Linear Algebra Notes

Contents

1. Vector Spaces

1.1. $\mathbb{R}^n$ and $\mathbb{C}^n$

$\mathbb{R}$ and $\mathbb{C}$ are defined as usual.

Example *(Complex Commutativity)*: $\alpha\beta = \beta\alpha$ for all $\alpha, \beta \in \mathbb{C}$.

Definition (Elements in $\mathbb{F}^n$ and Operations): $\mathbb{F}^n$ is all n-tuples

$$(x_1, x_2, \ldots, x_n)$$

with $x_i \in \mathbb{F}$.

Addition is pointwise. Scalar multiplication by λ multiplies each element by λ. If context is clear, $0 \in \mathbb{F}^n$ denotes $(0, 0, \ldots, 0)$, where there are n 0s.

1.2. Definition of Vector Space

Definition (Vector Space): A vector space V is a set V with addition and scalar multiplication with commutativity, associativity, additive identity, additive inverse, multiplicative identity, and distributive properties. Elements of a vector space are called vectors or points.

Example: $\mathbb{R}^n$ and $\mathbb{C}^n$ are vector spaces, just verify the properties hold. $\mathbb{F}^\infty$ is also a vector space.

Definition ($\mathbb{F}^S$): If S is a set, then $\mathbb{F}^S$ is the set of functions from S to $\mathbb{F}$.

Example *($\mathbb{F}^S$ is a vector space)*: The 0 function $0(x) = 0$ for all $x \in \mathbb{F}$ is the additive identity. The additive inverse is $(-f)(x) = -f(x)$. All other properties of vector spaces hold by spamming axioms. $\mathbb{F}^n$ is a special case of this, where $S = \{1, 2, ..., n\}$.

Proposition (Unique additive identity): A vector space has a unique additive identity.

Proof: $0, 0'$ are identities.

$$0' = 0' + 0 = 0 + 0' = 0.$$

∎

Proposition (Unique additive inverse): Every element in a vector space has a unique additive inverse.

Proof: $v \in V$, with w, w' as inverses.

$$w = w + 0 = w + (v + w') = (w + v) + w' = 0 + w' = w'.$$

∎

Proposition: $0v = 0$ for every $v \in V$.

Proof: $0v = (0 + 0)v = 0v + 0v.$ ∎

Remark: We have to use $0 = 0 + 0$ since we have to use the distributive property to connect scalar multiplication and vector addition.

Proposition: $a0 = 0$ for scalar a.

Proof: $a0 = a(0 + 0) = a0 + a0$. ■

Proposition: $(-1)v = -v$.

Proof:

$$v + (-1)v = 1v + (-1)v = (1 + (-1))v = 0v = 0.$$

■

1.2.1. Problems

Problem (Exercise 1): Prove that $-(-v) = v$ for every $v \in V$.

Solution:

$$-(-v) + (-v) = (-1)(-v) + (-1)v = (-1)(-v + v) = (-1)(0) = 0,$$

so $-(-v)$ is the additive inverse of $-v$, so $-(-v) = v$, as desired.

Problem (Exercise 3): Suppose $v, w \in V$. Explain why there exists a unique $x \in V$ such that $v + 3x = w$.

Solution: First existence. Adding $-v$ to both sides gives $3x = w - v$. Multiplying by $\frac{1}{3}$ on both sides gives $x = \frac{1}{3}(w - v)$.

Now uniquness. Suppose y, y' both satisfy. Then $y = \frac{1}{3}(w - v) = y'$.

Problem (Exercise 4): The empty set is a not a vector space. Why?

Solution: Doesn't satsify additive identity. There are no elements, so there cannot exist and additive identity.

Problem (Exercise 5): Show that the additive inverse condition in the definition of a vector space can be replaced with

$$0v = 0 \text{ for all } v \in V.$$

Solution:

$$0v = (1 + (-1))v = 1v + (-1)v = v + (-1)v = 0,$$

so there exists a w such that $v + w = 0$, as desired.

Problem (Exercise 6): Is $\mathbb{R} \cup \{\infty\} \cup \{-\infty\}$ a vector space over $\mathbb{R}$?

Solution: No. We have that $(2-1)\infty = (1)\infty = \infty$ and $2\infty - 1\infty = \infty + (-\infty)0$, so it is not distributive, so it in not a vector space.

1.3. Subspaces

Definition (Subspace): A subset U of V is called a subspace of V if U is also a vector space.

Example: $(x_1, x_2, 0)$ with $x_1, x_2 \in \mathbb{F}$ is a subspace of $\mathbb{F}^3$.

Proposition (Conditions for subspaces): A $U \subset V$ is a subspace of V if and only if U satisfies the conditions below.
- Additive identity: $0 \in U$
- Closed under addition
- Closer under scalar multiplication

Solution: If U is a subspace of V, then it satisfies the properties by definition.

First condition ensures additive identity. Second and third make sure addition and scalar work. Additive inverse holds by scalar multiplication by -1, and associativity and distributivity hold because that holds on V.

Example:

- If $b \in \mathbb{F}$ then
$$\{(x_1, x_2, x_3, x_4) \in \mathbb{F}^4 : x_3 = 5x_4 + b\}$$
is a subspace if and only if $b = 0$. When $b = 0$, we can easily verify all the subspace conditions hold. If we have a subspace, then $0 \in U$, so $0 = x_3 = 5(0) + b$ means b must be 0.
- The set of continuous real-valued functions on the interval $[0, 1]$ is a subspace of $\mathbb{R}^{[0,1]}$. $f(x) = 0$ is the additive identity for $\mathbb{R}^{[0,1]}$, and clearly addition and scalar multiplication are closed.
- The set of differentiable real-valued functions on $\mathbb{R}$ is a subspace of $\mathbb{R}^{\mathbb{R}}$.
- The set of differentiable real-valued functions f on the interval $(0, 3)$ such that $f'(2) = b$ is a subspace of $\mathbb{R}^{(0,3)}$ if and only if $b = 0$.
- The set of all sequences of comples numbers with limit 0 is a subspace of $\mathbb{C}^{\infty}$.

Definition: Suppose $U_1, U_2, ..., U_m$ are subsets of V. The sum of $U_1, U_2, ..., U_m$ is the set of all possible sums of elements in the subsets. So
$$U_1 + \cdots + U_m = \{u_1 + \cdots + u_m \mid u_i \in U_i\}.$$

Example: We have $U = \{(x, 0, 0) \in \mathbb{F}^3 \mid x \in \mathbb{F}\}$ and $W = \{(0, y, 0) \in \mathbb{F}^3 \mid y \in \mathbb{F}\}$. Then,
$$U + W = \{(x, y, 0) \in \mathbb{F}^3 \mid x, y, \in \mathbb{F}\}.$$

Example: U, W are subsets in $\mathbb{F}^4$, $U = (x, x, y, y)$, $W = (x, x, x, y)$. Then
$$U + W = (x, x, y, z)$$
since when adding two elements from U and W, the sum always has equal first and second components. The sum of the third components can be arbitrary, and same for the fourth.

Proposition (Minimality of subspace sums): If $U_1, ..., U_m$ are subspaces of V, then $U' = \sum U_i$ is the smallest subspace containing U_i.

Solution: Cleary U' is a subspace. All the U_i are contained in U'. Also, in any subspace with U_i, we must have U' by closed addition. Thus, we have minimmality.

Definition (Direct Sum): Suppose $U_1, ..., U_m$ are subspaces of V. Tehn $\sum U_i$ is called a direct sum if each element of $\sum U_i$ can be written in only one way as a sum of $\sum u_i$, where $u_i \in U_i$. If $\sum U_i$ is a direct sum, then we denote is with $U_1 \oplus U_2 \oplus \cdots \oplus U_m$.

Example: $U = (x, y, 0), W = (0, 0, z)$. Then $U \oplus W = \mathbb{F}^3$.

Example *(Nonexample)*: $U_1 = (x, y, 0), U_2 = (0, 0, z), U_3 = (0, y, y)$. We have that $\mathbb{F}^3 = \sum U_i$ since we can write every vector in $\mathbb{F}^3$ as the sum of three vectors from each of the subsets. But $(0, 0, 0)$ can be written in two different ways, so it's not a direct sum.

Proposition (Direct Sum Condition): U_i are subspaces of V. Then $W = \sum U_i$ is a direct sum if and only if the only way to write 0 is by writing 0 in each subspace.

Solution: If W is a direct sum, then by definition the only way to write 0 is by taking 0 from each U_i (0 is in each of these by subspace condition). Now suppose the only way to write 0 is to take $u_i = 0$. Now consider $v \in W$. Suppose there are two ways to write it,

$$v = \sum u_i = \sum v_i.$$

Subtracting gives $0 = \sum u_i - v_i$. We know that $u_i - v_i \in U_i$ because it is a subapce, and we also know that the only way to write 0 is having all components equal 0. Thus, $u_i = v_i$, so there is only one way to write each vector as a sum, as desired.

Proposition (Direct sum of two subspaces): U, W are subspaces of V. Then $U + W$ is a direct sum of and only if $U \cap W = \{0\}$.

Solution: If $U + W$ is a direct sum, then for $v \in U \cap W$, we have that $0 = v + (-v)$, with $v \in U$ and $-v \in W$. By unique representation, $v = 0$, so $U \cap W = \{0\}$. If $U \cap W = 0$, then for $u \in U$ and $w \in W$ we have $0 = u + w$. We need to show $u = w = 0$. The equation implies $u = -w \in W$, so $u \in U \cap W$, which means $u = 0$ as desired.

1.3.1. Problems

Problem (Exercise 1): For each of the following subsets of $\mathbb{F}^3$, determine whether it is a subapce of $\mathbb{F}^3$:
- $\{(x_1, x_2, x_3) : x_1 + 2x_2 + 3x_3 = 0\}$
- $\{(x_1, x_2, x_3) : x_1 + 2x_2 + 3x_3 = 4\}$
- $\{(x_1, x_2, x_3) : x_1 x_2 x_3 = 0\}$
- $\{(x_1, x_2, x_3) : x_1 = 5x_3\}$

Solution:
- Yes, it is closed under addition and scalar multiplication and has 0.
- No, does not have 0.
- Not closed under addition.
- Yes.

Problem (Exercise 3): Show that the set of differentiable real valued functions on $(-4, 4)$ such that $f'(-1) = 3f(2)$ is a subspace of $\mathbb{R}^{-4,4}$.

Solution: $f \equiv 0$ clearly satisfies the additive identity. It's also easy to see it's closed under addition and scaling.

Problem (Exercise 4): Show that the set of continuous real valued functions on the interval $[0, 1]$ such that $\int_0^1 = b$ is a subspace of $\mathbb{R}^{[0,1]}$ if and only if $b = 0$.

Solution: If $b = 0$, then the conclusion follows easily from verifying subspace conditions. If the set is a subspace, then for f we need $\int_0^1 2f = 2b = b$, which implies $b = 0$.

Problem (Exercise 5): Is $\mathbb{R}^2$ a subspace of $\mathbb{C}^2$.

Solution: No, it's not closed under scalar multiplication.

Problem (Exercise 6):
- Is $\{(a, b, c) \in \mathbb{R}^3 : a^3 = b^3\}$ a subspace of $\mathbb{R}^3$?
- Is $\{(a, b, c) \in \mathbb{C}^3 : a^3 = b^3\}$ a subspace of $\mathbb{C}^3$?

Solution:
- Yes, since over $\mathbb{R}$ we have $a^3 = b^3 \Rightarrow a = b$.
- No, since $(\omega, 1, 0) + (1, 1, 0) = (\omega + 1, 2, 0)$ is not in the set.

Problem (Exercise 7): Prove or give a counterexample: if U is a nonempty subset of $\mathbb{R}^2$ such that U is closed under addition and under taking additive inverses, then U is a subspace of $\mathbb{R}^2$.

Solution: $U = \{(x, y) : x, y, \in \mathbb{Z}\}$.

Problem (Exercise 8): Given an example of a nonempty subset U of $\mathbb{R}^2$ such that U is closed under scalar multiplication but is not a subspace.

Solution: $U = \{(x, y) : xy = 0\}$

Problem (Exercise 9): Is the set of periodic functions $f : \mathbb{R} \to \mathbb{R}$ a subspace of $\mathbb{R}^{\mathbb{R}}$?

Solution: No. Consider $\sin(x)$ and $\sin(\pi x)$. When added, they do not form a periodic function, so the set is not closed under addition.

Problem (Exercise 10): If U_1 and U_2 are subspaces of V, then $U_1 \cap U_2$ is a subspace of V.

Solution: Let $W = U_1 \cap U_2$. 0 is in U_1 and U_2, so $0 \in W$. Consider a vector $w \in W$. Since $w \in U_1$, $\lambda w \in U_1$. Similarly for U_2. Thus W is closed under scalar multiplication. A similar argument can be used for addition, so W is a subspace of V.

Problem (Exercise 12): Prove that the union of two subspaces of V is a subspace of V if and only if one of the subspaces contains the other two.

Solution: Let U_1, U_2, be subspaces of V. If $U_1 \subset U_2$, then $U_1 \cup U_2 = U_2$, which is a subspace. Now suppose $U_1 \cup U_2 = W$ is a subspace. Suppose for the sake of contradiction $U_1 \not\subset U_2$. Pick $x \in U_1$, $x \notin U_2$ and $y \notin U_1, y \in U_2$. We know that $x + y = w \in W$ must be in U_1 or U_2. Suppose its in U_1. Then, $y = w - x = w + (-1)x \in U_1$, a contradiction. The same applies to U_2. Thus, $U_1 \subseteq U_2$.

Problem (Exercise 19): Prove or give a counterexample: if U_1, U_2, W are subspaces of V such that

$$U_1 + W = U_2 + W,$$

then $U_1 = U_2$.

Solution: $U_1 = \{0\}, U_2 = W$, W is any subspace of V that's not $\{0\}$.

Problem (Exercise 20): Suppose $U = (x, x, y, y)$. Find a subspace W of $\mathbb{F}^4$ such that $\mathbb{F}^4 = U \oplus W$.

Solution: The subspace $W = (a, 0, 0, b)$. Note that we can write any element in $\mathbb{F}^4$ as a sum of elements from U and V, so $U + v = \mathbb{F}^4$. Note also that the only way two vectors in W and U are equal is when $x = y = a = b = 0$, or in other words, $U \cap W = \{0\}$. Then, $U \oplus W = \mathbb{F}^4$.

Problem (Exercise 21): Suppose $U = \{(x, y, x + y, x - y, 2x)\}$. Find a subspace W of $\mathbb{F}^5$ such that $\mathbb{F}^5 = U \oplus W$.

Solution: $W = \{(0, 0, a, b, c)\}$

Problem (Exercise 22): U is the same as the previous problem. Find $W_1, W_2, W_3 \neq \{0\}$ such that $\mathbb{F}^5 = U \oplus W_1 \oplus W_2 \oplus W_3$.

Solution: $W_1 = (0, 0, a, 0, 0), W_2 = (0, 0, 0, b, 0), W_3 = (0, 0, 0, 0, c)$

Problem (Exercise 23): Prove or give a counterexample: if U_1, U_2, W are subspaces of V such that

$$V = U_1 \oplus W \text{ and } V = U_2 \oplus W,$$

then $U_1 = U_2$.

Solution: $W = (x, x), U_1 = (a, 0), U_2 = (0, b)$

Problem (Exercise 24): U_o is the set of real-valued functions. U_e is defined similarly. Show that $\mathbb{R}^{\mathbb{R}} = U_o \oplus U_e$.

Solution: Note that both U_o and U_e are subspaces, and that $U_o \cap U_e = \{0\}$. Note that for any $f \in \mathbb{R}^{\mathbb{R}}$, we can write an even function $e(x) = \frac{f(x)+f(-x)}{2}$ and an odd function $o(x) = \frac{f(x)-f(-x)}{2}$, so $U_o + U_e = \mathbb{R}^{\mathbb{R}}$. But we have $U_o \cap U_e = \{0\}$, so $U_o + U_e$ is a direct sum, so we're done.

2. Finite Dimensional Vector Spaces

2.1. Span and Linear Independence

2.1.1. Linear Combinations and Span

Definition (Linear Combination): A linear combination of a list $v_1, \ldots, v_m$ of vectors in V is a vector of the form

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

where $a_i \in \mathbb{F}$.

Example: $(17, -4, 2)$ is a linear combination of $(2, 1, -3), (1, -2, 4)$ while $(17, -4, 5)$ is not.

Definition (Span): The set of all linear combinations of a list of vectors v_i in V is called the span of v_i, denoted by $\operatorname{span}(v_1, \ldots, v_m)$. In other words,

$$\operatorname{span}(v_1, \ldots, v_m) = \left\{\sum a_i v_i : a_i \in \mathbb{F}\right\}.$$

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

Proposition (Span is the smallest containing subspace): The span of a list of vectors in V is the smallest subspace of V containing all vectors in the list.

Proof: Suppose v_i is a list of vectors in V. Note that the span is indeed a subspace since 0 is in the span and the same is closed under addition and scalar multiplication. Note that each v_k is also in the span.

Because subspaces are closed under scalar multiplication and addition, every subspace of V that contains each v_k contains $\operatorname{span}(v_1, \ldots, v_m)$. Thus $\operatorname{span}(v_1, \ldots, v_m)$ is the smallest subspace of V containing all vectors v_i. ■

Definition: If $\operatorname{span}(v_1, \ldots, v_m)$ equals V, we say that the list $v_1, \ldots, v_m$ *spans* V.

Example:

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

spans $\mathbb{F}^n$, where there are n vectors in the list.

Definition (Finite-dimensional vector space): A vector space is called finite-dimensional if some list of vectors in it spans the space.

Definition: $\mathcal{P}(\mathbb{F})$ is the set of all polynomials with coefficients in $\mathbb{F}$. $\mathcal{P}_m(\mathbb{F})$ denotes the set of all polynomials with coefficients in $\mathbb{F}$ and degree at most m.

Definition (Infinite dimensional vector space): A vector space is called infinite-dimensional if it is not finite-dimensional.

2.1.2. Linear Independence

Definition (Linearly independent): A list $v_1, ..., v_m$ of vectors in V is called linearly independent if the only choice of $a_1, ..., a_m \in \mathbb{F}$ that makes

$$\sum a_i v_i = 0$$

is $a_i = 0$ for all i. The empty list is also declared to be linearly independent.

Proposition: If some vectors are removed from a linearly independent list, the remaining list is also linearly independent.

Proof: If a list is linearly independent, then the only a_is that work are all 0. Suppose we remove some of the vectors. If the new list wasn't linearly dependent, then we could just drop 0s in front of the vectors we got rid of and have a non linearly independent list of vectors, which is a contradiction. ■

Lemma (Linear dependence lemma): Suppose $v_1, ..., v_m$ is a linearly dependent list in V. Then there exists $k \in \{1, 2, ..., m\}$ such that

$$v_k \in \operatorname{span}(v_1, ..., v_{k-1}).$$

Furthermore, if k satsfies the condition above and the kth term is removed from $v_1, ..., v_m$, then the span of the remaining list equals $\operatorname{span}(v_1, ..., v_m)$.

Proof: Because the list $v_1, ..., v_m$ is linearly dependent, there exist numbers $a_1, ..., a_m \in \mathbb{F}$, not all 0, such that

$$\sum a_i v_i = 0.$$

Let k be the largest element of $\{1, ..., m\}$ such that $a_k \neq 0$. Then

$$v_k = -\frac{a_1}{a_k} v_1 - \cdots - \frac{a_{k-1}}{a_k} v_{k-1},$$

which proves that $v_k \in \operatorname{span}(v_1, ..., v_{k-1})$, as desired.

Now suppose k is any element of $\{1, ..., m\}$ such that $v_k \in \operatorname{span}(v_1, ..., v_{k-1})$. Let $b_1, ..., b_{k-1} \in \mathbb{F}$ such that

$$v_k = b_1 v_1 + \cdots + b_{k-1} v_{k-1}.$$

Suppose $u \in \operatorname{span}(v_1, ..., v_m)$. Then there exist $c_1, ..., c_m \in \mathbb{F}$ such that

$$u = c_1 v_1 + \cdots + c_m v_m.$$

In the equation above, we can replace v_k with the right side of the equation two above, which shows that u is in the span of the list obtained by removing the kth term from $v_1, ..., v_m$. Thus removing the kth term of the list $v_1, ..., v_m$ does not change the span of the list. ■

Lemma (length of linearly independent list $\leq$ length of spanning list): 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.

Proof: Suppose $u_1, ..., u_m$ is linearly dependent in V and $w_1, ..., w_n$ spans V.

First we append u_1 to the list of w_is. Since u_1 can be written as a linear combination of the w_i s, the new list is linearly dependent. By the linear dependence lemma, we can now take out one of the w_is (we can't take out u_1 since it's not 0).

We can continue this idea, appending u_k right after a_{k-1} and right before the w_is. Since the first k vectors are linearly independent, by the linear dependence lemma, we can remove one of the w_is and still have the same span. After adding all m vectors, the process stops. At each step as we add a u_i, the linear dependence lema implies that there is some w to remove. Thus there are at least as many w's as u's. ■

Proposition (Finite-dimensional subspaces): Every subspace of a finite-dimensional vector space is finite-dimensional.

Proof: Suppose V is finite-dimensional and U is a subspace of V.
- Step 1: If $U = \{0\}$, then U is finite dimmensional and we are done. If $U \neq \{0\}$, then choose a nonzero vector $u_i \in U$.
- Step k: If $U = \operatorname{span}(u_1, ..., u_{k-1})$, then we are done. If $U \neq \operatorname{span}(u_1, ..., u_{k-1})$, then a choose a vector $u_k \in U$ such that $u_k \notin \operatorname{span}\ (u_1, ..., u_k)$.

After each step, as long as the process continues, we have constructed a list of vectors such that no vector in this list is in the span of the previous vectors. Thus after each step we have constructed a linearly independent list. This linearly independent list cannot be longer than any spanning list of V. Thus the process eventually terminates, which means U is finite-dimensional. ■

2.1.3. Problems

Problem (Exercise 1): Find a list of four distinct vectors in $\mathbb{F}^3$ whose span equals

$$\{(x, y, z) \in \mathbb{F}^3 : x + y + z = 0\}.$$

Solution: $(1, -1, 0), (-1, 1, 0), (0, -1, 1), (0, 1, -1)$

Problem (Exercise 2): Prove or give a counterexample: if v_1, v_2, v_3, v_4 spans V, then the list

$$v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$$

also spans V.

Solution: Note that any linear combination of the four vectors given is equal to a linear combination of the four original vectors, so the new list spans V.

Problem (Exercise 3): Suppose $v_1, ..., v_m$ is a list of vectors in V. For $k \in \{1, ..., m\}$, let

$$w_k = v_1 + \cdots + v_k.$$

Show that $\operatorname{span}(v_1, ..., v_m) = \operatorname{span}(w_1, ..., w_m)$.

Solution:

$$\sum_{i=1}^{m} a_i w_i = \sum_{i=1}^{m} a_i \left(\sum_{j=1}^{i} v_j \right) = \sum_{k=1}^{m} \left(\sum_{i=1}^{m+1-k} a_{m+1-i} \right) v_k = \sum_{k=1}^{m} b_k v_k,$$

so any linear combination of w_is is a linear combination of v_is.

Problem (Exercise 11): Prove or give a counterexample: if $v_1, ..., v_m$ and $w_1, ..., w_m$ are linearly independent lists of vectors in V, then the list $v_1 + w_1, ..., v_m + w_m$ is linearly independent.

Solution: Take $w_i = -v_i$.

Problem (Exercise 17): Prove that V is infinite-dimensional if and only if there is a sequence $v_1, v_2, ...$ of vectors in V such that $v_1, ..., v_m$ is linearly independent for every positive integer m.

Solution: First suppose such a sequence exists. Then we know that $v_m \notin \operatorname{span}(v_1, ..., v_{m-1})$, since otherwise the list that included v_m would not be linearly independent. Thus, no list could span the entire space, which implies it is infinte-dimensional.

Now suppose the space is infinite-dimensional. This means no list spans the entire space. Thus, if we have a linearly independent list of size $m - 1$, there is some vector in V that's not in the span of the list, and appending that creates a new linearly independent list, which can keep going.

2.2. Bases

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

Proposition (Criterion for basis): A list $v_1, \ldots, v_n$ of vectors in V is a basis of V if and only if every $v \in V$ can be written uniquely in the form

$$a_1 v_1 + \cdots + a_n v_n,$$

where $a_1, \ldots, a_n \in \mathbb{F}$.

Proof: Basically uses ideas that led to linear independence. ■

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

Proof: If a vector v_k is in the span of the vectors $v_1, \ldots, v_{k-1}$, discard it. Keep doing this until you reach the end. The new list clearly must still span V, and by the linear dependence lemma, the new list is linearly independent, which means it's a basis. ■

This immediately implies every finite-dimensional vector space has a basis.

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

Proof: Adjoin a spanning list to the vectors, then use proposition 2.2.2 to reduce to a basis (none of the original vectors get deleted since they were linearly independent). ■

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

Proof: Because V is finite-dimensional, so is U. Thus there is a basis $u_1, \ldots, u_m$ of U. We can extend this list to a basis of V: $u_1, \ldots, u_m, w_1, \ldots, w_n$. Let $W = \operatorname{span}(w_1, \ldots, w_n)$.

To show $V = U + W$, note that for $v \in V$ we have

$$v = \sum a_i u_i + \sum b_i w_i.$$

The first sum is a vector in U, and the second is a vector in V, so we have $v \in U + W$, which means $V = U + W$.

Now suppose $v \in U \cap W$. Then we have

$$v = \sum a_i u_i = \sum b_i w_i \Rightarrow \sum a_i u_i - \sum b_i w_i = 0.$$

Since the list $u_1, ..., u_m, w_1, ..., w_n$ is linearly independent, this implies $a_i, b_i = 0$ for all i, so $U \cap W = \{0\}$. ∎

A simple consequence of this is that the extension of the list defines W needed to get $U \oplus W = V$.

2.2.1. Problems

Problem (Exercise 3): Let U be the subspace of $\mathbb{R}^5$ defined by

$$U = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 : x_1 = 3x_2 \text{ and } x_3 = 7x_4\}.$$

Find a basis of U. Then extend the basis to a basis in $\mathbb{R}^5$, then find a subspace W of $\mathbb{R}^5$ such that $\mathbb{R}^5 = U \oplus W$.

Solution: A basis is $(3, 1, 0, 0, 0), (0, 0, 7, 1, 0), (0, 0, 0, 0, 1)$. This is clearly linearly independent and clearly spans the subspace. We can extend the basis to

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

as a basis of $\mathbb{R}^5$. $W = \{(x, 0, 0, y, 0)\}$ is a W that works.

Problem (Exercise 4): Let U be the subspace of $\mathbb{C}^5$ defined by

$$U = \{(z_1, z_2, z_3, z_4, z_5) \in \mathbb{C}^5 : 6z_1 = z_2 \text{ and } z_3 + 2z_4 + 3z_5 = 0\}.$$

Repeat the procedure in the last problem.

Solution: A basis is $(1, 6, 0, 0, 0), \left(0, 0, 0, 1, -\frac{2}{3}\right), \left(0, 0, 1, 0, -\frac{1}{3}\right)$. We can extend this to

$$(1, 6, 0, 0, 0), \left(0, 0, 0, 1, -\frac{2}{3}\right), \left(0, 0, 1, 0, -\frac{1}{3}\right), (1, 0, 0, 0, 0), (0, 0, 0, 0, 1).$$

A W that work is $W = \{(x, 0, 0, 0, y)\}$.

Problem (Exercise 5): Suppose V is finite-dimensional and U, W are subspaces of V such that $V = U + W$. Prove that there exists a basis of V consisting of vectors in $U \cup W$.

Solution: Append a basis of W to a basis of V. Since $V = U + W$, any vector in V can be represented as a linear combination of vectors from this list, so the list spans V. Then just reduce the list down to a basis.

Problem (Exercise 6): Prove or give a counterexample: If $p_0, ..., p_3 \in \mathcal{P}_3(\mathbb{F})$ such that none of the p_i's have degree 2, then the p_i's are not a basis of $\mathcal{P}_3(\mathbb{F})$.

Solution:

$$1, x, x^3, x^3 + x^2$$

Problem (Exercise 7): Suppose v_1, v_2, v_3, v_4 is a basis of V. Prove that

$$v_1 + v_2, v_2 + v_3, v_3 + v_4, v_4$$

is also a basis of V.

Solution: Clearly the list spans V, since any linear combination of these vectors is just a linear combination of v_is. It also must be linearly independent since the v_is are linearly independent.

Problem (Exercise 11): Suppose V is a real vector space. Show that if $v_1, ..., v_n$ is a basis of V as a real vector space, then $v_1, ..., v_n$ is also a basis of $V_{\mathbb{C}}$.

Solution: Let $u + iv \in V_{\mathbb{C}}$. Then we can write

$$\sum a_j v_j + i \sum b_j v_j = \sum (a_j + i b_j) v_j.$$

Thus it spans $V_{\mathbb{C}}$. It's also clearly linearly independent, so it's a basis of $V_{\mathbb{C}}$.

2.3. Dimension

Proposition (basis length does not depend on basis): Any two bases of a finite-dimensional vector space have the same length.

Proof: Follows from len(linearly independent) $\leq$ len(spanning). ■

Definition (dimension): The dimension of a finite-dimensional vector space, denoted as $\dim V$, is the length of any basis of the vector space.

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

Proof: Pick a basis of U. This basis must also be linearly independent in V, so it can be extended to a basis of V, giving the desired inequality. ■

Proposition: Suppose V is finite-dimensional. Then evert linearly independent list of vectors in V of length $\dim V$ is a basis of V.

Proof: Suppose $\dim V = n$ and $v_1, ..., v_n$ is a linearly independent in V. Then it can be extended to a basis, but since a basis must have length n, no elements need to be added. Thus, the list is already a basis. ■

Corollary: Suppose that V is a finite-dimensional subspace and U is a subspace of V such that $\dim U = \dim V$. Then $U = V$.

Proof: Let $u_1, ..., u_n$ be a basis of U. Thus $n = \dim U = \dim V$. Thus $u_1, ..., u_n$ is a linearly independent list of vectors in V of length $\dim V$. Thus the list is a basis of V, which means every vector of V is a linear combination of vectors in the list, which means $U = V$. ■

Proposition: Suppose V is finite-dimensional. Then every spanning list of vectors in V of length $\dim V$ is a basis of V.

Proof: Suppose $\dim V = n$ and $v_1, ..., v_n$ spans V. The list $v_1, ..., v_n$ can be reduced to a basis of V. However, every basis of V has length n, so the reduction is trivial, thus the list is a basis of V. ■

Lemma (dimension of a sum): If V_1 and V_2 are subspaces of a finite-dimensional vector space, then

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

Proof: Let $v_1, ..., v_n$ be a basis of $V_1 \cap V_2$, so the $\dim(V_1 \cap V_2) = m$. Because v_is is a basis of $V_1 \cap V_2$, it is linearly independent in V_1, so we can extend it to a basis of V_1: $v_1, ..., v_n, u_1, u_j$. Similarly, extend it to a basis of V_2: $v_1, ..., v_n, w_1, ..., w_k$. Thus, $\dim V_1 = m + j$ and $\dim V_2 = m + k$. Showing that $v_i \cup u_i \cup w_i$ is basis of $V_1 + V_2$ will complete the proof.

Note that each vector in the list will be contained in $V_1 + V_2$, and that since part of the list is a basis of V_1 and the other part is a basis of V_2, every vector in $V_1 + V_2$ can be represented by the list, which means that list spans $V_1 + V_2$. All that remains to be shown is that the list is linearly independent.

Suppose

$$\sum a_i v_i + \sum b_i u_i + \sum c_i w_i = 0.$$

We can rewrite this as

$$\sum c_i w_i = -\sum a_i v_i - \sum b_i u_i.$$

The right side is a vector in V_1, and the left side is a linear combination of vectors in V_2, so both sides are in their intersection. Thus we can write

$$\sum c_i w_i = \sum d_i v_i.$$

However, v_i, w_i is the basis of V_2, so it's linearly independent, implying $c_i = d_i = 0$. Thus we have

$$\sum a_i v_i + \sum b_i u_i = 0.$$

Since these vectors are linearly independent in V_1, $a_i = b_i = 0$. ■

2.3.1. Problems

Problem (Exercise 9): Suppose m is a positive integer and $p_0, ..., p_m \in \mathcal{P}(\mathbb{F})$ are such that p_k has degree k. Prove that $p_0, ..., p_m$ is a basis of $\mathcal{P}_m(\mathbb{F})$.

Solution: Note that $\dim \mathcal{P}_m(\mathbb{F}) = m + 1$. Thus we just need to show that the list is linearly independent. This is easy, since the only way they can sum to 0 is if p_m's coefficient is 0 (to get rid of x^m) and so on.

Problem (Exercise 11): Suppose $\dim U = \dim W = 4$ are subspaces of $\mathbb{C}^6$. Prove that there are two vectors in $U \cap W$ such that they are linearly independent.

Solution: Note that

$$6 \geq \dim(U + W) = 8 - \dim(U \cap W),$$

so $U \cap W$ must have dimension at least 2, implying the result.

Problem (Exercise 14): Suppose $\dim V = 10$ and $\dim V_1 = \dim V_2 = \dim V_3 = 7$ are subspaces of V. Prove that $V_1 \cap V_2 \cap V_3 \neq \{0\}$.

