tv, \cdots, T^n v$. By the Fundamental Theorem of Algebra, $p(T)$ can be factorized as $c(T-\lambda_1 I)\cdots(T-\lambda_m I)v$, implying the existence of eigenvalue.

6.2.3 Upper-Triangular Matrices

Diagonal of a Matrix, Upper-Triangular Matrix 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.

Upper-triangular matrix is in the form of

$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

the 0 in the matrix above indicates that all entries below the diagonal equal 0.

Conditions for Upper-Triangular Matrix

(Theorem 5.26) Suppose $T \in \mathcal{L}(V)$ and $v_1, \cdots, v_n$ is a basis of V. Then the following are equivalent:

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

(b) $Tv_j \in \operatorname{span}(v_1, \cdots, v_j)$ for each $j = 1, \cdots, n$;

(c) $\operatorname{span}(v_1, \cdots, v_j)$ is invariant under T for each $j = 1, \cdots, n$.

Theorem 5.26 Sketch of Proof $(b) \Rightarrow (c)$: Obviously $Tv_k \in \operatorname{span}(v_1, \cdots, v_j)$, so $\operatorname{span}(v_1, \cdots, v_j)$ is invariant.

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

(Theorem 5.27) Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some basis of V.

Theorem 5.27 Sketch of Proof Mathematical Induction. Obviously desired result holds if dim $V = 1$. (Proof 2 given in the book) For $n =$ dim V > 1, let $U = \operatorname{span}(v_1)$, so we can construct $T/U \in \mathcal{L}(V/U)$ where dim $V/U = n-1$. Given $(T/U)(v_j+U) = Tv_j+U$ by definition and $(T/U)(v_j+U) \in \operatorname{span}(v_2+U, \cdots, v_j+U)$ by inductive hypothesis, we can prove $Tv_j \in \operatorname{span}(v_1, \cdots, v_j)$. Hence, by proving $v_1, \cdots, v_n$ is a basis by complement subspace, we complete the proof.

Determination of Invertibility from Upper-Triangular Matrix

(Theorem 5.30) Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V. Then T is invertible if and only if all the entries on the diagonal of that upper-triangular matrix are nonzero.

Theorem 5.30 Sketch of Proof $\Rightarrow$ Direction: Proof by contradiction. If $\lambda_1 = 0$, null $T \neq \{0\}$, so T is not injective. If $\lambda_j = 0 : j > 1$, range $T|_{U_j} \neq U_j$, so not surjective.

$\Rightarrow$ Direction: We need to prove for each j, $T(v_j/\lambda_j) = c_1 v_1 + \cdots + c_{j-1}v_{j-1} + v_j$, implying $v_j \in \text{range } T$, followed by range $T = V$.

Determination of Eigenvalues from Upper-Triangular Matrix

(Theorem 5.32) Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V. Then the eigenvalues of T are precisely the entries on the diagonal of that upper-triangular matrix.

Theorem 5.32 Sketch of Proof $M(T - \lambda I)$ is given by

$$M(T - \lambda I) = \begin{pmatrix} \lambda_1 - \lambda & & * \\ & \ddots & \\ 0 & & \lambda_n - \lambda \end{pmatrix}$$

T is not invertible iff some entries in the diagonal are zero, thus $\lambda_1, \cdots, \lambda_n$ are all eigenvalues.

6.3 Eigenspaces and Diagonal Matrices

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

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

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

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

Sum of Eigenspaces is a Direct Sum

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

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

is a direct sum. Furthermore,

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

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

Conditions Equivalent to Diagonalizability

(Theorem 5.41) Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, \cdots, \lambda_m$ 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 exits 1-dimensional subspaces $U_1, \cdots, U_n$ of V, each invariant under T, such that $V = U_1 \oplus \cdots \oplus U_n$;

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

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

Theorem 3.95 Sketch of Proof

- $(a) \Leftrightarrow (b)$: T has a diagonal matrix, with $\lambda_1, \cdots, \lambda_n$ on the diagonal, with respect to a basis $v_1, \cdots, v_n$ if and only if $Tv_j = \lambda v_j$.
- $(b) \Rightarrow (c)$: let $U_j = \text{span}(v_j)$ where $v_1, \cdots, v_n$ is a basis consisting eigenvectors of T. Hence, we can prove $V = U_1 \oplus \cdots \oplus U_n$
- $(c) \Rightarrow (b)$: choose non-zero vector v_j for each U_j. Each vector in V can be written as a sum $u_1 + \cdots + u_n$ where u_k is in U_j and thus a scalar multiple of v_j. Hence, $v_1, \cdots, v_n$ is a basis of V.
- $(b) \Rightarrow (d)$: $v_1, \cdots, v_n$ is a basis consisting of eigenvectors of T, so every vector in V is a scalar combination of the list. Note that $c_j v_j \in E(\lambda_j, T)$ for each j, so $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.
- $(d) \Rightarrow (e)$: by theorem 3.78 A sum if a direct sum iff dimensions add up.
- $(e) \Rightarrow (b)$: choose a basis of each $E(\lambda_j, T)$, and put all bases together to form a list $v_1, \cdots, v_n$. The list is linearly independent (To prove linear independence, let $u_j \in E(\lambda_j, T)$ be the collection of all $a_k v_k \in E(\lambda_j, T)$. $u_1, \cdots, u_m = 0$ implies each u equals 0, implying that each a equals 0.) and thus a basis of V

Enough Eigenvalues Implies Diagonalizability

(Theorem 5.44) Tf $T \in \mathcal{L}(V)$ has $\dim V$ distinct eigenvalues, then T is diagonalizable.

