Math 110 (Linear Algebra Done Right) Notes

Henry Yan

July 2021

Notes

Algorithmic proofs only works for finite objects.

Complex Numbers

Let $\lambda = a + bi = r(\cos\theta + i\sin\theta) = re^{i\theta} \in \mathbb{C}$. Then $|\lambda|^2 = \lambda\bar{\lambda}$.

Notation

F will stand for $\mathbb{R}$ or $\mathbb{C}$. But in most cases, F can be any field, so for almost any definition, theorem, or proof presented in the book to be true in "**F**", it will likely be true for any field.

F^S denotes the set of functions from S to F.

1 Vector Spaces

$\mathbb{R}^2 = \{(x, y)|x, y \in \mathbb{R}\}$ is a vector space, in particular, it is a plane of real numbers. $\mathbb{R}^3$ is another example of a vector space.

Given some non-negative integer n, a **list** of length n (often called an n-tuple) is an ordered collection of, basically, anything separated by commas and surrounded by parenthesis. Note that a list must be of finite length, and that a list of length 0 will be ().

Given some list $(x_1, x_2, \ldots, x_n)$, for $j \in \{1, 2, \ldots n\}$, we say x_j is the jth coordinate of the list above.

Lists can differ in 2 ways, the length of the list or the content of the coordinates. Note that a list differs from a set as order and repetitions matter for a list but don't for a set.

We often let 0 denote the list of length n for which all the coordinates are 0, where n can be determined from context.

A **vector** can be thought of as an arrow from the origin to some point. Vector addition, geometrically, can be understood as placing the end of one arrow at the tip of another (generalizes very intuitively to higher dimensions - more additions), the sum of the two vectors is called the **resultant vector**.

A set V together with the operations $+ : V \times V \to V$ and $\cdot : F \times V \to V$ is a **vector space** if the following conditions are met: [1]

---

[1]Note that by definition, V is closed under addition and scalar multiplication.

1. Commutativity - $u + v = v + u, \forall u, v \in V$;

2. Associativity - $(u + v) + w = u + (v + w), (ab)v = a(bv) \forall u, v, w \in V, a, b \in F$;

3. Additive Identity - $0 \in V$ such that $v + 0 = v, \forall v \in V$;

4. Additive Inverse - For $v \in V, \exists w \in V$ such that $v + w = 0$;

5. Multiplicative Identity - $\exists 1 \in F$, such that $1v = v, \forall v \in V$;

6. Distributive Properties - $a(u + v) = au + av, (a + b)u = au + bu, \forall a, b \in F, u, v \in V$.

More precisely, when we need to specify a set for where the constants a, b come from in the above definition, we say that V is a vector space over F, or an F-space, (or whatever the set a, b are in is). A result of the above definition is that a vector space forms an Abelian group under addition, though with more structure as it also has conditions regarding scalar multiplication.

A vector space over $\mathbb{R}$ is often called a **real vector space**, and a vector space over $\mathbb{C}$ is often called a **complex vector space**.

F^n is vector space over F. The simplest vector space contains only 1 point, 0; in other words, $\{0\}$ is a vector space. The largest vector space is F^∞ over F.

Definition 1.1. A function $p : F \to F$ is a **polynomial** with coefficients in F if $\exists a_0, \ldots a_n \in F$ such that $p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0, \forall z \in F$.

We define $P(F)$ to be the set of all polynomials with coefficients in F. Addition in $P(F)$ is defined as you would expect, $(p + q)(z) = p(z) + q(z), z \in F$. Scalar multiplication on $P(F)$ is defined as $(ap)(z) = ap(z)$, for $a \in F, p \in P(F)$. It follows that $P(F)$ is a vector space, which comes to show that a vector space doesn't need to consist of lists, it can also take the form of other algebraic objects, such as polynomials.

As a result of vector spaces being a group, many group theoretic properties carry over, such as the uniqueness of the additive identity and additive inverse.

*NOTE: From here onward, it will be understood that V will denote a vector space over F.

A subset U of V is called a **subspace**, sometimes called a **linear subspace**, of V if U is also a vector space. For example, $\{(x_1, x_2, 0)|x_1, x_2 \in F\}$ is a subspace of F^3. Conditions for a subspace are as follows:

1. Additive Identity: $0 \in U$;

2. Closed Under Addition: $u, v \in U \implies u + v \in U$;

3. Closed Under Scalar Multiplication: $a \in F, u \in U \implies au \in U$

For any vector space V, the smallest subspace of V is $\{0\}$, empty set doesn't work since the subspace needs to contain the additive identity, and the largest subspace of V is V itself.

The only subspaces of $\mathbb{R}^2$ are $\{0\}, \mathbb{R}^2$, and each of the lines in $\mathbb{R}^2$ through the origin. Similarly, the only subspaces of $\mathbb{R}^3$ are $\{0\}, \mathbb{R}^3$, each of the lines in $\mathbb{R}^3$ through the origin, and each of the planes in $\mathbb{R}^3$ through the origin.

Let $U_1, \ldots, U_m$ be subspaces of V. We define the **sum** of $U_1, \ldots, U_m$ to be $U_1 + \cdots + U_m = \{u_1 + \cdots + u_m | u_1 \in U_1, \ldots, u_m \in U_m\}$. $U_1 + \cdots + U_m$ is a subspace of V.
In the case that $V = U_1 + \cdots + U_m$ and every vector in V can be uniquely represented by $u_1 + \cdots + u_m, u_i \in U_i$, we say that V is the **direct sum** of $U_1, \ldots, U_m$, using the notation of $U_1 \oplus \cdots \oplus U_m$.

Example 1.1. Let $U = \{(x, y, 0) \in F^3 | x, y \in F\}, W = \{(0, 0, z) \in F^3 | z \in F\}$. Then $F^3 = U \oplus W$.

Proposition 1.8: Suppose that $U_1, \ldots, U_n$ are subspaces of V. Then $V = U_1 \oplus \cdots \oplus U_n$ if and only if both the following conditions hold:

1. $V = U_1 + \cdots + U_n$;

2. The only way to write 0 as a sum of $u_1 + \cdots + u_n, u_i \in U_i$, is by taking $u_i = 0, \forall i$.

Proposition 1.9: Suppose that U and W are subspaces of V. Then $V = U \oplus W$ if and only if $V = U + W$ and $U \cap W = \{0\}$.

2 Finite-Dimensional Vector Spaces

A **linear combination** of a list $(v_1, \ldots, v_m)$ of vectors in V is a finite-sum (resulting vector) of the form $a_1 v_1 + \cdots + a_m v_m, a_i \in F$. By definition, a list S of vectors in a vector space is closed under linear combination.

The set of all linear combinations of is called the **span** of $(v_1, \ldots, v_m)$, which we denote by $\mathrm{span}(v_1, \ldots, v_m) = \{a_1 v_1 + \cdots + a_m v_m | a_1, \ldots, a_m \in F\}$.

It is easy to show that the span of any list of vectors in V is a subspace of V. We consider that the span of the empty list () equals $\{0\}$, since the empty list is not a subspace of V.

Any vector in a vector space can be written as a linear combination of the vectors in that vector space. Thus, $\mathrm{span}(v_1, \ldots, v_m)$ contains $v_1, \ldots, v_m$. Conversely, because vector spaces are closed under scalar multiplication, V contains $\mathrm{span}(v_1, \ldots, v_m)$. It follows that $\mathrm{span}(v_1, \ldots, v_m)$ is the smallest subspace of V containing each of $v_1, \ldots, v_m$.

If $\mathrm{span}(v_1, \ldots, v_m) = V$, we say that $(v_1, \ldots, v_m)$ **spans** V.