Solution: First we have

$$\dim(V_1 \cap V_2 + V_3) = \dim(V_1 \cap V_2) + \dim V_3 - \dim(V_1 \cap V_2 \cap V_3).$$

Note that $\dim(V_1 + V_2) = 14 - \dim(V_1 \cap V_2)$. Plugging this in, we have

$$\dim(V_1 \cap V_2 + V_3) = 21 - \dim(V_1 + V_2) - \dim(V_1 \cap V_2 \cap V_3).$$

Note that the left side is at most 10, so we have

$$\dim(V_1 + V_2) + \dim(V_1 \cap V_2 \cap V_3) \geq 11.$$

If the second term on the left was zero, then $\dim(V_1 + V_2)$ would need to be at least 11, which is impossible since they are both subspaces of a 10 dimensional vector space. Thus $\dim(V_1 \cap V_2 \cap V_3) \geq 1$, which implies the intersection is not equal to $\{0\}$.

Problem (Exercise 17): Suppose that $V_1, ..., V_m$ are finite-dimensional subspaces of V. Prove that

$$\dim(V_1 + \cdots + V_m) \leq \dim V_1 + \cdots + \dim V_m.$$

Solution: Easy by induction.

Problem (Exercise 20): Prove that if V_1, V_2, and V_3 are finite-dimensional subspaces of a vector space, then

$$\dim(V_1 + V_2 + V_3) = \dim V_1 + \dim V_2 + \dim V_3$$
$$- \frac{\dim(V_1 \cap V_2) + \dim(V_2 \cap V_3) + \dim(V_3 \cap V_1)}{3}$$
$$- \frac{\dim((V_1 + V_2) \cap V_3) + \dim((V_2 + V_3) \cap V_1) + \dim((V_3 + V_1) \cap V_2)}{3}.$$

Solution: Note that

$$\dim(V_1 + V_2 + V_3) = \dim(V_1 + V_2) + \dim V_3 - \dim((V_1 + V_2) \cap V_3) =$$
$$\dim V_1 + \dim V_2 - \dim(V_1 \cap V_2) + \dim V_3 - \dim((V_1 + V_2) \cap V_3).$$

Sum cyclically and divide by 3 to get the desired result.

3. Linear Maps

General maps T are assumed to be in $\mathcal{L}(V, W)$.

Basis of V and W are $v_1, ..., v_n$ and $w_1, ..., w_m$ unless stated otherwise.

3.1. Vector Space of Linear Maps

3.1.1. Definition of Linear Maps

Definition (Linear map): A linear map from V to W is a function $T : V \to W$ such that $T(u + v) = Tu + Tv$ and $T(\lambda v) = \lambda(Tv)$. The set of linear maps from V to W is denoted by $\mathcal{L}(V, W)$, and if $W = V$, then it is denoted as $\mathcal{L}(V)$.

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

$$Tv_k = w_k$$

for each k.

Proof: First existence. Define $T : V \to W$ by

$$T(c_1 v_1 + \cdots + c_n v_n) = c_1 w_1 + \cdots + c_n w_n,$$

where c_i are arbitrary scalars. Since the vs form a basis, this is indeed a function from V to W. Setting $c_k = 1$ and everything else to 0 for each k shows $Tv_k = w_k$. Showing this is a linear map is very easy.

Now uniquness. Suppose $T \in \mathcal{L}(V, W)$ and that $Tv_k = w_k$. Then we have $T(c_k v_k) = c_k w_k$. Additivity implies

$$T(c_1 v_1 + \cdots + c_n v_n) = c_1 w_1 + \cdots + c_n w_n.$$

Thus T is uniquely determined on $\operatorname{span}(v_1, ..., v_n)$ by the equation above. Since the vs are a basis, T is uniquely determined on V. ■

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

Definition: Suppose $S, T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbb{F}$. Then $S + T$ is defined as

$$(S+T)(v) = Sv + Tv$$

and λT is defined as

$$(\lambda T)(v) = \lambda(Tv).$$

This definition of addition and scalar multiplication makes $\mathcal{L}(V, W)$ a vector space, which is easy to verify.

Definition (product of linear maps): If $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$, then the product $ST \in \mathcal{L}(U, W)$ is defined by

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

for all $u \in U$.

Associativity, identity, and distributivity all apply to the products of linear maps.

Proposition: Suppose T is a linear maps. Then $T(0) = 0$.

Solution:

$$T(0) = T(0+0) = T(0) + T(0) \Rightarrow T(0) = 0.$$

3.1.3. Problems

Problem (Exercise 4): Suppose $T \in \mathcal{L}(V, W)$ and $v_1, ..., v_m$ is a list of vectors in V such that $Tv_1, ..., Tv_m$ is a linearly independent list in W. Prove that $v_1, ..., v_m$ is linearly independent.

Solution:

$$\sum c_i T v_i = 0 \Rightarrow T\left(\sum c_i v_i\right) = 0.$$

Problem (Exercise 7): Prove that if $\dim V = 1$ and $T \in \mathcal{L}(v)$, then there exists $\lambda \in \mathbb{F}$ such that $Tv = \lambda v$ for all $v \in V$.

Solution: Let w be the basis of V. Then every vector in V can be written as cw for some $c \in \mathbb{F}$. Suppose $Tw = c_0 w$. Then multiplying by any scalar yields $Tv = c_0 v$ where $v \in V$, so we're done.

Problem (Exercise 11): Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T is a scalar multiple of the idenityt if and only if $ST = TS$ for all $S \in \mathcal{L}(V)$.

Solution: Suppose T is a scalar multiple of the identity. Then the equation easily follows. Now suppose T is not a scalar multiple of the identity. There exists v such that $Tv = u \neq \lambda v$ for any $\lambda \in \mathbb{F}$. Thus, u and v are linearly independent. The result holds trivially for $\dim V = 1$, so assume $\dim V \geq 2$. Then u and v are part of a basis of V. Let $Sv = v$ and $Su = 0$. Then

$$S(Tv) = Su = 0 \neq u = Tv = T(Sv).$$

Problem (Exercise 13): Show that a linear map can be extended from a subspace of V to V.

Solution: Let U be the subspace. Let $Tu_i = w_i$ be the outputs of the basis vectors of U. Extend the basis of U to V. Let the new vectors in the basis map to any vector in the output vector space, done.

3.2. Null Spaces and Ranges

3.2.1. Null Space and Injectivity

Definition (null space): For $T \in \mathcal{L}(V, W)$, the null space if T, denoted null T, is

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

Proposition: null T is a subspace of V.

Proof: Since $T(0) = 0$, $0 \in \text{null } T$. Note that if $u, v \in \text{null } T$ then $T(u + v) = Tu + Tv = 0$, and similarly with scalars, so it is indeed a subspace. ■

Definition (injective): Standard definition for injective functions.

Proposition: T is injective if and only if null $T = \{0\}$.

Proof: First suppose T is injective. We already know that $0 \in \text{null } T$. Now suppose $v \in \text{null } T$. Then $Tv = 0 = T(0)$, which implies that $v = 0$.

Now suppose null $T = 0$. Pick u, v such that $Tu = Tv$. Then $T(u - v) = 0$, which implies $u - v \in \text{null } T$, which implies $u = v$. ■

3.2.2. Range and Surjectivity

Definition (range): The range of T is the following subset of W:

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

Proposition: range T is a subspace of W.

Proof: Cleary $0 \in \text{range } T$. Note that for $w_1, w_2 \in W$ we have

$$T(v_1 + v_2) = Tv_1 + Tv_2 = w_1 + w_2$$

for some v_1, v_2, so $w_1 + w_2$ is in range T. Similarly for scalars. ■

Definition (surjective): $T \in \mathcal{L}(V, W)$ is surjective if range $T = W$.

3.2.3. Fundamental Theorem of Linear Maps

Theorem (fundamental theorem of linear maps): Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Then range T is finite-dimensional and

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

Proof: Let $u_1, \ldots, u_m$ be a basis of null T, extend it to a basis of V:

$$u_1, \ldots, u_m, v_1, \ldots, v_n.$$

Thus we need to show $\dim \operatorname{range} T = n$. We show that Tv_i form a basis of range T.

Pick $v \in V$. We can write it as

$$v = \sum c_i u_i + \sum d_i v_i.$$

Applying T to both sides yields

$$Tv = \sum c_i T v_i,$$

where all the u's disappear because they are in null T. This equation implies Tv_i spans range T, which also implies range T is finite dimensional.

To show the list is linearly independent, we have

$$\sum c_i T v_i = 0 \Rightarrow \sum c_i v_i \in \operatorname{null} T.$$

Thus we have

$$\sum c_i v_i = \sum d_i u_i,$$

and since $u_1, \ldots, u_m, v_1, \ldots, v_m$ is linearly independent, $c_i = d_i = 0$, done. ■

Proposition: Suppose V and W are finite-dimensional vector spaces.
- If $\dim V > \dim W$, then there is no injective linear map from V to W.
- If $\dim V < \dim W$, then there is no surjective linear map from V to W.

Proof:
- $\dim \operatorname{null} T = \dim V - \dim \operatorname{range} T \geq \dim V - \dim W > 0$.
- $\dim \operatorname{range} T = \dim V - \dim \operatorname{null} T \leq \dim V < \dim W$.

■

3.2.4. Problems

Problem (Exercise 11): Suppose V is finite dimensional. Prove that there exists a subspace U of V such that

$$U \cap \operatorname{null} T = \{0\} \quad \text{and} \quad \operatorname{range} T = \{Tu : u \in U\}.$$

Solution: Take U such that $V = U \oplus \operatorname{null} T$, which exists and satisfies the first condition. Let the basis of null T be represented by N, and extend it to a basis of V with the list M. Note that $\operatorname{span}(M) = U$. Thus, any vector in V will have the vectors in N that constitute it become 0, while every vector in M will map to Tu, implying range T is equal to the range of T on U.

Problem (Exercise 12): Suppose T is a linear map from $\mathbb{F}^4$ to $\mathbb{F}^2$ such that

$$\operatorname{null} T = \{(x_1, x_2, x_3, x_4) \in \mathbb{F}^4 : x_1 = 5x_2 \text{ and } x_3 = 7x_4\}.$$

Prove that T is surjective.

Solution: Note that $\dim \operatorname{null} T = 2$. Since $\dim \mathbb{F}^4 = 4$, we have $\dim \operatorname{range} T = 2$. Since range $T = 2$, $\dim \mathbb{F}^2 = 2$, and since range T is a subspace of $\mathbb{F}^2$, we must have range $T = \mathbb{F}^2$, so T is surjective.

Problem (Exercise 15): Suppose there exists a linear map on V whose null space and range are both finite-dimensional. Prove that V is finite-dimensional.

Solution: Let T be the map. Let $v_1, ..., v_m$ be the basis of null T, and let $w_1, ..., w_n$ be the basis of range T. Since w_i's are linearly independent, we have that $u_1, ..., u_n$ are linearly independent, where $Tu_i = w_i$. Now pick some vector $v \in V$ not in null T. We have $Tv \in \operatorname{range} T$, so $Tv = \sum a_i w_i$ for some choice of a's. Replacing the w's with Tu's yields $Tv = T(\sum a_i u_i)$. Since $v \notin \operatorname{null} T$, we've essentially changed T to T', where $\operatorname{null} T' = \{0\}$, so we have $v = \sum a_i v_i$, which implies there exists a spanning list for V, which means it's finite-dimensional.

Problem (Exercise 21): Suppose V is finite-dimensional, $T \in \mathcal{L}(V, W)$, and U is a subspace of W. Prove that $\{v \in V : Tv \in U\}$ is a subspace of V and

$$\dim\{v \in V : Tv \in U\} = \dim \operatorname{null} T + \dim(U \cap \operatorname{range} T).$$

Solution: Denote the set as U'. Checking it's a subspace is easy. Note that null $T \subseteq U'$, since $0 \in U$. Note also that $U \cap \operatorname{range} T$ is the subspace give by just applying T to U'. Now define $T' \in \mathcal{L}(U', W)$ such that $T'v = Tv$ for $v \in U'$ (essentially restricting T to U'). We have $\operatorname{null} T = \operatorname{null} T'$ and $\operatorname{range} T' = U \cap \operatorname{range} T$, so applying the fundamental theorem of linear maps to T' yields

$$\dim U' = \dim \operatorname{null} T + \dim(U \cap \operatorname{range} T).$$

Problem (Exercise 22): Suppose U and V are finite dimensional vector spaces and $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(U, V)$. Prove that

$$\dim \operatorname{null} ST \leq \dim \operatorname{null} S + \dim \operatorname{null} T.$$

Solution: Note that the basis of null T will be in null ST. Note also that at most $\dim \operatorname{null} S$ vectors can be wrriten as an output of T, so those vectors contribute at most $\dim \operatorname{null} S$ to $\dim \operatorname{null} ST$, yielding the desired inequality.

Problem (Exercise 23): Same conditions as last problem. Prove that

$$\dim \operatorname{range} ST \leq \min\{\dim \operatorname{range} S, \dim \operatorname{range} T\}.$$

Solution: If $\dim \operatorname{range} T \leq \dim \operatorname{range} S$, then $\dim \operatorname{range} S|_{\dim \operatorname{range} T} \leq \dim \operatorname{range} T$ by the fundamental theorem of linear maps. If $\dim \operatorname{range} S \leq \dim \operatorname{range} T$, then clearly the inequality holds.

Problem (Exercise 24):
(a) Suppose $\dim V = 5$ and $S, T \in \mathcal{L}(V)$ are such that $ST = 0$. Prove that $\dim \operatorname{range} TS \leq 2$.
(b) Give an example of $S, T \in \mathcal{L}(\mathbb{F}^5)$ with $ST = 0$ and $\dim \operatorname{range} TS = 2$.

Solution:
(a) Note that the condition implies $\dim \operatorname{null} ST = 5 \leq \dim \operatorname{null} S + \dim \operatorname{null} T \Rightarrow \dim \operatorname{range} S + \dim \operatorname{range} T \leq 5$. Then applying the previous exercise yields the desired result.
(b) $T(a, b, c, d, e) = (0, a, 0, c, 0)$, $S(a, b, c, d, e) = (a, 0, c, 0, e)$

3.3. Matrices

3.3.1. Representing a Linear Map by a Matrix

Definition (matrix): An m-by-n matrix is a rectangular array of elements of $\mathbb{F}$ with m rows and n columns:

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

$A_{j,k}$ denote the entry in row j, column k of A.

Definition (matrix of a linear map): Suppose $T \in \mathcal{L}(V, W)$ and $v_1, ..., v_n$ is a basis of V and $w_1, ..., w_m$ 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 are defined by

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

If the bases are note clear from context, then the notation $\mathcal{M}(T, (v_1, ..., v_n), (w_1, ..., w_m))$ is used.

3.3.2. Addition and Scalar Multiplication of Matrices

Definition (matrix addition and scalar multiplication): Matrices of the same dimensions are added pointwise, and a matrix being multiplied by a scalar has the scalar distributed to all entries.

Proposition: $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$ and $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

Proof: For the the first one, note that on a basis vector v_k, $(S + T)v_k$ will be equal to the sum of the representations of Tv_k and Sv_k by a basis of W. Thus their coefficients add up, which corresponds to the adding of elements in the matrices pointwise. Similarly for scalar mulitplication. ■

Proposition: Suppose m and n are positibe integers. Then $\mathbb{F}^{m,n}$, the set of all m by n matrices with entries in $\mathbb{F}$, is a vector space of dimension mn.

Proof: It's clear that $\mathbb{F}^{m,n}$ is a vector space. For the dimension, note that the list of the matrices that have 1 in one slot and 0 in the rest for all slots spans the space and is linearly independent. ∎

3.3.3. Matrix Multiplication

Definition (matrix multiplication): Suppose A is an m by n matrix and B is an n by p matrix. Then AB is defined to be the m by p matrix whose entry in row j, column k, is given by

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

Thus the entry in row j, column k, of AB is computed by taking row j of A and column k of B, multiplying together corresponding entries, and then summing.

Proposition: If $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$, then $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.

Proof: This follows because matrix multiplication was defined for this to be true. ∎

Definition ($A_{j,\cdot}$, $A_{\cdot,k}$): Suppose A is an m by n matrix. $A_{j,\cdot}$ represents the 1 by n matrix consisting of row j of A. $A_{\cdot,k}$ represents the 1 by k matrix consisting of column k of A.

Proposition: Suppose A is an m by n matrix and B is an n by p matrix. Then

$$(AB)_{j,k} = A_{j,\cdot} \cdot B_{\cdot,k}$$

Proof: By definition, we have

$$(AB)_{j,k} = A_{j,1}B_{1,k} + \cdots + A_{j,n}B_{n,k}.$$

By definition this is equal to $A_{j,\cdot} \cdot B_{\cdot,k}$. ∎

Proposition: Suppose A is an m by n matrix and B is an n by p matrix. Then

$$(AB)_{\cdot,k} = AB_{\cdot,k}$$

Proof: The entry in row j of $(AB)_{\cdot,k}$ is the left side of the previous result and the entry in row j is the right side if the previous result. ∎

Proposition (linear combination of columns): Suppose A is an m by n matrix and $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$. Then

$$Ab = b_1 A_{\cdot,1} + \cdots + b_n A_{\cdot,n}$$

Proof: The entry in row k of the m by 1 product is by definition

$$A_{k,1} b_1 + \cdots + A_{k,n} b_n.$$

This holds for all k, so we can replace $A_{k,i}$ with $A_{\cdot,i}$. ■

Analogous results of the last two propositions hold for the rows of matrices.

Proposition: Suppose C is an m by c matrix and R is a c by n matrix.
- If $k \in \{1, ..., n\}$, then column k of CR is a linear combination of the columns of C, with the coefficients of this linear combination coming from column k of R.
- If $k \in \{1, ..., m\}$, then row j of CR is a linear combination of the rows of R, with the coefficients of this linear combination coming from row j of C.

Proof: Column k of CR equals $CR_{\cdot,k}$ by 3.3.3.3, which equals the linear combination of the columns of C with coefficients coming from $R_{\cdot,k}$ by 3.3.3.4. Proving the second bullet uses the analogous row results. ■

3.3.4. Column-Row Factorization and Rank of a Matrix

Definition (column rank, row rank): Suppose A is an m by n matrix with entries in $\mathbb{F}$.
- The column rank of A is the dimension of the span of the columns of A in $\mathbb{F}^{m,1}$.
- The row rank of A is the dimension of the span of the rows of A in $\mathbb{F}^{1,n}$.

Example: Suppose

$$A = \begin{pmatrix} 4 & 7 & 1 & 8 \\ 3 & 5 & 2 & 9 \end{pmatrix}.$$

The column rank of A is the dimension of

$$\operatorname{span}\left(\begin{pmatrix} 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 8 \\ 9 \end{pmatrix}\right)$$

in $\mathbb{F}^{2,1}$. The first two vectors are linearly independent, so the span has dimension at least two. Since $\dim \mathbb{F}^{2,1} = 1$, the span of the list must be 2, so the column rank of A is two.

The row rank of A is the dimension if

$$\operatorname{span}((4 \;\; 7 \;\; 1 \;\; 8), (3 \;\; 5 \;\; 2 \;\; 9))$$

in $\mathbb{F}^{1,4}$. Both vectors are linearly independent, so the dimension of the span is 2.

Definition (transpose): The transpose of a matrix A, denoted by A^t, is the matrix obtained from A by interchanging rows and columns. 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

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

Example: If $A = \begin{pmatrix} 5 & -7 \\ 3 & 8 \\ -4 & 2 \end{pmatrix}$, then $A^t = \begin{pmatrix} 5 & 3 & -4 \\ -7 & 8 & 2 \end{pmatrix}$.

Proposition (column-row factorization): Suppose A is an m by n matrix with entries in $\mathbb{F}$ and column rank $c \geq 1$. Then there exist an m by c matrix and a c by n matrix R, both with entries in $\mathbb{F}$, such that $A = CR$.

Proof: Each column of A is an m by 1 matrix. The list $A_{\cdot,1}, \ldots, A_{\cdot,n}$ of columns of A can be reduced to a basis of the span of the columns of A. This basis has length c by the definition of column rank. The c columns in this basis can be put together to form an m by c matrix C.

If $k \in \{1, \ldots, n\}$, then column k of A is a linear combination of the columns of C. Make the coefficients of this linear combination into column k of a c by n matrix that we call R. Then $A = CR$. ■

Proposition (column rank equals row rank): Suppose $A \in \mathbb{F}^{m,n}$. Then the column rank of A equals the row rank of A.

Proof: Let c denote the column rank of A. Let $A = CR$ be the column-row factorization of A, where C is an m by c matrix and R is a c by n matrix. We know every row of A is a linear combination of the rows of R. Because R has c rows, this implies that the row rank of A is less than or equal to the column rank c of A. To prove the inequality in the other direction, apply the result in the previous paragraph to A^t, getting

$$\begin{aligned} \text{column rank of } A &= \text{row rank of } A^t \\ &\leq \text{column rank of } A^t \\ &= \text{row rank of } A. \end{aligned}$$

Thus the column rank equals the row rank. ■

3.3.5. Problems

Problem (Exercise 15): Prove that if A is an m by n matrix and C is an n by p matrix, then

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

Solution: Note that

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

Problem (Exercise 17): Suppose $T \in \mathcal{L}(V)$, and $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ are bases of V. Prove that the following are equivalent.
(a) T is injective
(b) The columns of $\mathcal{M}(T)$ are linearly independent in $\mathbb{F}^{n,1}$.
(c) The columns of $\mathcal{M}(T)$ span $\mathbb{F}^{n,1}$.
(d) The rows of $\mathcal{M}(T)$ are linearly independent in $\mathbb{F}^{1,n}$.
(e) The rows of $\mathcal{M}(T)$ span $\mathbb{F}^{1,n}$.

Here $\mathcal{M}(T)$ means $\mathcal{M}(T, (u_1, \ldots, u_n), (v_1, \ldots, v_n))$.

Solution: 2 and 3 are equivalent since a spanning/linearly independent list implies the list is a basis in a vector space where the list equals the dimension. Similarly for 4 and 5. Note also that the last 4 are equivalent since column rank = row rank. Now we just need to show $1 \Leftrightarrow 2$. Denote $\mathcal{M}(T) = A$.

First suppose the columns are linearly independent. Then the only linear combination of columns that equals the 0 matrix is when all the coefficients are 0. This means that $T(u)$ is only 0 for an input vector u when the linear combination of the basis vectors that make it up have coefficients all 0. This means that $\dim \operatorname{null} T = 0$, which means T is injective. Now suppose the columns were not linearly independent. Then some linear combination of some of the columns equals another column. WLOG let these columns be $A_{\cdot,1}, \ldots, A_{\cdot,c}$. Then

$$\sum_{i=1}^{c} a_i A_{\cdot,i} = A_{\cdot,k}$$

for some $n \geq k \geq c$. This is equivalent to saying that

$$T\left(\sum_{i=1}^{c} a_i v_i - v_k\right) = 0.$$

Since the v's are a basis, the vector on the left is nonzero, so $\dim \operatorname{null} T > 0$, so T is not injective.

3.4. Invertibility and Isomorphisms

3.4.1. Invertible Linear Maps

Definition (invertible): A linear map $T \in \mathcal{L}(V, W)$ is invertible if there exists a linear map $S \in \mathcal{L}(W, V)$ such that ST equals the indentity operator on V and TS equals the identity operator on W.

Proposition (inverse is unique): An invertible linear map has a unique inverse, denoted by T^{-1}.

Proof: If S_1, S_2 are inverses, then

$$S_1 = S_1 I = S_1(TS_2) = (S_1 T)S_2 = IS_2 = S_2.$$

■

Proposition: A linear map is invertible if and only if it is injective and surjective.

Proof: Suppose $T \in \mathcal{L}(V, W)$. First suppose T is invertible. Suppose $Tu = Tv$. Then we have

$$u = T^{-1}(Tu) = T^{-1}(Tv) = v,$$

which implies injectivity. Now let $w \in W$. We have $w = T(T^{-1}w)$, so range $T = W$, implying surjectivity.

Now suppose T is injective and surjective. For each $w \in W$, define $S(w)$ to be the unique element of V such that $T(S(w)) = w$ (exists by injective + surjective). This definition implies that $T \circ S$ equals the indentity operator on W. Also note that for $v \in V$ we have

$$T((S \circ T)v) = (T \circ S)(Tv) = I(Tv) = Tv,$$

so $(S \circ T)v = v$ by injectivity. Thus $S \circ T$ is the indentity operator on V.

To prove that S is a linear map, just throw $S(w)$ inside T and use linearity and scalar multiplication. ■

Proposition: Suppose that V and W are finite-dimensional vector spaces, $\dim V = \dim W$, and $T \in \mathcal{L}(V, W)$. Then

$$T \text{ is invertible} \Leftrightarrow T \text{ is injective} \Leftrightarrow T \text{ is surjective}.$$

Proof: If T is injective, by the fundamental theorem of linear maps we have

$$\dim \operatorname{range} T = \dim V - \dim \operatorname{null} T = \dim V = \dim W,$$

which means T is surjective. If T is surjective, then we can similarly show that $\dim \operatorname{null} T = 0$, so T is injective. Thus if we have injectivity or surjectivity, then both hold, which implies T is invertible. ■

Proposition: Suppose V and W are finite-dimensional vector spaces of the same dimension, $S \in \mathcal{L}(V, W)$, and $T \in \mathcal{L}(W, V)$. Then $ST = I$ if and only if $TS = I$.

Proof: First suppose $ST = I$. If $v \in V$ and $Tv = 0$, then

$$v = Iv = (ST)v = S(Tv) = S(0) = 0.$$

Thus T is injective, and since $\dim V = \dim W$, T is invertible. Multiplying $ST = I$ by T^{-1} on the right yields $S = T^{-1}$, which implies $TS = TT^{-1} = I$, as desired. The other direction is proved basically identically. ■

3.4.2. Isomoprhic Vector Spaces

Definition (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.

Proposition: Two finite-dimensional vector spaces over $\mathbb{F}$ are isomoprhic if and only if they have the same dimension.

Proof: First suppose V and W are isomoprhic. Then there exists an isomorphism T from V to W. Since T is invertible, we have $\operatorname{null} T = \{0\}$ and $\operatorname{range} T = 0$. Thus we have $\dim V = \dim W$ by the fundamental theorem of linear maps.

Now suppose V and W have the same dimension. Let $v_1, ..., v_n$ be a basis of V and $w_1, ..., w_n$ be a basis of W. Let

$$T\left(\sum c_i v_i\right) = \sum c_i w_i.$$

Note that T is surjective since the right side spans V, and $\operatorname{null} T = 0$ since the right side is linearly independent. Thus T is an isomorphism, meaning V and W are isomorphic. ■

Proposition ($\mathcal{L}(V, W)$ and $\mathbb{F}^{m,n}$ are isomoprhic): Suppose $v_1, ..., v_n$ and $w_1, ..., w_n$ are a basis of V and W respectively. Then $\mathcal{M}$ is an isomorphism between $\mathcal{L}(V, W)$ and $\mathbb{F}^{m,n}$, where $\mathcal{M}$ is the linear map that gives the matrix for T with respect to the bases.

Proof: First we show $\mathcal{M}$ is injective. If $T \in \mathcal{L}(V, W)$ and $\mathcal{M}(T) = 0$, then $Tv_i = 0$ for all i. Since v_i's are a basis of V, this implies $T = 0$, so $\operatorname{null} M = \{0\}$.

Now surjectivity. Pick $A \in \mathbb{F}^{m,n}$. By the linear map lemma, there exists $T \in \mathcal{L}(V, W)$ such that

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

for all $k = 1, \ldots, n$. Because $\mathcal{M}(T)$ equals A, the range of $\mathcal{M}$ equals $\mathbb{F}^{m,n}$ as desired. ∎

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

Proof: Follows since $\dim \mathbb{F}^{m,n} = mn$ and $\mathbb{F}^{m,n}$ is isomorphic to $\mathcal{L}(V, W)$. ∎

3.4.3. Linear Maps Thought of as Matrix Multiplication

Definition (matrix of vector): Suppose $v \in V$ and $v_1, \ldots, v_n$ is a basis of V. The matrix of v with respect to the basis is in n by 1 matrix

$$\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

where $b_1, \ldots, b_n$ are the scalars such that

$$v = b_1 v_1 + \cdots + b_n v_n.$$

This matrix is denoted $\mathcal{M}(v)$.

Proposition ($\mathcal{M}(T)_{\cdot,k} = \mathcal{M}(Tv_k)$): Suppose $T \in \mathcal{L}(V, W)$ and $v_1, \ldots, v_n$ is basis of V and $w_1, \ldots, w_m$ is basis of W. Let $1 \leq k \leq n$. Then k^{th} column of $\mathcal{M}(T)$, which is denoted by $\mathcal{M}(T)_{\cdot,k}$ equals $\mathcal{M}(Tv_k)$.

Proof: Note that the k^{th} column of $\mathcal{M}(T)$ is defined at the scalars that make up Tv_k in W, so this follows by definition. ∎

Proposition: Suppose $T \in \mathcal{L}(V, W)$ and $v \in V$. Then we have

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

Proof: Suppose $v = \sum b_i v_i$. Thus

$$Tv = \sum b_i T v_i.$$

Then we have

$$\begin{aligned}\mathcal{M}(Tv) &= b_1 \mathcal{M}(Tv_1) + \dots + b_n \mathcal{M}(Tv_n) \\ &= b_1 \mathcal{M}(T)_{\cdot,1} + \dots + b_n \mathcal{M}(T)_{\cdot,n} \\ &= \mathcal{M}(T)\mathcal{M}(v).\end{aligned}$$

The last equality comes from the linear combination of columns proposition earlier. ■

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

Proof: The linear map that takes $w \in W$ to $\mathcal{M}(w)$ is an isomorphism from W onto $\mathbb{F}^{m,1}$. The restriction of this isomorphism to range T (which equals span$(Tv_1, \dots, Tv_n)$) is an isomorphism from range T onto span$(\mathcal{M}(Tv_1), \dots, \mathcal{M}(Tv_n))$. For each $k \in \{1, \dots, n\}$, the m by 1 matrix $\mathcal{M}(Tv_k)$ equals column k of $\mathcal{M}(T)$. Thus

$$\dim \operatorname{range} T = \text{column rank of } \mathcal{M}(T),$$

as desired. ■

3.4.4. Change of Basis

If $T \in \mathcal{L}(V)$, and we are using the same basis for input and output, then we generally have $\mathcal{M}(T, (v_1, \dots, v_n)) = \mathcal{M}(T, (v_1, \dots, v_n), (v_1, \dots, v_n))$.

Definition (identity matrix): Suppose n is a positive integer. The n by n matrix

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

with 1's on the diagonal and 0's everywhere else is called the identity matrix and denoted as I.

Definition (invertible): A square matrix A is called invertible if there is a square matrix B of the same size such that $AB = BA = I$. B is called the inverse of A and denoted by A^{-1}.

Corollary:

$$(AC)^{-1} = C^{-1}A^{-1}$$

and

$$(A^{-1})^{-1} = A.$$

Proof: Easy to verify. ■

Proposition: Suppose $T \in \mathcal{L}(U, W)$ and $S \in \mathcal{L}(V, W)$. If $u_1, ..., u_m$ is a basis of U, $v_1, ..., v_n$ is a basis of V, and $w_1, ..., w_p$ is a basis of W, then

$$\mathcal{M}\big(ST, (u_1, ..., u_m), (w_1, ..., w_p)\big)$$
$$= \mathcal{M}\big(S, (v_1, ..., v_n), (w_1, ..., w_p)\big)\mathcal{M}(T, (u_1, ..., u_m), (v_1, ..., v_n)).$$

Proof: Follows from definition of matrix multiplication (this proposition was noted noted earlier, there the bases of the matrices are just explicitly written out). ■

Proposition (matrix of identity operator with respect to two bases): Suppose that $u_1, ..., u_n$ and $v_1, ..., v_n$ are bases of V. Then the matrices

$$\mathcal{M}(I, (u_1, ..., u_n), (v_1, ..., v_n)) \text{ and } \mathcal{M}(I, (v_1, ..., v_n), (u_1, ..., u_n))$$

are invertible, and each is the inverse of each other.

Proof: In the matrix multiplication proposition above, replace S and T with the identity map I, getting

$$I = \mathcal{M}(I, (u_1, ..., u_n), (v_1, ..., v_n))\mathcal{M}(I, (v_1, ..., v_n), (u_1, ..., u_n)),$$

where the I on the left is the identity matrix. Swapping the bases yields

$$I = \mathcal{M}(I, (v_1, ..., v_n), (u_1, ..., u_n))\mathcal{M}(I, (u_1, ..., u_n), (v_1, ..., v_n)).$$

By the definition of invertibility, we are done. ■

Example: Consifer the bases $(4, 2), (5, 3)$ and $(1, 0), (0, 1)$ of $\mathbb{F}^2$. Because $I(4, 2) = 4(1, 0) + 2(0, 1)$ and $I(5, 3) = 5(1, 0) + 3(0, 1)$, we have

$$\mathcal{M}(I, ((4, 2), (5, 3)), ((1, 0), (0, 1))) = \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix}.$$

The inverse of the matrix is

$$\begin{pmatrix} \frac{3}{2} & -\frac{5}{2} \\ -1 & 2 \end{pmatrix}.$$

Thus by the previous result we have

$$\mathcal{M}(I, ((1, 0), (0, 1)), ((4, 2), (5, 3))) = \begin{pmatrix} \frac{3}{2} & -\frac{5}{2} \\ -1 & 2 \end{pmatrix}.$$

Theorem (change of basis formula): Suppose $T \in \mathcal{L}(V)$. Suppose $u_1, ..., u_n$ and $v_1, ..., v_n$ are bases of V. Let

$$A = \mathcal{M}(T, (u_1, ..., u_n)) \text{ and } B = \mathcal{M}(T, (v_1, ..., v_n))$$

and $C = \mathcal{M}(T, (u_1, ..., u_n), (v_1, ..., v_n))$. Then

$$A = C^{-1}BC.$$

Proof: In the matrix multiplication proposition, replace w_k with u_k are replace S with I to obtain

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

where C^{-1} was obtained using the proposition above.