7 Operators on Complex Vector Spaces

7.1 Generalized Eigenvectors and Nilpotent Operators

7.1.1 Null Spaces of Powers of an Operator

Sequence of Increasing Null Spaces

(Theorem 8.2) Suppose $T \in \mathcal{L}(V)$. Then

$$\{0\} = \text{null } T^0 \subset \text{null } T^1 \subset \cdots \subset \text{null } T^k \subset \text{null } T^{k+1} \subset \cdots$$

Theorem 8.2 Sketch of Proof Proving $\forall\ v \in \text{null } T^k$, $v \in \text{null } T^{k+1}$ yields the statement.

Equality in the Sequence of Null Spaces

(Theorem 8.3) Suppose $T \in \mathcal{L}(V)$. Suppose m is a non-negative integer such that null $T^m \subset$ null T^{m+1}. Then

$$\text{null } T^m = \text{null } T^{m+1} = \text{null } T^{m+2} = \cdots$$

Theorem 8.3 Sketch of Proof We need to prove null $T^{m+k} =$ null T^{m+k+1}. One direction is obvious. For $v \in \text{null } T^{m+k+1}$, $T^k \in \text{null } T^{m+1} = \text{null } T^m$, which implies inclusion in another direction.

Null Spaces Stop Growing

(Theorem 8.4) Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then

$$\text{null } T^n = \text{null } T^{n+1} = \text{null } T^{n+2} = \cdots$$

Theorem 8.4 Sketch of Proof Contrapositive. Suppose null $T^n \neq$ null T^{n+1}, then null $T^0 \subsetneq$ null $T^1 \subsetneq \cdots \subset$ null $T^n \subsetneq$ null T^{n+1}, so $\dim T^{n+1} \geq n+1$, which clearly false.

V is the Direct Sum of null $\mathbf{T^{\dim V}}$ and range $\mathbf{T^{\dim V}}$

(Theorem 8.5) Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then

$$V = \text{null } T^n \oplus \text{range } T^n$$

Theorem 8.4 Sketch of Proof We first need to prove the sum is direct by null $T^n \cap$ range $T^n = \{0\}$. dim null $T^n \oplus$ range $T^n = \dim V$ by the Rank-Nullity Theorem gives the sum is equal to V.

7.1.2 Generalized Eigenvectors

Generalized Eigenvector 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 $j \in \mathbb{N}^+$.

Generalized Eigenspace, $\mathbf{G(\lambda, T)}$ Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. 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.

Note there is no concept of generalized eigenvalues because all "generalized eigenvalues" are eigenvalue of T.

Description of Generalized Eigenspaces

(Theorem 8.11) Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. Then $G(\lambda, T) = \text{null } (T - \lambda I)^{\dim V}$.

Theorem 8.11 Sketch of Proof By definition, $G(\lambda, T)$ is the intersection of all null $(T - \lambda I)^j$. $\forall j \in \mathbb{N}^+$, null $(T - \lambda I)^j \subset$ null $(T - \lambda I)^{\dim V}$ by 8.2 and 8.4, implying the desired statement.

Linearly Independent Generalized Eigenvectors

(Theorem 8.13) Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \cdots, \lambda_m$ are distinct eigenvalues of T and $v_1, \cdots, v_m$ are corresponding generalized eigenvectors. Then $v_1, \cdots, v_m$ is linearly independent.

Theorem 8.13 Sketch of Proof Let k be the largest integer s.t. $w := (T - \lambda_1 I)^k \neq 0$, then $(T - \lambda I)^k w = (\lambda_1 - \lambda)w$. Construct $S = (T - \lambda_1 I)^k (T - \lambda_2 I)^{\dim V} \cdots (T - \lambda_m I)^{\dim V}$ and apply S to $a_1 v_1 + \cdots + a_m v_m = 0$. We obtain $a_1(\lambda_1 - \lambda_2)^{\dim V} \cdots (\lambda_1 - \lambda_m)^{\dim V} w = 0$m implying $a_1 = 0$. Repeat the procedure so $a_1 = \cdots = a_m = 0$, implying $v_1, \cdots, v_m$ is linearly independent.

7.1.3 Nilpotent Operator

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

Nilpotent Operator Raised to Dimension of Domain is 0

(Theorem 8.18) Suppose $N \in \mathcal{L}(V)$ is nilpotent. Then $N^{\dim V} = 0$.

Theorem 8.18 Sketch of Proof $V = G(0, T) := \text{null } T^{\dim V}$, implying the desired statement.

Matrix of a Nilpotent Operator

(Theorem 8.19) Suppose N is a nilpotent operator on V. Then there is a basis of V with respect to which the matrix of N has the form

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

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

Theorem 8.19 Sketch of Proof First choose a basis of null N, extend it to the basis of null N^2, and so on to null N^k where k is the smallest integer s.t. $N^k = 0$. Suppose v_m is a vector add to extend the basis to null N^m. Then $N^m v_m = 0$, $Nv_m \in \text{null } N^{m-1}$, so v_m is in the span of v_1's to v_{m-1}'s (thus the entry on the diagonal in the corresponding column is 0). Hence, the matrix is in the desired form.

7.2 Decomposition of an Operator

7.2.1 Description of Operators on Complex Vector Spaces

The Null Space and Range of p(T) are Invariant Under T

(Theorem 8.20) Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbf{F})$. Then null $p(T)$ and range $p(T)$ are invariant under T.