A vector space is **finite-dimensional** if some list (by definition list implies finite) of vectors in it spans the space. Another way of defining this would be to say that a vector space is finite-dimensional if it is spanned by finitely many vectors. For example, $F^n, n \in \mathbb{N}$ is finite-dimensional because

$$((1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1))$$

spans F^n.

By convention, the zero polynomial is said to have degree $-\infty$.

We define $P_m(F)$ to be the set of all polynomials with degree at most m, with coefficients in F. Clearly, $P_m(F)$ is a subspace of $P(F)$. $P_m(F)$ is spanned by $(1, z, \ldots, z^m)$.

A vector space that is not finite-dimensional is **infinite-dimensional**. An example of an infinite-dimensional vector space is $P(F)$. This is easy to see by an extremal principle proof; consider any list of elements of $P(F)$, then consider the highest degree, m, polynomial in the list of elements. Then the highest degree polynomial that we can make as a linear combination of elements of the list is m. Clearly, this list cannot span $P(F)$. Thus, $P(F)$ is infinite-dimensional.

A list $(v_1, \ldots, v_m)$ of vectors in V is called **linearly independent** if the only choice of $a_1, \ldots, a_m \in F$ that makes $a_1 v_1 + \cdots + a_m v_m = 0$ is $a_i = 0, 1 \le i \le m$.

Note that a direct consequence of a list of vectors in V being linearly independent is that every vector $v \in \mathrm{span}(v_1, \ldots, v_m)$ can be written uniquely as a linear combinations of vectors in $(v_1, \ldots, v_m)$.

Proof. Assume for the sake of contradiction that $v = a_1 v_1 + \cdots + a_m v_m = a_1' v_1 + \cdots + a_m' v_m$. Subtracting, we get that $0 = (a_1 - a_1')v_1 + \cdots + (a_m - a_m')v_m$. By definition of linearly independence, $a_i = a_i', 1 \le i \le m$. $\square$

Example 2.1. $(1, z, \ldots, z^m)$ is linearly independent in $P(F)$. This is easy to see because for

$$a_0 + a_1 z + \cdots + a_m z^m = 0 \tag{1}$$

where at least 1 of the a_i's aren't equal to 0, then (1) is satisfied for at most m values of z. Clearly this implies that $a_i = 0, 0 \leq i \leq m$. By definition, $(1, z, \ldots, z^m)$ is linearly independent.

A list of vectors is called **linearly dependent** if it is not linearly independent. The easiest way to prove that a list of vectors is linearly dependent is by that there are 2 or more ways of representing the zero vector as a linear combination.

We consider the empty list, (), to be linearly independent.

Example 2.2. $((2, 3, 1), (1, -1, 2), (7, 3, 8))$ is linearly dependent because
$2(2, 3, 1) + 3(1, -1, 2) + (-1)(7, 3, 8) = (0, 0, 0)$.

Example 2.3. Any list of vectors containing the zero vector is linearly dependent. [2]

Proposition: A subset $\{v_1, \ldots, v_m\} \subseteq V$ is linearly independent iff none of the vectors is a linear combination of the other.

Linear Dependence Lemma: If $(v_1, \ldots, v_m)$ is linearly dependent in V and $v_i \neq 0, 1 \leq i \leq m$, then there exists $j \in \{1, \ldots, m\} \setminus \{i\}$ such that the following hold:

1. $v_j \in \text{span}(v_1, \ldots, v_{j-1})$;

2. if the jth term is removed from $(v_1, \ldots, v_m)$, the span of the remaining list equals $\text{span}(v_1, \ldots, v_m)$.

The intuition for this lemma is that if you consider the farthest right vector (vector with the greatest index, j) with a non-zero coefficient, then $v_j \in \text{span}(v_1, \ldots, v_{j-1})$ as all the coefficients with indexes greater than j are 0.

$$0 = a_1 v_1 + \cdots + a_j v_j + a_{j+1} v_{j+1} \cdots + a_m v_m$$

$$\implies 0 = a_1 v_1 + \cdots + a_j v_j \implies v_j = -\frac{a_1}{a_j} v_1 - \cdots - \frac{a_{j-1}}{a_j} v_{j-1}$$

Theorem 2.6: 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.

Corollary 2.6.1: Theorem 2.6 implies that if the length of a list of vector V is greater than the length of a spanning list of vectors, then V must be linearly dependent.

Proposition 2.7: Every subspace of a finite-dimensional vector space is finite-dimensional.

A **basis** is a set of vectors $B \subseteq V$ that is both linearly independent and spans V.

Grammar: Plural of bases (pronounced bay-seas) is basis. By theorem 2.6, a bases can only be 1 size, where it is spanning, but still linearly independent. So a bases must be the largest linearly independent list, while being the smallest spanning list.

The **standard basis** of F^3 is $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$. This generalizes very intuitive to higher dimensions of F. Similarly, $\{1, z, \ldots, z^m\}$ is a basis of $P_m(F)$. However, $V = \{(1, 0, 0, \ldots), (0, 1, 0, \ldots), \ldots\}$ is not a bases of F^∞ as the vector $(1, 1, 1, \ldots)$ cannot be formed by a linear combination of vectors in V as a linear combination is, by definition, a finite sum. Yet, $\{1, z, z^2, \ldots\}$ is a basis of $P(F)$, as a polynomial, by definition 1.1, is a finite sum of monomials. Note that a consequence of this definition is that a power series is not a polynomial, as power series are infinite.

---

[2]Hint: Consider the coefficient in front of the zero vector.

Proposition 2.8: A list $B = \{v_1, \ldots, v_n\}$ is a basis of V iff every $v \in V$ can be uniquely written in the form $v = a_1 v_1 + \cdots + a_n v_n, a_i \in F$.

Definition 2.1. If $B = \{v_1, \ldots, v_n\}$ is a basis of V and $v \in V$, then $\exists$ scalars $a_1, \ldots, a_n \in F$ s.t.

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

This representation is unique and the scalars are called the **coordinates** of v with respect to the basis B.

Definition 2.1 gives us a coordinate system to define a vector space.

Theorem 2.10: Every spanning list in a vector space can be reduced to a basis of the vector space.

Corollary 2.10.1: Every finite-dimensional vector space has a basis.

Theorem 2.12: Every linearly independent list in a vector space can be expanded to a basis of the vector space.

Proposition 2.13: Suppose V is finite-dimensional and U is a subspace of V. Then there is a subspace W of V such that $V = U \oplus W$.

Theorem 2.14: Any two bases of a finite-dimensional vector space have the same length.

Proof. We use the result that the length of any spanning list is $\geq$ than the length of any linearly independent list.

Consider the bases B_1 and B_2 of V. By definition, a bases is both linearly independent and spanning, so $|B_1| \geq |B_2|$ and $|B_2| \geq |B_1| \implies |B_1| = |B_2|$ □

The **dimension** of a vector space is the number of vectors in a bases and is denoted $\dim V$.

Example 2.4. $\{1, z, z^2, \ldots, z^m\}$ is a basis of $P_z(F)$. So $\dim P_z(F) = m + 1$.

*NOTE. The dimension of a vector space depends on the field F. $\mathbb{C}$ is a 1 dimensional vector space over $\mathbb{C}$. But $\mathbb{C}$ is a 2 dimensional vector space over $\mathbb{R}$ (where the list of vectors is $(1, i)$, for example).

More intuitively, dimension should be thought of as the size of a vector space.

Proposition 2.15: If V is finite-dimensional and U is a subspace of V, then $\dim U \leq \dim V$.

Proposition 2.16: If V is finite-dimensional, then every spanning list of vectors in V with length $\dim V$ is a basis of V.