Use the multiplication proposition again, this time replacing w_k with v_k. Also replace T with I and S with T to obtain

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

Subbing this into the first equation we got yields the desired result. ■

Proposition (matrix of inverse equals inverse of matrix): Suppose $v_1, ..., v_n$ is a basis of V and $T \in \mathcal{L}(V)$ is invertible. Then $\mathcal{M}(T^{-1}) = (\mathcal{M}(T))^{-1}$, where both matrices are with respect to the basis $v_1, ..., v_n$.

Proof: By the matrix multiplication proposition, we have

$$\mathcal{M}(T^{-1}T) = \mathcal{M}(T^{-1})\mathcal{M}(T) \Rightarrow \mathcal{M}(T)^{-1} = \mathcal{M}(T^{-1}).$$

■

3.4.5. Problems

Problem (Exercise 3): Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the following are equivalent:
(a) T is invertible.
(b) $Tv_1, ..., Tv_n$ is a basis of V for every basis $v_1, ..., v_n$ of V.
(c) $Tv_1, ..., Tv_n$ is a basis of V for some basis $v_1, ..., v_n$ of V.

Solution: 2 obviously implies 3. To show 3 implies 2, note that any other basis of V can be written in terms of $v_1, ..., v_n$, and thus the outputs on those new basis vectors must be a basis, otherwise we would get a contradiction about $v_1, ..., v_n$ being basis vectors.

Note that 1 implies 2 since that means T is injective and surjective, so range $T = V$, so the basis vector outputs will span V. 2 implies 1 since it implies null $T = \{0\}$, so T in injective, which means it's invertible.

Problem (Exercise 11): Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that

$$ST \text{ is invertible} \Leftrightarrow S \text{ and } T \text{ are invertible.}$$

Solution: Suppose ST is invertible. Then null $ST = \{0\}$ and range $ST = V$. Note that for nonzero v, we must have $Tv \neq 0$, so T is invertible. Since T is invertible, T is also surjective, so S must also be injective in order for range $ST = V$, giving the desired result. If both S and T are invertible, then the null space of both is 0, so range $ST = V$.

3.5. Products and Quotients of Vector Spaces

3.5.1. Products of Vector Spaces

Definition (product of vector spaces): Suppose $V_1, ..., V_m$ are vector spaces over $\mathbb{F}$. The product $V_1 \times \cdots \times V_m$ is defined by

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

Addition and scalar multiplication are defined pointwise.

Proposition: Product of vector spaces over $\mathbb{F}$ is a vector space over $\mathbb{F}$.

Proof: Easy to verify. ■

Proposition: Suppose $V_1, ..., V_m$ are finite-dimensional vector spaces. Then $V_1 \times \cdots \times V_m$ is finite-dimensional and

$$\dim(V_1 \times \cdots \times V_m) = \sum \dim V_i.$$

Proof: Choose a basis for each V_i. Consider the list of vectors in the product space that only consists of a single vector over all vectors from each basis. This is a basis of the product space, and has length $\sum \dim V_i$. ■

Proposition: Suppose $V_1, ..., V_m$ are subspaces of V. Define a linear map $\Gamma : V_1 \times \cdots \times V_m \to V_1 + \cdots + V_m$ by

$$\Gamma(v_1, ..., v_m) = v_1 + \cdots + v_m.$$

Then $V_1 + \cdots + V_m$ is a direct sum if and only if Γ is injective.

Solution: Γ is injective if and only if null $\Gamma = \{0\}$. Thus the only way to get 0 is if we take each $v_i = 0$, which means $V_1 + \cdots + V_m$ is a direct sum.

Proposition: Suppose V is finite-dimensional and $V_1, ..., V_m$ are subspaces of V. Then $V_1 + \cdots + V_m$ is a direct sum if and only if

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

Solution: Γ in the previous result is surjective. Thus Γ is injective if and only if

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

3.5.2. Quotient Spaces

Definition $(v+U)$: Suppose $v \in V$ and $U \subseteq V$. Then $v+U$ is the subset of V defined by

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

Definition (translate): For $v \in V$ and U a subset of V, the set $v+U$ is said to be a translate of U.

Definition (quotient space): Suppose U is subspace of V. Then the quotient space V/U is the set of all translates of U:

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

Example *(quotient spaces)*:
- If $U = \{(x, 2x) \in \mathbb{R}^2\}$, then $\mathbb{R}^2/U$ is the set of all lines in $\mathbb{R}^2$ that have slope 2.
- If U is a line in $\mathbb{R}^3$ containing the origin, then $\mathbb{R}^3/U$ is the set of all lines in $\mathbb{R}^3$ parallel to U.
- If U is a plane in $\mathbb{R}^3$ containing the origin, then $\mathbb{R}^3/U$ is the set of all planes in $\mathbb{R}^3$ parallel to U.

Proposition (two translates of a subspace are equal or disjoint): Suppose U is a subspace of V and $v, w \in V$. Then

$$v - w \in U \Leftrightarrow v + U = w + U \Leftrightarrow (v + U) \cap (w + U) \neq \emptyset.$$

Proof: First suppose $v - w \in U$. For $u \in U$ we have

$$v + u = w + ((v - w) + u) \in w + U.$$

Thus $v + U \subseteq w + U$. Similarly we have $w + U \subseteq v + U$. Thus $v + U = w + U$.

If the two sets are equal then they're obviously nondisjoint.

Now suppose the last statement is true. Thus there exist $u_1, u_2 \in U$ such that

$$v + u_1 = w + u_2.$$

Thus $v - w = u_2 - u_1 \in U$. ■

Definition (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 all $v, w \in V$ and all $\lambda \in \mathbb{F}$.

Proposition (quotient space is a vector space): Suppose U is a subspace of V. Then v/U, with the operations of addition and scalar multiplication defined above, is a vector space.

Proof: The potential problem with the definition above is that the translate representation is not unique. Suppose $v_1, v_2, w_1, w_2 \in V$ are such that

$$v_1+U=v_2+U \text{ and } w_1+U=w_2+U.$$

To show that the definition of addition on V/U above makes sense, we must show that $(v_1+w_1)+U=(v_2+w_2)+U$. We have

$$v_1-v_2, w_1-w_2 \in U.$$

Thus we have $(v_1-v_2)+(w_1-w_2)=(v_1+w_1)-(v_2+w_2) \in U$. This we have

$$(v_1+w_1)+U=(v_2+w_2)+U,$$

as desired. We show the same thing for scalar multiplication.

Note that identity of V/U is $0+U=U$, and the inverse of $v+U$ is $(-v)+U$. ■

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

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

for each $v \in V$.

Proposition (dimension of quotient space): Suppose V is finite-dimensional and U is a subspace of V. Then $\dim V/U=\dim V-\dim U$.

Proof: Let π denote the quotient map. If $v \in V$, then $v+U=0+U$ if and only if $v \in U$, which means null $\pi=U$. The definition of π implies range $\pi=V/U$ (V/U is the space of all translates of U, and the right side of the quotient map can take on all $v \in V$). By the fundamental theorem of linear maps, we have $\dim V=\dim U+\dim V/U$, which yields the desired result.

■

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

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

Proposition (null space and range of $\tilde{T}$): suppose $T \in \mathcal{L}(V, W)$. Then
(a) $\tilde{T} \circ \pi = T$, where π is the quotient map of V into $V/\,(\text{null } T)$;
(b) $\tilde{T}$ is injective;
(c) range $\tilde{T}$ = range T;
(d) $V/\,(\text{null } T)$ and range T are isomoprhic vector spaces.

Proof:
(a) For $v \in V$, we have $\tilde{T}(\pi(v))\pi = \tilde{T}(v + \text{null } T) = Tv$, as desired.
(b) Suppose $v \in V$ and $\tilde{T}(v + \text{null } T) = 0$. Then $Tv = 0$. Thus $v \in \text{null } T$. This means $v + \text{null } T = 0 + \text{null } T$. This implies that null $\tilde{T} = \{0 + \text{null } T\}$. Thus $\tilde{T}$ is injective.
(c) By definition range $\tilde{T}$ = range T.
(d) 2 and 3 imply that if we think of $\tilde{T}$ as mapping into range T, then $\tilde{T}$ is an isomorphism from $V/\,(\text{null } T)$ onto range T.

■

3.5.3. Problems

Problem (Exercise 1): Prove that

$$\text{graph of } T = \{(v, Tv) \in V \times W : v \in V\}$$

is a subspace of $V \times W$ if and only if T is linear.

Solution: If T is linear, then the conclusion is obvious. If graph of T is a subspace, then $(v, Tv) + (w, Tw) = (v + w, Tv + Tw) = (v + w, T(v + w)) \Rightarrow Tv + Tw = T(v + w)$ for all $v, w \in V$. Scalar multiplication follows similarly.

Problem (Exercise 2): Suppose that $V_1, \ldots, V_m$ are vector spaces such that $V_1 \times \cdots \times V_m$ is finite-dimensional. Prove that V_k is finite dimensional for each k.

Solution: WLOG we show V_1 is finite-dimensional, as the rest follow similarly. Note that $V_1 \times \{0\} \times \cdots \times \{0\}$ is a subspace of $V_1 \times \cdots \times V_m$, so it is finite dimensional. Consider

$$T : V_1 \times \{0\} \times \cdots \times \{0\} \to V_1$$

such that

$$T((v_1, 0, ..., 0)) = v_1,$$

where $v_1 \in V_1$. Note that the right side spans V_1, so range $T = V_1$, and also note that for the right side to be 0, $v_1 = 0$, so null $T = \{0\}$, meaning T is an isomorphism. Thus, since $V_1 \times \{0\} \times \cdots \times \{0\}$ is finite-dimensional, V_1 must also be finite-dimensional.

Problem (Exercise 6): Suppose $v, x \in V$ and U, W are subspaces of V such that $v + U = x + W$. Prove that $U = W$.

Solution: Note that $0 \in U$, so we have $v = x + w \to v - x = w \in W$. This implies $v + W = x + W$. Similarly we have $v + U = x + U$. Both of these imply that $v + W = v + U$. Since both subspaces are translated by the same vector, the initial two subspaces must be equal.

Problem (Exercise 9): Prove that a nonempty subset A of V is a translate of some subspace V if and only if $\lambda v + (1 - \lambda)w \in A$ for all $v, w \in A$ and all $\lambda \in \mathbb{F}$.

Solution: Suppose $\lambda v + (1 - \lambda)w \in A$ for all v, w, λ. Fix $x \in A$ and consider $(-x) + A$. Pick $u \in (-x) + A$ and note that

$$\lambda u = \lambda v - \lambda x = -x + \lambda v + (1 - \lambda)x \in (-x) + A,$$

so $(-x) + A$ is closed under scalar mulitplication.

Now consider $a, b \in (-x) + A$. We have $\frac{a+b}{2} = -x + \frac{v+w}{2} \in (-x) + A$ by taking $\lambda = \frac{1}{2}$. Note that $0 \in (-x) + A$ since $x \in A$. Thus $(-x) + A$ is a subspace of V, so A is a translate of a subspace.

For the other direction, let $A = x + U$ for some subspace U of V. Note that

$$\lambda(x + v) + (1 - \lambda)(x + w) = x + (\lambda v + \lambda w + w) \in A$$

for $v, w \in U$, as desired.

Problem (Exercise 13): Suppose U is a subspace of V such that V / U is finite dimensional. Prove that V is isomoprhic to $U \times (V / U)$.

Solution: Let $v_1 + U, ..., v_m + U$ be a basis of V / U. Then $v_i \notin U$. Let $W = \operatorname{span}(v_1, ..., v_m)$. Note $(\sum a_i v_i) + U = 0 + U \iff \sum a_i v_i = u$ for some $u \in U$. Since the list is a basis, the a_i's must be 0, which implies $W \cap U = \{0\}$. Thus $U + W$ is a direct sum.

Let $T : U \times (V / U) \to V$, where $T(u, v + U) = u + \sum a_i v_i$, where $(\sum a_i v_i) + U = v + U$. Suppose $T(u, v + U) = u + \sum a_i v_i = 0$. Since U and W are a direct sum, we must have $0 = u$ and $0 = \sum a_i v_i$, and since the v_i are linearly independent, $a_i = 0$. Thus T is injective.

Let $S : V \to U \times (V/U)$, where $Tv = (v - \sum a_i v_i, v + U)$, where $v + U = (\sum a_i v_i) + U$. Suppose $Tv = (v - \sum a_i v_i, v + U) = (0, 0 + U)$. The first slot implies $v \in W$, and the second slot implies $v \in U$, and these together imply $v = 0$. Thus S is injective.

Note that $TSv = T(v - \sum a_i v_i, v + U) = v - \sum a_i v_i + \sum a_i v_i = v$. Thus, T and S are inverses of each other, which implies V and $U \times (V/U)$ are isomoprhic.

Problem (Exercise 14): Suppose U and W are subspaces of V and $V = U \oplus W$. Suppose $w_1, ..., w_m$ is a basis of W. Prove that $w_1 + U, ..., w_m + U$ is a basis of V/U.

Solution: Note that $(\sum a_i w_i) + U = 0 + U \Rightarrow \sum a_i w_i = u \in U$, and since W and U are a direct sum, we have $a_i = 0$, so the list is linearly independent. Now note that $v = u + w \Rightarrow v + U = w + U = (\sum a_i w_i) + U$, so the list is spanning.

Problem (Exercise 16): Suppose $\varphi \in \mathcal{L}(V, \mathbb{F})$, and $\varphi \neq 0$. Prove that $\dim V/(\text{null } \varphi) = 1$.

Solution: Note that $V/(\text{null } \varphi)$ and range φ are isomorphic. Since $\varphi \neq 0$, range $\varphi = \mathbb{F}$, so $\dim V/(\text{null } \varphi) = \dim \text{range } \varphi = 1$.

Problem (Exercise 19): Suppose $T \in \mathcal{L}(V, W)$ and U is a subspace of V. Let π denote the quotient map from V onto V/U. Prove that there exists $S \in \mathcal{L}(V/U, W)$ such that $T = S \circ \pi$ if and only if $U \subseteq \text{null } T$.

Solution: Suppose $U \subseteq \text{null } T$. Let $S(v + U) = Tv$. We need to show that the choice of representative doesn't matter. Suppose $v_1 + U = v_2 + U \Rightarrow v_1 - v_2 \in U$. Then $T(v_1 - v_2) = 0 \Rightarrow S(v_1 + U) = Tv_1 = Tv_2 = S(v_2 + U)$. Thus S is a linear map. If $v \in V$, then $S(\pi(v)) = S(v + U) = Tv$, as desired.

Now suppose $T = S \circ \pi$. For $u \in U$, we have $\pi(u) = u + U = 0 + U$. Thus $S(\pi(u)) = Tu = 0 \Rightarrow U \subseteq \text{null } T$.

3.6. Duality

3.6.1. Dual space and Dual Map

Definition (linear functional): A *linear functional* on V is an element of $\mathcal{L}(V, \mathbb{F})$.

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

Proposition ($\dim V' = \dim V$): Suppose V is finite-dimensional. Then V' is also finite dimensional and

$$\dim V' = \dim V.$$

Proof: We have

$$\dim V' = \dim \mathcal{L}(V, \mathbb{F}) = (\dim V)(\dim \mathbb{F}) = \dim V.$$

■

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

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

Proposition: Suppose $v_1, ..., v_n$ is a basis of V and $\varphi_1, ..., \varphi_n$ is the dual basis. Then

$$v = \varphi_1(v)v_1 + \cdots + \varphi_{n(v)}v_n$$

for each $v \in V$.

Proof: We can write $v = \sum c_i v_i$. Applying φ_j to each side yields

$$\varphi_j(v) = c_j.$$

Thus we can replace c_j with $\varphi_j(v)$, giving us the desired result. ■

Proposition: Suppose V is finite-dimensional. Then the dual basis of V is a basis of V'.

Proof: Suppose $v_1, ..., v_n$ is a basis of V. Let $\varphi_1, ..., \varphi_n$ be the dual basis. Suppose

$$\sum a_i \varphi_i = 0.$$

We have

$$\left(\sum a_i \varphi_i\right)(v_k) = a_k$$

for each k. The first equation implies $a_k = 0$ for all k, so the φ's are linearly independent. Since $\dim V' = n$, $\varphi_1, ..., \varphi_n$ must be a basis. ■

Definition (dual map): Suppose $T \in \mathcal{L}(V, W)$. The *dual map* of T is the linear map $T' \in \mathcal{L}(W', V')$ defined for each $\varphi \in W'$ by

$$T'(\varphi) = \varphi \circ T.$$

Proposition (algebraic properties of dual maps): Suppose $T \in \mathcal{L}(V, W)$. Then
- $(S + T)' = S' + T'$ for all $S \in \mathcal{L}(V, W)$,
- $(\lambda T)' = \lambda T'$ for all $\lambda \in \mathbb{F}$,
- $(ST)' = T'S'$ for all $S \in \mathcal{L}(W, U)$.

Proof:
- $(S + T)'(\varphi) = \varphi \circ (S + T) = \varphi \circ S + \varphi \circ T = S'(\varphi) + T'(\varphi)$.
- $(\lambda T)'(\varphi) = \varphi \circ (\lambda T) = \lambda(\varphi \circ T) = \lambda T'(\varphi)$.
- Suppose $\varphi \in U'$. Then

  $$(ST)'(\varphi) = \varphi \circ (ST) = (\varphi \circ S) \circ T = T'(\varphi \circ S) = T'(S'(\varphi)) = (T'S')(\varphi).$$

■

3.6.2. Null Space and Range of Dual Linear map

Definition (annihilator): For $U \subseteq V$, the annihilator of U, denoted by U^0, is defined by

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

Proposition: Suppose $U \subseteq V$. Then U^0 is a subspace of V'.

Proof: Note that $0 \in U^0$.

Suppose $\varphi, \psi \in U^0$. For $u \in U$ we have

$$(\varphi + \psi)(u) = \varphi(u) + \psi(u) = 0 + 0 = 0,$$

so U^0 is closed under addition. We can similarly show that U^0 is closed under scalar multiplication. ■

Proposition (dimension of the annihlator): Suppose V is finite-dimensional and U is a subspace of V. Then

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

Below are two proofs, with the first being slicker.

Proof: Let $i \in \mathcal{L}(U, V)$ be the inclusion map defined by $i(u) = u$ for each $u \in U$. The fundamental theorem of linear maps on i' yields

$$\dim V' = \dim \operatorname{null} i' + \dim \operatorname{range} i'.$$

Suppose $i'(\varphi) = 0$. We have $i'(\varphi) = \varphi \circ i = 0$. Note that since range $i = U$, φ must be 0 on all of U, implying $\varphi \in U^0$. We also have $\dim V' = \dim V$, so we can write

$$\dim V = \dim U^0 + \dim \operatorname{range} i'.$$

If $\varphi \in U'$, then φ can be extended to a linear functional ψ on V. The definition of i' shows that $i'(\psi) = \varphi$. Thus $\varphi \in \operatorname{range} i'$, which implies that range $i' = U'$. Thus $\dim \operatorname{range} i' = \dim U' = \dim U$, and putting everthing together we have

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

■

Proof: Let $u_1, ..., u_m$ be a basis of U and extend it to $u_1, ..., u_m, ..., u_n$ a basis of V. let $\varphi_1, ..., \varphi_n$ be the dual basis. We will show that $\varphi_{m+1}, ..., \varphi_n$ is a basis of U^0. Note that these maps are linear independent since they are part of the dual basis, so we just need to show that they span U^0.

Suppose $\varphi \in \operatorname{span}(\varphi_{m+1}, ..., \varphi_n)$. We can write $\varphi = \sum_{i=m+1}^{n} c_i \varphi_i$. For any $\sum_{i=1}^{m} a_i u_i = u \in U$, we have

$$\left(\sum_{i=m+1}^{n} c_i \varphi_i\right)\left(\sum_{i=1}^{m} a_i u_i\right) = 0,$$

so $\operatorname{span}(\varphi_{m+1}, ..., \varphi_n) \subseteq U^0$.

Now suppose $\varphi \in U^0$. We can write it as $\varphi = \sum_{i=1}^{n} c_i \varphi_i$. Note that since $u_i \in U$ for $i = 1, ..., m$ and since $\varphi \in U^0$, we have

$$0 = \varphi(u_i) = c_i.$$

Thus $\varphi = \sum_{i=m+1}^{n} c_i \varphi_i$, so $\varphi \in \operatorname{span}\left(\varphi_{m_1}, ..., \varphi_n\right)$. Thus $U^0 \subseteq \operatorname{span}(\varphi_{m+1}, ...\varphi(n))$, implying the desired result. ■

Proposition: Suppose V is finite-dimensional and U is a subspace of V. Then
- $U^0 = \{0\} \iff U = V$;
- $U^0 = V' \iff U = \{0\}$.

Proof: For the first bullet point, we have

$$\begin{aligned} U^0 = \{0\} &\iff \dim U^0 = 0 \\ &\iff \dim U = \dim V \\ &\iff U = V. \end{aligned}$$

Similarly, we have

$$\begin{aligned} U^0 = V' &\iff \dim U^0 = \dim V' \\ &\iff \dim U^0 = \dim V \\ &\iff \dim U = 0 \\ &\iff U = \{0\}. \end{aligned}$$

■

Proposition (null space of T'): Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then
- $\text{null } T' = (\text{range } T)^0$;
- $\dim \text{null } T' = \dim \text{null } T + \dim W - \dim V$.

Proof: First suppose $\varphi \in \text{null } T'$. Thus $0 = T'(\varphi) = \varphi \circ T$. Hence

$$0 = (\varphi \circ T)(v) = \varphi(Tv) \text{ for every } v \in V.$$

Thus $\varphi \in (\text{range } T)^0$, which implies $\text{null } T' \subseteq (\text{range } T)^0$. Now suppose $\varphi \in (\text{range } T)^0$. Thus $\varphi(Tv) = 0$ for every vector $v \in V$. Hence $0 = \varphi \circ T = T'(\varphi)$, whcih implies $\varphi \in \text{null } T'$, which again implies $(\text{range } T)^0 \subseteq \text{null } T'$, so we're done.

The second bullet follows easily:

$$\begin{aligned} \dim \text{null } T' &= \dim (\text{range } T)^0 \\ &= \dim W - \dim \text{range } T \\ &= \dim W - (\dim V - \dim \text{null } T) \\ &= \dim \text{null } T + \dim W - \dim V. \end{aligned}$$

■

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

$$T \text{ is surjective} \iff T' \text{ is injective}.$$

Proof:

$$\begin{aligned} T \in \mathcal{L}(V, W) \text{ is surjective} &\iff \text{range } T = W \\ &\iff (\text{range } T)^0 = \{0\} \\ &\iff \text{null } T' = \{0\} \\ &\iff T' \text{ is injective.} \end{aligned}$$

■

Proposition (range of T'): Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then
- $\dim \text{range } T' = \dim \text{range } T$;
- $\text{range } T' = (\text{null } T)^0$.

Proof: We have

$$\begin{aligned} \dim \text{range } T' &= \dim W' - \dim \text{null } T' \\ &= \dim W - \dim (\text{range } T)^0 \\ &= \dim \text{range } T. \end{aligned}$$

For the second bullet, first suppose $\varphi \in \text{range } T'$. Thus there exists $\psi \in W'$ such that $\varphi = T'(\psi)$. If $v \in \text{null } T$, then

$$\varphi(v) = (T'(\psi))(v) = (\varphi \circ T)(v) = \psi(Tv) = \psi(0) = 0.$$

Thuse $\varphi \in (\text{null } T)^0$. This implies that $\text{range } T' \subseteq (\text{null } T)^0$.

To complete the proof, we show that $\text{range } T'$ and $(\text{null } T)^0$ have the same dimension. Note that

$$\begin{aligned} \dim \text{range } T' &= \dim \text{range } T \\ &= \dim V - \dim \text{null } T \\ &= \dim (\text{null } T)^0, \end{aligned}$$

where the first equality comes from the first bullet. ■

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

$$T \text{ is injective} \iff T' \text{ is surjective.}$$

Proof:

$$\begin{aligned} T \text{ is injective} &\iff \text{null } T = \{0\} \\ &\iff (\text{null } T)^0 = V' \\ &\iff \text{range } T' = V'. \end{aligned}$$

■

3.6.3. Matrix of Dual of Linear map

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

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

Proof: Let $A = \mathcal{M}(T)$ and $C = \mathcal{M}(T')$. Suppose $1 \leq j \leq m$ and $1 \leq k \leq n$. From the definition of $\mathcal{M}(T')$ we have

$$T'(\psi_j) = \sum_{r=1}^{n} C_{r,j} \varphi_r.$$

The left side is $\psi_j \circ T$. Thus applying both sides of the equation to v_k (a basis vector) yields

$$\begin{aligned}(\psi_j \circ T)(v_k) &= \sum_{r=1}^{n} C_{r,j} \varphi_{r(v_k)} \\ &= C_{k,j}.\end{aligned}$$

We also have

$$\begin{aligned}(\psi_j \circ T)(v_k) &= \psi_{j(T v_k)} \\ &= \psi_j \left(\sum_{r=1}^{m} A_{r,k} w_r \right) \\ &= \sum_{r=1}^{m} A_{r,k} \psi_j(w_r) = A_{j,k}.\end{aligned}$$

Thus we have $C_{k,j} = A_{j,k}$, which means $C = A^t$, giving us the desired result. ∎

Heres another proof that column rank equals row rank.

Proposition: Suppose $A \in \mathbb{F}^{m,n}$. Then the column rank of A equals the row rank of A.

Proof: Define $T : \mathbb{F}^{n,1} \to \mathbb{F}^{m,1}$ by $Tx = Ax$. Thus $\mathcal{M}(T) = A$, where $\mathcal{M}(T)$ is with respect to the standard bases of $\mathbb{F}^{n,1}$ and $\mathbb{F}^{m,1}$. Now

$$\begin{aligned}\text{column rank of } A &= \text{column rank of } \mathcal{M}(T) \\ &= \dim \operatorname{range} T \\ &= \dim \operatorname{range} T' \\ &= \text{column ranke of } \mathcal{M}(T') \\ &= \text{column rank of } A^t \\ &= \text{row rank of } A.\end{aligned}$$

∎

3.6.4. Problems

Problem (Exercise 1): Explain why each linear functional is surjective or is the zero map.

Solution: If a linear functional is not the 0 map, then $\dim \operatorname{range} \varphi \geq 1$, and since $\dim \mathbb{F} = 1$, range $\varphi = \mathbb{F}$.

Problem (Exercise 3): Suppose V is finite-dimensional and $v \in V$ with $v \neq 0$. Prove that there exists $\varphi \in V'$ such that $\varphi(v) = 1$.

Solution: Pick a surjective map φ (must exist by exercise 1). Suppose $\varphi(v) = a$. Then the map $\frac{1}{a}\varphi$ works.

Problem (Exercise 4): Suppose V is finite-dimensional and U is a subspace of V such that $U \neq V$. Prove that there exists $\varphi \in V'$ such that $\varphi(u) = 0$ for every $u \in U$ but $\varphi \neq 0$.

Solution: Note that since $U \neq V$, $\dim U < \dim V$. Thus, $\dim U^0 = \dim V - \dim U > 0$, so there exists nonzero maps in the annihilator of U.

Problem (Exercise 5): Suppose $T \in \mathcal{L}(V, W)$ and $w_1, ..., w_m$ is a basis of range T. Hence for each $v \in V$, there exist unique numbers $\varphi_1(v), ..., \varphi_{m(v)}$ such that

$$Tv = \varphi_1(v)w_1 + \cdots + \varphi_m(v)w_m,$$

thus defining functions $\varphi_1, ..., \varphi_m$ from V to $\mathbb{F}$. Show that each of the functions $\varphi_1, ..., \varphi_m$ is a linear functional on V.

Solution: Clearly $\varphi_{i(0)} = 0$, since $T(0) = 0 = \sum \varphi_i(0)w_i$.

We have

$$T(v + w) = \sum \varphi_i(v + w)w_i = \sum(\varphi_i(v) + \varphi_i(w))w_i = Tv + Tw,$$

so φ_i is additive. Similarly, its easy to show that φ_i is homogenous, so they are indeed linear functionals.

Problem (Exercise 10): Suppose m is a positive integer.
- Show that $1, x - 5, ..., (x - 5)^m$ is a basis of $\mathcal{P}_m(\mathbb{R})$.
- What is the dual basis in the first bullet?

Solution: Clearly the list is a basis since the degree of every polynomial is different. We have

$$\varphi_k(p) = \frac{p^k(5)}{k!}.$$

Problem (Exercise 11): Suppose $v_1, ..., v_n$ is basis of V and $\varphi_1, ..., \varphi_n$ is the corresponding dual basis of V'. Suppose $\psi \in V'$. Prove that

$$\psi = \psi(v_1)\varphi_1 + \cdots + \psi(v_n)\varphi_n.$$

Solution: Since linear maps are entirely determined on the basis, we just need to show this equation holds for each basis vector. Plug in v_i to both sides of the equation to obtain

$$\psi(v_i) = \psi(v_1)\varphi_1(v_i) + \cdots + \psi(v_n)\varphi_n(v_i) = \psi(v_i)\varphi_i(v_i) = \psi_{v_i}.$$

Problem (Exercise 13): Show that the dual map of the identity operator on V is the identity operator on V'.

Solution: We have

$$I'(\varphi) = \varphi \circ I = \varphi.$$

Problem (Exercise 14): Define $T : \mathbb{R}^3 \to \mathbb{R}^2$ by

$$T(x, y, z) = (4x + 5y + 6z, 7x + 8y + 9z).$$

Suppose φ_1, φ_2 denotes the dual basis of the standard basis of $\mathbb{R}^2$ and ψ_1, ψ_2, ψ_3 denotes the dual basis of the standard basis of $\mathbb{R}^3$.

- Describe the linear functionals $T'(\varphi_1)$ and $T'(\varphi_2)$.
- Write $T'(\varphi_1)$ and $T'(\varphi_2)$ as linear combinations of ψ_1, ψ_2, ψ_3.

Solution: Note that $T'(\varphi_1)(x, y, z) = 4x + 5y + 6z$. Similarly, $T'(\varphi_2)(x, y, z) = 7x + 8y + 9z$. Thus, we can write

$$T'(\varphi_1) = 4\psi_1 + 5\psi_2 + 6\psi_3 \text{ and } T'(\varphi_2) = 7\psi_1 + 8\psi_2 + 9\psi_3.$$

Problem (Exercise 15): Define $T : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ by

$$Tp = x^2 p(x) + p''(x).$$

- Suppose $\varphi \in \mathcal{P}(\mathbb{R})'$ is defined by $\varphi(p) = p'(4)$. Describe the linear functional $T'(\varphi)$ on $\mathcal{P}(\mathbb{R})$.
- Suppose $\varphi \in \mathcal{P}(\mathbb{R})'$ is defined by $\varphi(p) = \int_0^1 p$. Evaluate $(T'(\varphi))(x^3)$.

Solution: For the first bullet, we have

$$T'(\varphi)(p) = (\varphi \circ T)(p) = \varphi(x^2 p(x) + p''(x)) = 8p(4) + 16p'(4) + p'''(4).$$

For the second bullet, we have

$$T'(\varphi)(x^3) = (\varphi \circ T)(x^3) = \varphi(x^5 + 6x) = \int_0^1 x^5 + 6x \, \mathrm{d}x = \frac{19}{6}.$$

Problem (Exercise 16): Suppose W is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that

$$T' = 0 \Longleftrightarrow T = 0.$$

Solution: Suppose $T' = 0$. Thus we have

$$\varphi \circ T = 0$$

for any $\varphi \in W'$. If $(\text{range } T)^0 = W'$, then range $T = \{0\}$, which implies $T = 0$, so now we can assume $(\text{range } T)^0 \neq W'$. Pick $\varphi \notin (\text{range } T)^0$. Thus if T is nonzero, there exists some v such that $\varphi(Tv) \neq 0$, but this is impossible. Thus, $T = 0$.

Now suppose $T = 0$. Then we have

$$T'(\varphi) = \varphi \circ T = \varphi(0) = 0$$

for all $\varphi \in W'$, so $T' = 0$.

Problem (Exercise 32): The *double dual space* of V, denote V'', is defined to be the dual space of V'. Define $\Lambda : V \to V''$ by

$$\Lambda v = \psi_v,$$

where $\psi_v(\varphi) = \varphi(v)$ for $\varphi \in V'$.

(a) Show that Λ is a linear map from V to V''.
(b) Show that if $T \in \mathcal{L}(V)$, then $T'' \circ \Lambda = \Lambda \circ T$, where $T'' = (T')'$.
(c) Show that if V is finite dimensional, then Λ is an isomorphism from V onto V''.

Solution:

(a) Easy to check
(b) We have $(T'' \circ \Lambda)v = T''(\psi_v) = \psi_v \circ T'$ and $(\Lambda \circ T)v = \psi_{Tv}$. We have $(\psi_v \circ T')(\varphi) = \psi_v \circ \varphi \circ T = (\varphi \circ T)v = \varphi(Tv) = \psi_{Tv}(\varphi)$.
(c) Pick T invertible. Suppose $\Lambda v = \psi_v = 0$ and v is nonzero. Thus we have $0 = (T'' \circ \Lambda v)\varphi = (\Lambda \circ Tv)\varphi = \psi_{Tv}(\varphi) = \varphi(Tv)$. Since T is invertible, $Tv \neq 0$. Note this holds for all φ, so define pick φ such that $\varphi(Tv) \neq 0$. Then clearly the statement doesn't hold. Thus v must be 0, so Λ is injevtive. Since $\dim V = \dim V''$, Λ is invertible, and thus an isomorphism.

Problem (Exercise 33): Suppose U is a subspace of V. Let $\pi : V \to V/U$ be the quotient map. Thus $\pi' \in \mathcal{L}\big((V/U)', V'\big)$.

- Show that π' is injective.
- Show theat range $\pi' = U^0$.
- Conclude that π' is an isomorphism from $(V/U)'$ onto U^0.