Theorem 8.20 Sketch of Proof Suppose $v \in \text{null } p(T)$, then $p(T)(Tv) = 0$, so $Tv \in \text{null } p(T)$. Suppose $w \in \text{range } p(T)$ where $p(T)v = w$, then $Tw = p(T)(Tv)$, so $Tw \in \text{range } p(T)$.

Description of Operators on Complex Vector Spaces

(Theorem 8.21) Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \cdots, \lambda_m$ 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 8.21 (a) Sketch of Proof** Mathematical Induction. For $n = \dim V > 1$. (1) Decomposition: suppose $V = \text{null } (T - \lambda_1 I)^n \oplus \text{range } (T - \lambda_1 I)^n = G(\lambda_1, T) \oplus U$ (U is invariant). By inductive hypothesis, $U = G(\lambda_2, T|_U) \oplus \cdots \oplus G(\lambda_m, T|_U)$. (2) Prove $G(\lambda_j, T|_U) = G(\lambda_j, T)$: Let $v := v_1 + u = v_1 + v_2 + \cdots + v_m$, linearly independence implies (in particular) $v_1 = 0$. Therefore, $v \in U$, so $G(\lambda_m, T) \subset G(\lambda_m, T|_U)$, completing the proof.

A Basis of Generalized Eigenvectors

(Theorem 8.23) 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.

Theorem 8.23 Sketch of Proof Choose a basis of each $G(\lambda_j, T)$, putting all bases together yields the basis of V by (8.21).

7.2.2 Multiplicity of an Eigenvalue

Multiplicity Suppose $T \in \mathcal{L}(V)$. The ***multiplicity*** of an eigenvalue of T is defined to be the dimension of the corresponding generalized eigenspace $G(\lambda, T)$, equivalently $\dim \operatorname{null} (T - \lambda I)^{\dim V}$.

Sum of the Multiplicities Equals dim V

(Theorem 8.26) Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Then the sum of the multiplicities of all the eigenvalues of T equals $\dim V$.

Theorem 8.26 Sketch of Proof Directly follows from (8.21).

7.2.3 Block Diagonal Matrices

Block Diagonal Matrix A block diagonal matrix is a square matrix of the form

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

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

Block Diagonal Matrix with Upper-Triangular Blocks

(Theorem 8.29) Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \cdots, \lambda_m$ be the distinct eigenvalues of T, with multiplicities $d_1, \cdots, d_m$. Then there is a basis of V with respect to 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-by-d_j upper-triangular matrix of the form

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

Theorem 8.29 Sketch of Proof (1) Construct A_j: Each $(T - \lambda_j I)|_{G(\lambda_j, T)}$ is nilpotent, so $T|_{G(\lambda_j,T)} = (T - \lambda_j I)|_{G(\lambda_j,T)} - \lambda_j I|_{G(\lambda_j,T)}$ is in the desired form. (2) Construct block diagonal matrix with A's: Each $G(\lambda_j, T)$ is invariant. Putting all bases of $G(\lambda_j, T)$ together, we have the desired matrix.

7.2.4 Square Roots

Identity Plus Nilpotent has a Square Root

(Theorem 8.31) Suppose $N \in \mathcal{L}(V)$ is nilpotent. Then $I + N$ has a square root.

Theorem 8.31 Sketch of Proof Guessing their is a square root of the form $I + a_1 N + \cdots a_{m-1} N^{m-1}$ (based on Taylor series of $\sqrt{1+x}$). By $I + N = (I + a_1 N + \cdots a_{m-1} N^{m-1})^2$, we can solve for all coefficients and verify some choice of a_j's gives a square root.

Over $\mathbb{C}$ Invertible Operators Have Square Roots

(Theorem 8.33) Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ is invertible. Then T has a square root.

Theorem 8.33 Sketch of Proof (1) Construct R_j: $(T - \lambda_j)$ is nilpotent, so $\exists\, N_j : T|_{G(\lambda_j, T)} = N_j - \lambda_j I = \frac{1}{\lambda_j}(I + N_j/\lambda_j)$. Define R_j to be the product of $\sqrt{\lambda_j}$ and square root of $I + N_j/\lambda_j$ (guaranteed by 8.31). (2) Construct a square root R: Define $R = R_1 u_1 + \cdots + R_m u_m : u_j \in G(\lambda_j, T)$, perform R on both sides we obtain $R^2 = T$.

7.3 Characteristic and Minimal Polynomial

7.3.1 The Carley-Hamilton Theorem

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

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

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

Degree and Zeros of Characteristic Polynomial

(Theorem 8.36) 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 8.36 Sketch of Proof Directly follows from 8.26 and the definition.

Cayley-Hamilton Theorem

(Theorem 8.37) Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let q denote the characteristic polynomial of T. Then $q(T) = 0$.

Theorem 8.37 Sketch of Proof $(T - \lambda_j I)|_{G(\lambda_j, T)}$ is nilpotent (8.21). Given the multiplicity $d_j = \dim G(\lambda_j, T)$, $(T - \lambda_j I)^{d_j}|_{G(\lambda_j, T)} = 0$, for all j. Decomposing V to the direct sum of generalized eigenspaces (8.21), $q(T)v = 0$, completing the proof.

7.3.2 The Minimal Polynomial

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

Minimal Polynomial

(Theorem 8.40) Suppose $T \in \mathcal{L}(V)$. Then there is a unique monic polynomial p of smallest degree such that $p(T) = 0$.

Theorem 8.40 Sketch of Proof $I, T, \cdots, T^{n^2}$ is linearly dependent, so there exists smallest m s.t. $I, \cdots, T^m$ is linearly dependent. Then we can construct $p : p(T) := c_0 I + \cdots + c_{m-1}T^{m-1} + T^m = 0$. Verifying p is monic, smallest degree, and unique will complete the proof.