Proposition 2.17: If V is finite-dimensional, then every linearly independent list of vectors in V with length $\dim V$ is a basis of V.

Theorem 2.18: $\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$ (exact same form as Principle of Inclusion Exclusion)

Corollary 2.18.1: $\dim(U_1 \oplus U_2) = \dim U_1 + \dim U_2$

Proposition 2.19: Suppose V is finite-dimensional and $U_1, \ldots, U_m$ are subspaces of V such that

$$V = U_1 + \cdots + U_m$$

and

$$\dim V = \dim U_1 + \cdots + \dim U_m,$$

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

3 Linear Maps

When working with linear maps, we often replace the longhand of a function, $T(v)$, with Tv.

A linear map from vector spaces 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 - $T(av) = a(Tv), \forall a \in F, v \in V$.

The set of all linear maps from V to W, denoted by $\mathcal{L}(V, W)$, is a vector space. The following shorthand is also used: $\mathcal{L}(V) = \mathcal{L}(V, V)$.

Addition and Scalar Multiplication on $\mathcal{L}(V, W)$: For $S, T \in \mathcal{L}(V, W), v \in V, a \in F$

$$(S + T)(v) = Sv + Tv, (aT)(v) = a(Tv)$$

Zero function: $0 \in \mathcal{L}(V, W)$ is defined by:
$$0(v) = 0v = 0.$$

The 0 on the left side of the above equation is the function 0, but the 0 on the right side of the equation is the additive identity 0.

The **identity map**: Denoted by I, the identity map is the function on some vector space that takes each element to itself. $I \in \mathcal{L}(V, W)$ is defined by
$$Iv = v.$$

For $T \in \mathcal{L}(V, W)$,
$$TI = IT = T,$$

where the first I is the identity operator on V and the second I is the identity operator on W, as $(TI)v = T(Iv)$, so I is operating on $v \in V$.

Product of Linear Maps: If $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(U, V)$, then $ST \in \mathcal{L}(U, W)$, where ST is the standard function concatenation, being $(ST)(v) = S(Tv) = S \circ T(v)$.

3.4 (*linear map lemma*): Suppose $v_1, \ldots, v_n$ is a basis of V and $w_1, \ldots, w_n \in W$, then there exists a unique linear map $T : V \to W$ such that
$$Tv_k = w_k,$$

for each $k = 1, \ldots, n$.

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

In basic words, the null space of a function is the set of inputs that will produce an output of 0.

Proposition 3.1: If $T \in \mathcal{L}(V, W)$, then $\operatorname{null} T$ is a subspace of V.

Proof. Using additivity, $T(0) = T(0 + 0) = T(0) + T(0)$, which implies that $T(0) = 0, 0 \in \operatorname{null} T$.

For $u, v \in \operatorname{null} T, T(u + v) = T(u) + T(v) = 0 + 0 = 0$, hence $u + v \in \operatorname{null} T$ ($\operatorname{null} T$ is closed under addition).

If $a \in F$ and $u \in \operatorname{null} T, T(au) = a(Tu) = a(0) = 0$, hence $au \in \operatorname{null} T$, so T is closed under scalar

multiplication.

We have shown that T contains 0 and is closed under addition and scalar multiplication, so we have shown that null T is a subspace of V. $\qquad\square$

A linear map $T : V \to W$ is said to be **injective** if whenever $u, v \in V$ and $Tu = Tv, u = v$.

Proposition 3.2: Let $T \in \mathcal{L}(V, W)$, T is injective iff null $T = \{0\}$.

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

Proposition 3.3: If $T \in \mathcal{L}(V, W)$, then range T is a subspace of W.

A linear map $T : V \to W$ is called **surjective** if it's range equals W. In math terms, range $T = W$.

Theorem 3.4 (*fundamental theorem of linear maps*): If V is finite-dimensional and $T \in \mathcal{L}(V, W)$, then range T is a finite-dimensional subspace of W and
$$\dim V = \dim \text{null } T + \dim \text{range } T.$$

Corollary 3.5: If V and W are finite-dimensional vector spaces such that $\dim V > \dim W$, then no linear map from V to W is injective.

$\dim W \geq \dim \text{range } T$

Corollary 3.6: If 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 means that the constant term on the right side of each equation equals 0.

3.24: A homogeneous system of linear equations in which there are more variables than equations must have non-zero solutions.

3.26: An **inhomogeneous** system of linear equations in which there are more equations than variables has no solution for some choice of the constant terms.

Matrices of a Linear Map

An m by n **matrix** A is a rectangular array of elements of F with m rows and n columns. $A_{j,k}$ is the element in the jth row and kth column position of A.

When determining the matrix of a linear map, the image of each basis vector is placed in order of the basis vectors, left to right.

The **column and row rank** of a matrix A is the dimension of the span of the columns and rows of A, respectively. The rank of a matrix A is the column rank of A.

3.117 (dimension of range T equals column rank of $\mathcal{M}(T)$): For $T \in \mathcal{L}(V, W)$, then $\dim \text{range } T$ equals the column rank of $\mathcal{M}(T)$.

3.118: The column rank of A equals the row rank of A.

Definition: Matrix of a linear map, $\mathcal{M}(T)$

Suppose $T \in \mathcal{L}(V, W)$ and $v_1, \ldots, v_n$ is a basis of V and $w_1, \ldots, w_m$ 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 context, then we write $\mathcal{M}(T, (v_1, \ldots, v_n), (w_1, \ldots, w_m))$.

Another way to write this is

$$Tv_k = \sum_{j=1}^{m} A_{j,k}w_j$$

Notation: Let $F^{m,n}$ be the set of all m by n matrices with entries in F.

$\dim F^{m,n} = mn$, where $F^{m,n}$ is a vector space.

Consider linear maps $T : U \to V$ and $S : V \to W$, and let $v_1, \ldots, v_n$ be a basis of V, $w_1, \ldots, w_m$ be a basis of W, and $u_1, \ldots, u_p$ be a basis of U. Let $\mathcal{M}(S) = A$ and $\mathcal{M}(T) = C$. Then $\mathcal{M}(ST)$ is the m by p matrix whose entry in row j, column k, is

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

In simpler words, the entry in row j, column k, of AC is computed by taking row j of A, column k of C, and summing the product of the corresponding entries.

3.41 (The matrix of the product of linear maps): If $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$, then $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.

3.33 (matrix of the sum of linear maps): Suppose $S, T \in \mathcal{L}(V, W)$, then $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$

3.36 (the matrix of a scalar times a linear map): Suppose $\lambda \in F$ and $T \in \mathcal{L}(V, W)$, then $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

Notation: Suppose A is an m by n matrix.

1. $A_{j,\cdot}$ denotes the 1 by n matrix consisting of row j of A, for $j = 1, \ldots, m$.

2. $A_{\cdot,k}$ denotes the m by 1 matrix consisting of column k of A, for $k = 1, \ldots, n$.

3.44 (Entry of matrix product equals row times column): 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 $1 \le j \le m$ and $1 \le k \le p$.

3.45 (Column of matrix product equals matrix times column): Suppose A is an m by n matrix and C is an n by p matrix. Then

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

for $1 \le k \le p$.

Row of matrix product equals row times matrix: Suppose A is an m by n matrix and C is an n by p matrix. Then

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

for $1 \leq j \leq m$.

3.47: **Linear combination of columns**: Suppose A is an m by n matrix and $c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ is an n by 1 matrix. Then

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

Linear combination of rows: Suppose $a = (a_1 \ldots a_n)$ is a 1 by n matrix and c is an n by p matrix. Then

$$aC = a_1 C_{1,.} + \cdots + a_n C_{n,.}.$$

Invertibility and Isomorphic Vector Spaces

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 TS 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.