Solution: Suppose $\pi'(\varphi) = \varphi \circ \pi = 0$. Thus for any $v \in V$, we have $\varphi(\pi(v)) = \varphi(v + U) = 0$. Thus null $\varphi = V/U \Rightarrow \varphi = 0$, so π' is injective.

Suppose $\varphi \in (V/U)'$. For any $u \in U$, we have

$$\pi'(\varphi)(u) = (\varphi \circ \pi)(u) = \varphi(u + U) = \varphi(0 + U) = 0.$$

Thus for any $\varphi \in (V/U)$, $\pi'(\varphi) \in U^0$, so range $\pi' \subseteq U^0$.

Now suppose $\psi \in U^0$. Thus, $\psi(u) = 0$ for any $u \in U$. Define $\varphi \in (V/U)'$ as $\varphi(v + U) = \psi(v)$. It's easy to verify that φ is a consistent linear map. For any $v \in V$, we have

$$\pi'(\varphi)(v) = (\varphi \circ \pi)(v) = \varphi(v + U) = \psi(v).$$

Thus there exists $\varphi \in (V/U')$ such that $\pi'(\varphi) = \psi$. This implies $U^0 \subseteq \text{range } \pi'$, giving us range $\pi' = U^0$.

If we consider π' as a map from $(V/U)'$ to U^0, then π' is both injective and surjective. Thus π' is an isomorphism from $(V/U)'$ onto U^0.

4. Polynomials

Pretty easy stuff.

Below is a linear algebra proof of the division "algorithm" for polynomials.

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

$$p = sq + r$$

and $\deg r < \deg s$.

Proof: Let $n = \deg p$ and let $m = \deg s$. If $n < m$, then take $q = 0$ and $r = p$, which satisfty all the necessary conditions. Now we can assume $n \geq m$.

The list

$$1, z, ..., z^{m-1}, s, zs, ..., z^{n-m} s$$

is linearly independent in $\mathcal{P}_n(\mathbb{F})$ because each polynomial has a different degree. Since $\dim \mathcal{P}_n(\mathbb{F}) = n + 1$ and the list has length $n + 1$, the list is a basis of $\mathcal{P}_n(\mathbb{F})$.

Because $p \in \mathcal{P}_n(\mathbb{F})$ and the list is a basis, there exist unique constans $a_0, a_1, ..., a_{m-1}, b_0, b_1, ..., b_{n-m}$ such that

$$\begin{aligned} p &= a_0 + a_1 z + \cdots + a_{m-1} z^{m-1} + b_0 s + b_1 z s + \cdots + b_{n-m} z^{n-m} s \\ &= \underbrace{a_0 + a_1 z + \cdots + a_{m-1} z^{m-1}}_{r} + s\underbrace{\left(b_1 + b_1 z + \cdots + b_{n-m} z^{n-m}\right)}_{q}. \end{aligned}$$

These choices of r and q satsify all the conditions. ■

4.1. Problems

Problem (Exercise 7): Suppose that m is a nonnegative integer, $z_1, ..., z_{m+1}$ are distinct elements of $\mathbb{F}$, and $w_1, ..., w_{m+1} \in \mathbb{F}$. Prove that there exists a unique polynomial $p \in \mathcal{P}_{m(\mathbb{F})}$ such that

$$p(z_k) = w_k$$

for each $k = 1, ..., m + 1$.

Solution: Define the map $T : \mathcal{P}_{m(\mathbb{F})} \to \mathbb{F}^{m+1}$ by

$$T(p) = (p(z_1), ..., p(z_{m+1})).$$

Note that $T(0) = 0$ and that T is both additive and homogenous. Next we show T is injective. Suppose $T(p) = (0, ..., 0)$ and p is nonzero. Thus implies that p has $m + 1$ distinct roots. However, this is impossible, since p has degree at most m. Thus $p = 0$, meaning T is injective. Since the input and output space have the same dimension, this implies that T is invertible. Thus, given a unique output $(w_1, ..., w_{m+1})$, there exists a unique polynomial p such that $T(p) = (w_1, ..., w_{m+1})$.

Problem (Exercise 13): Suppose $p \in \mathcal{P}(\mathbb{F})$ with $p \neq 0$. Let $U = \{pq : q \in \mathcal{P}(\mathbb{F})\}$. Show that $\dim \mathcal{P}(\mathbb{F})/U = \deg p$ and find a basis of $\mathcal{P}(\mathbb{F})/U$.

Solution: Consider the list

$$1 + U, z + U, ..., z^{\deg p - 1} + U.$$

Suppose

$$\left(c_0 + c_1 z + \cdots + c_{\deg p - 1} z^{\deg p - 1}\right) + U = 0 + U.$$

Note that in order for the left to be 0, it would need to be contained in U. However, U only contains polynomials of degree at least $\deg p$, which the left is degree at most $\deg p - 1$. Thus, all the c's must be 0, meaning the list is linearly independent.

To show the list is spanning, pick $f + U \in \mathcal{P}(\mathbb{F})/U$. Note that we can reduce f by elements in U, since $f - u = g \Rightarrow f - g = u$, which would mean $f + U = g + U$. Thus we can drop f to a $\deg p - 1$ polynomial. After that, it's easy to get a representation using our list. Thus the list spans $\mathcal{P}(\mathbb{F})/U$, which means it's a basis, so $\dim \mathcal{P}(\mathbb{F})/U = \deg p$.

Problem (Exercise 14): Suppose $p, q \in \mathcal{P}(\mathbb{C})$ are nonconstant polynomials with no zeros in common. Let $m = \deg p$ and $n = \deg q$. use linear algebra to prove there exist $r \in \mathcal{P}_{n-1}(\mathbb{C})$ and $s \in \mathcal{P}_{m-1}(\mathbb{C})$ such that

$$rp + sq = 1.$$

Solution: Define $T : \mathcal{P}_{n-1}(\mathbb{C}) \times \mathcal{P}_{m-1}(\mathbb{C}) \to \mathcal{P}_{m+n-1}(\mathbb{C})$ by

$$T(r, s) = rp + sq.$$

If we show T is injective, then

$$\dim \operatorname{range} T + \dim \operatorname{null} T = \dim \mathcal{P}_{n-1}(\mathbb{C}) \times \mathcal{P}_{m-1}(\mathbb{C}) \Rightarrow \dim \operatorname{range} T = m + n,$$

which would imply range $T = \mathcal{P}_{m+n-1}(\mathbb{C})$, meaning T is surjective, which would imply the desired result.

Suppose $T(r, s) = 0$. Thus we have $rp = -sq$. This implies that both sides have the same roots. Since p and q share no roots, all of p's roots must be in s. However, s has degree at most $m - 1$, while p has degree p, meaning this is impossible if r and s are nonzero. Thus they must be 0, implying T is injective.

5. Eigenvalues and Eigenvectors

5.1. Invariant Subspaces

5.1.1. Eigenvalues

Definition (operator): A linear map from a vector space to itself is called an *operator*.

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

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

Proposition (equivalent conditions to be an eigenvalue): Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{F}$. 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.

Proof: 1 and 2. are equivalent since $Tv = \lambda v \Rightarrow Tv - \lambda I v = (T - \lambda I)(v) = 0$. The last three are all equivalent by previous results. ■

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

Proposition (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.

Proof: Suppose this is false. Then there exists a smallest positive integer m such that there exists a linearly dependent list $v_1, \dots, v_m$ of eigenvectors of T corresponding to distinct eigenvalues

$\lambda_1, \ldots, \lambda_m$ of T. Thus there exist $a_1, \ldots, a_m \in \mathbb{F}$, none of which are 0 (because of the minimmality of m) such that

$$a_1 v_1 + \cdots + a_m v_m = 0.$$

Apply $T - \lambda_m I$ to both sides to obtain

$$a_1(\lambda_1 - \lambda_m)v_1 + \cdots + a_{m-1}(\lambda_{m-1} - \lambda_m)v_{m-1} = 0.$$

Because the eigenvalues are distinct, none of the coefficients of the vectors are 0, so this contradicts the minimmality of m. ■

Proposition: Suppose V is finite-dimensional. Then each operator on V has at most $\dim V$ distinct eigenvalues.

Proof: Suppose $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of $T \in \mathcal{L}(V)$. Then the corresponding eigenvectors $v_1, \ldots, v_m$ are linearly independent, which implies $m \leq \dim V$. ■

5.1.2. Polynomials Applied to Operators

T^m denotes $\underbrace{T \cdots T}_{m \text{ times}}$, with $T^0 = I$ and $T^{-m} = (T^{-1})^m$ if T is invertible.

For a $p \in \mathcal{P}(\mathbb{F})$, $p(T)$ denotes

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m.$$

We have $(pq)(T) = p(T)q(T) = q(T)p(T)$ (formal proof omitted because obvious). What this implies is that a polynomial with an operator plugged in can be factored like normal and the order of the factors doesn't matter (e.g. $T^2 - 9I = (T - 3I)(T + 3I) = (T + 3I)(T - 3I)$).

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

Proof: Suppose $u \in \operatorname{null} p(T)$. Then $p(T)u = 0$. Thus

$$(p(T))(Tu) = T(p(T)u) = T(0) = 0.$$

Thus $Tu \in \operatorname{null} p(T)$, so it is invariant under T.

Suppose $u \in \operatorname{range} p(T)$. Then there exists $v \in V$ such that $u = p(T)v$. Thus

$$Tu = T(p(T)v) = p(T)(Tv),$$

so it is invariant under T. ■

5.1.3. Problems

Problem (Exercise 1): Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V.
- Prove that if $U \subseteq \text{null } T$, then U is invariant under T.
- Prove that is $\text{range } T \subseteq$, then U is invariant under T.

Solution:
- Let $u \in U$. Then $Tu = 0 \in U$.
- Let $u \in U$. Then $Tu \in \text{range } T \subseteq U$.

Problem (Exercise 2): Suppose that $T \in \mathcal{L}(V)$ and $V_1, ..., V_m$ are subspaces of V invariant under T. Prove that $V_1 + \cdots + V_m$ is invariant under T.

Solution: Let V' be the sum of the subspaces, and suppose $v' \in V'$. Then we have writw $v' = v_1 + \cdots + v_m$, where $v_i \in V_i$. Applying T yields $\sum T v_i$. Note that each v_i under T stays in the same subspace, so the sum of them all will still be in V'.

Problem (Exercise 3): Suppose $T \in \mathcal{L}(V)$. Prove that the intersection of every collection of subspaces of V invariant under T is invariant under T.

Solution: Suppose v is in every invariant subspace under T. Since each subspace is invariant, Tv will also be in each subspace, meaning it is also in the intersection.

Problem (Exercise 4): Prove or give a counterexample: If V is finite-dimensional and U is a subspace of V that is invariant under every operator on V, then $U = \{0\}$ or $U = V$.

Solution: Since V is finite-dimensional, let $u_1, ..., u_m$ be a basis of U. Suppose $0 < m < \dim V$. Then we can extend the basis to a basis of V and let $T u_i = v$ for a basis vector v outside U, which means the subspace is not invariant under this operator. Thus $U = \{0\}$ or $U = V$.

Problem (Exercise 5): Suppose $T \in \mathcal{L}(\mathbb{R}^2)$ is defined by $T(x, y) = (-3y, x)$. Find the eigenvalues of T.

Solution: We need $T(x, y) = \lambda(x, y)$. We have the following system of equations:

$$\begin{cases} \lambda x = -3y \\ \lambda y = x \end{cases}.$$

Subbing in the second equation into the first, we have $\lambda^2 y = -3y$. If $y = 0$, then $x = 0$, which we can't have since $(x, y) \neq 0$. Thus we can divide by y to obtain $\lambda^2 = -3$. This has no real solutions, so T has no eigenvalues.

Problem (Exercise 6): Define $T \in \mathcal{L}(\mathbb{F}^2)$ by $T(w, z) = (z, w)$. Find all eigenvalues and eigenvectors of T.

Solution: We need solutions to

$$\begin{cases} \lambda w = z \\ \lambda z = w \end{cases}.$$

From this we get $\lambda^2 w = w$. w can't be 0, so we have $\lambda^2 = 1$. Thus, $\lambda = \pm 1$. This yields eigenvectors of (w, w) and $(w, -w)$.

Problem (Exercise 7): Define $T \in \mathcal{L}(\mathbb{F}^3)$ by $T(z_1, z_2, z_3) = (2z_2, 0, 5z_3)$. Find all eigenvalues and eigenvectors of T.

Solution: We need solutions to

$$\begin{cases} \lambda z_1 = 2z_2 \\ \lambda z_2 = 0 \\ \lambda z_3 = 5z_3 \end{cases}.$$

If $z_2 = 0$, then $z_1 = 0$, so $\lambda = 5$ is an eigenvalue, with associated eigenvector $(0, 0, z_3)$. Otherwise $\lambda = 0$ with associated eigenvector $(z_1, 0, 0)$.

Problem (Exercise 8): Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that if λ is an eigenvalue of P, then $\lambda = 0$ or $\lambda = 1$.

Solution: We need λ such that $Pv = \lambda v$. Applying P to both sides yields $P^2 v = \lambda P v \Rightarrow Pv = \lambda^2 v$. Thus $\lambda^2 v = \lambda v$. This implies $\lambda = 0, 1$.

Problem (Exercise 9): Define $T : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ by $Tp = p'$. Find all eigenvalues and eigenvectors of T.

Solution: We need $\lambda p = p'$. Solving this differential equation yields $Ce^{\lambda x} = p$. The left is only a polynomial when $\lambda = 0$, which has associated eigenvector $p = c$.

Problem (Exercise 10): Define $T \in \mathcal{P}_4(\mathbb{R})$ by $Tp = xp'$. Find all eigenvalues and eigenvectors of T.

Solution: We need $\lambda p = xp'$. Solving the differential equation yields $Cx^\lambda = p$. Since p must have degree at most 4, we have $\lambda = 0, 1, 2, 3, 4$ with associated eigenvectors c, cx, cx^2, cx^3, cx^4.

Problem (Exercise 11): Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\alpha \in \mathbb{F}$. Prove that there exists $\delta > 0$ such that $T - \lambda I$ is invertible for all $\lambda \in \mathbb{F}$ such that $0 < |\alpha - \lambda| < \delta$.

Solution: If $T - \lambda I$ is invertible, then λ is not an eigenvalue, so we just need to find an open interval with α as its midpoint not containing any eigenvalue of T. Since V is finite-dimensional, there are finitely many eigenvalues, so a sufficiently small choice of delta ($\delta = \min\{|\alpha - \lambda_i|\}$) will suffice.

Problem (Exercise 12): Suppose $V = U \oplus W$, where U and W are nonzero subspaces of V. Define $P \in \mathcal{L}(V)$ by $P(u + w) = w$. Find all eigenvalues and eigenvectors of P.

Solution: We need $u = \lambda u + \lambda w$. If $\lambda = 0$, then $u = 0$ and w can be anything, so $E(0, P) = W$. Similarly, if $\lambda = 1$, then $E(1, P) = U$. If $\lambda \neq 0, 1$ then $\frac{1-\lambda}{\lambda} u = w$, but this is impossible unless both u and w are 0.

Problem (Exercise 13): Suppose $T \in \mathcal{L}(V)$. Suppose $S \in \mathcal{L}(V)$ is invertible. Prove that T and $S^{-1}TS$ have the same eigenvalues, and describe the relationship betweent the eigenvectors of the two maps.

Solution: Suppose λ is an eigenvalue of T. Then we have $Tv = \lambda v$. Since S is invertible, there exists $u \in V$ such that $Su = v$. Replacing v with Su yields

$$TSu = \lambda Su.$$

Apply S^{-1} to both sides to obtain

$$S^{-1}TSu = \lambda u,$$

so lambda is indeed an eigenvector of $S^{-1}TS$. We can reverse this process to show that every eigenvalue of $S^{-1}TS$ is also an eigenvalue of T. From this, it's clear that the eigenvectors of $S^{-1}TS$ are $S^{-1}v$, where v is an eigenvector of T.

Problem (Exercise 14): Give an example of an operator on $\mathbb{R}^4$ that has no real eigenvalues.

Solution: $T(w, x, y, z) = (x, y, z, -w)$ has characteristic polynomial $\lambda^4 + 1 = 0$.

Problem (Exercise 15): Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{F}$. Show that λ is an eigenvalue of T if and only if λ is an eigenvalue of the dual operator $T' \in \mathcal{L}(V')$.

Solution: Suppose λ is an eigenvalue of T'. We have $T'(\varphi) = \lambda\varphi$. Restrict φ to the subspace that's not null φ (exists since $\varphi \neq 0$). This means φ is injective. Pick v from this subspace and plug it in to both sides to obtain

$$\varphi(Tv) = \varphi(\lambda v).$$

Injectivity implies $Tv = \lambda v$, so λ is an eigenvalue of T.

Now suppose λ is an eigenvalue of T. Then $\dim(T - \lambda I) \neq \dim V$. Thus $\dim(T' - \lambda I') = \dim(T - \lambda I) \neq \dim V$, so it does have an eigenvalue.

Problem (Exercise 16): Suppose $v_1, ..., v_n$ is a basis of V and $T \in \mathcal{L}(V)$. Prove that if λ is an eigenvalue of T, then

$$|\lambda| \leq n \max\left\{\left|\mathcal{M}(T)_{j,k}\right| : 1 \leq j, k, \leq n\right\}.$$

Solution: Let M be the max in the problem. If v is an eigenvector of T, we can write

$$T(a_1 v_1 + \cdots + a_n v_n) = \lambda a_1 v_1 + \cdots + \lambda a_n v_n.$$

We can write the left as

$$a_1 \left(A_{1,1} v_1 + \cdots + A_{n,1} v_n\right) + \cdots + a_n \left(A_{1,n} v_1 + \cdots + A_{n,n} v_n\right)$$

Pick r such that $|a_r|$ is maximized. For that r we have

$$\sum_{c=1}^{n} a_c A_{r,c} = \lambda a_r.$$

Taking absolute values on both sides and using the triangle inequality on the left, we have

$$\sum_{c=1}^{n} \left|a_c A_{r,c}\right| \geq |\lambda a_r|.$$

By definition, we have

$$|M| \sum_{c=1}^{n} |a_c| \geq |\lambda a_r|.$$

We can divide by $|a_r|$, and since $|a_r|$ is maximized, we have

$$n|M| \geq |M| \sum_{c=1}^{n} \left|\frac{a_c}{a_r}\right| \geq |\lambda|,$$

so we're done.

Problem (Exercise 17): Suppose $\mathbb{F} = \mathbb{R}$, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{R}$. Prove that λ is an eigenvalue of T if and only if λ is an eigenvalue of $T_{\mathbb{C}}$.

Solution: If $Tv = \lambda v$, then clearly $T(v + 0i) = Tv + T(0)i = \lambda v$. If $T(v + ui) = \lambda(v + ui)$, then we must have $Tv = \lambda v$ and $Tu = \lambda u$.

Problem (Exercise 18): Suppose $\mathbb{F} = \mathbb{R}$, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{C}$. Prove that λ is an eigenvalue of $T_{\mathbb{C}}$ if and only if $\overline{\lambda}$ is an eigenvalue of $T_{\mathbb{C}}$.

Solution: Suppose $T(v + ui) = (a + bi)(v + ui) = (av - bu) + i(bv + au)$. Then $Tv = av - bu$ and $Tu = bv + au$. Thus $T(v - ui) = (av - bu) - i(bv + au) = (a - bi)(v - ui)$.

Problem (Exercise 19): Show that the forward shift operator $T \in \mathcal{L}(\mathbb{F}^\infty)$ defined by

$$T(z_1, z_2, ...) = (0, z_1, z_2, ...)$$

has no eigenvalues.

Solution: We need $(\lambda z_1, \lambda z_2, \lambda z_3, ...) = (0, z_1, z_2, ...)$. λ can't be 0, otherwise $v = 0$. This we need $z_1 = 0$. However, this implies $z_2 = 0$, and so on, so there are no eigenvalues.

Problem (Exercise 20): Define the backward shift operator $S \in \mathcal{L}(\mathbb{F}^\infty)$ defined by

$$S(z_1, z_2, z_3, ...) = (z_2, z_3, ...).$$

Show that every element of $\mathbb{F}$ is an eigenvalue of S and find all eigenvectors of S.

Solution: Each of the following is the eigenvector for the eigenvalue λ: $(z, \lambda z, \lambda^2 z, ...)$.

Problem (Exercise 21): Suppose $T \in \mathcal{L}(V)$ is invertible. Suppose $\lambda \in \mathbb{F}$ with $\lambda \neq 0$. Prove that λ is an eigenvalue of T if and only if $\frac{1}{\lambda}$ is an eigenvalue of T^{-1}, and prove that T and T^{-1} have the same eigenvectors.

Solution: We have $Tv = \lambda v$. Applying T^{-1} to both sides yields

$$v = \lambda T^{-1} v \Rightarrow \frac{1}{\lambda} v = T^{-1} v,$$

implying $\frac{1}{\lambda}$ is an eigenvalue of T^{-1} and that v is an eigenvector of T^{-1}. We can go in the other direction to show the reverse inclusion. Thus both sets are equal for both maps.

Problem (Exercise 22): Suppose $T \in \mathcal{L}(V)$ and there exist nonzero vectors u and w in V such that

$$Tu = 3w \text{ and } Tw = 3u.$$

Prove that 3 and -3 is an eigenvalue of T.

Solution: Apply T to both sides of the first equation to get $T^2u = 3Tw$. Subbing the second equation into this new equation yields $T^2u = 9u$. Factoring this yields

$$(T - 3I)(T + 3I)u = 0.$$

If $(T - 3I)u = 0$, then 3 is an eigenvalue. Otherwise if $(T - 3I)u \neq 0$, then $(T + 3I)v = 0$, so -3 is an eigenvalue.

Problem (Exercise 23): Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that ST and TS have the same eigenvalues.

Solution: Suppose $STv = \lambda v$. Then we have $TS(Tv) = T(STv) = T(\lambda v) = \lambda Tv$, so λ is an eigenvalue of TS. We can do the same thing in reverse to get the reverse inclusion, so both maps have the same set of eigenvectors.

Problem (Exercise 24): Suppose A is an n by n matrix with entries in $\mathbb{F}$. Define $T \in \mathcal{L}(\mathbb{F}^n)$ by $Tx = Ax$.

- Suppose the sum of the entries in each row of A equals 1. Prove that 1 is an eigenvalue of T.
- Suppose the sum of the entries in each column of A equals 1. Prove that 1 is an eigenvalue of T.

Solution: For the first bullet, take $x = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$.

For the second bullet, we can set up a system of equations that when added together yields $0 = 0$. Thus there are infinitely many solutions, any of which is an eigenvector.

Alternatively note that A^t is the matrix of T', and since T has eigenvalue 1, T' must also have eigenvalue 1.

Problem (Exercise 25): Suppose $T \in \mathcal{L}(V)$ and u, w are eigenvectors of T such that $u + w$ is also an eigenvector of T. Prove that u and w are eigenvectors of T corresponding to the same eigenvalue.

Solution: Suppose they don't correspond to the same eigenvalue. Thus one is not a scalar multiple of the other. We have $Tv = \lambda_1 v$ and $Tw = \lambda_2 w$. Then we have $\lambda_1 v + \lambda_2 w = T(u + w) = \lambda_3(u + w)$. Rewriting yields

$$(\lambda_3 - \lambda_1)u = (\lambda_2 - \lambda_3)w.$$

If $\lambda_2 = \lambda_3$, then $\lambda_1 = \lambda_3$, but this is impossible since $\lambda_1 \neq \lambda_2$. Thus we can divide by $\lambda_2 - \lambda_3$. However, that implies w is a scalar multiple of u, meaning that $w = \operatorname{span}(v)$, so they doth both correspond to the same eigenvalue, a contradiction.

Problem (Exercise 26): Suppose $T \in \mathcal{L}(V)$ is such that every nonzero vector in V is an eigenvector of T. Prove that T is a scalar multiple of the identity operator.

Solution: Using the previous exercise, we can conclude that $Tv = \lambda I v$ for every $v \in V$.

Problem (Exercise 28): Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has at most $1 + \dim \operatorname{range} T$ distinct eigenvalues.

Solution: Let $\lambda_1, ..., \lambda_m$ be distinct nonzero eigenvalues. Since a list of eigenvectors corresponding to distinct eigenvalues is linearly independent, we have $\lambda_1 v_1, ..., \lambda_m v_n$ is linearly independent. Each of these is in range T, so $m \leq \dim \operatorname{range} T$. Then we add an extra possible eigenvalue for when the eigenvalue is 0.

Problem (Exercise 30): Suppose $T \in \mathcal{L}(V)$ and $(T - 2I)(T - 3I)(T - 4I) = 0$. Suppose λ is an eigenvalue of T. Prove that $\lambda = 2$ or 3 or 4.

Solution: Let v be an eigenvector with eigenvalue λ. We have

$$\begin{aligned}(T - 2I)(T - 3I)(T - 4I)v &= (T - 2I)(T - 3I)((\lambda - 4)v) \\ &= (T - 2I)((\lambda - 3)(\lambda - 4)v) \\ &= (\lambda - 2)(\lambda - 3)(\lambda - 4)v = 0.\end{aligned}$$

Thus we must have $\lambda = 2$ or 3 or 4.

Problem (Exercise 31): Give an example of $T \in \mathcal{L}(\mathbb{R}^2)$ such that $T^4 = -I$.

Solution: 45° rotation map.

Problem (Exercise 32): Suppose $T \in \mathcal{L}(V)$ has no eigenvalues and $T^4 = I$. Prove that $T^2 = -I$.

Solution: Note that $T^4 - I = 0 \Rightarrow (T - I)(T + I)(T^2 + I) = 0$. Since T has no eigenvalues, $T - I$ nor $T + I$ can be 0 on any vector. Thus we must have $T^2 + I = 0$.

Problem (Exercise 33): Suppose $T \in \mathcal{L}(V)$ and m is a positive integer.
- Prove that T is injective if and only if T^m is injective.
- Prove that T is surjective if and only if T^m is surjective.

Solution:
- If T is injective, then obviously any successive applications of T on a nonzero vector will still be nonzero. If T is not injective, and if $Tv = 0$ for nonzero v, then $T^m v = T^{m-1}(Tv) = T^{m-1}(0) = 0$.
- If T is surjective, then T^2 is also surjective, since T outputs every vector in V, which the next T then takes in. By induction T^m is surjective. If T is not surjective, then there is some vector that cannot be reached regardless of input vector, which implies that T^m is not surjective.

Problem (Exercise 35): Suppose that $\lambda_1, ..., \lambda_n$ is a list of distinct real numbers. Prove that the list $e^{\lambda_1 x}, ..., e^{\lambda_n x}$ is linearly independent in the vector space of real values functions on $\mathbb{R}$.

Solution: Let $V = \operatorname{span}(e^{\lambda_1 x}, ..., e^{\lambda_n x})$ and let D be the differentiation operator on V. Then it's easy to see that λ_i is an eigenvalue with eigenvector $e^{\lambda_i}x$. Since eigenvectors that correspond to distinct eigenvalues are linearly independent, we have our desired conclusion.

Problem (Exercise 36): Suppose $\lambda_1, ..., \lambda_n$ is a list of real numbers with distinct absolute values. Prove that $\cos(\lambda_1 x), ..., \cos(\lambda_n x)$ is linearly independent in the vector space of real valued functions on $\mathbb{R}$.

Solution: Define $V = \operatorname{span}(\cos(\lambda_i x))$ and $T(\sum c_i \cos(\lambda_i x)) = \sum c_i \lambda_i \cos(\lambda_i x)$. It can easily be seen that T is a linear map. It has eigenvalues λ_i with eigenvector $\cos(\lambda_i x)$. Like the previouos, this implies the list is linearly independent.

Problem (Exercise 37): Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Define $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ by

$$\mathcal{A}(S) = TS$$

for each $S \in \mathcal{L}(V)$. Prove that the set of eigenvalues of T equals the set of eigenvalues of $\mathcal{A}$.

Solution: Suppose $TS = \lambda S$. This implies $(T - \lambda I)S = 0$. Since $S \neq 0$, $T - \lambda I$ must have a nonzero null space, which means λ is an eigenvalue of T.

Now suppose $(T - \lambda I)v = 0$. Choose S such that range $S = \operatorname{span}(v)$. Then we have $(T - \lambda I)(Sv) = 0$ for all $v \in V$, which implies $TS = \lambda S$, so λ is an eigenvalue of $\mathcal{A}$.

Problem (Exercise 38): Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V invariant under T. The *quotient operator* $T/U \in \mathcal{L}(V/U)$ is defined by

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

for each $v \in V$.

- Show that the definition of T/U makes sense and show that T/U is an operator on V/U.
- Show that each eigenvalue of T/U is an eigenvalue of T.

Solution: First we show that choice of representative doesn't matter. Suppose $v + U = w + U$. Then $v - w \in U$. This implies $T(v - w) \in U$, which further implies $Tv + U = Tw + U$, so indeed the choice of representative doesn't matter. From here it's easy to see T/U is an operator.

Suppose λ is an eigenvalue of T/U. Then we have

$$Tv + U = \lambda v + U.$$

Thus $Tv - \lambda v \in U$. Thus we have $Tv - \lambda v = u$ for some $u \in U$. If λ is an eigenvalue of T, then we need $T(u' + v) = \lambda(u' + v) \Rightarrow Tv - \lambda v = \lambda u' - Tu'$. Thus we need $(T - \lambda I)u' = -u$. If $T - \lambda I$ is not invertible, then $\operatorname{null}(T - \lambda I) \neq \{0\}$, so λ is an eigenvalue. Otherwise u' does exist and so $u' + v$ is an eigenvector with eigenvalue λ.

Problem (Exercise 40): Suppose $S, T \in \mathcal{L}(V)$ and S is invertible. Suppose $p \in \mathcal{P}(\mathbb{F})$ is a polynomial. Prove that

$$p(STS^{-1}) = Sp(T)S^{-1}.$$

Solution: Note that $(STS^{-1})^n = STS^{-1}STS^{-1} \cdots STS^{-1} = ST^nS^{-1}$. Next we have

$$aST^nS^{-1} + bST^mS^{-1} = S(aT^nS^{-1} + bT^mS^{-1}) = S((aT^n + bT^m)(S^{-1})) = S(aT^n + bT^m)S^{-1}.$$

Thus the conclusion follows easily.

5.2. The Minimal polynomial

5.2.1. Existence of Eigenvalues on Complex Vector Spaces

Theorem (existence of eigenvalues): Every operator on a finite-dimensional nonzero complex vector space has an eigenvalue.

Proof: Let $\dim V = n$ and $T \in \mathcal{L}(V)$. Choose $v \in V$ with $v \neq 0$. Then

$$v, Tv, T^2v, ..., T^nv$$

is not linearly independent since it has length $n + 1$. Thus there exists a nonconstant polynomial p of smallest degree such that

$$p(T)v = 0.$$

By the funamental theorem of algebra, we can write

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

This implies

$$0 = p(T)v = (T - \lambda I)(q(T)v).$$

Because q has smaller degree than p, $q(T)v \neq 0$. Thus the equation above implies that λ is an eigenvalue with eigenvector $q(T)v$. ∎

5.2.2. Eigenvalues and the Minimal Polynomial

Proposition: Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then there is a unique monic polynomial $p \in \mathcal{P}(\mathbb{F})$ of smallest degree such that $p(T) = 0$. Furthermore, $\deg p \leq \dim V$.

Proof: If $\dim V = 0$, then I is the zero operator on V and thus we take p to be the constant polynomial 1.

Now we induct on $\dim V$. Suppose the result is true for every integer less than $\dim V$. Let $0 \neq v \in V$. The list $v, Tv, ..., T^{\dim V}v$ has length $1 + \dim V$ and thus is linearly dependent. By the linear dependence lemma, there is a smallest m such that $T^m v$ is a linear combination of the vectors in the list before it. Thus we have scalars c_i such that

$$c_0 v + c_1 T v + \cdots + c_{m-1} T^{m-1} v + T^m v = 0.$$

Define $q \in \mathcal{P}_m(\mathbb{F})$ by

$$c_1 + c_1 z + \cdots + c_{m-1} z^{m-1} + z^m.$$

Thus we have $q(T)v = 0$.

For nonnegative k we have

$$q(T)\left(T^k v\right) = T^k(q(T)v) = T^k(0) = 0.$$

By the linear dependence lemma $v, Tv, ..., T^{m-1}v$ are linearly independent. Thus $\dim \operatorname{null} q(T) \geq m$. Thus we have

$$\dim \operatorname{range} q(T) = \dim V - \dim \operatorname{null} q(T) \leq \dim V - m.$$

Because range $q(T)$ is invariant under T, we can apply yhe induction hypothesis to the opetator $T|_{\operatorname{range} q(T)}$ on the vector space range $q(T)$. Thus there is a monic polynomial $s \in \mathcal{P}(\mathbb{F})$ with

$$\deg s \leq \dim V - m \text{ and } s\left(T|_{\operatorname{range} q(T)}\right) = 0.$$

Hence for all $v \in V$ we have

$$(sq)(T)(v) = s(T)(q(T)v) = 0$$

because $q(T)v \in \operatorname{range} q(T)$ and $s(T)|_{\operatorname{range} q(T)} = s\left(T|_{\operatorname{range} q(T)}\right) = 0$. Thus sq is a monic polynomial such that $\deg sq \leq \dim V$ and $(sq)(T) = 0$. This completes existence.

To prove uniqueness, suppose p and r are two distinct monic minimal polynomials. Then we have $(p - r)(T) = 0$. However, note that $\deg(p - r) < \deg p$, and we can divide by the leading coefficient of $p - r$ to a get a monic polynomial, contradicting minimmality. Thus $p - r = 0$. ■

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

Proposition (eigenvalues are the zeroes of the minimal polynomial): Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then the zeroes of the minimal polynomial of T are the eigenvalues of T. Further, if V is a complex vector space, then the minimal polynomial T has the form

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

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

Proof: Let p be the minimal polynomial of T. First suppose $\lambda \in \mathbb{F}$ is a zero of p. Then p can be written in the form

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

where q is a monic polynomial with coefficients in $\mathbb{F}$. Because $p(T) = 0$, we have

$$0 = (T - \lambda I)(q(T)v)$$

for all $v \in V$. Because the degree of q is less than the degree of the minimal polynomial p, there exists at least one vector $v \in V$ such that $q(T)v \neq 0$. The equation above thus implies that λ is an eigenvalue of T, as desired.