Minimal Polynomial Suppose $T \in \mathcal{L}(V)$. Then the minimal polynomial of T is the unique monic polynomial p of smallest degree such that $p(T) = 0$.

q(T) = 0 Implies q is a Multiple of the Minimal Polynomial

(Theorem 8.46) Suppose $T \in \mathcal{L}(V)$ and $q \in \mathcal{P}(\mathbf{F})$. Then $q(T) = 0$ if and only if q is a polynomial multiple of the minimal polynomial of T.

Theorem 8.46 Sketch of Proof $\Leftarrow$ direction is obvious. We can prove $\Rightarrow$ direction by division algorithm.

Characteristic Polynomial is a Multiple of Minimal Polynomial

(Theorem 8.48) Suppose $\mathbf{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then the characteristic polynomial of T is a polynomial multiple of the minimal polynomial of T.

Theorem 8.48 Sketch of Proof Directly follows from 8.46 and Cayley-Hamilton Theorem (8.37).

Eigenvalues are the Zeros of the Minimal Polynomial

(Theorem 8.49) Let $T \in \mathcal{L}(V)$. Then the zeros of the minimal polynomial of T are precisely the eigenvalues of T.

Theorem 8.49 Sketch of Proof $\Rightarrow$ direction can be proven by $p(T)v = (T - \lambda I)(q(T)v) = 0$ given $p(z) = (z - \lambda)q(z)$. To prove $\Leftarrow$ direction, we can show $T^j v = \lambda^j v$ if v is an eigenvector correspond to λ. Therefore, $0 = p(T)v = p(\lambda)v$, implying $p(\lambda) = 0$.

7.4 Jordan Form

Degree and Zeros of Characteristic Polynomial

(Theorem 8.36) Suppose $N \in \mathcal{L}(V)$ is nilpotent. Then there exist vectors $v_1, \cdots, v_n \in V$ and nonnegative integers $m_1, \cdots, m_n$ such that

(a) $N^{m_1}v_1, \cdots, Nv_1, v_1, \cdots, N^{m_n}v_n, \cdots, Nv_n, v_n$ is a basis of V;

(b) $N^{m_1+1}v_1 = \cdots = N^{m_n+1}v_n = 0$.

Theorem 8.55 Sketch of Proof Mathematical Induction. The result holds for $\dim V = 1$. For $k = \dim N > 1$,

(a) We can construct a basis of range N (by induction hypothesis),

$$\mathfrak{B}_1 = N^{m_1}v_1, \cdots, Nv_1, v_1, \cdots, N^{m_n}v_n$$

(b) Given all element in $\mathfrak{B}_1 \in$ range N, we can extend $\mathfrak{B}_1$ to

$$\mathfrak{B}_2 := N^{m_1+1}u_1, \cdots, Nu_1, u_1, \cdots, N^{m_n+1}u_n, \cdots, Nu_n, u_n$$

where each u_j satisfy $Tu_j = v_j$. We then can prove $\mathfrak{B}_2$ is linearly independent.

(c) Then, we further extend $\mathfrak{B}_2$ to a basis of V, $\mathfrak{B}_3 := \mathfrak{B}_2, w_1, \cdots, w_p$. Clearly $Nw_j \in$ range $N =$ span$(\mathfrak{B}_1)$, so $\exists\ x_j \in$ span$(\mathfrak{B}_2) : Nw_j = Nx_j$. We can construct $u_{n+j} = w_j - x_j$, and note $Nu_{n+j} = N(w_j - x_j) = 0$. Hence,

$$\mathfrak{B}_4 = N^{m_1+1}u_1, \cdots, Nu_1, u_1, \cdots, N^{m_n+1}u_n, \cdots, Nu_n, u_n, u_{n+1}, \cdots, u_{n+p}$$

$\mathfrak{B}_4$ can be expressed as the linear combination of $\mathfrak{B}_2$, thus the list spans V and is a basis of V.

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

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

where each A_j is an upper-triangular matrix of the form

$$\begin{pmatrix} \lambda_1 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_j \end{pmatrix}$$

Jordan Form

(Theorem 8.60) 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.

Theorem 8.60 Sketch of Proof For each j, $\exists\ N_j = (T - \lambda_j I)|_{G(\lambda_j, T)} : T|_{G(\lambda_j,T)} = N_j|_{G(\lambda_j,T)} + \lambda_j I$ and N is nilpotent (8.21). Nilpotent operator N on $v_k, Nv_k, \cdots, N^{m_k}v_k$ (as described in 8.55) has the form

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

for all k. Putting all lists together, the matrix of $T|_{G(\lambda_j,T)}$ has the desired form (A_j). Putting bases of each V_j together, by $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_p, T)$, $M(T)$ is in the desired form.

8 Inner Product Spaces

8.1 Multilinear Forms

Bilinear Form A ***bilinear form*** on V is a function $B : V \times V \to F$ which is linear in each variable separately, that is, if one variable is fixed, the functions is linear.

Note that we are not viewing $V \times V$ as a vector space, because a bilinear form is not linear on the vector space $V \times V$.

A bilinear form on $\mathbf{F}^n$ has the form

$$B(v, w) = v^T A w$$

for a unique $n \times n$ matrix A, where v, w are column vectors and v^T denotes the transpose of v (the corresponding row vector).