An invertible inverse map T has a unique inverse $T^{-1} = S$. If $T \in \mathcal{L}(V, W)$, then $T^{-1} \in \mathcal{L}(W, V)$.

*** A linear map is invertible iff it is bijective.

For linear maps $S \in \mathcal{L}(V, W)$ and $T = \mathcal{L}(W, V)$, where W, V are the same dimension, $ST = I \iff TS = I$.

Isomorphism

An **isomorphism** is an invertible linear map. An isomorphism can be thought of as a relabeling of basis vectors, and thus the vector space.

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

3.59: Two finite-dimensional vector spaces over F are isomorphic iff they have the same dimension.

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

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

Definition 3.62: matrix of a vector, $\mathcal{M}(u)$
Suppose $u \in V$ and $v_1, \ldots, v_n$ is a basis of V. The matrix of u with respect to this basis is the n-by-1 matrix

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

where $c_1, \ldots, c_n$ are scalars such that $u = c_1 v_1 + \cdots + c_n v_n$.

3.64: $\mathcal{M}(T)_{.,k} = \mathcal{M}(Tv_k)$.

Linear maps act like matrix multiplication: Suppose $T \in \mathcal{L}(V, W)$ and $u \in V$. Suppose $v_1, \ldots, v_n$ is a basis of V and $w_1, \ldots, w_m$ is a basis of W, then

$$\mathcal{M}(Tu) = \mathcal{M}(T)\mathcal{M}(u).$$

$\mathcal{L}(V)$ is called an **operator**.

Injectivity is equivalent to surjectivity in finite dimensions: 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.

3D Problem 10: Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Then ST is invertible iff S and T are invertible.

Dual Space and Annihilators

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

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

Theorem: $\dim V' = \dim V$

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 in V', where each φ_i is the linear functional on V such that

$$\varphi_i(v_k) = \begin{cases} 1 \text{ if } k = i \\ 0 \text{ if } k \neq i \end{cases}$$

3.94: Suppose $v_1, \ldots, v_n$ is a basis of V and $\varphi_1, \ldots, \varphi_n$ is the corresponding dual basis, then

$$v = \varphi_1(v)v_1 + \cdots + \varphi_n(v)v_n$$

for every $v \in V$.

3.96: The dual basis is a basis of the dual space. In other words, the dual basis of a specific 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 all $\varphi \in W'$.

Properties of dual maps:

1. $(S + T)' = S' + T'$, for all $S, T \in \mathcal{L}(V, W)$

2. $(\lambda T)' = \lambda T'$, for all $\lambda \in F$ and $T \in \mathcal{L}(V, W)$

3. $(ST)' = T'S'$, for all $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$

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

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

U^0 is a subspace of V'.

3.104: Suppose V is finite-dimensional, then $\dim U + \dim U^0 = \dim V$.

3.107 (the null space of T'):

1. $\operatorname{null} T' = (\operatorname{range} T)^0$;

2. $\operatorname{range} T' = (\operatorname{null} T)^0$.

Let $T \in \mathcal{L}(V, W)$, then

1. T is surjective iff T' is injective.

2. T is injective iff T' is surjective.

$\dim \operatorname{range} T = \dim \operatorname{range} T'$

The **transpose** of a matrix A, denoted A^t, is the matrix obtained from A by interchanging the rows and columns. Visually, this is flipping the matrix across the diagonal of square matrix inside A. In mathematical terms, $(A^t)_{j,k} = A_{k,j}$

$\mathcal{M}(T') = (\mathcal{M}(T))^t$.

The **row rank** of a matrix A is the dimension of the span of the rows of A in $F^{1,n}$.

The **column rank** of a matrix A is the dimension of the span of the columns of A in $F^{m,1}$.

Let $T \in \mathcal{L}(V, W)$. Then $\dim \operatorname{range} T$ equals the column rank of $\mathcal{M}(T)$.

For $A \in F^{m,n}$, the row rank of A equals the column rank of A.

Eigenvalues and Eigenvectors

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

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

The idea of an eigenvalue is that $\operatorname{span} v$ is a one-dimensional subspace of W invariant under T, where v is an eigenvector of T.

5.7 - (equivalent conditions to be an eigenvalue): Suppose $T \in \mathcal{L}(V)$ and $\lambda \in F$, then the following are equivalent:

1. λ is an eigenvalue of T;

2. $T - \lambda I$ is not injective;

3. $T - \lambda I$ is not surjective;

4. $T - \lambda I$ is not invertible.

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

5.11 - (linearly independent eigenvectors): Suppose $T \in \mathcal{L}(V)$. Then every list of eigenvectors of T corresponding to distinct eigenvalues of T is linearly independent.

5.12: Each operator on V has at most $\dim(V)$ distinct eigenvalues.

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

1. T^m is T being applied m times.

2. T^0 is the identity operator I on V.

3. If T is invertible with inverse T^{-1}, then T^{-m} is defined by $T^{-m} = (T^{-1})^m$.

Let $p(z) = a_m z^m + \cdots + a_1 z + a_0 \in \mathcal{P}(F)$ and $T \in \mathcal{L}(V)$, then the operator $p(T)$ is defined by $p(T) = a_m T^m + \cdots + a_1 T + a_0$.

5.17 - (multiplicative properties): Suppose $p, q \in \mathbb{P}(F)$ and $T \in \mathcal{L}(W)$, then

1. $(pq)(T) = p(T)q(T)$

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

5.18 - (null space and range of $p(T)$ are invariant under T): Suppose $p \in \mathbb{P}(F)$ and $T \in \mathcal{L}(V)$. Then $\operatorname{null} p(T)$ and $\operatorname{range} p(T)$ are invariant under T.

5.19 (operators on complex vector spaces have an eigenvalue): Every operator on a finite-dimensional, nonzero, complex vector space has an eigenvalue.

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

5.22 - (existence, uniqueness, and degree of minimal polynomial): Suppose $T \in \mathcal{L}(V)$. Then there exists a unique monic polynomial $p \in \mathbb{P}(F)$ of smallest degree such that $p(T) = 0$. Furthermore, $\deg p \leq \deg V$.

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

To compute the minimal polynomial of an operator $T \in \mathcal{L}(V)$, the smallest positive integer m such that the equation
$$c_0 I + c_1 T + \cdots + c_{m-1} T^{m-1} = -T^m$$
is the degree of the minimal polynomial.

5.27 - (eigenvalues are the zeros of the minimal polynomial): Suppose $T \in \mathcal{L}(V)$,

1. The zeros of the minimal polynomial of T are the eigenvalues of T.

2. If V is a complex vector space, then the minimal polynomial of T has the form

$$(z - \lambda_1) \ldots (z - \lambda_m),$$

where $\lambda_1, \ldots, \lambda_m$ is the list of all the eigenvalues of T, possibly with repetitions.

5.29: Suppose $q \in \mathbb{P}(F)$ and $T \in \mathcal{L}(V)$. Then $q(T) = 0$ iff q is a polynomial multiple of the minimal polynomial of T.

5.31 - (minimal polynomial of a restriction operator): Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V that is invariant under T. Then the minimal polynomial of T is a polynomial multiple of the minimal polynomial of $T|_U$.

5.32: T is not invertible $\iff$ the constant term of the minimal polynomial of T is 0.

5.33: Every operator on a finite-dimensional nonzero vector space has an invariant subspace with dimension 1 or dimension 2.

5.34: Every operator on an odd-dimensional vector space has an eigenvalue.

Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V, then $\mathcal{M}(T) = \begin{pmatrix} A_{1,1} \ldots A_{1,n} \\ \vdots \\ A_{n,1} \ldots A_{n,n} \end{pmatrix}$ denotes the $n \times n$ **matrix of** T with respect to the basis. The entries $A_{j,k}$ are defined by $Tv_k = A_{1,k}v_1 + \cdots + A_{n,k}v_n$. If the basis is not clear, then the notation $\mathcal{M}(T, (v_1, \ldots, v_n))$. If no basis is specified, the standard basis can be assumed.

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

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

5.39 - (conditions for upper-triangular matrix): Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V, then the following are equivalent:

1. the matrix of T with respect to $v_1, \ldots, v_n$ is upper triangular;

2. $\text{span}(v_1, \ldots, v_k)$ is invariant under T for each $k = 1, \ldots, n$;

3. $Tv_k \in \text{span}(v_1, \ldots, v_k)$ for each $k = 1, \ldots, n$.

5.40 - (equation satisfied by operator with upper-triangular matrix): Suppose $T \in \mathcal{L}(V)$ and V has a basis with respect to T has an upper triangular matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$. Then

$$(T - \lambda_1 I) \ldots (T - \lambda_n I) = 0.$$

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

5.44 - (necessary and sufficient condition to have an upper-triangular matrix): Suppose $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some basis of V if and only if the minimal polynomial of T equals $(z - \lambda_1) \ldots (z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in F$.

5.47 - (if $F = \mathbb{C}$, then every operator on V has an upper-triangular matrix): Suppose $T \in \mathcal{L}(V)$.

1. If $F = \mathbb{C}$, then T has an upper triangular matrix with respect to some basis of V.

2. If $F = \mathbb{R}$, then T has an upper-triangular matrix with respect to some basis of V if and only if every zero of the minimal polynomial of T, thought of as a polynomial with complex coefficients, is real.

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

An operator on V is called **diagonalizable** if the operator has a diagonal matrix with respect to some basis of V.

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in F$. The **eigenspace** of T corresponding to λ is the subspace $E(\lambda, T)$ of W 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.

5.54: Suppose $T \in \mathcal{L}(V)$ and $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T. Then

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

is a direct sum. Furthermore,

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

5.55 - (conditions equivalent to diagonalizability): Suppose $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. Then the following are equivalent.

1. T is diagonalizable

2. V has a basis consisting of eigenvectors of T

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

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

5.58 - (enough eigenvalues implies diagonalizability): Suppose $T \in \mathcal{L}(V)$ has $\dim V$ distinct eigenvalues. Then T is diagonalizable.

5.62 - (necessary and sufficient condition for diagonalizability): Suppose $T \in \mathcal{L}(V)$. Then T is diagonalizable iff the minimal polynomial of T equals $(z - \lambda_1) \ldots (z - \lambda_m)$ for some list of distinct numbers $\lambda_1, \ldots, \lambda_m \in F$.

5.65 - (restriction of diagonalizable to invariant subspace is diagonalizable): Suppose $T \in \mathcal{L}(V)$ is diagonalizable and U is a subspace of V that is invariant under T. Then $T|_U$ is a diagonalizable operator on U.

Two operators S and T on the same vector space **commute** if $ST = TS$.

5.69 - (commuting operators correspond to commuting matrices)

5.70 - (eigenspace is invariant under commuting operator): Suppose $S, T \in \mathcal{L}(V)$ commute and $\lambda \in F$. Then $E(\lambda, S)$ is invariant under T.

5.71: Two diagonalizable operators on the same vector space have diagonal matrices with respect to the same basis if and only if the two operators commute.

5.73: Every pair of commuting operators on a finite-dimensional, nonzero, complex vector space has a common eigenvector.

5.75: Suppose V is a complex vector space and S, T are commuting operators on V. Then there is a

basis of V with respect to which both S and T have upper triangular matrices.

5.76 - (eigenvalues of sum and product of commuting operators): Suppose V is a complex vector space and S, T are commuting operators on V. Then

1. every eigenvalue of $S + T$ is an eigenvalue of S plus an eigenvalue of T

2. every eigenvalue of ST is an eigenvalue of S times an eigenvalue of T.

Inner Product Spaces

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

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

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

Dot product properties: For $x, y \in \mathbb{R}^n$,

1. $x \cdot x \geq 0$,

2. $x \cdot x = 0$ iff $x = 0$,

3. $x \cdot y = y \cdot x$.

For $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$, we define the **norm** of

$$||z|| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}.$$

The Euclidean inner product of vector $w = (w_1, \ldots, w_n) \in \mathbb{F}^n$ with $z = (z_1, \ldots, z_n) \in \mathbb{F}^n$ is

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

NOTATION *** For $\lambda \in \mathbb{C}$, the notation $\lambda \geq 0$ means λ is real and non-negative.

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

1. $\langle w, w \rangle \geq 0$,

2. $\langle w, w \rangle = 0$ iff $w = 0$,

3. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$,

4. $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$ for $\lambda \in F$,

5. $\langle v, w \rangle = \overline{\langle w, v \rangle}$.

An **inner product space** is a vector space W along with an inner product on W.

*** For the rest of this chapter, V and W denote inner product spaces over F and V is finite-dimensional.

6.6 - (basic properties of an inner product): Let $v, w \in W$, then

1. For each fixed w, the function that takes v to $\langle v, w \rangle$ is a linear map from W to F,

2. $\langle 0, w \rangle = 0$,

3. $\langle w, 0 \rangle = 0$,

4. $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$,

5. $\langle v, \lambda w \rangle = \overline{\lambda} \langle v, w \rangle$.

For $w \in W$, the **norm** of w, denoted $||w||$, is defined by

$$||w|| = \sqrt{\langle w, w \rangle}.$$

6.9 - (basic properties of the norm): Suppose $w \in W$ and $\lambda \in F$,

1. $||w|| = 0$ iff $w = 0$,

2. $||\lambda w|| = |\lambda| \, ||w||$.

Two vectors $v, w \in W$ are called **orthogonal** if $\langle v, w \rangle = 0$.

$\vec{0}$ is orthogonal to everything, it is also the only vector orthogonal to itself.

6.12 - (Pythagorean theorem): If $v, w \in W$ are orthogonal, then $||v + w||^2 = ||v||^2 + ||w||^2$.

6.13 - (an orthogonal decomposition): Suppose $v, w \in W$ and $w \neq 0$, set

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

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

6.14 - (Cauchy–Schwarz inequality): Suppose $v, w \in W$. Then

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

where equality occurs iff one of v, w is a scalar multiple of the other.

6.17 - (triangle inequality): Suppose $v, w \in W$. Then

$$||v + w|| \leq ||v|| + ||w||,$$

where equality occurs iff one of v, w is a non-negative real multiple of the other.

6.21 - (parallelogram equality): Suppose $v, w \in W$. Then

$$||v + w||^2 + ||v - w||^2 = 2(||v||^2 + ||w||^2).$$

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 $e_1, \ldots, e_m$ of vectors in W is **orthonormal** if

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

6.24 - (norm of an orthonormal linear combination): Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in W, then

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

for all $a_i \in F$.

6.25 - (orthonormal lists are linearly independent)

6.26 - (Bessel's inequality): Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in W. If $w \in W$, then

$$|\langle w, e_1 \rangle|^2 + \cdots + |\langle w, e_m \rangle|^2 \leq ||w||^2.$$

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

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

6.30 - (writing a vector as linear combination of orthonormal basis): Suppose $e_1, \ldots, e_m$ is an orthonormal basis of V and $v, w \in V$, then

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

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

3. $\langle v, w \rangle = \langle v, e_1 \rangle \overline{\langle w, e_1 \rangle} + \cdots + \langle v, e_n \rangle \overline{\langle w, e_n \rangle}$.