Now suppose $\lambda \in \mathbb{F}$ is an eigenvalue of T. Thus there exists $v \in V$ with $v \neq 0$ such taht $Tv = \lambda v$. Repeated applications of T to both sides yield $T^k v = \lambda^k v$. Thus

$$p(T)v = p(\lambda)v.$$

Because p is the minimal polynomial, the left side is 0, which then implies $p(\lambda) = 0$, so λ is a zero of p.

The last part of the proposition follows from the fundamental theorem of algebra. ■

Proposition: Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $q \in \mathcal{P}(\mathbb{F})$. Then $q(T) = 0$ if and only if q is a multiple of the minimal polynomial of T.

Proof: Let p be the minimal polynomial.

Suppose $q(T) = 0$. Thus we have

$$q = ps + r.$$

We also have

$$0 = q(T) = p(T)s(T) + r(T) = r(T).$$

Since $\deg r < \deg p$, we must have $r = 0$, so q is indeed a multiple of p.

Now suppose q is a multiple of p. Then we have

$$q(T) = p(T)s(T) = 0.$$

■

Corollary: Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, 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 $T|_U$.

Proof: Suppose p is the minimal polynomial of T. Thus $p(T)v = 0$ for all $v \in V$, which implies $p(T)u = 0$ for all $u \in U$. Thus $p(T|_U) = 0$. Thus, applying the previous proposition to $T|_U$ yields that p is a polynomial multiple of the minimal polynomial of $T|_U$. ■

Proposition: Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then T is not invertible if and only if the constant term of the minimal polynomial of T is 0.

Proof: Suppose $T \in \mathcal{L}(V)$ and p is the minimal polynomial of T. Then

$$\begin{aligned} T \text{ is not invertible} &\iff 0 \text{ is an eigenvalue of } T \\ &\iff 0 \text{ is a zero of } p \\ &\iff \text{the constant term of } p \text{ is } 0. \end{aligned}$$

■

5.2.3. Eigenvalues on Odd-Dimensional Real Vector Spaces

Proposition (even-dimensional null space): Suppose $\mathbb{F} = \mathbb{R}$ and V is finite-dimensional. Suppose also that $T \in \mathcal{L}(V)$ and $b, c \in \mathbb{R}$ with $b^2 < 4c$. Then $\dim \operatorname{null} (T^2 + bT + cI)$ is an even number.

Proof: We have that null $(T^2 + bT + cI)$ is invariant under T. We can replace V with $\operatorname{null}(T^2 + bT + cI)$ and replacing T with T restricted to $\operatorname{null}(T^2 + bT + cI)$, we can assume $T^2 + bT + cI = 0$. Now we need to prove that $\dim V$ is even.

Suppose $\lambda \in \mathbb{R}$ and $v \in V$ are such that $Tv = \lambda v$. Then

$$0 = (T^2 + bT + cI)v = (\lambda^2 + b\lambda + c)v = \left(\left(\lambda + \frac{b}{2}\right)^2 + c - \frac{b^2}{4}\right)v.$$

The scalar on the right is positive, so this implies $v = 0$, which means T has no eigenvectors.

Let U be a subspace of V that is invariant under T and has the largest dimension among all subspaces of V that are invariant under T and have even dimension. If $U = V$, then we're done, otherwise assume there exists $w \in V$ such that $w \notin U$.

Let $W = \operatorname{span}(w, Tw)$. Then W is invariant under T because $T(Tw) = -bTw - cw$. Furthermore, $\dim W = 2$ because otheriwse w would be an eigenvector of T. Now we have

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W) = \dim U + 2,$$

where $U \cap W = \{0\}$ because otherwise $U \cap W$ would be a one-dimensional subspace of V that is invariant under T.

Because $U + W$ is invariant under T, the equation shows that there exists a subspace of V invariant under T of even dimension larger than $\dim U$. Thus we have a contradiction. ∎

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

Proof: Let odd $n = \dim V$ and $T \in \mathcal{L}(V)$. We use induction on n to show T has an eigenvalue. Note that the result holds for $n = 1$ since every vector in V is an eigenvector of T.

Suppose $n \geq 3$ and the result holds for all odd integers less than n. Let p be the minimal polynomial of T. If p is a polynomial multiple of $x - \lambda$ for some $\lambda \in \mathbb{R}$, then λ is an eigenvalue of T and we are done. Thus we can assume p is a multiple of $x^2 + bx + c$ that has no real roots.

We have $p = q(x^2 + bx + c)$ for some polynomial q. We have

$$0 = p(T) = (q(T))(T^2 + bT + cI),$$

which means that $q(T)$ equals 0 on $\operatorname{range}(T^2 + bT + cI)$. Because $\deg p < \deg p$ and p is the minimal polynomial of T, this implies that $\operatorname{range}(T^2 + bT + cI) \neq V$.

We have

$$\dim V = \dim \operatorname{null} (T^2 + bT + cI) + \dim \operatorname{range}(T^2 + bT + cI).$$

Because $\dim V$ is odd and $\dim \operatorname{null}(T^2 + bT + cI)$ is even, the range of the map is odd.

We know range$(T^2 + bT + cI)$ is invariant under T and has odd dimension less than $\dim V$. Thus by our hypothesis, T restricted to range$(T^2 + bT + cI)$ has an eigenvalue, which means T has an eigenvalue. ■

5.2.4. Problems

Problem (Exerise 1): Suppose $T \in \mathcal{L}(V)$. Prove that 9 is an eigenvalue of T^2 if and only if 3 or −3 is an eigenvalue of T.

Solution: Suppose 9 is an eigenvalue of T^2. Then $T^2 - 9I = (T - 3I)(T + 3I) = 0$. This implies 3 or −3 is an eigenvalue.

Now suppose 3 or −3 is an eigenvalue. Then we have $(T \pm 3I)v = 0$. Applying $T \mp 3I$ to both sides yields $(T^2 - 9I)v = 0$, so 9 is an eigenvalue of T^2.

Problem (Exercise 2): Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ has no eigenvalues. Prove that every subspace of V invariant under T is either $\{0\}$ or infinte dimensional.

Solution: If V was nonzero finite dimensional and invariant, then $T\mid_U$ would necessarily have an eigenvalue, which would mean T would have an eigenvalue. Thus the subspace must either be $\{0\}$ or infinite dimensional.

Problem (Exercise 3): Suppose n is a positive integer and $T \in \mathbb{F}^n$ is defined by

$$T(x_1, ..., x_n) = (x_1 + \cdots + x_n, ..., x_1 + \cdots + x_n).$$

- Find all eigenvalues and eigenvectors of T.
- Find the minimal polynomial of T.

Solution: The minimal polynomial is $T^2 - nT$, and since $T - aI$ cannot be equal to 0, it is indeed minimal. Thus T has eigenvalues n and 0. n has the corresponding eigenvector $(1, ..., 1)$. For 0, there are $n - 1$ eigenvectors: $(1, 0, 0, ..., 0, -1, 0, ..., 0)$, where −1 takes each of the $n - 1$ spots after 1.

Problem (Exercise 6): Suppose $T \in \mathcal{L}(\mathbb{F}^2)$ is defined by $T(w, z) = (-z, w)$. Find the minimal polynomial of T.

Solution: $T^2 + 1$

Problem (Exercise 8): Suppose $T \in \mathcal{L}(\mathbb{R}^2)$ is the operator of counterclockwise rotation by 1°. Find the minimal polynomial of T.

Solution: Everything here will be in degrees. The minimal polynomial is $z^2 - 2\cos(1)z + 1$. Note that the matrix of T is

$$\begin{pmatrix} \cos 1 & -\sin 1 \\ \sin 1 & \cos 1 \end{pmatrix}.$$

We can calculate the this matrix on the minimal polynomial and confirm that it is 0. Note that the polynomial has no real roots, so this polynomial is indeed minimal.

Problem (Exercise 12): Define $T \in \mathcal{L}(\mathbb{F}^n)$ by $T(x_1, x_2, x_3, ..., x_n) = (x_1, 2x_2, 3x_3, ..., nx_n)$. Find the minimal polynomial of T.

Solution: The eigenvalues of T are $1, 2, ..., n$. Thus the minimal polynomial is

$$(z-1)(z-2)\cdots(z-n).$$

Problem (Exercise 13): Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbb{F})$. Prove that there exists a unique $r \in \mathcal{P}(\mathbb{F})$ such that $p(T) = r(T)$ and $\deg r$ is less than the degree of the minimal polynomial of T.

Solution: Let m be the minimal polynomial. If $\deg p < \deg m$, then let $r = p$. Otherwise choose r such that $p - r = ms$, where s is some polynomial. Thus r exists and is unique (by division algorithm), and necessarily has degree at most $\deg m$.

Problem (Exercise 14): Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$ has minimal polynomial $4 + 5z - 6z^2 - 7z^3 + 2z^4 + z^5$. Find the minimal polynomial of T^{-1}.

Solution: Note that the constant term is nonzero, so T does have an inverse. We have

$$T^5 + 2T^4 - 7T^3 - 6T^2 + 5T + 4I = 0.$$

Applying T^{-5} to both sides yields

$$I + 2T^{-1} - 7T^{-2} - 6T^{-3} + 5T^{-4} + 4T^{-5}.$$

Thus $p = z^5 + \frac{5}{4}z^4 - \frac{3}{2}z^3 - \frac{7}{4}z^2 + \frac{1}{2}z + \frac{1}{4}$ is 0 on T^{-1}. Note that if the minimal polynomial of T^{-1} had degree less than, we could do the exact reverse procedure and conclude that T has minimal polynomial with degree less than 5, which is a contradiction. Thus p is the minimal polynomial of T^{-1}.

Problem (Exercise 15): Suppose V is a finite-dimensional complex vector space with $\dim V > 0$ and $T \in \mathcal{L}(V)$. Define $f : \mathbb{C} \to \mathbb{R}$ by

$$f(\lambda) = \dim \operatorname{range}(T - \lambda I).$$

Prove that f is not a continuous function.

Solution: We can rewrite $f(\lambda) = \dim V - \dim \operatorname{null}(T - \lambda I)$. Note that on eigenvalues of T, $\dim \operatorname{null}(T - \lambda I)$ changes from 0 to some positive integer, and since there are finitely many eigenvalues of T, it's easy to see that f is not continuous.

Problem (Exercise 21): Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the minimal polynomial of T has degree at most $1 + \dim \operatorname{range} T$.

Solution: Note that $v, Tv, ..., T^{\dim \operatorname{range} T} v$ is linearly dependent, since it's a list of length $1 + \dim \operatorname{range} T$ in $\operatorname{range} T$. From here continue as we did in the original proof.

Problem (Exercise 22): Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T is invertible if and only if $I \in \operatorname{span}(T, T^2, ..., T^{\dim V})$.

Solution: If T is invertible, then its minimal polynomial has nonzero constant term. Thus we can easily write I in terms of $T, T^2, ..., T^{\dim V}$. Now suppose I is some linear combination of the list. Then write all the operators on one side to see that it's equal to 0, so it's a multiple of the minimal polynomial of T. Since this polynomial has nonzero constant term, the minimal polynomial of T must also have nonzero constant term, which means T is invertible.

Problem (Exercise 25): Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V that is invariant under T.

- Prove that the minimal polynomial of T is a multiple of the minimal polynomial of T/U.
- Prove that

  $$(\text{minimal polynomia of } T|_U) \times (\text{minimal polynomial of } T/U)$$

  is a polynomial multiple of T.

Solution:

- Let m be the minimal polynomial of T. Then we have $m(T/U)(v + U) = m(T)v + U = 0 + U$. Thus $m(T/U) = 0$, which implies m is a multiple of the minimal polynomial fo T/U.
- Let p be the minimal polynomial of T/U and q be the minimal polynomial of $T|_U$. We have $p(T/U)(v + U) = p(T)v + U = 0 + U$, which implies $p(T)v = u$ for some $u \in U$. Now apply-

ing q to both sides yields $q(T)p(T)v = q(T)u = 0$. Thus $(pq)(T) = 0$, so the product is a multiple of m.

Problem (Exercise 26): Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V that is invariant under T. Prove that the set of eigenvalues of T equals the unoin of the set of eigenvalues of $T|_U$ and the set of eigenvalues of T/U.

Solution: Follows from the previous problem.

Problem (Exercise 28): Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the minimal polynomial of $T' \in \mathcal{L}(V')$ equals the minimal polynomial of T.

Solution: Note that $(T')^n(\varphi) = \varphi \circ T^n$. Thus $p(T')\varphi = \varphi(p(T))$. If p is the minimal polynomial of T', then $\varphi(p(T)) = 0$ for all φ. Thus $p(T) = 0$ (otherwise we could pick φ not zero on range $p(T)$). This implies p is a multiple of the minimal polynomial of T.

Now suppose $q(T) = 0$ is the minimal polynomial of T. Then we have $\varphi(q(T)) = q(T')(\varphi) = 0$ for all φ, so q is a multiple of the minimal polynomial of T'. Thus the minimal polynomial of both must be the same.

5.3. Upper-Triangular Matrices

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

Definition (upper triangular matrix): A square matrix is called upper triangular if all entries below the diagonal are 0.

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

Proposition (conditions for upper triangular matrix): Suppose $T \in \mathcal{L}(V)$ and $v_1, ..., v_n$ is a basis of V. Then the following are equivalent:
- The matrix of T with respect to $v_1, ..., v_n$ is upper triangular.
- $\text{span}(v_1, ..., v_k)$ is invariant under T for each k.
- $Tv_k \in \text{span}(v_1, ..., v_k)$ for each k.

Proof: Suppose the first bullet holds. Then for any $v \in \text{span}(v_1, ..., v_k)$, each basis vector in v will get mapped to a vector in $\text{span}(v_1, ..., v_k)$, since the matrix is upper triangular, so overall v will stay in $\text{span}(v_1, ..., v_k)$.

If the second bullet holds, then the third obviously follows. Now suppose the third bullet holds. This means that when writing Tv_k as a linear combination of basis vectors, we only require $v_1, ..., v_k$. Thus, all entries under the diagonal of $\mathcal{M}(T)$ are 0, meaning the matrix is upper triangular. ■

Proposition: Suppose $T \in \mathcal{L}(V)$ and V has a basis with respect to which T has an upper triangular matrix with diagonal entries $\lambda_1, ..., \lambda_n$. Then

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

Proof: Let $v_1, ..., v_n$ denote a basis of V with respect to which T has an upper triangular matrix with diagonal entries $\lambda_1, ..., \lambda_n$. Then $Tv_1 = \lambda_1 v_1$, which means $(T - \lambda_1 I)v_1 = 0$, which implies $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_1 = 0$ for all m.

Now note that $(T - \lambda_2 I)v_2 \in \text{span}(v_1)$. Thus $(T - \lambda_1 I)(T - \lambda_2 I)v_2 = 0$, which implies $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_2 = 0$ for $m = 2, ..., n$.

Continuing this pattern, we have $(T - \lambda_1 I) \cdots (T - \lambda_n I)v_k = 0$ for all k, which implies $(T - \lambda_1 I) \cdots (T - \lambda_n I)$ is the 0 operator. ■

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

Proof: Suppose $v_1, ..., v_n$ is a basis of V with respect to which T has an upper triangular matrix

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Because $Tv_1 = \lambda_1 v_1$ we have that λ_1 is an eigenvalue of T.

Pick $k \neq 1$. Then $(T - \lambda_k I_k)v_k \in \operatorname{span}(v_1, ..., v_{k-1})$. Thus $T - \lambda_k I$ maps $\operatorname{span}(v_1, ..., v_k)$ to $\operatorname{span}(v_1, ..., v_{k-1})$. Since

$$k = \dim \operatorname{span}(v_1, ..., v_k) > \dim \operatorname{span}(v_1, ..., v_{k-1}) = k - 1,$$

$(T - \lambda_k I)$ restricted to $\operatorname{span}(v_1, ..., v_k)$ is not injective, which implies there exists $v \in \operatorname{span}(v_1, ..., v_k)$ such that $(T - \lambda_k I)v_k = 0$. Thus λ_k is an eigenvalue of T, and this implies every value on the diagonal of $\mathcal{M}(T)$ is an eigenvalue.

To prove T has no other eigenvalues, let $q(z) = (z - \lambda_1) \cdots (z - \lambda_n)$. By the previous result, $q(T) = 0$. Thus q is a multiple of the minimal polynomial of T. Since every 0 of the minimal polynomial is an eigenvalue of T, every eigenvalue of T is a zero of q. Thus all eigenvalues of T are in the list $\lambda_1, ..., \lambda_n$. ■

Proposition: Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then T has an upper triangular matrix with respect to some basis of V if and only if the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, ..., \lambda_m \in \mathbb{F}$.

Proof: First suppose T has an upper triangular matrix with respect to some basis of V. Let $\alpha_1, ..., \alpha_n$ denote the diagonal entries. Let $q(z) = (z - \alpha_1) \cdots (z - \alpha_n)$. Then $q(T) = 0$ by a previous proposition. Thus q is a multiple of the minimal polynomial, which means the minimal polynomial equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some subset $\{\lambda_i\}$ of $\{\alpha_i\}$.

Now suppose the minimal polynomial of T is $(z - \lambda_1) \cdots (z - \lambda_m)$. We use induction on m. If $m = 1$, then $z - \lambda_1$ is the minimal polynomial of T, which implies $T = \lambda_1 I$, which is upper triangular.

Now suppose the result holds for all integers less than m. Let

$$U = \operatorname{range}(T - \lambda_m I).$$

Then U is invariant under T. Thus $T|_U$ is an operator on U. If $u \in U$, then $u = (T - \lambda_m I)v$ for some $v \in V$ and

$$(T - \lambda_1 I) \cdots (T - \lambda_{m-1} I)u = (T - \lambda_1 I) \cdots (T - \lambda_m I)v = 0.$$

Thus $(z - \lambda_1) \cdots (z - \lambda_{m-1})$ is a polynomial multiple of the minimal polynomial of $T|_U$. Thus the minimal polynomial of $T|_U$ is the product of at most $m - 1$ terms of the form $z - \lambda_k$.

By our induction hypothesis, there is a basis $u_1, \dots, u_M$ of U with respect to which $T|_U$ has an upper triangular matrix. Thus for each $k \in \{1, \dots, M\}$ we have

$$Tu_k = (T|_U)(u_k) \in \text{span}(u_1, \dots, u_k).$$

Extend the u_i's to $u_1, \dots, u_M, v_1, \dots, v_n$, a basis of V. For each $k \in \{1, \dots, N\}$ we have

$$Tv_k = (T - \lambda_m I)v_k + \lambda_m v_k.$$

The definition of U shows that $(T - \lambda_m I)v_k \in U = \text{span}(u_1, \dots, u_M)$. Thus we have

$$Tv_k = \text{span}(u_1, \dots, u_M, v_1, \dots, v_k)$$

, which implies T has an upper triangular matrix. ■

Corollary: 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.

Proof: Follows from the previous result and the fundamental theorem of algebra. ■

5.3.1. Problems

Problem (Exercise 4): Give an example of an operator whose matrix with respect to some basis contains only 0′s on the diagonal, but ht operator is invertible.

Solution:

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

Problem (Exercise 5): Give an example of an operator whose matrix with respect to some basis contains only nonzero numbers on the diagonal, but the operator is not invertible.

Solution:

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

5.4. Diagonalizable Operators

5.4.1. Diagonal Matrices

Definition (diagonal matrix): A *diagonal matrix* is a square matrix that is 0 everywhere except possibly on the diagonal.

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

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

$$E(\lambda, T) = \operatorname{null}(T - \lambda I) = \{v \in V : Tv = \lambda v\}.$$

Proposition (summ 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, if V is finite-dimensional, then

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

Proof: Suppose

$$v_1 + \cdots + v_m = 0,$$

where each v_k is in $E(\lambda_k, T)$. Because eigenvectors corresponding to distinct eigenvalues are linearly independent, this implies each v_k equals 0. Thus $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ is a direct sum.

Now suppose V is finite-dimensional. Then

$$\begin{aligned} \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) &= \dim(E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)) \\ &\leq \dim V. \end{aligned}$$

■

5.4.2. Conditions for Diagonalizability

Proposition (conditions equivalent to diagonalizability): Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, ..., \lambda_m$ denote the distinct eigenvalues of T. Then the following are equivalent:
- T is diagonalizable.
- V has a basis consisting of eigenvectors of T.
- $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.
- $\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$.

Proof: An operator $T \in \mathcal{L}(V)$ has a diagonal matrix

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

with respect to a basis $v_1, ..., v_n$ of V if and only if $Tv_k = \lambda_k v_k$ for each k. Thus the first and second bullet are equivalent.

Suppose the second bullet holds. Then every vector in V is a linear combination of eigenvectors of T, which implies

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

The previous result now shows the third bullet is true.

The third bullet immediately implies the fourth bullet.

Now suppose the fourth bullet is true. Choose a basis of each $E(\lambda_k, T)$. Put all these bases together to form a list $v_1, ..., v_n$ of eigenvectors of T, where $n = \dim V$. To show that this list is linearly independent, suppose

$$a_1 v_1 + \cdots + a_n v_n = 0.$$

For each $k = 1, ..., m$ let u_k denote the sum of all terms $a_j v_j$ such that $v_j \in E(\lambda_k, T)$. Thus each u_k is in $E(\lambda_k, T)$, and

$$u_1 + \cdots + u_m = 0.$$

Because eigenvectors corresponding to distinct eigenvalues are linearly independent, this implies that each u_k equals 0. Because each u_k is a sum of terms $a_j v_j$, where the v_j's were chosen to be a basis of $E(\lambda_k, T)$, this implies that all the a_j's equal 0. Thus $v_1, ..., v_n$ is linearly independent and thus is a basis of V. This proves the second bullet. ■

Corollary: Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$ has $\dim V$ distinct eigenvalues. Then T is diagonalizable.

Proof: Suppose T has distinct eigenvalues $\lambda_1, ..., \lambda_{\dim V}$. Let $v_k \in V$ be an eigenvector corresponding to the eigenvalue λ_k. Because eigenvectors corresponding to distinct eigenvalues are linearly independent, $v_1, ..., v_{\dim V}$ is linearly dependent, and thus forms a basis in V. With respect to this matrix, T has a diagonal matrix. ■

The above corollary is sufficient, but not necessary. Below is a necessary and sufficient condition for diagonalizability.

Proposition: Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then T is diagonalizable if and only if the minimal polynomial of T equals $(z - \lambda_1)\cdots(z - \lambda_m)$ for some list of distinct numbers $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$.

Proof: First suppose T is diagonalizable. Then there is a basis $v_1, \ldots, v_n$ of V consisting of eigenvectors of T. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. Then for each v_j, there exists λ_k with $(T - \lambda_k I)v_j = 0$. Thus

$$(T - \lambda_1 I)\cdots(T - \lambda_m I)v_j = 0,$$

which implies that the minimal polynomial of T equals $(z - \lambda_1)\cdots(z - \lambda_m)$.

Now suppose the minimal polynomial of T is $(z - \lambda_1)\cdots(z - \lambda_m)$ for some list of distinct $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$. Thus

$$(T - \lambda_1 I)\cdots(T - \lambda_m I) = 0.$$

We w prove T is diagonalizable by induction on m. If $m = 1$, then $T - \lambda_1 I = 0$, so T is a scalar multiple of the identity, which means it's diagonalizable.

Now suppose $m > 1$ and the result holds for all smaller integer. Note that range $(T - \lambda_m I)$ is invariant under T. Thus T restricted to range $(T - \lambda_m I)$ is an operator.

If $u \in \operatorname{range}(T - \lambda_m I)$, then $u = (T - \lambda_m I)v$ for some $v \in V$. Thus we have

$$(T - \lambda_1 I)\cdots(T - \lambda_{m-1} I)u = (T - \lambda_1 I)\cdots(T - \lambda_m I)v = 0.$$

Thus $(z - \lambda_1)\cdots(z - \lambda_{m-1})$ is a polynomial multiple of the minimal polynomial of the restricted T. By our induction hypothesis, there is a basis of $\operatorname{range}(T - \lambda_m I)$ consisting of eigenvectors of T.

Suppose $u \in \operatorname{range}(T - \lambda_m I) \cap \operatorname{null}(T - \lambda_m I)$. Then $(T - \lambda_m I)u = 0 \Rightarrow Tu = \lambda_m u$. Then we have

$$\begin{aligned} 0 &= (T - \lambda_1)\cdots(T - \lambda_{m-1})u \\ &= (\lambda_m - \lambda_1)\cdots(\lambda_m - \lambda_{m-1})u. \end{aligned}$$

Since the λ's are distinct, the coefficients are nonzero, so u must be 0. Thus $\operatorname{range}(T - \lambda_m I) + \operatorname{null}(T - \lambda_m I)$ is a direct sum whose dimension is $\dim V$ (by the fundamental theorem of linear maps and direct sum dimension). Thus $\operatorname{range}(T - \lambda_m I) \oplus \operatorname{null}(T - \lambda_m I) = V$. Every vector in the nullspace is an eigenvector of T with eigenvalue of λ_m. Adjoining the basis of the null space with the basis of the range space yields a basis of eigenvectors of T, as desired. ■

Proposition: 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.

Proof: Because T is diagonalizable, the minimal polynomial of T only has roots of multiplicity 1. Since the restriction operator's minimal polynomial is a factor of T's minimal polynomial, it must also only have roots with multiplicity 1, which implies it's diagonalizable. ■

5.4.3. Gershgorin Disk Theorem

Definition (Gershgorin disks): Suppose $T \in \mathcal{L}(V)$ and $v - 1, ..., v_n$ is a basis of V. Let A denote the matrix of T with respect to this basis. A *Gershgorin disk* of T with respect to the basis $v_1, ..., v_n$ is a set of the form

$$\left\{ z \in \mathbb{F} : |z - A_{j,j}| \leq \sum_{\substack{k=1 \\ k \neq j}}^{n} |A_{j,k}| \right\},$$

where $j \in \{1, ..., n\}$.

If $\mathbb{F} = \mathbb{C}$, then each disk a closed disk centered at $A_{j,j}$. If $\mathbb{F} = \mathbb{R}$, then each disk is a closed interval centered at $A_{j,j}$.

Theorem (Gershgorin disk theorem): Suppose $T \in \mathcal{L}(V)$ and $v_1, ..., v_n$ is a basis of V. Then each eigenvalue of T is contained in some Gershgorin disk of T with respect to the basis $v_1, ..., v_n$.

Proof: Suppose $\lambda \in \mathbb{F}$ is an eigenvalue of T. Let $w \in V$ be a corresponding eigenvector. There exist $c_1, ..., c_n \in \mathbb{F}$ such that

$$w = c_1 v_1 + \cdots + c_n v_n.$$

Let A denote the matrix of T with respect to the chosen basis. Applying T to both sides yields

$$\begin{aligned} \lambda w &= \sum_{k=1}^{n} c_k T v_k \\ &= \sum_{k=1}^{n} c_k \sum_{j=1}^{n} A_{j,k} v_j \\ &= \sum_{j=1}^{n} \left(\sum_{k=1}^{n} A_{j,k} c_k \right) v_j. \end{aligned}$$

Pick j such that

$$|c_j| = \max\{|c_1|, ..., |c_n|\}.$$

Since w is an eigenvector, we have that

$$\lambda c_j = \sum_{k=1}^{n} A_{j,k} c_k.$$

Subtract $A_{j,j} c_j$ from each side of the equation and divide by c_j to get

$$|\lambda - A_{j,j}| = \left| \sum_{\substack{k=1 \\ k \neq j}}^{n} A_{j,k} \frac{c_k}{c_j} \right| \leq \sum_{\substack{k=1 \\ k \neq j}}^{n} |A_{j,k}|.$$

Thus λ is in the j^{th} Gershgorin disk with respect to the basis $v_1, ..., v_n$. ■

5.4.4. Problems

Problem (Exercise 1): Suppoes V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$.
- Prove that if $T^4 = I$, then T is diagonalizable.
- Prove that if $T^4 = T$, then T is diagonalizable.
- Give an example of an operator $T \in \mathcal{L}(\mathbb{C}^2)$ such that $T^4 = T^2$ and T is not diagonalizable.

Solution:
- $(T - I)(T + I)(T - iI)(T + iI) = 0$, so the minimal polynomial of T only has roots with multiplicity 1, meaning it's diagonalizable
- $T(T - I)(T - \omega I)(T - \omega^2 I)$, where ω is a third root fo unity not equal to 1.
- $T(x, y) = (x - y, x - y)$, and the minimal polynomia is T^2.

Problem (Exercise 2): Suppose $T \in \mathcal{L}(V)$ has a diagonal matrix A with respect to some basis of V. Prove that if $\lambda \in \mathbb{F}$, then λ appears on the diagonal of A precisely $\dim E(\lambda, T)$ times.

Solution: Since A is diagonal, the basis of V is eigenvectors of T. Then $E(\lambda, T)$ contains $\dim E(\lambda, T)$ linearly independent eigenvectors, these eigenvectors make up part of the basis. Thus, there corresponding diagonal entries will be λ.

Problem (Exercise 3): Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that if the operator T is diagonalizable, then $V = \text{null } T \oplus \text{range } T$.

Solution: Note that null $T = E(0, T)$. Then note that $T(\sum c_i v_i) = c_i \lambda_i v_i$, so range T will consist of the rest of the eigenspaces. Since the sum of the eigenspaces is a direct sum, we have our desired result.

Problem (Exercise 4): Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the following are equivalent:
- $V = \text{null } T \oplus \text{range } T$
- $V = \text{null } T + \text{range } T$
- $\text{null } T \cap \text{range } T = \{0\}$

Solution: Follow from fundamental theorem of linear maps and subspace sums.

Problem (Exercise 7): Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that

$$E(\lambda, T) = E\left(\frac{1}{\lambda}, T^{-1}\right)$$

for every $\lambda \in \mathbb{F}$ with $\lambda \neq 0$.

Solution: $Tv = \lambda v \iff \frac{1}{\lambda} v = T^{-1} v$, so both $E(\lambda, T)$ and $E\left(\frac{1}{\lambda}, T^{-1}\right)$ include the same vectors.

Problem (Exercise 8): Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, ..., \lambda_m$ denote the distinct nonzero eigenvalues of T. Prove that

$$\dim E(\lambda_1, T) + \dots + \dim E(\lambda_m, T) \leq \dim \operatorname{range} T.$$

Solution: Choose a basis for each eigenspace. We have that $Tv_1, ..., Tv_n$ must be linearly independent (v_i's are eigenvectors), so $n \leq \dim \operatorname{range} T$.

Problem (Exercise 9): Suppose $R, T \in \mathcal{L}(\mathbb{F}^3)$ each have 2, 6, 7 as eigenvalues. Prove that there exists and invertible operator $S \in \mathcal{L}(\mathbb{F}^3)$ such that $R = S^{-1}TS$.

Solution: It's easy to see R and T are both diagonalizable, so let A represent the diagonal matrix with entries 2, 6, 7. Thus we can write

$$S_1^{-1} T S_1 = A \text{ and } S_2^{-1} R S_2 = A.$$

We can then rewrite this as

$$S_2 S_1^{-1} T S_1 S_2^{-1} = R.$$

Note that $\left(S_2 S_1^{-1}\right)^{-1} = S_2 S_2^{-1}$, so we let $S_1 S_2^{-1} = S$ and we're done.

Problem (Exercise 18): Suppose that $T \in \mathcal{L}(V)$ is diagonalizable and U is a subspace of V that is invariant under T. Prove that the quotient operator T/U is a diagonalizable operator on V/U.

Solution: Since the minimal polynomial of T/U is a divisor of the minimal polynomial of T, the result follows.

Problem (Exercise 20): Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if T' is diagonalizable.

Solution: Follows easily since both operators have the same minimal polynomial.

Problem (Exercise 21): Define $T \in \mathcal{L}(\mathbb{R}^2)$ by $T(x, y) = (y, x + y)$. Use this map to find an explicity formula for the Fibonacci sequence.

Solution: It's easy to see that $T^n(0, 1) = (F_n, F_{n+1})$. Next we find that the eigenvalues of T are $\lambda_1 = \frac{1+\sqrt{5}}{2}$ and $\lambda_2 = \frac{1-\sqrt{5}}{2}$ with associated eigenvectors

$$v_1 = \begin{pmatrix} 1 \\ \frac{1+\sqrt{5}}{2} \end{pmatrix} \text{ and } v_2 = \begin{pmatrix} 1 \\ \frac{1-\sqrt{5}}{2} \end{pmatrix}.$$

We have

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{5}}\left(\begin{pmatrix} 1 \\ \frac{1+\sqrt{5}}{2} \end{pmatrix} - \begin{pmatrix} 1 \\ \frac{1-\sqrt{5}}{2} \end{pmatrix}\right).$$

Applying T^n on both sides yields

$$\begin{pmatrix} F_n & F_{n+1} \end{pmatrix} = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n \begin{pmatrix} 1 \\ \frac{1+\sqrt{5}}{2} \end{pmatrix} - \left(\frac{1-\sqrt{5}}{2}\right)^n \begin{pmatrix} 1 \\ \frac{1-\sqrt{5}}{2} \end{pmatrix}\right).$$

Thus we have

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

Problem (Exercise 22): Suppose $T \in \mathcal{L}(V)$ and A is an n by n matrix that is the matrix of T with respect to some basis V. Prove that if

$$|A_{j,j}| > \sum_{\substack{k=1 \\ k \neq j}}^{n} |A_{j,k}|$$

for each $j \in \{1, ..., n\}$, then T is invertible.