Multilinear Form Let m be a positive integer. A ***m-linear form*** is a function $B : V \times \cdots \times V \to F$ which is separately linear in each variable. The set of m-linear forms is denoted as $V' \otimes \cdots \otimes V'$, or simply $V'^{\otimes^m}$.

The sum and scalar products as for linear forms are defined by:

$$(B_1 + B_2)v = B_1 v + B_2 v$$
$$(cB)v = cBv$$

where v denotes a m-tuple.

Symmetric Multilinear Form A ***symmetric bilinear form*** on a vector space V is a bilinear form $B : V \times V \to F$ which has the property that $B(v, w) = B(w, v)$ for all $v, w \in V$. More generally, a ***symmetric m-linear form*** is one which is invariant under all permutations of its argument, and the set of symmetric m-linear forms on V is denoted as $S^m V'$.

Quadratic Forms Lemma

The function $q_B : V \to F$, defined by $q_B(v) = B(v, v)$ where $B \in S^2 V'$, has the following properties:

1. $q_B(cv) = c^2 q_B(v)$ for all $c \in F, v \in V$.
2. The function $B_q = \frac{1}{2}(q_B(v + w) - q_B(v) - q_B(w))$ is a symmetric bilinear form; in fact, it is the bilinear form B.

A function $q : V \to F$ satisfying the above lemma is called a ***quadratic form***.

8.2 Inner Product and Norms

8.2.1 Inner Products

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

$$x \cdot y = x_1 y_1 + \cdots + x_n y_n c$$

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

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

1. Positivity: $\langle u, v\rangle \geq 0$ for all $v \in \mathbf{F}$;
2. Definiteness: $\langle u, v\rangle = 0$ if and only if $v = 0$;
3. Additivity in First Slot: $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$;
4. Homogeneity in Fist Slot: $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$ for all $\lambda \in \mathbf{F}$ and all $u, v \in V$;
5. Conjugate Symmetry: $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$.

The ***Euclidean Inner Product*** on F^n is defined by $\langle (w_1, \cdots, w_n), (z_1, \cdots, z_n)\rangle = w_1\overline{z_1} + \cdots + w_n\overline{z_n}$.

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

Basic Properties of an Inner Product

(Theorem 6.7)

(a) For each fixed $u \in V$, the function that takes v to $\langle v, u\rangle$ is a linear map from V to $\mathbf{F}$.

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

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

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

(e) $\langle u, \lambda v\rangle = \bar{\lambda}\langle u, v\rangle$ for all $\lambda \in \mathbf{F}$ and $u, v \in V$.

8.2.2 Norms

Norm, $\|\mathbf{v}\|$ For $v \in V$, the ***norm*** of v, denoted $\|v\|$, is defined by

$$\|v\| = \sqrt{\langle v, v, \rangle}$$

Basic Properties of the Norm

(Theorem 6.10) Suppose $V \in V$,

1. $\|v\| = 0$ if and only if $v = 0$.
2. $\|\lambda v\| = |\lambda|\|v\|$ for all $\lambda \in \mathbf{F}$.

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

Orthogonality and 0

(Theorem 6.12)

(a) 0 is orthogonal to every vector in V.

(b) 0 is the only vector in V that is orthogonal to itself.

Pythagorean Theorem

(Theorem 6.13) Suppose u and v are orthogonal vectors in V. Then

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

An Orthogonal Decomposition

(Theorem 6.14) Suppose $u, v \in V$, with $v \neq 0$. Set

$$c = \frac{\langle u, v\rangle}{\|v\|^2} \quad \text{and} \quad w = u - \frac{\langle u, v\rangle}{\|v\|^2}$$

Then $\langle w, v\rangle = 0 \quad$ and $\quad u = cv + w$.

Cauchy-Schwarz Inequality

(Theorem 6.15) Suppose $u, v \in V$. Then

$$|\langle u, v\rangle| \leq \|u\|\|v\|$$

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

Triangle Inequality

(Theorem 6.18) Suppose $u, v \in V$. Then

$$|\langle u + v\rangle| \leq \|u\| + \|v\|$$

This inequality is an equality if and only if one of u, v is a nonnegative multiple of the other.

Parallelogram Equality

(Theorem 6.22) Suppose $u, v \in V$. Then

$$|\langle u+v\rangle^2 + |\langle u-v\rangle| = 2(\|u\|^2 + \|v\|^2)$$

8.3 Orthonormal Bases

8.3.1 Orthonormal Bases

Orthonormal 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.

In other words, a list $e_1, \cdots, e_m$ of vectors in V is orthonormal if

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

The Norm of an Orthonormal Linear Combination

(Theorem 6.25) If $e_1, \cdots, e_m$ is an orthonormal list of vectors in V, then

$$\|a_1e_1 + \cdots + a_me_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

fo all $a_1, \cdots, a_m \in \mathbf{F}$.

An Orthonormal List is Linearly Independent

(Theorem 6.26) Every orthonormal list of vectors is linearly independent.

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

An Orthonormal List of the Right Length is an Orthonormal Basis

(Theorem 6.28) Every orthonormal list of vectors in V with length $\dim V$ is an orthonormal basis of V.

Writing a Vector as linear combination of orthonormal basis

(Theorem 6.30) Suppose $e_1, \cdots, e_n$ 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

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

Gram-Schmidt Procedure

(Theorem 6.31) Suppose $v_1, \cdots, v_m$ is a linearly independent list of vectors in V. Let $e_1 = v_1/\|v\|$. For $j = 2, \cdots, m$, define e_j inductively by

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

Then $e_1, \cdots, e_m$ is an orthonormal list of vectors in V such that

$$\operatorname{span}(v_1, \cdots, v_j) = \operatorname{span}(e_1, \cdots, e_j)$$

for $j = 1, \cdots, m$.