6.32 - (Gram–Schmidt procedure): Suppose $w_1, \ldots, w_m$ is a linearly independent list of vectors in W. Let $f_1 = w_1$. For $k = 2, \ldots, m$, define f_k inductively by

$$f_k = w_k - \frac{\langle w_k, f_1 \rangle}{||f_1||^2} f_1 - \cdots - \frac{\langle w_k, f_{k-1} \rangle}{||f_{k-1}||^2} f_{k-1}.$$

For each $k = 1, \ldots, m$, let $e_k = \dfrac{f_k}{||f_k||}$. Then $e_1, \ldots, e_m$ is an orthonormal list of vectors in W such that $\text{span}(w_1, \ldots, w_k) = \text{span}(e_1, \ldots, e_k)$, for $k = 1, \ldots, m$.

6.35 - (existence of orthonormal basis)

6.36: Every orthonormal list of vectors in V can be extended to an orthonormal basis of V.

6.38 - (upper-triangular matrix with respect to orthonormal basis): Suppose $T \in \mathcal{L}(V)$, then T has an upper-triangular matrix with respect to some orthonormal basis of V iff the minimal polynomial of T equals $(z - \lambda_1) \ldots (z - \lambda_m)$, for some $\lambda_1, \ldots, \lambda_m \in F$.

6.39 - (Schur's theorem): Every operator on a finite-dimensional complex inner product space has an upper-triangular matrix with respect to some orthonormal basis.

Linear Functionals on Inner Product Spaces

A **linear functional** on W is a linear map from W to F.

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

6.43 - (Riesz representation theorem): Suppose φ is a linear functional on V. Then there is a unique vector $w \in V$ such that
$$\varphi(v) = \langle v, w \rangle,$$
for every $v \in V$.

Orthogonal Complements and Minimization Problems

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

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

6.49 - (properties of orthogonal complement):

1. If U is a subset of W, then $U^\perp$ is a subspace of W,

2. $\{0\}^\perp = W$,

3. $W^\perp = \{0\}$,

4. If U is a subset of W, then $U \cap U^\perp \subset \{0\}$,

5. If G and H are subsets of W and $G \subset H$, then $H^\perp \subset G^\perp$.

6.50 - (direct sum of a subspace and its orthogonal complement): Suppose U is a finite-dimensional subspace of V, then
$$V = U \oplus U^\perp.$$

6.53 - (dimension of orthogonal complement): Suppose U is a subspace of V, then
$$\dim U^\perp = \dim V - \dim U.$$

6.54 - (orthogonal complement of the orthogonal complement): Suppose U is a finite-dimensional subspace of W, then
$$U = (U^\perp)^\perp.$$

6.56: Suppose U is a finite-dimensional subspace of W, then
$$U^\perp = \{0\} \iff U = W.$$

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

6.59 - (properties of orthogonal projection P_U): Suppose U is a finite-dimensional subspace of W and $w \in W$. Then, for every $u \in U$ and $v \in U^\perp$,

1. $P_U \in \mathcal{L}(W)$;

2. $P_U u = u$;

3. $P_U v = 0$;

4. $\text{range } P_U = U$;

5. $\text{null } P_U = U^\perp$;

6. $w - P_U w \in U^\perp$;

7. $P_U{}^2 = P_U$;

8. $||P_U w|| \leq ||w||$;

9. if $e_1, \ldots, e_m$ is an orthonormal basis of U, then

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

6.60 - (Riesz representation theorem): For $v \in V$, define $\varphi_v \in V'$ by

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

for $u \in V$. Then $v \to \varphi_v$ is a one-to-one function from V onto V'.

6.63 - (minimizing distance to a subspace): Suppose U is a finite-dimensional subspace of W, $w \in W$ and $u \in U$. Then

$$||w - P_U w|| \leq ||w - u||,$$

where equality occurs iff $u = P_U w$.

6.69 - (restriction of a linear map to obtain a one-to-one and onto map): Suppose $T \in \mathcal{L}(V, W)$. Then $T|_{(\text{null } T)^\perp}$ is a one-to-one map of $(\text{null } T)^\perp$ onto $\text{range } T$.

Suppose $T \in \mathcal{L}(V, W)$. The **pseudoinverse**, or **Moore-Penrose inverse**, $T^\dagger \in \mathcal{L}(W, V)$ of T is the linear map from W to V defined by $T^\dagger w = (T|_{(\text{null } T)^\perp})^{-1} P_{\text{range } T} w$, for some $w \in W$.

If $w \in (\text{range } T)^\perp$, then $T^\dagger w = 0$. If $w \in \text{range } T$, then $T^\dagger w$ is the unique element of $(\text{null } T)^\perp$ such that $T(T^\dagger w) = w$.

6.71 - (algebraic properties of the pseudoinverse): Suppose $T \in \mathcal{L}(V, W)$,

1. If T is invertible, then $T^\dagger = T^{-1}$.

2. $T^\dagger T = P_{(\text{null } T)^\perp}$ = the orthogonal projection of V onto $(\text{null } T)^\perp$.

3. $TT^\dagger = P_{\text{range } T}$ = the orthogonal projection of W onto $\text{range } T$.

Result #2 above implies that if T is injective, then $T^\dagger T$ is the identity operator on V. Result #3 above implies that if T is surjective, then $TT^\dagger$ is the identity operator on W.

6.72 - (pseudoinverse provides best approximate solution or best solution): Suppose $T \in \mathcal{L}(V, W)$ and $b \in W$.

1. If $x \in V$, then

$$||T(T^\dagger b) - b)|| \leq ||Tx - b||,$$

with equality iff $x \in T^\dagger b + \text{null } T$.

2. If $x \in T^\dagger b + \text{null } T$, then

$$||T^\dagger b|| \leq ||x||,$$

with equality iff $x = T^\dagger b$.

Operators on Inner Product Spaces

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 $w \in W$.

In other words, $u = T^*w$ is the unique vector in V such that $\langle Tv, w \rangle = \langle v, u \rangle$ for every $v \in V$.

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

7.5 - (properties of the adjoint): Suppose $T \in \mathcal{L}(V,W)$, then

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

2. $(\lambda T)^* = \overline{\lambda} T^*$;

3. $(T^*)^* = T$;

4. $(ST)^* = T^*S^*$ for all $S \in \mathcal{L}(W,U)$, where U is an inner product space over F;

5. $I^* = I$, for the identity operator I on V;

6. If T is invertible, then T^* is inevitable and $(T^*)^{-1} = (T^{-1})^*$.

7.6 - (null space and range of T^*): Suppose $T \in \mathcal{L}(V,W)$, then

1. $\text{null}\, T^* = (\text{range}\, T)^{\perp}$

2. $\text{range}\, T^* = (\text{null}\, T)^{\perp}$

3. $\text{null}\, T = (\text{range}\, T^*)^{\perp}$

4. $\text{range}\, T = (\text{null}\, T^*)^{\perp}$

The **conjugate transpose** of an m-by-n matrix A is the n-by-m matrix A^* obtained by interchanging the rows and columns and then taking the complex conjugate of each entry. In other words, $(A^*)_{j,k} = \overline{A_{k,j}}$.

7.9 - (matrix of T^* equals conjugate transpose of matrix of T): Let $T \in \mathcal{L}(V,W)$. Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V and $f_1, \ldots, f_m$ is an orthonormal basis of W. Then $M(T^*, (f_1, \ldots, f_m), (e_1, \ldots, e_n))$ is the conjugate transpose of $M(T, (e_1, \ldots, e_n), (f_1, \ldots, f_m))$. In other words,

$$M(T^*) = (M(T))^*.$$

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 iff

$$\langle Tv, w \rangle = \langle v, Tw \rangle,$$

for al $v, w \in V$.