Solution: Note that T is invertible if and only if 0 is not an eigenvalue of T. Suppose that 0 is an eigenvalue of T. Then by the Gershgorin disk theorem, 0 should be contained in some disk. However, by the equation in the problem, we can see that 0 is not contained in any disk. Thus 0 cannot be an eigenvalue of T, implying T is invertible.

5.5. Commuting Operators

Definition (commute): Two operators S and T on the same vector space *commute* if $ST = TS$. Two square matrices A and B of the same size *commute* if $AB = BA$.

Proposition: Suppose $S, T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V. Then S and T commute if and only if $\mathcal{M}(S, (v_1, \dots, v_n))$ and $\mathcal{M}(T, (v_1, \dots, v_n))$ commute.

Proof:

$$\begin{aligned} S \text{ and } T \text{ commute} &\iff ST = TS \\ &\iff \mathcal{M}(ST) = \mathcal{M}(TS) \\ &\iff \mathcal{M}(S)\mathcal{M}(T) = \mathcal{M}(T)\mathcal{M}(S) \\ &\iff \mathcal{M}(S) \text{ and } \mathcal{M}(T) \text{ commute.} \end{aligned}$$

■

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

Proof: Suppose $v \in E(\lambda, S)$. Then

$$S(Tv) = (ST)v = (TS)v = T(Sv) = T(\lambda v) = \lambda Tv.$$

Thus $Tv \in E(\lambda, S)$, as desired. ■

Proposition: 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.

Proof: If $S, T \in \mathcal{L}(V)$ has diagonal matrices with respect to the same basis, then the result easily follows, since when multiplying diagonal matrices you simply multiply corresponding diagonal entries.

Now suppose $S, T \in \mathcal{L}(V)$ are diagonalizable operators that commute. Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of S. Because S is diagonalizable, we have

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

For each $k = 1, \dots, m$, the subspace $E(\lambda_k, S)$ is invariant under T. Because T is diagonalizable, $T|_{E(\lambda_k, S)}$ is diagonalizable for each k. Thus for each k there is a basis of $E(\lambda_k, S)$ consisting of eigenvectors of T. Note each of these is necessarily an eigenvector of S as well, since they are in an eigenspace of S. Putting these bases together gives a basis of V, with each vector being an eigenvector of both S and T. Thus S and T both have diagonal matrices with respect to this basis, as desired. ■

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

Proof: Suppose V is a finite-dimensional nonzero complex vector space and $S, T \in \mathcal{L}(V)$ commute. Let λ be an eigenvalue of S. Thus $E(\lambda, S) \neq \{0\}$. $E(\lambda, S)$ is also invariant under T. Thus $T|_{E(\lambda,S)}$ has an eigenvector, which is an eigenvector for both S and T. ■

Proposition: Suppose V is a finite-dimensional 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.

Proof: Let $n = \dim V$. We use induction on n. The result clearly holds for $n = 1$, so suppose $n > 1$ and the result holds for $n - 1$.

Let v_1 be a common eigenvector of S and T. Thus we have $Sv_1, Tv_i \in \operatorname{span}(v_1)$. Let W be a subspace of V such that

$$V = \operatorname{span}(v_1) \oplus W.$$

Define $P : V \to W$ by

$$P(av_1 + w) = w$$

for each $a \in \mathbb{C}$ and each $w \in W$. Define $\hat{S}, \hat{T} \in \mathcal{L}(W)$ by

$$\hat{S}w = P(Sw) \text{ and } \hat{T}w = P(Tw)$$

for each $w \in W$. Now we show $\hat{T}$ and $\hat{S}$ commute. Suppose $w \in W$. Then there exists $a \in \mathbb{C}$ such that

$$\left(\hat{S}\hat{T}\right)w = \hat{S}(P(Tw)) = \hat{S}(Tw - av_1) = P(S(Tw - av_1)) = P((ST)w),$$

where the last equality holds because v_1 is an eigenvector of S and $Pv_1 = 0$. Similarly

$$\left(\hat{T}\hat{S}\right)w = P((TS)w).$$

Because S and T commute, we have $\hat{S}$ and $\hat{T}$ commute.

By out induction hypothesis, there exists a basis $v_2, ..., v_n$ of W such that $\hat{S}$ and $\hat{T}$ both have upper triangulat matrices with respect to this basis. The list $v_1, ..., v_n$ is a basis of V.

If $k \in \{2, ..., n\}$, then there exist $a_k, b_k \in \mathbb{C}$ such that

$$Sv_k = a_k v_1 + \hat{S}v_k \text{ and } Tv_k = b_k v_1 + \hat{T}v_k.$$

Because $\hat{S}$ and $\hat{T}$ have upper triangulat matrices with respect to $v_2, ..., v_n$, we know that $\hat{S}v_k, \hat{T}v_k \in \operatorname{span}(v_2, ..., v_n)$. Thus the equations above imply that

$$\hat{S}v_k, \hat{T}v_k \in \operatorname{span}(v_1, ..., v_n).$$

Thus S and T has upper triangular matrices with respect to $v_1, ..., v_n$, as desired. ■

Proposition (eigenvalues of sum and product of commuting operators): Suppose V is a finite dimensional complex vector space and S, T are commuting operators on V. Then
- every eigenvalue of $S + T$ is an eigenvalue of S plus an eigenvalue of T
- every eigenvalue of ST is an eigenvalue of S times an eigenvalue of T

Proof: There is a basis of V with respect to which both S and T have upper triangular matrices. With respect to that basis,

$$\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T) \text{ and } \mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T).$$

Since for upper triangular matrices are closed under addition and multiplication, we can indeed verify that the eigenvalues satisfy the relationship stated in the proposition. ■

5.5.1. Problems

6. Inner Product Spaces

6.1. Inner Products and Norms

6.1.1. Inner Products

Definition (inner product): And *inner product* on V is a function that takes each ordered pair (u, v) of elements of V to a number $\langle u, v\rangle \in \mathbb{F}$ and has the following properties:
- $\langle v, v\rangle \geq 0$ for all $v \in V$.
- $\langle v, v\rangle = 0$ if and only if $v = 0$.
- $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$.
- $\langle \lambda u, w\rangle = \lambda\langle u, v\rangle$ for all $\lambda \in \mathbb{F}$ and all $u, w \in V$.
- $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$.

Definition (inner product space): And *inner product space* is a vector space V along with an inner product on V.

Proposition:
- For each fixed $v \in V$, the function that takes $u \in V$ to $\langle u, v\rangle$ is a linear map from V to $\mathbb{F}$.
- $\langle 0, v\rangle = 0$ for every $v \in V$.
- $\langle v, 0\rangle = 0$ for every $v \in V$.
- $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$ for all $u, v, w \in V$.
- $\langle u, \lambda v\rangle = \overline{\lambda}\langle u, v\rangle$ for all $\lambda \in \mathbb{F}$ and all $u, v \in V$.

Proof: Easy to verify. ■

6.1.2. Norms

Definition (norm): For $v \in V$, the *norm* of v, denoted by $\|v\|$, is defined by

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

Proposition: Suppose $v \in V$.
- $\|v\| = 0$ if and only if $v = 0$.
- $\|\lambda v\| = |\lambda|\|v\|$ for all $\lambda \in \mathbb{F}$.

Proof: Also easy to verify. ■

Definition (orthogonal): Two vectors $u, v \in V$ are called *orthogonal* if $\langle u, v\rangle = 0$.

Proposition:
- 0 is orthogonal to every vector in V.
- 0 is the only vector in V that is orthogonal to itself.

Proof:
- $\langle 0, v\rangle = 0$
- $\langle v, v\rangle = 0 \Rightarrow v = 0$

■

Theorem (Pythagorean Theorem): Suppose $u, v \in V$. If u and v are orthogonal, then

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

Solution:

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

Proposition (orthogonal decomposition): Suppose $u, v \in V$ with $v \neq 0$. Set $c = \frac{\langle u,v\rangle}{\|v\|^2}$ and $w = u - \frac{\langle u,v\rangle}{\|v\|^2}v$. Then

$$u = cv + w \text{ and } \langle w, v\rangle = 0.$$

Proof: Easy to verify. ■

Theorem (Cauchy-Schwarz inequality): 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.

Proof: If $v = 0$, then both sides are 0, so we can assume $v \neq 0$. Consider the orthogonal decomposition

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

By the Pythagorean theorem,

$$\begin{aligned} \|u\|^2 &= \left\| \frac{\langle u, v\rangle}{\|v\|^2} v \right\|^2 + \|w\|^2 \\ &= \frac{|\langle u, v\rangle|^2}{\|v\|^2} + \|w\|^2 \\ &\geq \frac{|\langle u, v\rangle|^2}{\|v\|^2}. \end{aligned}$$

Multiplying and taking squares roots yields the obvious. We have equality when $\|w\| = 0$, which implies that u is a scalar multiple of v. ■

Theorem (triangle inequality): Suppose $u, v \in V$. Then

$$\|u + v\| \leq \|u\| + \|v\|.$$

Equality holds if and only if one is a nonnegative real multiple of the other.

Proof:

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

where the second to last line follows from Cauchy-Schwarz. Taking square roots yields the desired inequality. Equality occurs when there's equality in Cauchy-Schwarz, and the scalar must be a nonnegative real since we need $\langle u, v\rangle = \|u\|\|v\|$. ■

Theorem (parallelogram equality): Suppose $u, v \in V$. Then

$$\|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right).$$

Proof:

$$\begin{aligned}\|u+v\|^2 + \|u-v\|^2 &= \langle u+v, u+v\rangle + \langle u-v, u-v\rangle \\ &= \|u\|^2 + \|v\|^2 + \langle u,v\rangle + \langle v,u\rangle + \|u\|^2 + \|v\|^2 - \langle u,v\rangle - \langle v,u\rangle \\ &= 2\left(\|u\|^2 + \|v\|^2\right).\end{aligned}$$

■

6.1.3. Problems

Problem (Exercise 1): Prove or give a counterexample: If $v_1, ..., v_m \in V$, then

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\langle v_j, v_k\rangle \geq 0.$$

Solution:

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\langle v_j, v_k\rangle = \sum_{j=1}^{m}\left\langle v_j, \sum_{k=1}^{m} v_k\right\rangle = \left\langle \sum_{j=1}^{m} v_j, \sum_{k=1}^{m} v_k\right\rangle \geq 0.$$

Problem (Exercise 4): Suppose $T \in \mathcal{L}(V)$ is such that $\|Tv\| \leq \|v\|$ for every $v \in V$. Prove that $T - \sqrt{2}I$ is injective.

Solution: If $\sqrt{2}$ is not an eigenvalue of T, then we're done. Otherwise suppose it is. Pick an eigenvector v corresponding to $\sqrt{2}$. Then we have $\|Tv\| = \left\|\sqrt{2}v\right\| = \sqrt{2}\|v\|$, which is not less than $\|v\|$ for nonzero v, which is impossible. Thus $\sqrt{2}$ is not an eigenvalue of T.

Problem (Exercise 5): Suppose V is a real inner product space.
- Show that $\langle u+v, u-v\rangle = \|u\|^2 - \|v\|^2$ for every $u, v \in V$.
- Show that if $u, v \in V$ have the same norm, then $u + v$ is orthogonal to $u - v$.
- Show that the diagonals of a rhombus are perpendicular to each other.

Solution:

$$\langle u+v, u-v\rangle = \langle u,u\rangle - \langle u,v\rangle + \langle v,u\rangle - \langle v,v\rangle = \|u\|^2 - \|v\|^2.$$

The second bullet follows from this. Now consider a rhombus with sides u, v. Then the diagonals are $u + v$ and $u - v$, which are orthogonal, so they're perpendicular.

Problem (Exercise 8): Suppose $a, b, c, x, y \in \mathbb{R}$ and $a^2 + b^2 + c^2 + x^2 + y^2 \leq 1$. Prove that $a + b + c + 4x + 9y \leq 10$.

Solution: By Cauchy-Schwarz, we have

$$100 \geq (a^2 + b^2 + c^2 + x^2 + y^2)(1^2 + 1^2 + 1^2 + 4^2 + 9^2) \geq (a + b + c + 4x + 9y)^2.$$

Taking square roots yields the desired result.

Problem (Exercise 9): Suppose $u, v \in V$ and $\|u\| = \|v\| = 1$ and $\langle u, v\rangle = 1$. Prove that $u = v$.

Solution: Note that

$$\langle u - v, u - v\rangle = \langle u, u\rangle - \langle u, v\rangle - \langle v, u\rangle + \langle v, v\rangle = 1 - 1 - 1 + 1 = 0,$$

so $u - v = 0$.

Problem (Exercise 10): Suppose $u, v \in V$ and $\|u\|, \|v\| \leq 1$. Prove that

$$\sqrt{1 - \|u\|^2}\sqrt{1 - \|v\|^2} \leq 1 - |\langle u, v\rangle|.$$

Solution: By AM-GM,

$$\frac{\|u\|^2 + \|v\|^2}{2} \geq \|u\|\|v\|.$$

This can be rewritten as

$$\left(1 - \|u\|^2\right)\left(1 - \|v\|^2\right) \leq (1 - \|u\|\|v\|)^2.$$

Since the left side is positive, we can take square roots to get

$$\sqrt{1 - \|u\|^2}\sqrt{1 - \|v\|^2} \leq 1 - \|u\|\|v\| \leq 1 - |\langle u, v\rangle|,$$

where the last inequality comes from Cauchy-Schwarz.

Problem (Exercise 11): Find vectors $u, v \in \mathbb{R}^2$ such that u is a scalar multiple of $(1, 3)$, v is orthogonal to $(1, 3)$, and $(1, 2) = u + v$.

Solution: $u = \left(\frac{7}{10}, \frac{21}{10}\right), v = \left(\frac{3}{10}, -\frac{1}{10}\right)$

Problem (Exercise 14): Suppose $v \in V$ and $v \neq 0$. Prove that if $u \in V$ $\|u\| = 1$, then

$$\left\| v - \frac{v}{\|v\|} \right\| \leq \|v - u\|,$$

with equality if anf only if $u = v/\|v\|$.

Solution: We show the solution when $\|v\| > 1$. The other cases follow similarly. Note that

$$\left\| v - \frac{v}{\|v\|} \right\| = \left| 1 - \frac{1}{\|v\|} \right| \|v\| = \|v\| - 1 = \|v\| - \|u\| \leq \|v - u\|,$$

where the last line comes from the triangle inequality. Equality only occurs when

$$cu = v - u \Rightarrow c = -1 \pm \|v\|,$$

and $c = 1 + \|v\|$ is the only case where equality holds.

Problem (Exercise 15): Suppose u, v are nonzero vectors in $\mathbb{R}^2$. Prove that

$$\langle u, v \rangle = \|u\| \|v\| \cos \theta,$$

where θ is the angle between u and v.

Solution: By law of cosines, we have

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\| \cos \theta.$$

We can rewrite this as

$$\|u\|^2 + \|v\|^2 - 2\langle u, v \rangle = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\| \cos \theta.$$

Simiplfying yields

$$\langle u, v \rangle = \|u\| \|v\| \cos \theta.$$

Problem (Exercise 16): The angle between two vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ can be defined geometrically. However, geometry is not as clear in $\mathbb{R}^n$ for $n > 3$. Thus the angle between two nonzero bectors $x, y \in \mathbb{R}^n$ os defined to be

$$\arccos \frac{\langle x, y \rangle}{\|x\| \|y\|}.$$

Explain why the Cauchy-Schwarz inequality is needed to show this definition makes sense.

Solution: Note that we need $\left|\frac{\langle x,y\rangle}{\|x\|\|y\|}\right| \leq 1$ for all $x, y \in \mathbb{R}^n$ in order for the angle to be defined for every pair of vectors. This is equivalent to $|\langle x, y\rangle| \leq \|x\|\|y\|$, which is just CS.

Problem (Exercise 18): Suppose $f : [1, \infty) \to [0, \infty)$ is continuous. Show that

$$\left(\int_1^\infty f(x)\,\mathrm{d}x\right)^2 \leq \int_1^\infty x^2 f(x)^2\,\mathrm{d}x$$

and determine the equality case.

Solution: Let V be the vector space of continuous functions on $\mathbb{R}$ with inner product

$$\langle f, g\rangle = \int_1^\infty f(x)g(x)\,\mathrm{d}x.$$

Let $f = xf(x)$ and let $g = \frac{1}{x}$. Then by CS we have

$$\left(\int_1^\infty f(x)\,\mathrm{d}x\right)^2 \leq \left(\int_1^\infty x^2 f(x)^2\,\mathrm{d}x\right)\left(\int_1^\infty \frac{1}{x^2}\,\mathrm{d}x\right) = \int_1^\infty x^2 f(x)^2\,\mathrm{d}x.$$

Equality occurs when $f(x) = cg(x) \Rightarrow f(x) = \frac{c}{x^2}$.

Problem (Exercise 20): Prove that if $u, v \in V$, then $|\|u\| - \|v\|| \leq \|u - v\|$.

Solution: Suppose $\|u\| \geq \|v\|$. Then we have

$$\|u - v\| + \|v\| \geq \|u\| \Rightarrow \|u - v\| \geq \|u\| - \|v\|.$$

Now suppose $\|u\| \leq \|v\|$. Then we have

$$\|v - u\| + \|u\| \geq \|v\| \Rightarrow \|u - v\| = \|v - u\| \geq \|v\| - \|u\|.$$

Thus we get the desired result.

Problem (Exercise 21): Suppose $u, v \in V$ are such that

$$\|u\| = 3, \|u + v\| = 4, \|u - v\| = 6.$$

Determine $\|v\|$.

Solution: By the parallelogram equality,

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

Problem (Exercise 22): Show that if $u, v \in V$, then

$$\|u+v\|\|u-v\| \leq \|u\|^2 + \|v\|^2.$$

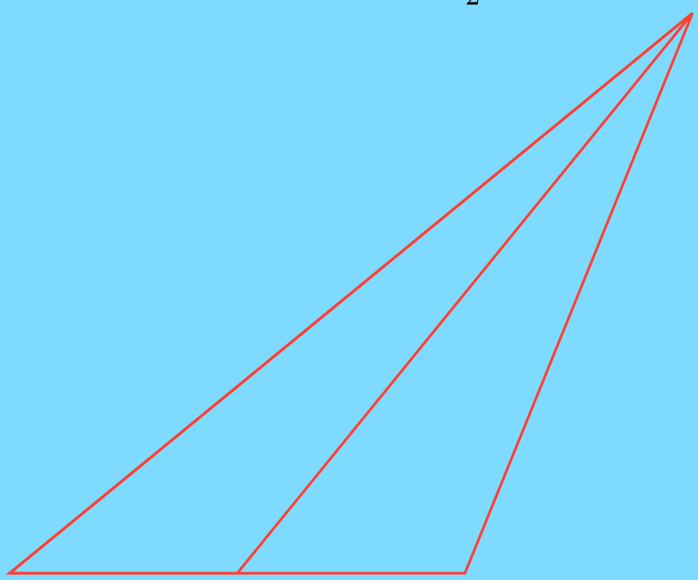
Solution:

$$\|u+v\|\|u-v\| \leq \frac{\|u+v\|^2 + \|u-v\|^2}{2} = \|u\|^2 + \|v\|^2,$$

where the last line comes from the parallelogram equality.

Problem (Exercise 34): Use inner products to prove Apollonius' identity: In a triangle with sides of length a, b, c, let d be length of the median to side c. Then

$$a^2 + b^2 = \frac{1}{2}c^2 + 2d^2.$$



Solution: Let a and c emanate from the origin. Then we have $\|b\| = \|a-c\|$ and $\|d\| = \left\|a - \frac{c}{2}\right\|$. Thus we have

$$\|b\|^2 = \|a\|^2 + \|c\|^2 - \langle a, c\rangle - \langle c, a\rangle \text{ and } \|d\|^2 = \|a\|^2 + \frac{1}{4}\|c\|^2 - \frac{1}{2}\langle a, c\rangle - \frac{1}{2}\langle c, a\rangle.$$

Thus we have

$$\|b\|^2 - \|a\|^2 - \|c\|^2 = 2\|d\|^2 - 2\|a\|^2 - \frac{1}{2}\|c\|^2.$$

Rearraning yields the desired result.

6.2. Orthonormal Bases

6.2.1. Orthonormal Lists and the Gram-Schmidt Procedure

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

Proposition: Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in V. Then

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

for all $a_i \in \mathbb{F}$.

Proof: Follows from repeated use of Pythagorean theorem. ■

Proposition: Every orthonormal list of vectors is linearly independent.

Solution: Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in V and we have

$$\sum a_i e_i = 0.$$

Then we have $\|\sum a_i e_i\| = \sum |a_i|^2 = 0$, which implies $a_i = 0$.

Theorem (Bessel's inequality): Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in V. If $v \in V$ then

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

Proof: Suppose $v \in V$. Then

$$v = \underbrace{\langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m}_{u} + \underbrace{v - \langle v, e_1\rangle e_1 - \cdots - \langle v, e_m\rangle e_m}_{w}.$$

For each k we have $\langle w, e_k\rangle = \langle v, e_k\rangle - \langle v, e_k\rangle\langle e_k, e_k\rangle = 0$. Thus implies $\langle w, v\rangle = 0$, so by the Pythagorean theorem, we have $\|v\|^2 = \|u\|^2 + \|w\|^2 \geq \|u\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2$. Note that equality occurs when $w = 0 \Longleftrightarrow v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m \Rightarrow v \in \operatorname{span}(e_1, \ldots, e_m)$.
■

Definition (orthonormal basis): An *orthonormal basis* of V is an orthonormal list of vectors in V that is also a basis of V.

Proposition: Suppose V is finite-dimensional. If an orthonormal list has length $\dim V$, then it is a basis of V.

Proof: Every orthonormal list is linearly independent, and since this list has length $\dim V$, it must be a basis. ■

Proposition: Suppose $e_1, …, e_n$ is an orthonormal basis of V, and $u, v \in V$. Then

- $v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n$
- $\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots |\langle v, e_n\rangle|^2$
- $\langle u, v\rangle = \langle u, e_1\rangle \overline{\langle v, e_1\rangle} + \cdots + \langle u, e_n\rangle \overline{\langle v, e_n\rangle}$

Proof: We can write

$$v = \sum a_i e_i.$$

Taking the inner product of both sides with e_k for any k yields

$$\langle v, e_k\rangle = a_k,$$

so the first bullet holds. The second bullet follows from an earlier proposition. The third bullet follows from taking the inner product of the first bullet with u and using conjuagte symmetry. ■

Theorem (Gram-Schmidt procedure): Suppose $v_1, …, v_m$ is a linearly independent list of vectors in V. Let $f_1 = v_1$. For $k = 2, …, m$, define f_k inductively by

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

For each $K = 1, …, m$ let $e_k = \frac{f_k}{\|f_k\|}$. Then $e_1, …, e_m$ is an orthonormal list of vectors in V such that

$$\operatorname{span}(v_1, …, v_k) = \operatorname{span}(e_1, …, e_k)$$

for each k.

Proof: We show this by induction. Note that $\|e_1\| = 1$, and we also have $\operatorname{span}(e_1) = \operatorname{span}\left(\frac{v_1}{\|v_1\|}\right) = \operatorname{span}(v_1)$.

Now suppose $e_1, …, e_{k-1}$ is an orthonormal list generated by the procedure, so the spans of v_i's and e_i's are the same.

Because $v_1, \ldots, v_m$ are linearly independent, we have $v_k \notin \operatorname{span}(v_1, \ldots, v_{k-1}) = \operatorname{span}(e_1, \ldots, e_{k-1}) = \operatorname{span}(f_1, \ldots, f_{k-1})$, which implies $f_k \neq 0$. Thus we can safely conclude $\|e_k\| = 1$.

For $j < k$, we have

$$\begin{aligned}\langle e_k, e_j \rangle &= \frac{1}{\|f_k\|\|f_j\|}\langle f_k, f_j \rangle \\ &= \frac{1}{\|f_k\|\|f_j\|}\left\langle v_k - \frac{\langle v_k, f_1 \rangle}{\|f_1\|^2} f_1 - \cdots - \frac{\langle v_k, f_{k-1} \rangle}{\|f_{k-1}\|^2} f_{k-1}, f_j \right\rangle \\ &= \frac{1}{\|f_k\|\|f_j\|}\left(\langle v_k, f_j \rangle - \langle v_k, f_j \rangle\right) \\ &= 0.\end{aligned}$$

Thus $e_1, \ldots, e_k$ is orthonormal. From the definition of e_k, we have $v_k \in \operatorname{span}(e_1, \ldots, e_k)$. Thus we have

$$\operatorname{span}(v_1, \ldots, v_k) \subseteq \operatorname{span}(e_1, \ldots, e_k).$$

Both lists above are linearly independent. Thus both subspaces have dimension k, which implies they are equal. ■

Corollary: Every finite dimensional inner product space has an orthonormal basis.

Proof: Suppose V is finite dimensional and choose a basis of V. Then apply the Gram-Schmidt procedure on it to obtain an orthonormal list of length of $\dim V$, which means it must be an orthonormal basis. ■

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

Proof: Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in V. Then $e_1, \ldots, e_m$ is linearly independent. Thus this list can be extended to a basis $e_1, \ldots, e_m, v_1, \ldots, v_n$ of V. Now apply Gram-Schmidt to obtain $e_1, \ldots, e_m, f_1, \ldots, f_n$, where the initial m vectors are unchanged because they're already orthonormal. Thus this list is an orthonormal basis. ■

Proposition: Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then T has an upper triangular matrix with respect to some orthonormal basis of V if and only if the minimal polynomial of T equals $(z - \lambda_1)\cdots(z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$.

Proof: Since T has an upper triangular matrix for some basis $v_1, \ldots, v_n$, we have $\operatorname{span}(v_1, \ldots, v_k)$ invariant under T for each k. Using the Gram-Schmidt procedure on $v_1, \ldots, v_n$ produces an orthonormal basis, where each sublist starting at e_1 will have the same span as the sublist of the same length starting at v_1. Thus T has an upper triangular matrix with respect to $e_1, \ldots, e_n$, and the result follows. ■

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

Proof: Follows from the last result. ∎

6.2.2. Linear Functionals on Inner Product Spaces

Theorem (Riesz representation theorem): Suppose V is finite-dimensional and φ is a linear functional on V. Then there is a unique vector $v \in V$ such that

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

for every $u \in V$.

Proof: First we show existence. Let $e_1, ..., e_n$ be an orthonormal basis of V. Then

$$\begin{aligned}\varphi(u) &= \varphi(\langle u, e_1 \rangle e_1 + \cdots + \langle u, e_n \rangle e_n) \\ &= \langle u, e_1 \rangle \varphi(e_1) + \cdots + \langle u, e_n \rangle \varphi(e_n) \\ &= \left\langle u, \overline{\varphi(e_1)} e_1 + \cdots + \overline{\varphi(e_n)} e_n \right\rangle.\end{aligned}$$

Thus we can let $v = \overline{\varphi(e_1)} e_1 + \cdots + \overline{\varphi(e_n)} e_n$.

Now we prove uniqueness. Suppose we have

$$\varphi(u) = \langle u, v_1 \rangle = \langle u, v_2 \rangle$$

for distinct v_1, v_2. Thus we have

$$0 = \langle u, v_1 \rangle - \langle u, v_2 \rangle = \langle u, v_1 - v_2 \rangle.$$

Taking $u = v_1 - v_2$ implies $v_1 = v_2$. ∎

6.2.3. Problems

Problem (Exercise 1): Suppose $e_1, ..., e_m$ is a list of vectors in V such that

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

Prove that $e_1, ..., e_m$ is orthonormal.

Solution: Setting $a_j = 0$ and everything else to 0 shows that each e_j has norm 1. Setting $a_j, a_k = 1$ and everything else to 0 yields

$$\left\|e_j + e_k\right\|^2 = 2 + \left\langle e_j, e_k \right\rangle + \left\langle e_k, e_j \right\rangle = 2.$$

If V is a real product space, then we're done. If V is a complex vector space, then let $\langle e_j, e_k \rangle = x$. Then we have $x + \overline{x} = 0$.

Now set $a_j = 1, a_k = i$ and everything else 0 to obtain

$$\left\| e_j + ie_k \right\|^2 = 2 - ix + i\overline{x} = 2.$$

This yields $x - \overline{x} = 0$, which thus implies $x = 0$.

Problem (Exercise 3): Suppose $e_1, ..., e_m$ is an orthonormal list in V and $v \in V$. Prove that

$$\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_m \rangle|^2 \Longleftrightarrow v \in \operatorname{span}(e_1, ..., e_m).$$

Solution: Bessel's inequality equality case.

Problem (Exercise 5): Suppose $f : [-\pi, \pi] \to \mathbb{R}$ is continuous. For each nonngeative integer k, define

$$a_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x) \cos(kx) \, \mathrm{d}x \text{ and } b_k = \frac{1}{\sqrt{x}} \int_{-\pi}^{\pi} f(x) \sin(kx) \, \mathrm{d}x.$$

Prove that

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) \leq \int_{-\pi}^{\pi} f^2.$$

Solution: Follows from Bessel's inequality using the trig orthonormal basis.

Problem (Exercise 7): Suppose $T \in \mathcal{L}(\mathbb{R}^3)$ has an upper triangular matrix with respect to the basis $(1, 0, 0), (1, 1, 1), (1, 1, 2)$. Find an orthonormal basis of $\mathbb{R}^3$ with respect to which T has an upper triangular matrix.

Solution: Applying the Gram-Schmidt procedure yields

$$(1, 0, 0), \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right).$$

Problem (Exercise 8): Find an orthonormal basis for $\mathcal{P}_2(\mathbb{R})$ with inner product $\langle p, q \rangle = \int_0^1 pq$.

Solution:

$$1, 2\sqrt{3}x - \sqrt{3}, 6\sqrt{5}x^2 - 6\sqrt{5}x + \sqrt{5}.$$

Problem (Exercise 10): Suppose $v_1, ..., v_m$ is a linearly independent list in V. Explain why the orthonormal list produced by the formulas of the Gram-Schmidt procedure is the only orthonormal list $e_1, ..., e_m$ in V such that $\langle v_k, e_k \rangle > 0$ and $\operatorname{span}(v_1, ..., v_k) = \operatorname{span}(e_1, ..., e_k)$ for each $k = 1, ..., m$.

Solution: We can show this via induction. Clearly it holds for $k = 1$. Suppose it's true up to $k - 1$. We needed

$$f_k = a_k v_k + \sum_{i=1}^{k-1} a_i' v_i.$$

We can replace a_i for $i < k$ with their representation from the orthonormal list generated so far. Thus we need

$$f_k = a_k v_k + \sum_{i=1}^{k-1} a_i v_i.$$

Since we need $\langle f_k, e_j \rangle = 0$, this forces $a_j = -a_k \langle v_k, e_j \rangle$. We also need $\langle v_k, f_k \rangle$ to be real and positive. Thus we need

$$a_k \|v_k\|^2 - \sum_{i=1}^{k-1} \langle v_k, a_k \langle v_k, e_i \rangle e_i \rangle = a_k \|v_k\|^2 - \overline{a_k} \sum_{i=1}^{k-1} |\langle v_k, e_i \rangle|^2 > 0.$$

Note that by Bessel's inequality, $\|v_k\|^2 > \sum_{i=1}^{k-1} |\langle v_k, e_i \rangle|^2$, since $v_k \notin \operatorname{span}(v_1, ..., v_{k-1})$. Thus the imaginary part of a_k must be 0. Because of Bessel's inequality, we also need the real part of a_k to be positive. Thus f_k is unique up to some positive real scalar, which means e_k is uniquely determined.

Problem (Exercise 11): Find a polynomial $q \in \mathcal{P}_2(\mathbb{R})$ such that $p\left(\frac{1}{2}\right) = \int_0^1 pq$ for every $p \in \mathcal{P}(\mathbb{R})$.

Solution: By Riesz representation theorem, $q = -15x^2 + 15x - \frac{3}{2}$.

Problem (Exercise 12): Find a polynomial $q \in \mathcal{P}_2(\mathbb{R})$ such that

$$\int_0^1 p(x) \cos(\pi x) \, dx = \int_0^1 p(x) q(x) \, dx.$$

Solution: By Riesz representation theorem, $q = \frac{12}{\pi^2} - \frac{24x}{\pi^2}$

6.3. Orthogonal Complements and Minimization Problems

6.3.1. Orthogonal Complements

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

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

Proposition:
- If U is subset of V, then $U^{\perp}$ is a subspace of V.
- $\{0\}^{\perp} = V$.
- $V^{\perp} = \{0\}$.
- If U is a subset of V, then $U \cap U^{\perp} \subseteq \{0\}$.
- If G and H are subsets of V and $G \subseteq H$, then $H^{\perp} \subseteq G^{\perp}$.

Proof:
- We have $\langle u, 0\rangle = 0$ for every $u \in U$, so $0 \in U^{\perp}$. We also have $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle = 0 + 0 = 0$, which implies $U^{\perp}$ is clsoed under addition. Similarly, it's closed under scalar multiplication.
- For $v \in V$, we have $\langle 0, v\rangle = 0$, which means $v \in \{0\}^{\perp}$. Thus, $\{0\}^{\perp} = V$.
- Essentially same reasoning as last statement.
- If $u \in U \cap U^{\perp}$, then $\langle u, u\rangle = 0$, which implies $u = 0$. Thus $U \cap U^{\perp} \subseteq \{0\}$.
- Suppoes $v \in H^{\perp}$. Then $\langle u, v\rangle = 0$ for every $u \in H$, which implies $\langle u, v\rangle = 0$ for every $u \in G$, which implies $v \in G^{\perp}$. Thus $H^{\perp} \subseteq G^{\perp}$.