Existence of Orthonormal Basis

(Theorem 6.34) Every finite-dimensional inner product space has an orthonormal basis.

Orthonormal List Extends to Orthonormal Basis

(Theorem 6.35) Suppose V is finite-dimensional. Then every orthonormal list of vectors in V can be extended to an orthonormal basis of V.

Upper-triangular Matrix With Respect to Orthonormal Basis

(Theorem 6.37) Suppose $T \in \mathcal{L}(V)$. If T has an upper-triangular matrix with respect to some basis of V, then T has an upper-triangular matrix with respect to some orthonormal basis of V.

Schur's Theorem

(Theorem 6.38) Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some orthonormal basis of V.

8.3.2 Linear Functional on Inner Product Spaces

Riesz Representation Theorem

(Theorem 6.42) Suppose V is finite-dimensional and φ is a linear functional on V. Then there is a unique vector $u \in V$ such that

$$\varphi(v) = \langle v, u\rangle$$

for every $v \in V$.

The vector u in above theorem is given by

$$u = \overline{\varphi(e_1)}e_1 + \cdots + \overline{\varphi(e_n)}e_n$$

8.4 Orthogonal Complements and Minimization Problems

8.4.1 Orthogonal Complements

Orthogonal Complement, $\mathbf{U}^{\perp}$ 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:

$$U^{\perp} = \{v \in V : \langle v, u \rangle = 0 \text{ for every } u \in U\}$$

Basic Properties of Orthogonal Complement

(Theorem 6.46)

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

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

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

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

(e) If U and W are subsets of V and $U \subset W$, then $W^{\perp} \subset U^{\perp}$.

Direct Sum of a Subspace and Its Orthogonal Complement

(Theorem 6.47) Suppose U is a finite-dimensional subspace of V. Then

$$V = U \oplus U^{\perp}$$

Dimension of the Orthogonal Complement

(Theorem 6.50) Suppose V is finite-dimensional and U is a subspace of V. Then

$$\dim U^{\perp} = \dim V - \dim U$$

The Orthogonal Complement of the Orthogonal Complement

(Theorem 6.51) Suppose U is a finite-dimensional subspace of V. Then

$$U = (U^{\perp})^{\perp}$$

8.4.2 Orthogonal Projection

Orthogonal Projection, $\mathbf{P_U}$ Suppose U is a finite-dimensional subspace of V. The orthogonal projection of V onto U is the operator $P_U v = u$ defined as follows: For $v \in V$, write $v = u + w$, where $u \in U$ and $w \in U$. Then $P_U v = u$.

Properties of the Orthogonal Projection $\mathbf{P_U}$

(Theorem 6.55) Suppose U is a finite-dimensional subspace of V and $v \in V$. Then

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

(b) $P_U u = u$ for every $u \in U$;

(c) $P_U w = 0$ for every $w \in U^\perp$;

(d) range $P_U = U$;

(e) null $P_U = U^\perp$;

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

8.4.3 Minimization Problems

Minimizing the Distance to a Subspace

(Theorem 6.56) 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 above is an equality if and only if $u = P_U v$.

In other words, $P_U v$ is the closest point in U to v.

9 Operators on Inner Product Spaces

9.1 Self-Adjoint and Normal Operator

9.1.1 Adjoints

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

$$\langle Tv, w \rangle = \langle v, T^* w \rangle$$

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

The adjoint is a linear map

(Theorem 7.5) If $T \in \mathcal{L}(V, W)$, then $T^* \in \mathcal{L}(V, W)$.

Properties of the adjoint

(Theorem 7.6)

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

(b) $(\lambda T)^* = \bar{\lambda} T^*$ for all $\lambda \in \mathbf{F}$ and $T \in \mathcal{L}(V, W)$;

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

(d) $I^* = I$ where I is the identity operator on V;

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

Null space and range of T^*

(Theorem 7.7) Suppose $T \in \mathcal{L}(V, W)$. Then

(a) null $T^* = (\text{range } T)^\perp$;

(b) range $T^* = (\text{null } T)^\perp$;

(c) null $T = (\text{range } T^*)^\perp$;

(d) range $T = (\text{null } T^*)^\perp$;

Theorem 7.7 Sketch of Proof We can first prove (a): $w \in \text{null } T^* \Leftrightarrow \langle Tv, w \rangle = \langle v, T^* w \rangle = 0 \Leftrightarrow w \in (\text{range } T)^\perp$. By replacing T by T^*, and/or take orthogonal complement, we obtain the other statements.

The matrix of T^*

(Theorem 7.10) Let $T \in \mathcal{L}(V, W)$. Suppose $e_1, \cdots, e_n$ is an orthonormal basis of V and $f_1, \cdots, f_m$ is an orthonormal basis of W. Then $\mathcal{M}(T^*, (f_1, \cdots, f_m), (e_1, \cdots, e_n))$ is the conjugate transpose of $\mathcal{M}(T, (e_1, \cdots, e_n), (f_1, \cdots, f_m))$.

Theorem 7.10 Sketch of Proof Based on 6.30, $Te_k = \langle Te_k, f_1 \rangle f_1 + \cdots + \langle Te_k, f_m \rangle f_m \Rightarrow \mathcal{M}(T)_{j,k} = \langle Te_k, f_j \rangle$. Similarly, $\mathcal{M}(T^*)_{k,j} = \langle T_j^f, e_k \rangle$. Therefore, $\mathcal{M}(T) = \overline{\mathcal{M}(T^*)}^t$.