*** **7.12** - (eigenvalues of self-adjoint operators are real).

7.13 - (Tv is orthogonal to v for all $v \iff T = 0$ (assuming $F = \mathbb{C}$)): Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$, then

$$\langle Tv, v \rangle = 0, \text{ for every } v \in V \iff T = 0.$$

7.14 - ($\langle Tv, v \rangle$ is real for all $v \iff T$ is self-adjoint (assuming $F = \mathbb{C}$)): Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$, then

$$T \text{ is self-adjoint} \iff \langle Tv, v \rangle \in \mathbb{R} \text{ for every } v \in V.$$

7.16 - (T self-adjoint and $\langle Tv, v \rangle = 0$ for all $V \iff T = 0$ (this is true over $\mathbb{R}$ as well, not just $\mathbb{C}$)): Suppose T is a self-adjoint operator on V, then

$$\langle Tv, v \rangle = 0 \text{ for every } v \in V \iff T = 0.$$

*** **Important Factorization**:

$$\langle Tv, w \rangle = \frac{\langle T(u + v), u + v \rangle - \langle T(u - v), u - v \rangle}{4}.$$

An operator on an inner product space is called **normal** if it commutes with it's adjoint. In other words, $T \in \mathcal{L}(V)$ is normal if $TT^* = T^*T$.

7.20 - (T is normal $\iff ||Tv|| = ||T^*v||$ for all v)

7.21 - (range, null space, and eigenvectors of a normal operator): Suppose $T \in \mathcal{L}(V)$ is normal, then

1. $\text{null } T = \text{null } T^*$;

2. $\text{range } T = \text{range } T^*$;

3. $V = \text{null } T \oplus \text{range } T$;

4. $T - \lambda I$ is normal for every $\lambda \in F$;

5. $Tv = \lambda v$ iff $T^*v = \overline{\lambda}v$, for $v \in V$ and $\lambda \in F$.

Part 5 of the above result, 7.21, implies that a normal operator and it's adjoint have the same eigenvectors and same eigenvalues (up to complex conjugates). The correspondence between eigenvalues and eigenvectors are the same between a normal operator and it's adjoint.

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

7.23 - (T is normal $\iff$ the real and imaginary parts of T commute): Suppose $F = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then T is normal iff there exist commuting self-adjoint operators Q and S such that $T = Q + iS$.

7.26 - (invertible quadratic expressions): Suppose $T \in \mathcal{L}(V)$ is self-adjoint and $b, c \in \mathbb{R}$ are such that $b^2 < 4c$. Then

$$T^2 + bT + cI$$

is an invertible operator.

7.27 - (minimal polynomial of self-adjoint operator): Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Then the minimal polynomial of T equals $(z - \lambda_1) \ldots (z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$.

7.29 - (real spectral theorem): Suppose $F = \mathbb{R}$ and $T \in \mathcal{L}(V)$, then the following are equivalent

1. T is self-adjoint.

2. T has a diagonal matrix with respect to some orthonormal basis of V.

3. V has an orthonormal basis consisting of eigenvectors of T.

7.31 - (complex spectral theorem): Suppose $F = \mathbb{C}$ and $T \in \mathcal{L}(V)$, then the following are equivalent

1. T is normal.

2. T has a diagonal matrix with respect to some orthonormal basis of V.

3. V has an orthonormal basis consisting of eigenvectors of T.

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

$$\langle Tv, v \rangle \geq 0,$$

for all $v \in V$.

An operator R is called a **square root** of an operator T if $R^2 = T$.

7.38 - (characterization of positive operators): Let $T \in \mathcal{L}(V)$. Then the following are equivalent:

1. T is a positive operator

2. T is self-adjoint and all eigenvalues of T are non-negative

3. with respect to some orthonormal basis of V the matrix of T is a diagonal matrix with only non-negative numbers on the diagonal

4. T has a positive square root

5. T has a self-adjoint square root

6. $T = R^*R$ for some $R \in \mathcal{L}(V)$

7.39 - (Every positive operator on V has a unique positive square root)

***: For some positive operator T, $\sqrt{T}$ denotes the unique positive square root of T.

7.43 - (T positive and $\langle Tv, v \rangle = 0 \implies Tv = 0$)

A linear map $S \in \mathcal{L}(V, W)$ is called an **isometry** if

$$||Sv|| = ||v||$$

for every $v \in V$. In other words, a linear map is an isometry if it preserves norms.

7.49 - (characterization of isometries): Let $S \in \mathcal{L}(V, W)$. Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V and $f_1, \ldots, f_m$ is an orthonormal basis of W. Then the following are equivalent:

1. S is an isometry

2. $S^*S = I$

3. $\langle Su, Sv \rangle = \langle u, v \rangle$ for all $u, v \in V$

4. $Se_1, \ldots, Se_n$ is an orthonormal list in W

5. The columns of $M(S, (e_1, \ldots, e_n), (f_1, \ldots, f_m))$ form an orthonormal list in F^m with respect to the Euclidean inner product.

An operator $S \in \mathcal{L}(V)$ is called **unitary** if S is an invertible isometry.

7.53 - (characterization of unitary operators): Let $S \in \mathcal{L}(V)$. Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V. Then the following are equivalent:

1. S is a unitary operator

2. $S^*S = SS^* = I$

3. S is invertible and $S^{-1} = S^*$

4. $Se_1, \ldots, Se_n$ is an orthonormal basis of V

5. The rows of $M(S, (e_1, \ldots, e_n))$ form an orthonormal list of F^n with respect to the Euclidean inner product.

6. S^* is a unitary operator

Analogy: A complex number z can be thought of as corresponding to an operator $S \in \mathcal{L}(V)$, and $\bar{z}$ corresponding to S^*. If z is real, then $S = S^*$, as we know. Similarly, the non-negative numbers correspond to positive operators. The values of $|z| = 1$ corresponds to isometries, as $S^*S = I$.

7.54 - (eigenvalues of unitary operators have absolute value 1)

7.55 - (description of unitary operators on complex inner product spaces): Suppose $F = \mathbb{C}$ and $S \in \mathcal{L}(V)$. Then the following are equivalent:

1. S is a unitary operator

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

An n-by-n matrix is called **unitary** if it's columns (or rows) form an orthonormal basis in F^n.

7.57 - (characterizations of unitary matrices): Suppose Q is an n-by-n matrix. Then the following are equivalent:

1. Q is a unitary matrix.

2. The rows of Q form an orthonormal list in F^n.

3. $||Qv|| = ||v||$ for every $v \in F^n$.

4. $Q^*Q = QQ^* = I$, where I is the n-by-n identity matrix with 1's on the diagonal and 0's elsewhere.

7.58 - (QR factorization): Suppose A is a square matrix with linearly independent columns. Then there exist unique matrices Q and R such that Q is unitary, R is upper triangular with all positive entries on its diagonal, and
$$A = QR.$$

7.61 - (properties of T^*T): Suppose $T \in \mathcal{L}(V, W)$. Then

1. T^*T is a positive operator on V;

2. $\operatorname{null} T^*T = \operatorname{null} T$;

3. $\operatorname{range} T^*T = \operatorname{range} T^*$;

4. $\dim \operatorname{range} T = \dim \operatorname{range} T^* = \dim \operatorname{range} T^*T$.

Suppose $T \in \mathcal{L}(V, W)$. The **singular values** of T are the non-negative square roots of the eigenvalues of T^*T listed in decreasing order, each included as many times as the dimension of the corresponding eigenspace of T^*T.