■

Proposition: Suppose U is a finite-dimensional subspace of V. Then

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

Proof: First we show $V = U + U^{\perp}$. Suppose $v \in V$, and let $e_1, ..., e_m$ be an orthonormal basis of U. We have

$$v = \underbrace{\langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m}_{u} + \underbrace{v - \langle v, e_1\rangle e_1 - \cdots - \langle v, e_m\rangle e_m}_{w}.$$

Clearly $u \in U$, since $e_1, ..., e_m$ is basis of U. We also have

$$\langle w, e_k\rangle = \langle v, e_k\rangle - \langle v, e_k\rangle = 0$$

for each k, so $w \in U^{\perp}$. Thus we can write each vector in v as a vector from U and $U^{\perp}$, so we have $V = U + U^{\perp}$.

We know that $U \cap U^\perp = \{0\}$ (since U is a subspace), so we have that $U + U^\perp$ is a direct sum, as desired. ■

Corollary: Suppose V is finite-dimensional and U is a subspace of V. Then

$$\dim U^\perp = \dim V - \dim U.$$

Proof: Follows from dimension of direct sums. ■

Proposition: Suppose U is a finite-dimensional subspace of V. Then

$$U = \left(U^\perp\right)^\perp.$$

Proof: Note that for $u \in U$, $\langle u, w \rangle = 0$ for every $w \in U^\perp$. Since u is orthogonal to every vector in $U^\perp$, by definition $u \in \left(U^\perp\right)^\perp$, which shows $U \subseteq \left(U^\perp\right)^\perp$.

Now suppose $v \in \left(U^\perp\right)^\perp$. By the previous proposition, we can write $v = u + w$, where $u \in U$ and $w \in U^\perp$. We have $v - u = w \in U^\perp$. Because $v, u \in \left(U^\perp\right)^\perp$ (u is in it because of the previous paragraph), we have $v - u \in \left(U^\perp\right)^\perp$. Thus $v - u \in U^\perp \cap \left(U^\perp\right)^\perp$, which implies $v - u = 0$, which further implies $v \in U$. Thus $\left(U^\perp\right)^\perp \subseteq U$, as desired. ■

Proposition: Suppose U is a finite-dimensional subspace of V. then

$$U^\perp = \{0\} \Longleftrightarrow U = V.$$

Proof:

$$U^\perp = \{0\} \Longleftrightarrow \left(U^\perp\right)^\perp = \{0\}^\perp \Longleftrightarrow U = V.$$

■

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

Proposition (properties of orthogonal projection): Suppose U is a finite dimensional subspace of V.
- $P_U \in \mathcal{L}(V)$
- $P_U u = u$ for every $u \in U$
- $P_U w = 0$ for every $w \in U^\perp$
- range $P_U = U$
- null $P_U = U^\perp$
- $v - P_U v \in U^\perp$ for every $v \in V$
- $P_U^2 = P_U$
- $\|P_U v\| \leq \|v\|$ for every $v \in V$
- If $e_1, \ldots, e_m$ is an orthonormal basis of U and $v \in V$, then $P_U v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_m \rangle e_m$

Proof:
- Easy to see
- $P_U u = P_U(u + 0) = u$
- $P_U w = P_U(0 + w) = 0$
- By definition range $P_U \subseteq U$, and by the second bullet $U \subseteq$ range P_U
- $U^\perp \subseteq$ null P_U follows from the third bullet. If $v \in$ null P_U, then we must have $v \in U^\perp$, which implies null $P_U \subseteq U^\perp$.
- If $v = u + w$, then $v - P_U v = v - u = w \in U^\perp$.
- If $v = u + w$, then $P_U(P_U v) = P_U u = u = P_U v$.
- If $v = u + w$, then $\|P_U v\|^2 = \|u\|^2 \leq \|u\|^2 + \|w\|^2 = \|v\|^2$, where the last line comes from the Pythagorean theorem.
- Look at first line of proof of $V = U \oplus U^\perp$.

■

Theorem (Riesz representation theorem): Suppose V is finite dimensional. For each $v \in V$, define $\varphi_v \in V'$ by

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

for each $u \in V$. Then $v \to \varphi_v$ is a surjective functiona from V to V'.

Proof: Suppose $0 \neq \varphi \in V'$ ($\varphi = 0 = \varphi_0$). Thus null $\varphi \neq V \Rightarrow (\text{null } \varphi)^\perp \neq \{0\}$. Let $0 \neq w \in (\text{null } \varphi)^\perp$. Let

$$v = \frac{\overline{\varphi(w)}}{\|w\|^2} w.$$

Then $v \in (\text{null } \varphi)^\perp$ and is nonzero. Taking the norm of both sides yields

$$\|v\| = \frac{|\varphi(w)|}{\|w\|}.$$

Apply φ to both sides of the first equation to get

$$\varphi(v) = \frac{\overline{\varphi(w)}}{\|w\|^2}\varphi(w) = \frac{|\varphi(w)|^2}{\|w\|^2} = \|v\|^2.$$

Now suppose $u \in V$. We have

$$u = \left(u - \frac{\varphi(u)}{\varphi(v)}v\right) + \frac{\varphi(u)}{\|v\|^2}v.$$

The term in the parantheses is in null φ and so is orthogonal to v. Thus we have

$$\langle u, v\rangle = \frac{\varphi(u)}{\|v\|^2}\langle v, v\rangle = \varphi(u).$$

Thus $\varphi = \varphi_u$, as desired. ∎

6.3.2. Minimization Problems

Proposition: Suppose U is a finite-dimensional subspace of V, $v \in V$, and $u \in U$. Then

$$\|v - P_U v\| \leq \|v - u\|.$$

Equality occurs if and only if $u = P_U v$.

Proof:

$$\begin{aligned}
\|v - P_U v\|^2 &\leq \|v - P_U v\|^2 + \|P_U v - u\|^2 \\
&= \|(v - P_U v) + (P_U v - u)\|^2 \\
&= \|v - u\|^2.
\end{aligned}$$

Equality occurs when $\|P_U v - u\|^2 = 0 \iff P_U v = u$. ∎

6.3.3. Pseudoinverse

Proposition: Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Then $T|_{(\text{null } T)^\perp}$ is an injective map of $(\text{null } T)^\perp$ onto range T.

Proof: Let R be the restriction map in the proposition. If $Rv = 0$ for $v \in (\text{null } T)^\perp$, then $v \in \text{null } T \Rightarrow v \in \text{null } T \cap (\text{null } T)^\perp$, which implies $v \in \{0\}$. Thus null $R = \{0\}$, which means R is injective.

Clearly range $R \subseteq$ range T. For the other direction, suppose $w \in$ range T. Thus there exist $v \in V$ such that $w = Tv$. We have $v = u + x$, where $u \in \text{null } T$ and $x \in (\text{null } T)^\perp$. thus

$$Rx = Tx = Tv - Tu = w - 0 = w,$$

which shows $w \in$ range R, giving us the inclusion in the other direction. Thus range $R =$ range T. ∎

Definition (pseudoinverse): Suppose V is finite dimensional and $T \in \mathcal{L}(V, W)$. The *pseudoinverse* $T^\dagger \in \mathcal{L}(W, V)$ of T is the linear map from W to V defined by

$$T^\dagger w = \left(T|_{(\text{null } T)^\perp}\right)^{-1} P_{\text{range } T} w$$

for each $w \in W$.

Proposition (properties of pseudoinverse): Suppose V is finite dimensional and $T \in \mathcal{L}(V, W)$.
- If T is invertible, then $T^\dagger = T^{-1}$.
- $TT^\dagger = P_{\text{range } T} =$ the orthogonal projection of W onto range T.
- $T^\dagger T = P_{(\text{null } T)^\perp} =$ the orthogonal projection of V onto $(\text{null } T)^\perp$.

Proof:
- Suppose T is invertible. Then $(\text{null } T)^\perp = V$ and range $T = W$. Thus $T|_{(\text{null } T)^\perp} = T$ and $P_{\text{range } T}$ is the identity operator on W. Thus $T^\dagger = T^{-1}$.
- Suppose $w \in$ range T. Then

$$TT^\dagger w = T\left(T|_{(\text{null } T)^\perp}\right)^{-1} w = w = P_{\text{range } T} w.$$

If $w \in (\text{range } T)^\perp$, then $T^\dagger w = 0 \Rightarrow TT^\dagger w = 0 = P_{\text{range } T} w$. Thus $TT^\dagger$ and $P_{\text{range } T}$ agree on range T and $(\text{range } T)^\perp$, which both collectively make up V. Thus the two maps are equal.
- Suppose $v \in (\text{null } T)^\perp$. Because $Tv \in$ range T, by definition we have

$$T^{\dagger(Tv)} = \left(T|_{(\text{null } T)^\perp}\right)^{-1} (Tv) = v = P_{(\text{null } T)^\perp} v.$$

If $v \in$ null T, then $T^\dagger T v = 0 = P_{(\text{null } T)^\perp} v$. This $T^\dagger T$ and $P\left((\text{null } T)^\perp\right)$ agree on $(\text{null } T)^\perp$ and null T, implying they are equal.

■

Proposition: Suppose V is finite-dimensional, $T \in \mathcal{L}(V, W)$, and $w \in W$.
- If $v \in V$, then

$$\left\|T\left(T^\dagger w\right) - w\right\| \leq \|Tv - w\|,$$

with equality if and only if $v \in T^\dagger w + \text{null } T$.
- If $v \in T^\dagger w + \text{null } T$, then

$$\left\|T^\dagger v\right\| \leq \|v\|,$$

with equality if and only if $v = T^\dagger w$.

Proof:

- Suppose $v \in V$. Then

$$Tv - w = \left(Tv - TT^{\dagger}w\right) + \left(TT^{\dagger}w - w\right).$$

The first term is in range T, and by the previous proposition the second term is in $(\text{range } T)^{\perp}$. Thus by the Pythagorean theorem the second term is less than or equal to $\|Tv - w\|$, with equality if and only if $Tv - TT^{\dagger}w = 0 \Rightarrow v - T^{\dagger}w \in \text{null } T$, which is eqivalent to $v \in T^{\dagger}w + \text{null } T$.

- Suppose $v \in T^{\dagger}w + \text{null } T$. Thus $v - T^{\dagger}w \in \text{null } T$. We have

$$v = \left(v - T^{\dagger}w\right) + T^{\dagger}w.$$

By definition $T^{\dagger}w \in (\text{null } T)^{\perp}$. Thus by the Pythagorean theorem, we have $\left\|T^{\dagger}w\right\| \leq \|v\|$, with equality if and only if $v = T^{\dagger}w$.

∎

6.3.4. Problems

Problem (Exercise 5): Suppose V is finite dimensional and U is a subspace of V. Show that $P_{U^{\perp}} = I - P_U$.

Solution: For $v \in V$, write $v = u + w$, where $u \in U$ and $w \in U^{\perp}$. Then we have

$$P_{U^{\perp}}v = w = v - u = (I - P_U)v.$$

Problem (Exercise 6): Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Show that

$$T = TP_{(\text{null } T)^{\perp}} = P_{\text{range } T}T.$$

Solution: For $v \in V$, we have $Tv \in \text{range } T$, so $P_{\text{range } T}Tv = Tv$, which implies $T = P_{\text{range } T}T$. We can also write $v = u + w$, where $u \in \text{null } T$ and $w \in (\text{null } T)^{\perp}$. Thus $Tv = Tu + Tw = Tw = TP_{(\text{null } T)^{\perp}}v$, which implies $T = TP_{(\text{null } T)^{\perp}}$.

Problem (Exercise 15): In $\mathbb{R}^4$, let

$$U = \text{span}((1, 1, 0, 0), (1, 1, 1, 2)).$$

Find $u \in U$ such that $\|u - (1, 2, 3, 4)\|$ is as small as possible.

Solution: We know that the u that minimizes $\|u - (1, 2, 3, 4)\|$ is $P_U(1, 2, 3, 4)$. By the Gram-Schmidt procedure, we have that $e_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), e_2 = \left(0, 0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$ is an orthonormal basis of U. Letting $v = (1, 2, 3, 4)$, we also know that

$$P_Uv = \langle v, e_1\rangle e_1 + \langle v, e_2\rangle e_2.$$

Computing this yields $u = \left(\frac{3}{2}, \frac{3}{2}, \frac{11}{5}, \frac{22}{5}\right)$.

Problem (Exercise 20): Suppose V is fintie dimensional and $T \in \mathcal{L}(V, W)$. Show that

$$\text{null } T^\dagger = (\text{range } T)^\perp \quad \text{and} \quad \text{range } T^\dagger = (\text{null } T)^\perp.$$

Solution: Suppose $w \in \text{null } T^\dagger$. Then we must have $P_{\text{range } T} w = 0$, since the other map in the definition (the invertible part) is injective. This implies $w \in (\text{range } T)^\perp$, so $\text{null } T^\dagger \subseteq (\text{range } T)^\perp$. For $w \in (\text{range } T)^\perp$, note that $P_{\text{range } T} w = 0$, so $(\text{range } T)^\perp \subseteq \text{null } T^\dagger$.

By definition, $T|_{(\text{null } T)^\perp}$ has domain $(\text{null } T)^\perp$, so clearly $\text{range } T^\dagger \subseteq (\text{null } T)^\perp$. Now pick $u \in (\text{null } T)^\perp$. Then $T|_{(\text{null } T)^\perp} u = Tu \in \text{range } T$. Thus $T^\dagger(Tu) = u$, so $(\text{null } T)^\perp \subseteq \text{range } T^\dagger$.

Problem (Exercise 23): Suppose V and W are finite dimensional and $T \in \mathcal{L}(V, W)$. Prove that

$$\left(T^\dagger\right)^\dagger = T.$$

Solution: We have

$$\left(T^\dagger\right)^\dagger v = \left(T^\dagger|_{(\text{null } T^\dagger)^\perp}\right)^{-1} P_{\text{range } T^\dagger} v.$$

By exercise 20, the right side is equal to

$$\left(T^\dagger|_{\text{range } T}\right)^{-1} P_{(\text{null } T)^\perp} v.$$

Now let $v = u + w$, where $u \in (\text{null } T)^\perp$ and $w \in \text{null } T$. Then we have

$$\left(T^\dagger\right)^\dagger v = \left(T^\dagger|_{\text{range } T}\right)^{-1} u.$$

Note that if we restrict $T^\dagger$ to range T, then we can drop the projection part of the definition of the pseudoinverse, which yields $T^\dagger|_{\text{range } T} = \left(T|_{(\text{null } T)^\perp}\right)^{-1} \iff \left(T^\dagger|_{\text{range } T}\right)^{-1} = T|_{(\text{null } T)^\perp}$. Thus we have

$$\left(T^\dagger\right)^\dagger v = \left(T|_{(\text{null } T)^\perp}\right) u = Tu + Tw = Tv.$$

7. Operators on Inner Product Spaces

7.1. Self-Adjoint and Normal Operators

7.1.1. Adjoints

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

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

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

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

Proof: Suppose $T \in \mathcal{L}(V, W)$. If $v \in V$ and $w_1, w_2 \in W$, then

$$\begin{aligned} \langle Tv, w_1 + w_2 \rangle &= \langle Tv, w_1 \rangle + \langle Tv, w_2 \rangle \\ &= \langle v, T^* w_1 \rangle + \langle v, T^* w_2 \rangle \\ &= \langle v, T^* w_1 + T^* w_2 \rangle. \end{aligned}$$

Thus we have $T^*(w_1 + w_2) = T^* w_1 + T^* w_2$. Similarly we can show that $T^*(\lambda w) = \lambda T^* w$. ∎

Proposition (properties of the adjoint): Suppose $T \in \mathcal{L}(V, W)$. Then
- $(S + T)^* = S^* + T^*$ for all $S \in \mathcal{L}(V, W)$
- $(\lambda T)^* = \overline{\lambda} T^*$ for all $\lambda \in \mathbb{F}$
- $(T^*)^* = T$
- $(ST)^* = T^* S^*$ for all $S \in \mathcal{L}(W, U)$ (U is a finite dimensional inner product space)
- $I^* = I$, where I is the identity operator on V
- If T is invertible, then T^* is invertible and $(T^*)^{-1} = (T^{-1})^*$.

Proof: Suppose $V \in V$ and $w \in W$.
- We have $\langle (S + T)v, w \rangle = \langle Sv, w \rangle + \langle Tv, w \rangle = \langle v, S^* w \rangle + \langle v, T^* w \rangle = \langle v, S^* w + T^* w \rangle$. Thus $(S + T)^* w = S^* w + T^* w$.
- If $\lambda \in \mathbb{F}$, then $\langle (\lambda T)v, w \rangle = \lambda \langle Tv, w \rangle = \lambda \langle v, T^* w \rangle = \left\langle v, \overline{\lambda} T^* w \right\rangle$. Thus $(\lambda T)^* w = \overline{\lambda} T^* w$.
- We have $\langle T^* w, v \rangle = \overline{\langle v, T^* w \rangle} = \overline{\langle Tv, w \rangle} = \langle w, Tv \rangle$. Thus $(T^*)^* v = Tv$.
- Suppose $S \in \mathcal{L}(W, U)$ and $u \in U$. Then $\langle (ST)v, u \rangle = \langle S(Tv), u \rangle = \langle Tv, S^* u \rangle = \langle v, T^*(S^* u) \rangle$. Thus $(ST)^* u = (T^* S^*) u$.
- Suppose $u \in V$. Then $\langle Iu, v \rangle = \langle u, v \rangle$, so $I^* v = v$.

- Suppose T is invertible. Taking adjoints of both sides of $T^{-1}T = I$ yields $T^*\left(T^{-1}\right)^* = I$. Similarly, we have $TT^{-1} = I \Rightarrow \left(T^{-1}\right)^* T^* = I$. Thus $\left(T^{-1}\right)^*$ is the inverse of T^*.

■

Proposition (null space and range of T^*): Suppose $T \in \mathcal{L}(V, W)$. Then
- null $T^* = (\text{range } T)^\perp$
- range $T^* = (\text{null } T)^\perp$
- null $T = (\text{range } T^*)^\perp$
- range $T = (\text{null } T^*)^\perp$

Proof: First we prove the first bullet. Let $w \in W$. Then

$$\begin{aligned} w \in \text{null } T^* &\iff T^* w = 0 \\ &\iff \langle v, T^* w\rangle = 0 \text{ for all } v \in V \\ &\iff \langle Tv, w\rangle = 0 \text{ for all } v \in V \\ &= w \in (\text{range } T)^\perp. \end{aligned}$$

Taking the orthogonal complement of both sides of the first bullet yields the fourth bullet. Replace T with T^* in the first to get the third, and do the same to the fourth to get the second. ■

Definition (conjugate transpose): The *conjugate transpose* of an m by n matrix A is the n by m matrix A^* obtained by transposing the matrix and taking the complex conjugate of every entry. In other words, if $j \in \{1, ..., n\}$ and $k \in \{1, ..., m\}$, then

$$A^*_{j,k} = \overline{A_{k,j}}.$$

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

$$\mathcal{M}(T^*) = (\mathcal{M}(T))^*.$$

Proof: The kth column of $\mathcal{M}(T)$ comes from writing Te_k as a linear combination of f_i's. Because the f's form an orthonormal basis, we have

$$Te_k = \langle Te_k, f_1\rangle f_1 + \cdots + \langle Te_k, f_m\rangle f_m.$$

Thus, entry j, k is $\langle Te_k, f_j\rangle$. Changing T to T^* and swapping the bases yields that j, k in $\mathcal{M}(T^*)$ is $\langle T^* f_k, e_j\rangle = \langle f_k, Te_j\rangle = \overline{\langle Te_j, f_k\rangle}$, which is the complex conjugate of k, j in $\mathcal{M}(T)$. ■

7.1.2. Self-Adjoint Operators

Definition (self-adjoint): An operator $T \in \mathcal{L}(V)$ is called *self-adjoint* if $T = T^*$.

Proposition: Every eigenvalue of a self-adjoint operator is real.

Proof: Suppose T is self-adjoint and let λ be an eigenvalue of T. Then for an eigenvector v corresponding to λ, we have

$$\lambda \|v\|^2 = \langle \lambda v, v \rangle = \langle Tv, v \rangle = \langle v, Tv \rangle = \langle v, \lambda v \rangle = \overline{\lambda} \|v\|^2,$$

which implies $\lambda = \overline{\lambda}$. Thus λ is real. ∎

Proposition: Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Then

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

Proof: If $u, w \in V$, then

$$\begin{aligned} \langle Tu, w \rangle = & \frac{\langle T(u+w), u+w \rangle - \langle T(u-w), u-w \rangle}{4} \\ & + \frac{\langle T(u+iw), u+iw \rangle - \langle T(u-iw), u-iw \rangle}{4} i. \end{aligned}$$

(can be checked by computing the right side) Now suppose $\langle Tv, v \rangle = 0$ for every $v \in V$. Then the above equation implies $\langle Tu, w \rangle = 0$ for all $u, w \in V$. Taking $w = Tu$ implies $Tu = 0$ for all $u \in U$, which implies $T = 0$. ∎

Proposition: Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Then

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

Proof: If $v \in V$, then

$$\langle T^* v, v \rangle = \overline{\langle v, T^* v \rangle} = \overline{\langle Tv, v \rangle}.$$

Now we have

$$\begin{aligned} T \text{ is self adjoint} & \iff T - T^* = 0 \\ & \iff \langle (T - T^*) v, v \rangle = 0 \text{ for every } v \in V \\ & \iff \langle Tv, v \rangle - \overline{\langle Tv, v \rangle} = 0 \text{ for every } v \in V \\ & \iff \langle Tv, v \rangle \in \mathbb{R} \text{ for every } v \in V. \end{aligned}$$

The second equivalence comes from the previous proposition, and the third comes from the equation at the beginning. ∎

Proposition: Suppose T is self adjoint. Then

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

Proof: We already proved this for complex inner product spaces, so can assume V is a real inner product space. If $u, w \in V$, then

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

which can be proven by computing the right side and using the equation $\langle Tw, u\rangle = \langle w, Tu\rangle = \langle Tu, w\rangle$.

Now suppose $\langle Tv, v\rangle = 0$ for all $v \in V$. Then we can conclude $\langle Tu, w\rangle = 0$ for all $u, w \in V$. Thus taking $w = Tu$ implies $Tu = 0$ for all $u \in V$, which implies $T = 0$. ■

7.1.3. Normal Operators

Definition (normal): An operator on an inner product space is called *normal* if it commutes with its adjoint:

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

Proposition: Suppose $T \in \mathcal{L}(V)$. then

$$T \text{ is normal} \Longleftrightarrow \|Tv\| = \|T^*v\| \text{ for every } v \in V.$$

Proof:

$$\begin{aligned} T \text{ is normal} &\Longleftrightarrow T^*T - TT^* = 0 \\ &\Longleftrightarrow \langle (T^*T - TT^*)v, v\rangle = 0 \text{ for every } v \in V \\ &\Longleftrightarrow \langle T^*Tv, v\rangle = \langle TT^*v, v\rangle \text{ for every } v \in V \\ &\Longleftrightarrow \langle Tv, Tv\rangle = \langle T^*v, T^*v\rangle \text{ for every } v \in V \\ &\Longleftrightarrow \|Tv\|^2 = \|T^*v\|^2 \text{ for every } v \in V \\ &\Longleftrightarrow \|Tv\| = \|T^*v\| \text{ for every } v \in V. \end{aligned}$$

(The second equivalence is allowed because $T^*T - TT^*$ is self adjoint.) ■

Proposition: Suppose $T \in \mathcal{L}(V)$ is normal. Then
- null $T =$ null T^*
- range $T =$ range T^*
- $V =$ null $T \oplus$ range T
- $T - \lambda I$ is normal for every $\lambda \in \mathbb{F}$
- If $v \in V$ and $\lambda \in \mathbb{F}$, then $Tv = \lambda v$ if and only if $T^* v = \overline{\lambda} v$.

Proof:
- Suppose $v \in V$. Then
$$v \in \text{null } T \iff \|Tv\| = 0 \iff \|T^* v\| = 0 \iff v \in \text{null } T^*.$$
- We have range $T = (\text{null } T^*)^{\perp} = (\text{null } T)^{\perp} =$ range T^*.
- $V =$ null $T \oplus (\text{null } T)^{\perp} =$ null $T \oplus$ range $T^* =$ null $T \oplus$ range T.
- We have
$$\begin{aligned}(T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)\left(T^* - \overline{\lambda} I\right) \\ &= TT^* = \overline{\lambda} T - \lambda T^* + |\lambda|^2 I \\ &= T^* T - \overline{\lambda} T - \lambda T^* + |\lambda|^2 I \\ &= (T^* - \lambda I)(T - \lambda I) \\ &= (T - \lambda I)^* (T - \lambda I).\end{aligned}$$
- We have $\|(T - \lambda I)v\| = \left\|(T - \lambda I)^* v\right\| = \left\|\left(T^* - \overline{\lambda} I\right)v\right\|$. Thus $\|(T - \lambda I)v\| = 0$ if and only if $\left\|\left(T^* - \overline{\lambda} I\right)v\right\| = 0$. Thus $Tv = \lambda v$ if and only if $T^* v = \overline{\lambda} v$.

■

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

Proof: Suppose α, β are distinct eigenvalues of T with corresponding eigenvectors u, v. Thus $Tu = \alpha u$ and $T^* v = \overline{\beta} v$. Then we have
$$\begin{aligned}(\alpha - \beta)\langle u, v\rangle &= \langle \alpha u, v\rangle - \left\langle u, \overline{\beta} v\right\rangle \\ &= \langle Tu, v\rangle - \langle u, T^* v\rangle = 0.\end{aligned}$$
Since $\alpha \neq \beta$, we must have $\langle u, v\rangle = 0$. ■

Proposition: Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then T is normal if and only if there exist commuting self adjoint operators A and B such that $T = A + iB$.

Proof: First suppose T is normal. Let $A = \frac{T + T^*}{2}$ and $B = \frac{T - T^*}{2i}$. A and B are self adjoint and $T = A + iB$. Then we have
$$AB - BA = \frac{T^* T - TT^*}{2i}.$$

Because T is normal, the right side is 0, which means A and B commute.

Now suppose A and B exist as in the proposition. Then $T^* = A - iB$. Adding this to $T = A + iB$ and dividing by 2 yields our expression for A at the beginning of the proof. Subtracting and dividing by 2 yields B. Both of these imply the equation at the end of the forward direction, and since A and B commute, this implies T is normal. ■

7.1.4. Problems

Problem (Exercise 1): Suppose n is a positive integer. Define $T \in \mathcal{L}(\mathbb{F}^n)$ by

$$T(z_1, ..., z_n) = (0, z_1, ..., z_{n-1}).$$

Find a formula for $T^*(z_1, ..., z_n)$.

Solution: We have

$$\begin{aligned}\langle (0, z_1, ..., z_{n-1}), (x_1, x_2, ..., x_n) \rangle &= z_1 x_2 + z_2 x_3 + \cdots + z_{n-1} x_n \\ &= \langle (z_1, z_2, ..., z_n), (x_2, x_3, ..., x_n, 0) \rangle.\end{aligned}$$

Thus we have $T^*(x_1, x_2, ..., x_n) = (x_2, x_3, ..., x_n, 0)$.

Problem (Exercise 2): Suppose $T \in \mathcal{L}(V, W)$. Prove that

$$T = 0 \Longleftrightarrow T^* = 0 \Longleftrightarrow T^*T = 0 \Longleftrightarrow TT^* = 0.$$

Solution: We have $\langle v, T^*w \rangle = \langle Tv, w \rangle = \langle 0, w \rangle = 0$ for all $v \in V$ and $w \in W$. Thus we can pick $v = T^*v$ to obtain $\langle T^*w, T^*w \rangle = 0 \Rightarrow T^*w = 0$ for all w, which implies $T^* = 0$.

Now suppose $TT^* = 0$, and suppose $Tv \neq 0$. We have $Tv \in \text{range } T$ and $Tv \in \text{null } T^* = (\text{range } T)^\perp$. Thus, $Tv \in \text{range } T \cap (\text{range } T)^\perp = \{0\}$, which is a contradiction. Thus, $T = 0$.

Problem (Exercise 3): Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. Prove that

$$\lambda \text{ is an eigenvalue of } T \Longleftrightarrow \overline{\lambda} \text{ is an eigenvalue of } T^*.$$

Solution: Suppose λ is not an eigenvalue of T. Then $T - \lambda I$ is invertible, which implies $(T - \lambda I)^* = T^* - \overline{\lambda} I$ is invertible, which implies $\overline{\lambda}$ is not an eigenvalue of T^*. The reverse direction follows similarly.

Problem (Exercise 4): Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V. Prove taht

$$U \text{ is invariant under } T \Longleftrightarrow U^\perp \text{ is invariant under } T^*.$$

Solution: Suppose $u, Tu \in U$ and $u' \in U^{\perp}$. We have $\langle u, T^* u' \rangle = \langle Tu, u' \rangle = 0$ for all $u \in U$. This $T^* u' \in U^{\perp}$, so $U^{\perp}$ is indeed invariant under T^*. The reverse direction follows simlarly.

Problem (Exercise 5): Suppose $T \in \mathcal{L}(V, W)$. Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V and $f_1, \ldots, f_m$ is an orthonormal basis of W. Prove that

$$\|Te_1\|^2 + \cdots + \|Te_n\|^2 = \|T^* f_1\|^2 + \cdots + \|T^* f_m\|^2.$$

Solution: Recall that for an orthonormal basis, we have $\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2$.

Proving this amount to a simple calculation:

$$\begin{aligned}
\sum_{i=1}^{n} \|Te_i\|^2 &= \sum_{i=1}^{n} \sum_{j=1}^{m} |\langle Te_i, f_j \rangle|^2 \\
&= \sum_{j=1}^{m} \sum_{i=1}^{n} |\langle e_i, T^* f_j \rangle|^2 \\
&= \sum_{j=1}^{m} \sum_{i=1}^{n} |\langle T^* f_j, e_i \rangle|^2 = \sum_{j=1}^{m} \|T^* f_j\|^2.
\end{aligned}$$

Problem (Exercise 6): Suppose $T \in \mathcal{L}(V, W)$. Prove that
(a) T is injective $\Longleftrightarrow T^*$ is surjective
(b) T is surjective $\Longleftrightarrow T^*$ is injective

Solution:
(a) $\text{null } T = \{0\} \Rightarrow (\text{range } T^*)^{\perp} = \{0\} \Rightarrow \text{range } T^* = W$
(b) $\text{range } T = W \Rightarrow (\text{null } T^*)^{\perp} = W \Rightarrow \text{null } T^* = \{0\}$

Problem (Exercise 7): Prove that if $T \in \mathcal{L}(V, W)$, Then
(a) $\dim \text{null } T^* = \dim \text{null } T + \dim W - \dim V$
(b) $\dim \text{range } T^* = \dim \text{range } T$

Solution:
(a)

$$\begin{aligned}
\dim \text{null } T^* &= \dim (\text{range } T)^{\perp} \\
&= \dim W - \dim \text{range } T \\
&= \dim W - (\dim V - \dim \text{null } T) = \dim \text{null } T + \dim W - \dim W
\end{aligned}$$

(b)

$$\begin{aligned}\dim \operatorname{null} T^* + \dim \operatorname{range} T^* = \dim W \Rightarrow \dim \operatorname{range} T^* &= \dim W - (\dim \operatorname{null} T + \dim W - \dim V) \\ &= \dim V - \dim \operatorname{null} T = \dim \operatorname{range} T\end{aligned}$$

Problem (Exercise 10): Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Prove that T is self adjoint if and only if

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

for all $v \in V$.

Solution: Clearly if T is self adjoint the equation holds. Now suppose the equation holds for all $v \in V$. We have $\langle Tv, v\rangle = \langle T^* v, v\rangle = \langle v, Tv\rangle = \overline{\langle Tv, v\rangle}$. Thus, $\langle Tv, v\rangle \in \mathbb{R}$ for all $v \in V$, which implies T is self adjoint.

Problem (Exercise 11): Define an operator $S : \mathbb{F}^2 \to \mathbb{F}^2$ by $S(w, z) = (-z, w)$.
(a) Find a formula for S^*.
(b) Show that S is normal but not self adjoint.
(c) Find all eigenvalues of S.

Solution:
(a) $\langle S(a_1, a_2), (b_1, b_2)\rangle = -a_2 b_1 + a_1 b_2 = \langle (a_1, a_2), (b_2, -b_1)\rangle$, so $S^*(x, y) = (y, -x)$.
(b) S and its adjoint commute, but they're not equal.
(c) $\lambda = \pm i$.

Problem (Exercise 12): An operator $B \in \mathcal{L}(V)$ is called *skew* if

$$B^* = -B.$$

Suppose $T \in \mathcal{L}(V)$. Prove that T is normal if and only if there exist commuting operators A and B such that A is self adjoint, B is a skew operator, and $T = A + B$.

Solution: Suppose we can represent T as described in the problem. Then

$$\begin{aligned} TT^* &= (A + B)(A^* + B^*) \\ &= (A + B)(A - B) \\ &= AA - AB + BA - BB \\ &= AA - BA + AB - BB \\ &= (A - B)(A + B) \\ &= (A^* + B^*)(A + B) = T^* T, \end{aligned}$$

so T is normal.

Now suppose T is normal. Then we have $A = \frac{T + T^*}{2}$ is self adjoint and $B = \frac{T - T^*}{2}$ is skew, and $AB = \frac{1}{4}(TT - T^* T^*) = BA$.

Problem (Exercise 13): Suppose $\mathbb{F} = \mathbb{R}$. Define $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ by $\mathcal{A}T = T^*$ for all $T \in \mathcal{L}(V)$. Find all eigenvalues of $\mathcal{A}$ and the minimal polyomial of A.

Solution: Note that $\mathcal{A}^2 - I = 0$, so $z^2 - 1$ is a multiple of the minmal polynomial if $\mathcal{A}$. Note also that $\lambda = \pm 1$ are eigenvalues of $\mathcal{A}$ (corresponding to self adjoint and skew operators). Thus the minimal polynomial of $\mathcal{A}$ is a multiple of $z^2 - 1$. Both of these imply that the minimal polynomial of $\mathcal{A}$ is $z^2 - 1$.