9.1.2 Self-Adjoint Operators

Self-adjoint An operator $T \in \mathcal{L}(V)$ is called ***self-adjoint*** if $T = T^*$. In other words, $T \in \mathcal{L}(V)$ is self-adjoint if and only if $\langle Tv, w\rangle = \langle v, Tw\rangle$ for all $v, w \in V$.

Eigenvalues of self-adjoint operators are real

(Theorem 7.13) Every eigenvalues of a self-adjoint operator is real.

Over $\mathbb{C}$, Tv is orthogonal to v for all v only for the 0 operator

(Theorem 7.14) Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Suppose $\langle Tv, v\rangle = 0$ for all $v \in V$. Then $T = 0$.

Over $\mathbb{C}$, $\langle Tv, v\rangle$ is real for all v only for self-adjoint operator

(Theorem 7.15) Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Then T is self-adjoint if and only if $\langle Tv, v, \rangle \in \mathbb{R}$ for every $v \in V$.

If T is normal and $\langle Tv, v\rangle$ for all v then T=0

(Theorem 7.16) Suppose T is a self-adjoint operator on V such that $\langle Tv, v\rangle = 0$ for all $v \in V$. Then $T = 0$.

9.1.3 Normal Operators

Normal 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 $TT^* = T^*T$.

T is normal if and only if $\|Tv\|$=$\|T^*v\|$ are equal for all v

(Theorem 7.20) An operator $T \in \mathcal{L}(V)$ is normal if and only if

$$\|Tv\| = \|T^*v\|$$

for all $v \in V$.

For T normal, T and T^* have the same eigenvectors

(Theorem 7.21) 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 $\bar{\lambda}$.

Orthogonal eigenvectors for normal operators

(Theorem 7.22) Suppose $T \in \mathcal{L}(V)$ is normal. Then eigenvectors of T corresponding to distinct eigenvalues are orthogonal.

9.2 The Spectral Theorem

Complex Spectral Theorem

(Theorem 7.24) Suppose $\mathbf{F} = \mathbb{C}$ 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.

Real Spectral Theorem

(Theorem 7.29) Suppose $\mathbf{F} = \mathbb{R}$ 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.

Theorem 7.29 (a) ⇒ (b) Sketch of Proof Mathematical Induction. For $\dim V > 1$, assume $(a) \Rightarrow (b)$ for real spaces of smaller dimension. There exists an eigenvector $u : \|u\| = 1$. [(7.27) guarantees the existence of eigenvector for self-adjoint operator.] Then $U = \operatorname{span}(u)$ is invariant under T, and thus $T|_{U^\perp}$ is self-adjoint [(7.28)(c) states that T is self-adjoint and U is T-invariant implies $T|_{U^\perp}$ is self-adjoint.] By inductive hypothesis, there is an orthonormal basis $U^\perp$ consisting of eigenvectors. Adjoining u to the basis gives the desired basis.

9.3 Isometry

Isometry An operator $S \in \mathcal{L}(V)$ is called an ***isometry*** if $\|Sv\| = \|v\|$ for all $v \in V$. In other words, an operator is an isometry if it preserves norms.

Characterization of isometries

(Theorem 7.42) Suppose $S \in \mathcal{L}(V)$. Then the following are equivalent:

(a) S is an isometry;

(b) $\langle Su, Sv \rangle = \langle u, v \rangle$ for all $u, v \in V$;

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

(d) there exists an orthonormal basis $e_1, \cdots, e_n$ of V such that $Se_1, \cdots, Se_n$ is orthonormal;

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

(f) $SS^* = I$;

(g) S^* is an isometry;

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

Description of isometries when F=$\mathbb{C}$

(Theorem 7.43) 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.

10 Trace and Determinant

10.1 Trace

10.1.1 Change of Basis

Identity Matrix, I Suppose n is a positive integer. The n-by-n diagonal matrix

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

is called the ***identity matrix*** and is denoted I.

Invertible, Inverse, $\mathbf{A^{-1}}$ A square matrix A is called ***invertible*** if there is a square matrix B of the same size such that $AB = BA = I$; we call B the ***inverse*** of A and denote it by A^{-1}.

The matrix of the product of linear maps

(Theorem 10.4) Suppose $u_1, \cdots, u_n$ and $v_1, \cdots, v_n$ and $w_1, \cdots, w_n$ are all bases of V. Suppose $S, T \in \mathcal{L}(V)$. Then

$$\begin{aligned} &\mathcal{M}(ST,(u_1, \cdots, u_n), (w_1, \cdots, w_n)) \\ &\qquad = \mathcal{M}(S, (v_1, \cdots, v_n), (w_1, \cdots, w_n))\mathcal{M}(T, (u_1, \cdots, u_n), (v_1, \cdots, v_n)) \end{aligned}$$

Matrix of the identity with respect to two bases

(Theorem 10.5) Suppose $u_1, \cdots, u_n$ and $v_1, \cdots, v_n$ are bases of V. Then the matrices $\mathcal{M}(I, (u_1, \cdots, u_n), (v_1, \cdots, v_n))$ and $\mathcal{M}(I, (v_1, \cdots, v_n), (u_1, \cdots, u_n))$ are invertible, and each is the inverse of the other.