5.54 - (sum of eigenspaces is a direct sum): Suppose $T \in \mathcal{L}(V)$ and $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T. Then
$$E(\lambda_1, T) + \cdots + E(\lambda_m, T)$$
is a direct sum. Furthermore,
$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim V.$$

7.65 - (role of positive singular values): Suppose $T \in \mathcal{L}(V, W)$. Then

1. T is injective $\iff$ 0 is not a singular value of T;

2. the number of positive singular values of T equals $\dim \operatorname{range} T$;

3. T is surjective $\iff$ number of positive singular values of T equals $\dim W$.

7.66 - (isometries characterized by having all singular values equal 1): Suppose that $S \in \mathcal{L}(V, W)$. Then S is an isometry $\iff$ all the singular values of S equal 1.

7.67 - (singular value decomposition): Suppose $T \in \mathcal{L}(V, W)$ and the positive singular values of T are $s_1, \ldots, s_m$. Then there exist orthonormal lists $e_1, \ldots, e_m$ in V and $f_1, \ldots, f_m$ in W such that
$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$
for every $v \in V$.

The defintion of a diagonal matrix extends very intuitively to non-square matrices.

7.72 - (singular value decomposition of adjoint and pseudoinverse): Suppose $T \in \mathcal{L}(V, W)$ and the positive singular values of T are $s_1, \ldots, s_m$. Suppose $e_1, \ldots, e_m$ and $f_1, \ldots, f_m$ are orthonormal lists in V and W, respectively, such that
$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$
for every $v \in V$. Then
$$T^*w = s_1 \langle w, f_1 \rangle e_1 + \cdots + s_m \langle w, f_m \rangle e_m$$
and
$$T^\dagger w = \frac{\langle w, f_1 \rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_m \rangle}{s_m} e_m$$

7.77 - (matrix version of singular value decomposition): Suppose A is an M-by-N matrix. Then there exist an M-by-N diagonal matrix D with non-negative entries, an M-by-M unitary matrix Q_1, and an N-by-N unitary matrix Q_2 such that
$$A = Q_1 D Q_2.$$

Operators on Complex Vector Spaces

8.1 - (sequence of increasing null spaces): Suppose $T \in \mathcal{L}(V)$. Then

$$\{0\} = \operatorname{null} T^0 \subset \operatorname{null} T^1 \subset \dots .$$

8.2 - (equality in the sequence of null spaces): Suppose $T \in \mathcal{L}(V)$ and m is a non-negative integer such that $\operatorname{null} T^m = \operatorname{null} T^{m+1}$. Then

$$\operatorname{null} T^m = \operatorname{null} T^{m+1} = \operatorname{null} T^{m+2} = \dots .$$

8.3 - (null spaces stop growing): Suppose $T \in \mathcal{L}(V)$. Then

$$\operatorname{null} T^{\dim V} = \operatorname{null} T^{\dim V+1} = \dots .$$

8.4 - (V is the direct sum of $\operatorname{null} T^{\dim V}$ and $\operatorname{range} T^{\dim V}$): Suppose $T \in \mathcal{L}(V)$. Then

$$V = \operatorname{null} T^{\dim V} \oplus \operatorname{range} T^{\dim V}.$$

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

$$(T - \lambda I)^k v = 0,$$

for some positive integer $k \in \mathbb{N}$.
Note that every eigenvector of T is a generalized eigenvector, simply considering the $k = 1$ case in the above equation.

A nonzero vector $v \in V$ is a **generalized eigenvector** of T corresponding to λ if and only if

$$(T - \lambda I)^{\dim V} v = 0.$$

8.9 - (a basis of generalized eigenvectors): Suppose $F = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then there is a basis of V consisting of generalized eigenvectors of T.
However, if $F = \mathbb{R}$ and $\dim V > 1$, then not all operators on V have the property that there exists a basis of V consisting of generalized eigenvectors of the operator.

8.11 - (generalized eigenvector corresponds to a unique eigenvalue): Suppose $T \in \mathcal{L}(V)$, then each generalized eigenvector of T corresponds to only one eigenvalue of T.

8.12 - (linearly independent generalized eigenvectors): Suppose $T \in \mathcal{L}(V)$. Then every list of generalized eigenvectors of T corresponding to distinct eigenvalues of T is linearly independent.

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

8.16 - (nilpotent operator raised to dimension of domain is 0): Suppose $T \in \mathcal{L}(V)$ is nilpotent. Then $T^{\dim V} = 0$.

8.17 - (eigenvalues of nilpotent operator): Suppose $T \in \mathcal{L}(V)$.

1. If T is nilpotent, then 0 is an eigenvalue of T and T has no other eigenvalues.

2. If $F = \mathbb{C}$ and 0 is the only eigenvalue of T, then T is nilpotent.

8.18 - (minimal polynomial and upper-triangular matrix of nilpotent operator): Suppose $T \in \mathcal{L}(V)$, then the following are equivalent

1. T is nilpotent.

2. The minimal polynomial of T is z^m for some positive integer m.

3. There is a basis of V with respect to which the matrix of T is upper-triangular with all 0's on the diagonal.

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in F$. 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.

8.20 - (description of generalized eigenspaces): Suppose $T \in \mathcal{L}(V)$ and $\lambda_k \in F$. Then

$$G(\lambda_k, T) = \operatorname{null}(T - \lambda_k I)^{\dim V}.$$

8.22 - (generalized eigenspace decomposition): Suppose $F = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of T. Then

1. $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$;

2. Each $G(\lambda_k, T)$ is invariant under T;

3. Each $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent.

Suppose $T \in \mathcal{L}(V)$, the **multiplicity**, or algebraic 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

$$\dim \operatorname{null}(T - \lambda I)^{\dim V}.$$

The **geometric multiplicity** of an eigenvalue is the number of eigenvectors it has (not generalized eigenvectors, just normal eigenvectors!)

8.26 - (sum of the multiplicities equals $\dim V$)

Suppose $F = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T, with multiplicities $d_1, \ldots, d_m$, respectively. Then

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

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

8.29 - (degree and zeros of characteristic polynomial): Suppose $F = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then

1. the characteristic polynomial of T has degree $\dim V$;

2. the zeros of the characteristic polynomial of T are the eigenvalues of T.

8.30 - (Cayley–Hamilton theorem): Suppose $F = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Let q denote the characteristic polynomial of T. Then $q(T) = 0$.

8.31 - (characteristic polynomial is a multiple of minimal polynomial)

8.31 implies that if the minimal polynomial of an operator T has degree $\dim V$, then the characteristic

polynomial of T is the same as the minimal polynomial of T.

A **block diagonal matrix** is a diagonal square matrix with square matrices (not necessarily the same size) on the diagonal (a number can be thought of as a 1×1 square matrix).

8.36 - (identity plus nilpotent has a square root): Suppose $T \in \mathcal{L}(V)$ is nilpotent. Then $I + T$ has a square root.

8.38 - (over $\mathbb{C}$, invertible operators have square roots): Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ is invertible. Then T has a square root.

8.41 - (basis corresponding to nilpotent operator): Suppose $T \in \mathcal{L}(V)$ is nilpotent. Then there exist vectors $v_1, \ldots, v_n \in V$ and non-negative integers $m_1, \ldots, m_n$ such that

1. $T^{m_1}v_1, \ldots, Tv_1, v_1, \ldots, T^{m_n}v_n, \ldots, Tv_n, v_n$ is a basis of V;

2. $T^{m_1+1}v_1 = \cdots = T^{m_n+1}v_n = 0$.

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, where each matrix on the diagonal is an upper-triangular matrix with the eigenvalue corresponding to that inner matrix on the main diagonal, and all 1's on the diagonal right above that. (rest are 0's).

8.46 - (Jordan form): Suppose $F = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then there is a basis of V that is a Jordan basis for T.