Change of basis formula

(Theorem 10.7) Suppose $T \in \mathcal{L}(V)$. Let $u_1, \cdots, u_n$ and $v_1, \cdots, v_n$ be bases of V. Let $A = \mathcal{M}(I, (u_1, \cdots, u_n), (v_1, \cdots, v_n))$. Then

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

10.1.2 Trace

Trace of an Operator Suppose $T \in \mathcal{L}(V)$. The *trace* of T is the sum of the eigenvalues of T (or $T_{\mathbb{C}}$ if in real vector space) with each eigenvalue repeated according to its multiplicity. The trace of T is denoted trace T.

Trace and characteristic polynomial

(Theorem 10.12) Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then trace T equals the negative of the coefficient of z^{n-1} in the characteristic polynomial of T.

Trace of a Matrix The trace of a square matrix A, denoted trace A, is defined to be the sum of the diagonal entries of A.

Trace of AB equals trace of BA

(Theorem 10.14) If A and B are square matrices of the same size, then

$$\operatorname{trace}(AB) = \operatorname{trace}(BA)$$

Trace of matrix of operator dies not depend on basis

(Theorem 10.15) Let $T \in \mathcal{L}(V)$. Suppose $u_1, \cdots, u_n$ and $v_1, \cdots, v_n$ are bases of V. Then

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

Trace of an operator equals trace of its matrix

(Theorem 10.16) Suppose $T \in \mathcal{L}(V)$. Then trace $T = \operatorname{trace} \mathcal{M}(T)$.

Trace is additive

(Theorem 10.18) Suppose $S, T \in \mathcal{L}(V)$. Then $\operatorname{trace}(S + T) = \operatorname{trace} S + \operatorname{trace} T$.

The identity is not the difference of ST and TS

(Theorem 10.19) There do not exist operators $S, T \in \mathcal{L}(V)$ such that $ST - TS = I$.

10.2 Determinant

10.2.1 Determinant of an Operator

Determinant of an Operator, det T Suppose $T \in \mathcal{L}(V)$. The ***determinant*** of T is the product of the eigenvalues of T (or $T_{\mathbb{C}}$ if in real vector space), with each eigenvalue repeated according to its multiplicity. The determinant of T is denoted by det T.

Determinant and characteristic polynomial

(Theorem 10.22) 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.

Invertible is equivalent to nonzero determinant

(Theorem 10.24) An operator on V is invertible if and only if its determinant is nonzero.

Characteristic polynomial of T equals $\det(zI - T)$

(Theorem 10.25) Suppose $T \in \mathcal{L}(V)$. Then the characteristic polynomial of T equals $\det(zI - T)$.

10.2.2 Determinant of a Matrix

Permutation, perm n, $\mathbf{S_n}$ A ***permutation*** of $(1, \cdots, n)$ is a list $(m_1, \cdots, m_n)$ that contains each numbers $1, \cdots, n$ once. The set of all permutations of $(1, \cdots, n)$ is denoted perm n or S_n.

Sign of Permutation The sign of a permutation $(m_1, \cdots, m_n)$ is defined to be 1 if the number of pairs of integers (j, k) with $1 \leq j < k \leq n$ such that j appears after k in the list $(m_1, \cdots, m_n)$ is even and 1 if the number of such pairs is odd.

In other words, the sign of a permutation equals 1 if the natural order has been changed an even number of times and equals -1 if the natural order has been changed an odd number of times.

Interchanging two entries in a permutation

(Theorem 10.32) Interchanging two entries in a permutation multiplies the sign of the permutation by -1.

Determinant of a Matrix, det A Suppose A is an n-by-n matrix

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

The ***determinant*** of A, denoted $\det A$, is defined by

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

Interchanging two columns in a matrix

(Theorem 10.36) Suppose A is a square matrix and B is the matrix obtained from A by interchanging two columns. Then

$$\det A = -\det B$$

Matrices with two equal columns

(Theorem 10.37) If A is a square matrix that has two equal columns, then $\det\ A = 0$.

Permuting the columns of a matrix

(Theorem 10.38) Suppose $A = (A_{\cdot,1}, \cdots, A_{\cdot,n})$ is an n-by-n matrix and $(m_1, \cdots, m_n)$ is a permutation. Then

$$\det(A_{\cdot,m_1}, \cdots, A_{\cdot,m_n}) = (\text{sign}(m_1, \cdots, m_n)) \det\ A$$

Determinant is a linear function of each column

(Theorem 10.39) Suppose k, n are positive integers with $1 \leq k \leq n$. Fix n-by-1 matrices $A_{\cdot,1}, \cdots, A_{\cdot,n}$ except $A_{\cdot,k}$. Then the function that takes an n-by-1 column vector $A_{\cdot,k}$ to

$$\det(A_{\cdot,1}, \cdots, A_{\cdot,k}, \cdots, A_{\cdot,n})$$

is a linear map from vector space of n-by-1 matrices with entries $\mathbf{F}$ to $\mathbf{F}$.

Determinant is multiplicative

(Theorem 10.40) Suppose A and B are square matrices of the same size. Then

$$\det(AB) = (\det\ A)(\det\ B)$$

Determinant of matrix of operator does not depend on basis

(Theorem 10.41) Let $T \in \mathcal{L}(V)$. Suppose $u_1, \cdots, u_n$ and $v_1, \cdots, v_n$ are bases of V. Then

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

Determinant of an operator equals determinant of its matrix

(Theorem 10.42) Suppose $T \in \mathcal{L}(V)$. Then $\det A = \det \mathcal{M}(A)$.

Determinant is multiplicative

(Theorem 10.44) Suppose $S, T \in \mathcal{L}(V)$. Then

$$\det(ST) = \det(TS) = (\det S)(\det T)$$