Problem (Exercise 15): Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that
(a) T is self adjoint $\Longleftrightarrow T^{-1}$ is self adjoint
(b) T is normal $\Longleftrightarrow T^{-1}$ is normal

Solution:
(a) $T = T^* \Rightarrow T^{-1} = (T^*)^{-1} \Rightarrow T^{-1} = (T^{-1})^*$
(b) $TT^* = T^*T \Rightarrow (T^*)^{-1}T^{-1} = T^{-1}(T^*)^{-1} \Rightarrow (T^{-1})^*T^{-1} = T^{-1}(T^{-1})^*$

Problem (Exercise 16): Suppose $\mathbb{F} = \mathbb{R}$. Show that the set of self adjoint operators on V is a subspace of $\mathcal{L}(V)$, and find the dimension of that subspace.

Solution: It follows the set is a subspace by rules of adjoints. To find the dimension, note that with respect to the standard basis of $\mathcal{L}(V)$, the upper triagnular portion of the matrix can be picked freely, but then the bottom part of the matrix is fixed (the matrix must equal its transpose). Thus the dimension is $\frac{(\dim V)(\dim V - 1)}{2}$.

Problem (Exercise 23): Suppose T is a normal operator on V. Suppose also that $v, w \in V$ satisfy the equations

$$\|v\| = \|w\| = 2, \quad Tv = 3v, \quad Tw = 4w.$$

Show that $\|T(v + w)\| = 10$.

Solution: Note that $\langle Tv, Tw \rangle = 12\langle v, w \rangle = 0$, since v and w are distinct eigenvectors of a normal operator. We then have

$$\|T(v + w)\|^2 = \langle Tv + Tw, Tv + Tw \rangle = \|Tv\| + \|Tw\| = 100,$$

as desired.

Problem (Exercise 30): Suppose that $T \in \mathcal{L}(\mathbb{F}^3)$ is normal and $T(1,1,1) = (2,2,2)$. Suppose $(z_1, z_2, z_3) \in \text{null } T$. Prove that $z_1 + z_2 + z_3 = 0$.

Solution: Note that $T(1,1,1) = (2,2,2) \Rightarrow T^*(1,1,1) = (2,2,2)$. Suppose $T(z_1, z_2, z_3) = 0$. Then we have

$$\begin{aligned} 0 &= \langle T(z_1, z_2, z_3), (1,1,1) \rangle \\ &= \langle (z_1, z_2, z_3), T^*(1,1,1) \rangle \\ &= 2z_1 + 2z_2 + 2z_3, \end{aligned}$$

which implies the desired result.

Problem (Exercise 32): Suppose $T : V \to W$ is a linear map. Show that under the standard identification of V with V' and the corresponding identification of W with W', the adjoint map $T^* : W \to V$ corresponds to the dual map $T' : W' \to V'$. More precisely, show that

$$T'(\varphi_w) = \varphi_{T^* w}$$

for all $w \in W$.

Solution:

$$\varphi_{T^* w} u = \langle u, T^* w \rangle = \langle Tu, w \rangle = (\varphi_w \circ T)u = T'(\varphi_w)u.$$

7.2. Spectral Theorem

7.2.1. Real Spectral Theorem

Proposition (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.

Proof: For nonzero v we have

$$\begin{aligned}\langle (T^2 + bT + cI)v, v\rangle &= \langle T^2 v, v\rangle + b\langle Tv, v\rangle + c\langle v, v\rangle \\ &= \langle Tv, Tv\rangle + b\langle Tv, Tv\rangle + c\|v\|^2 \\ &\geq \|Tv\|^2 - |b|\|Tv\|\|v\| + c\|v\|^2 \\ &= \left(\|Tv\| - \frac{|b|\|v\|}{2}\right)^2 + \left(c - \frac{b^2}{4}\right)\|v\|^2 \\ &> 0,\end{aligned}$$

where the third line follows from CS. The last inequality implies $(T^2 + bT + cI)v \neq 0$, which implies the map is injective, and thus invertible. ∎

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

Proof: First suppose $\mathbb{F} = \mathbb{C}$. We know that all of T's eigenvalues are real, and since the eigenvalues of T are the zeroes of the minimal polynomial, the result follows.

Now suppose $\mathbb{F} = \mathbb{R}$. We can write the minimal polynomial as

$$p(z) = (z - \lambda_1)\cdots(z - \lambda_m)(z^2 + b_1 z + c_1)\cdots(z^2 + b_N z + c_N),$$

where the quadratics are irriducible over $\mathbb{R}$. We have $p(T) = 0$. If $N > 0$, then we can multiply $p(T)$ by the inverse of one of the quadratic factors (they're invertible by the previous result). However, this will reduce the degree of the minimal polynomial by 2, which contradicts minimality. Thus $N = 0$, and the minimal polynomial has the desired form. ∎

Theorem (real spectral theorem): Suppose $\mathbb{F} = \mathbb{R}$ and $T \in \mathcal{L}(V)$. THen the following are equivalent:
(a) T is self-adjoint.
(b) T has a diagonal matrix with respect to some orthonormal basis of V.
(c) V has an orthonormal basis consisting of eigenvectors of T.

Proof: First suppose (a) holds. Then we know that T has minimal polynomial with factors of degree 1, which implies that there exists an orthonormal basis for which T has an upper-triangular matrix. With respect to this basis, $T^* = T^t = T$. This implies that everything above the diagonal of T must be 0 as well, which implies the matrix is diagonal.

Now suppose (b) holds. Then the diagonal matrix equals its transpose. Thus, with respect to the orthonormal basis, $T^* = T$.

The equivalence of (b) and (c) follows from earlier propositions (eigenbases create diagonal matrices). ∎

7.2.2. Complex Spectral Theorem

Theorem (complex spectral theorem): Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then the following are equivalent:
(a) T is normal.
(b) T has a diagonal matrix with respect to some orthonormal basis of V.
(c) V has an orthonormal basis consisting of eigenvectors of T.

Proof: First suppose (a) holds. By Schur's theorem, there is an orthonormal basis $e_1, ..., e_n$ for which T is upper triangular. Thus we have

$$\mathcal{M}(T, (e_1, ..., e_n)) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ & \ddots & \vdots \\ 0 & & a_{n,n} \end{pmatrix}.$$

From the matrix, we have

$$\begin{aligned} \|Te_1\|^2 &= |a_{1,1}|^2 \\ \|T^*e_1\|^2 &= |a_{1,1}|^2 + |a_{1,2}|^2 + \cdots + |a_{1,n}|^2. \end{aligned}$$

Since T is normal, we have $\|Te_1\| = \|T^*e_1\|$, which implies that everything after $a_{1,1}$ in the second equation is 0. We can repeat for $e_2, e_3 ... e_n$ and get that the matrix is diagonal.

Now suppose (b) holds. Then T^* is also a diagonal matrix. Since diagonal matrices commute, T is normal.

The equivalence of (b) and (c) follows from earlier propositions (eigenbases create diagonal matrices). ∎

7.2.3. Problems

Problem (Exercise 1): Prove that a normal operator on a complex inner product space is self adjoint if and only if all its eigenvalues are real.

Solution: First suppose the normal operator has all real eigenvalues. By the complex spectral theorem, there exists an orthonormal basis for which the operator has a diagonal matrix. Since the matrix must be equal to its conjugate transpose, all the elements on the diagonal are real. Since the diagonal elements are the set of eigenvalues of the operator, we have desired result. The other direction follows since the eigenvalues of a self adjoint map are all real.

Problem (Exercise 2): Suppose $\mathbb{F} = \mathbb{C}, T \in \mathcal{L}(V)$ is normal and only has one eigenvalue. Prove that T is a scalar multiple of the identity operator.

Proof: By the complex spectral theorem, there exists an orthonormal basis for which T has a diagonal basis. Since all the eigenvalues of a map appear on the diagonal in a diagonal matrix, the diagonal of the map contains that one eigenvalue, while the rest of the matrix is 0. The matrix is a scalar multiple of the identity matrix, which implies the map is a scalar multiple of the identity map. ∎

Problem (Exercise 3): Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$ is normal. Prove that the set of eigenvalues of T is contained in $\{0, 1\}$ if and only if there is a subspace U of V such taht $T = P_U$.

Solution: If $T = P_U$, then $P^2 = P$, so $P(P - I) = 0$, which implies that the eigenvalues of T are in $\{0, 1\}$.

Now suppose the set of eigenvalues is in $\{0, 1\}$. By the complex spectral theorem, there's an orthonormal basis for which the matrix of T is diagonal, and this only contains 0s and 1s on the diagonal. Let $e_1, \ldots, e_n$ be the orthonormal basis, and let $e_{i_1}, \ldots, e_{i_k}$ be the vectors who have associated eigenvalue 1. We claim $T = P_U$, where $U = \operatorname{span}\left(e_{i_1}, \ldots, e_{i_k}\right)$. Let $v = \sum a_j e_j$. We have

$$Tv = \sum a_{i_j} e_{i_j} = P_U v,$$

where the first equality comes by matrix multiplication and the second one comes from the definition of the projection map.

Problem (Exercise 4): Prove that a normal operator on a complex inner product space is skew if and only if all its eigenvalues are pure imaginary.

Solution: First suppose a normal operator has pure imaginary eigenvalues. By CST we can write the operator as a diagonal matrix. Since all the eigenvalues are pure imaginary, its adjoint (conjugate transpose) will be the negative of the original matrix.

Now suppose the operator is skew. By CST, we can write it as a diagonal matrix. Then we have $T^* = -T$. Looking at one diagonal entry, we need $\overline{\lambda} = -\lambda$, which implies λ is pure imaginary. Since the diagonal contains all the eigenvalues, every eigenvalue of T is pure imaginary.

Problem (Exercise 6): Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$ is a normal operator such that $T^9 = T^8$. Prove that T is self adjoint and $T^2 = T$.

Solution: Note that $T^8(T - I) = 0$, so the eigenvalues of T are in $\{0, 1\}$. By CST we can write T as a diagonal matrix, and there will only be 0s and 1s on the diagonals. Thus the matrix of T equals its conjugate transpose, so T is self adjoint. Now consider $T(T - I)$ under that orthonormal basis. Since every basis vector is an eigenvector, and each one has associated eigenvalue either 0 or 1, $T(T - I) = 0$ on all basis vectors, which means it's 0 on V, whivh gives us the desired result.

Problem (Exercise 10): Suppose V is a complex inner product space. Prove that every normal operator on V has a square root.

Solution: Let T be a normal operator, and write T is a diagonal matrix with respect to the appropriate orthonormal basis. Thus we have

$$\mathcal{M}(T, (e_1, ..., e_n)) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Let S be the operator such that

$$\mathcal{M}(S, (e_1, ..., e_n)) = \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}.$$

Clearly $S^2 = T$, so we're done.

Problem (Exercise 11): Prove that every self adjoint operator on V has a cube root.

Solution: Basically the same as the previous problem.

Problem (Exercise 12): Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ is normal. Prove that if S is an operator on V that commutes with T, then S commutes with T^*.

Solution: We work in an orthonormal basis for which T is a diagonal operator. Suppose $Te_i = \lambda_i e_i \Rightarrow T^* e_i = \overline{\lambda_i} e_i$. Then we have $STe_i = \lambda_i Se_i = TSe_i \Rightarrow T^* Se_i = \overline{\lambda_i} Se_i = ST^* e_i$. Since this holds for every basis vector, S and T^* commute.

Problem (Exercise 14): Suppose $\mathbb{F} = \mathbb{R}$ and $T \in \mathcal{L}(V)$. Prove that T is self adjoint if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, where $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

Solution: First suppose T is self adjoint. Then we already know that eigenvectors corresponding to distinct eigenvalues are orthogonal, and the second part follows since T is diagonalizable.

Now suppose the last two implications hold. Let ℓ_i represent the basis of each eigenspace. We can use the Gram-Schmidt procedure on each to obtain an orthonormal basis for each eigenspace ℓ_i'. Combining all these lists and using the assumption that every eigenvector corresponding to distinct eigenvalues are orthogonal, we have an orthonormal eigenbasis, which means T is self adjoint.

Problem (Exercise 15): Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Prove that T is normal if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, where $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

Solution: First suppose T is normal. Then we already know that eigenvectors corresponding to distinct eigenvalues are orthogonal, and the second part follows since T is diagonalizable.

Now suppose the last two implications hold. Let ℓ_i represent the basis of each eigenspace. We can use the Gram-Schmidt procedure on each to obtain an orthonormal basis for each eigenspace ℓ_i'. Combining all these lists and using the assumption that every eigenvector corresponding to distinct eigenvalues are orthogonal, we have an orthonormal eigenbasis, which means T is normal.

Problem (Exercise 18): Give an example of a real inner product space V, an operator $T \in \mathcal{L}(V)$, and real numbers b, c with $b^2 < 4c$ such that

$$T^2 + bT + cI$$

is not invertible.

Solution: The 90° rotation map with $p(z) = z^2 + 1$ is not invertible.

Problem (Exercise 19): Suppose $T \in \mathcal{L}(V)$ is self adjoint and U is a subspace of V that is invariant under T.
(a) Prove that $U^\perp$ is invariant under T.
(b) Prove that $T|_U \in \mathcal{L}(U)$ is self adjoint.
(c) Prove that $T|_{U^\perp} \in \mathcal{L}(U)$ is self adjoint.

Solution:
(a) Suppose $u \in U$ and $u' \in U^\perp$. We have $\langle Tu, u' \rangle = 0 = \langle u, Tu' \rangle$. Thus $Tu' \in U^\perp$, which means $U^\perp$ is invariant under T.
(b) Suppose $u, v \in U$. We have $\langle Tu, v \rangle = \langle u, Tv \rangle$ for all u, v. Since $Tu, Tv \in U$, the restricted operator is clearly self adjoint.
(c) Same reasoning as previous bullet.

7.3. Positive operators

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

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

for all $v \in V$.

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

Proposition (characterization of positive operators): Let $T \in \mathcal{L}(V)$. Then the following are equivalent:
(a) T is a positive operator.
(b) T is self adjoint and all eigenvalues of T are nonnegative.
(c) With respect to some orthonormal basis of V, the matrix of T is a diagonal matrix with only nonnegative numbers on the diagonal.
(d) T has a positive square root.
(e) T has a self adjoint square root.
(f) $T = R^* R$ for some $R \in \mathcal{L}(V)$.

Proof: Suppose (a) holds. Then T is obviously self adjoint. Suppose λ is an eigenvalue of T with corresponding eigenvector v. Then

$$0 \leq \langle Tv, v\rangle = \langle \lambda v, v\rangle = \lambda\langle v, v\rangle,$$

which implies λ is nonnegative. Thus (b) holds, which means (a) implies (b).

Suppose (b) holds. By the spectral theorem, there is an orthonormal basis for which T is a diagonal operator. Since the matrix will contain the eigenvalues on the diagonal, every entry will be nonnegative. Thus (b) implies (c).

Suppose (c) holds. Let $e_1, ..., e_n$ be the orthonormal basis for which T is diagonal, and let $\lambda_1, ..., \lambda_n$ be the corresponding eigenvalues. Define $R \in \mathcal{L}(V)$ by $Re_k = \sqrt{\lambda_k} e_k$. Note that R has an is diagonal with respect to an orthonormal basis, so it's self adjoint. We also have

$$\langle Rv, v\rangle = \left\langle \sum a_i \sqrt{\lambda_i} e_i, \sum a_i e_i \right\rangle = \sum a_i^2 \sqrt{\lambda_i} \geq 0,$$

so R is positive. Note also that $R^2 e_k = \lambda_k e_k = Te_k$ for all basis vectors, so we have $R^2 = T$. Thus (c) implies (d).

Suppose (d) holds. Then (e) holds by definition.

Suppose (e) holds. Suppose $T = R^2$, where R is self adjoint. Then, since $R = R^*$, we have $T = R^* R$. This (e) implies (f).

Now suppose (f) holds. Note that $T = R^* R = T^*$, so T is self adjoint. We also have

$$\langle Tv, v\rangle = \langle R^* Rv, v\rangle = \langle Rv, Rv\rangle \geq 0.$$

Thus T is positive, so (f) implies (a). ■

Proposition: Every positive operator on V has a unique positive square root.

Proof: Suppose $T \in \mathcal{L}(V)$ is positive and $v \in V$ is an eigenvector of T. Thus there exists a real number $\lambda \geq 0$ such that $Tv = \lambda v$. Let R be a positive square root of T. We prove $Rv = \sqrt{\lambda} v$. This will imply that R on the eigenvectors of T is uniquely determined. Since by the spectral theorem there is a basis of V consisting of eigenvectors of T, R will be uniquely determined.

By the spectral theorem, there's an orthonormal basis $e_1, ..., e_n$ of V consisting of eigenvectors of R. Because R is positive, all its eigenvalues are nonnegative. Thus there exist nonnegative numbers $\lambda_1, ..., \lambda_n$ such that $Re_k = \sqrt{\lambda_k} e_k$.

Because $e_1, ..., e_n$ is a basis of V, we have

$$v = \sum a_i e_i.$$

Thus

$$Rv = \sum a_i \sqrt{\lambda_i} e_i.$$

This implies

$$\sum a_i \lambda e_i = \lambda v = Tv = R^2 v = \sum a_i \lambda_i e_i.$$

Thus $a_k(\lambda - \lambda_k) = 0$. Thus

$$v = \sum_{\{i : \lambda_i = \lambda\}} a_i e_i.$$

Thus

$$Rv = \sum_{\{i : \lambda_i = \lambda\}} a_i \sqrt{\lambda_i} e_i = \sum_{\{i : \lambda_i = \lambda\}} a_i \sqrt{\lambda} e_i = \sqrt{\lambda} v,$$

as desired. ■

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

Proposition: Suppose T is a positive operator on V and $v \in V$ is such that $\langle Tv, v\rangle = 0$. Then $Tv = 0$.

Proof: We have

$$0 = \langle Tv, v\rangle = \left\langle \sqrt{T}\sqrt{T}v, v\right\rangle = \left\langle \sqrt{T}v, \sqrt{T}v\right\rangle = \left\|\sqrt{T}v\right\|^2,$$

so $\sqrt{T}v = 0$. Thus $Tv = \sqrt{T}\left(\sqrt{T}v\right) = 0$. ■

7.3.1. Problems

Problem (Exercise 1): Suppose $T \in \mathcal{L}(V)$. Prove that if both T and $-T$ are positive operators, then $T = 0$.

Solution: We have $\langle Tv, v\rangle \geq 0 \Rightarrow \langle -Tv, v\rangle \leq 0$. However, we also have $\langle -Tv, v\rangle \geq 0$, so in fact $\langle Tv, v\rangle = 0$ for all $v \in V$. By the previous proposition, we have $Tv = 0$ for all $v \in V$, which implies $T = 0$.

Problem (Exercise 2): Suppose $T \in \mathcal{L}(\mathbb{F}^4)$ is the operator whose matrix with respect to the standard basis is

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}.$$

Show that T is an invertible operator.

Solution: The matrix of T is equal to its conjugate transpose, so T is self adjoint. Note that $\langle Tv, v\rangle = (2w - x)w + (-w + 2x - y)x + (-x + 2y - z)y + (-y + 2z)z = w^2 + (w - x)^2 + (x - y)^2 + (y - z)^2 + z^2 \geq 0$, so T is indeed positive.

Problem (Exercise 3): Suppose n is a postiive integer and $T \in \mathcal{L}(\mathbb{F}^n)$ is the operator whose matrix with respect to the standard basis consists of all 1's. Show that T is a positive operator.

Solution: Note that the matrix is equal to its conjugate transpose, so T is self adjoint. Note also that the minimal polynomial of T is $z^2 - nz = z(z - n)$. Thus all the eigenvalues of T are nonnegative, which means T is positive.

Problem (Exrcise 6): Prove that the sum of two positive operators on V is a positive operator.

Solution: Let S and T be positive operators on V. Then we have $\langle (S + T)v, v\rangle = \langle Sv, v\rangle + \langle Tv, v\rangle \geq 0$.

Problem (Exercise 7): Suppose $S \in \mathcal{L}(V)$ is an invertible operator and $T \in \mathcal{L}(V)$ is a positive operator. Prove that $S + T$ is invertible.

Solution: We have $\langle Sv, v\rangle > 0$ for all nonzero v and $\langle Tv, v\rangle \geq 0$. Thus $\langle (S+T)v, v\rangle > 0$ for all nonzero v, which implies $S + T$ is invertible positive.

Problem (Exercise 9): Suppose $T \in \mathcal{L}(V)$ is a positive operator and $S \in \mathcal{L}(W, V)$. Prove that S^*TS is a positive operator on W.

Solution: Clearly the map is self adjoint. Then we have $\langle S^*TSw, w\rangle = \langle TSw, Sw\rangle \geq 0$.

Problem (Exercise 10): Suppose T is a positive operator on V. Suppose $v, w \in V$ are such that

$$Tv = w \text{ and } Tw = v.$$

Prove that $v = w$.

Solution: The two equations imply

$$T(v-w) = -(v-w) \Rightarrow \langle T(v-w), (v-w)\rangle = \langle -(v-w), v-w\rangle = -\|v-w\|^2 \geq 0.$$

This implies $v - w = 0$.

Problem (Exercise 11): Suppose T is a positive operator on V and U is a subspace of V invariant under T. Prove that $T|_U \in \mathcal{L}(U)$ is a positive operator on U.

Solution: Clearly T_U is self adjoint on U, and for any $u \in U$ we have $\langle Tu, u\rangle \geq 0$, since T is positive. Thus T_U is positive.

Problem (Exercise 12): Suppose $T \in \mathcal{L}(V)$ is a positive operator. Prove that T^k is a positive operator for every positive integer k.

Solution: Clearly $\left(T^k\right)^* = \left(T^*\right)^k$. For even k, we have $\langle T^k v, v\rangle = \left\langle T^{\frac{k}{2}}v, T^{\frac{k}{2}}v\right\rangle = \left\|T^{\frac{k}{2}}v\right\|^2 \geq 0$. For odd k, we have $\langle T^k v, v\rangle = \left\langle T^{\frac{k+1}{2}}v, T^{\frac{k-1}{2}}v\right\rangle = \left\langle T\left(T^{\frac{k+1}{2}}v\right), T^{\frac{k+1}{2}}v\right\rangle \geq 0$.

Problem (Exercise 13): Suppose $T \in \mathcal{L}(V)$ is self adjoint and $\alpha \in \mathbb{R}$.

(a) Prove that $T - \alpha I$ is a positive operator if and only if α is less than or equal to every eigenvalue of T.

(b) Prove that $\alpha I - T$ is a positive operator if and only if α is greater than or equal to every eigenvalue of T.

Solution:

(a) First suppose $T - \alpha I$ is positive. Let $\lambda_{\min}$ be the smallest eigenvalue of T, with associated eigenvector v. Then we have $\langle (T - \alpha I)v, v\rangle = \langle Tv, v\rangle - \alpha\langle v, v\rangle = (\lambda_{\min} - \alpha)\|v\|^2 \geq 0$. thus we have $\lambda_{\min} \geq \alpha$.

Now suppose $\alpha \leq \lambda_{\min}$. Clearly $T - \alpha I$ is self adjoint. Since T is self adjoint, let $e_1, ..., e_n$ be an orthonormal eigenbasis. We have $\langle (T - \alpha I)(a_i e_i), a_i e_i\rangle = \langle T(a_i e_i), a_i e_i\rangle - \alpha\langle a_i e_i, a_i e_i\rangle = (\lambda_i - \alpha)\|a_i e_i\|^2 \geq 0$. We also have $\left\langle (T - \alpha I)(a_i e_i), a_j e_j\right\rangle = \left\langle T(a_i e_i), a_j e_j\right\rangle - \alpha\left\langle a_i e_i, a_j e_j\right\rangle = \lambda_i\left\langle a_i e_i, a_j e_j\right\rangle = 0$. Thus for any $v = \sum a_i e_i$, we have $\langle (T - \alpha I)v, v\rangle \geq 0$. Thus $T - \alpha I$ is positive.

(b) Basically the same as the last bullet.

Problem (Exercise 14): Suppose T is a positve operator on V and $v_1, ..., v_m \in V$. Prove that

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\langle Tv_k, v_j\rangle \geq 0.$$

Solution:

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\langle Tv_k, v_j\rangle = \sum_{j=1}^{m}\left\langle T\left(\sum_{k=1}^{m} v_k\right), v_j\right\rangle = \left\langle T\left(\sum_{k=1}^{m} v_k\right), \sum_{j=1}^{m} v_j\right\rangle \geq 0.$$

Problem (Exercise 15): Suppose $T \in \mathcal{L}(V)$ is self adjoint. Prove that there exist positivbe operators $A, B \in \mathcal{L}(V)$ such that

$$T = A - B \text{ and } \sqrt{T^* T} = A + B \text{ and } AB = BA = 0.$$

Solution: We claim

$$\frac{\sqrt{T^* T} + T}{2} = A \text{ and } \frac{\sqrt{T^* T} - T}{2} = B$$

work. Since T is self adjoint, we work in an orthonormal eigenbasis. Note that $Te_i = \lambda_i e_i \Rightarrow \sqrt{T^2} e_i = |\lambda_i| e_i$. First we show A is positive, and B will follow similarly.

Note that A is self adjoint, so we just need to show it's positive. We have

$$\langle Te_i, e_i\rangle + \left\langle \sqrt{T^2} e_i, e_i\right\rangle = \lambda_i + |\lambda_i| \geq 0 \text{ and } \left\langle Te_i, e_j\right\rangle + \left\langle \sqrt{T^2} e_i, e_j\right\rangle = 0.$$

Thus $\langle Av, v\rangle \geq 0$.

Next we show $AB = 0$, and BA follows simlarly. We have

$$4AB = T^* T - T^2 - \sqrt{T^* T}T + T\sqrt{T^* T} = T\sqrt{T^2} - \sqrt{T^2}T.$$

Note that $\left(T\sqrt{T^2} - \sqrt{T^2}T\right)e_i = \lambda_i|\lambda_i| - \lambda_i|\lambda_i| = 0$, so $AB = 0$.

Problem (Exercise 16): Suppose T is a positive operator on V. Prove that

$$\text{null } \sqrt{T} = \text{null } T \text{ and range } \sqrt{T} = \text{range } T.$$

Solution: We prove the first equations first. Note that $\sqrt{T}v = 0 \Rightarrow Tv = 0$, so null $\sqrt{T} \subseteq$ null T. We also have $Tv = 0 \Rightarrow 0 = \langle Tv, v\rangle = \left\langle \sqrt{T}v, \sqrt{T}v \right\rangle = \left\| \sqrt{T}v \right\|^2$, so $v \in$ null $\sqrt{T}$, so null $T \subseteq$ null $\sqrt{T}$.

Note that $Tu = v \Rightarrow \sqrt{T}\left(\sqrt{T}u\right) = v$, so range $T \subseteq$ range $\sqrt{T}$. Then fundamental theorem of linear maps finishes.

Problem (Exercise 17): Suppose that $T \in \mathcal{L}(V)$ is a positive operator. Prove that there exists a polynomial p with real coefficients such that $\sqrt{T} = p(T)$.

Solution: We work in an orthonormal eigenbasis. We need $\sqrt{T}e_i = p(T)e_i \Rightarrow \sqrt{\lambda_i}e_i = p(\lambda_i)e_i \Rightarrow \sqrt{\lambda_i} = p(\lambda_i)$. We can use polynomial interpolation to gurantee the existence of such a polynomial.

Remark: This problem implies that $\sqrt{T}$ commutes with $q(T)$ for any polynomial q. In fact, for any $p \in \mathcal{P}(\mathbb{F})$ such that $p(T)$ is positive, $\sqrt{p(T)}$ will commute with $q(T)$.

Problem (Exericise 19): Show that the identity operator on $\mathbb{F}^2$ has infinitely many self adjoint square roots.

Solution: $T(x, y) = (ax + by, bx - ay)$ works as long as $a^2 + b^2 = 1$, which has infinitely many solutions.

Problem (Exercise 24): Suppose S and T are positive operators on V. Prove that

$$\text{null}(S + T) = \text{null } S \cap \text{null } T.$$

Solution: Suppose $v \in \text{null}(S + T)$. Then we have

$$\langle (S + T)v, v\rangle = 0 \Rightarrow \langle Sv, v\rangle + \langle Tv, v\rangle = 0 \Rightarrow \langle Sv, v\rangle = \langle Tv, v\rangle = 0 \Rightarrow v \in \text{null } S \cap \text{null } T.$$

If $v \in$ null $S \cap$ null T, then obviously $v \in \text{null}(S + T)$, so we're done.

7.4. Isometries, Unitary Operators, and Matrix Factorization

7.4.1. Isometries

Definition (isometry): 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.

Remark: Note that $Sv = 0 \Rightarrow 0 = \|v\| = \|Sv\|$, so an isometry is injective.

Proposition (characterization of isometries): Suppose $S \in \mathcal{L}(V, W)$. 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:

(a) S is an isometry.
(b) $S^* S = I$.
(c) $\langle Su, Sv \rangle = \langle u, v \rangle$ for all $u, v \in V$.
(d) $Se_1, \ldots, Se_n$ is an orthonormal list in W.
(e) The columns of $\mathcal{M}(S, (e_1, \ldots, e_n), (f_1, \ldots, f_m))$ form an orthonormal list in $\mathbb{F}^m$ with respect to the Euclidean inner product.

Proof: Suppose (a) holds. For $v \in V$ we have

$$\langle (I - S^* S)v, v \rangle = \langle v, v \rangle - \langle S^* Sv, v \rangle = \|v\|^2 - \langle Sv, Sv \rangle = \|v\|^2 - \|Sv\|^2 = 0.$$

Since $I - S^* S$ is self adjoint, we have $I - S^* S = 0$. Thus (b) holds.

Now suppose (b) holds. Then $\langle u, v \rangle = \langle S^* Su, v \rangle = \langle Su, Sv \rangle$. Thus (c) holds.

Now suppose (c) holds. We have $\langle Se_i, Se_i \rangle = \langle e_i, e_i \rangle = 1$ and $\langle Se_i, Se_j \rangle = \langle e_i, e_j \rangle = 0$. Thus (d) holds.

Now suppose (d) holds. Letting A be the matrix, for $k, r \in \{1, \ldots, n\}$ we have

$$\sum_{j=1}^{m} A_{j,k} \overline{A_{j,r}} = \left\langle \sum_{j=1}^{m} A_{j,k} f_j, \sum_{j=1}^{m} A_{j,r} f_j \right\rangle = \langle Se_k, Se_r \rangle = \begin{cases} 1 \text{ if } k = r. \\ 0 \text{ if } k \neq r. \end{cases}$$

Thus (e) holds.

Now suppose (e) holds. By the previous equation, we have $Se_1, \ldots, Se_n$ is an orthonormal list in W. Thus

$$v = \sum \langle v, e_i \rangle e_i \Rightarrow Sv = \sum \langle v, e_i \rangle Se_i \Rightarrow \|v\|^2 = \|Sv\|^2 = \sum |\langle v, e_i \rangle|^2.$$

Thus (a) holds. ∎

7.4.2. Unitary operators

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

Remark: The invertible condition can be dropped since as we remarked before, isometries are injective and we're working on finite dimensional vector spaces.

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

(a) S is a unitary operator.
(b) $S^*S = SS^* = I$.
(c) S is invertible and $S^{-1} = S^*$.
(d) $Se_1, \ldots, Se_n$ is an orthonormal basis in V.
(e) The rows of $\mathcal{M}(S, (e_1, \ldots, e_n))$ form an orthonormal basis in $\mathbb{F}^n$ with respect to the Euclidean inner product.
(f) S^* is a unitary operator.

Proof: Suppose (a) holds. Thus $S^*S = I$, since S is an isometry. Since S is unitary, we have $S^*S = I \Rightarrow S^* = S^{-1} \Rightarrow SS^* = SS^{-1} = I$. Thus (b) holds.

Suppose (b) holds. Then obviously (c) holds.

Suppose (c) holds. Then $S^*S = I$, so by the previous proposition, $Se_1, \ldots, Se_n$ is an orthonormal list. Since it has length $\dim V$, it's an orthonormal basis. Thus (d) holds.

Suppose (d) holds. The equivalence of (a) and (d) in the previous proposition shows that S is unitary. Since we already showed (a) $\Rightarrow$ (b), we have $SS^* = I$, so by the previous proposition, S^* is an isometry. Again by the previous proposition, the columns of $\mathcal{M}(S^*, (e_1, \ldots, e_n))$ form an orthonormal basis in $\mathbb{F}^n$. Since $\mathcal{M}(S)$ is the complex conjugate of $\mathcal{M}(S^*)$, the rows of $\mathcal{M}(S)$ form an orthonormal basis in $\mathbb{F}^n$. Thus (e) holds.

Now suppose (e) holds. The columns of $\mathcal{M}(S^*)$ form an orthonormal basis in $\mathbb{F}^n$, so by the previous proposition, S^* is an isometry. Thus (f) holds.

Now suppose (f) holds. Then we can change S to S^* in the proposition statement and conclude that (a) holds. ∎

Proposition: Suppose λ is an eigenvalue of a unitary operator. Then $|\lambda| = 1$.

Proof: Suppose $Sv = \lambda v$. Then $|\lambda v| = \|\lambda v\| = \|Sv\| = \|v\|$, so $|\lambda| = 1$. ∎

Proposition: Suppose $\mathbb{F} = \mathbb{C}$ and $S \in \mathcal{L}(V)$. Then the following are equivalent.

(a) S is a unitary operator.

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

Proof: Suppose (a) holds. Then, since $S^*S = SS^* = I$, S is normal, so by the complex spectral theorem, there's an orthonormal eigenbasis of V. Since every eigenvalue of S has absolute value 1, (b) holds.

Now suppose (b) holds. We have

$$\langle Se_j, Se_k \rangle = \langle \lambda_j e_j, \lambda_k e_k \rangle = \lambda_j \overline{\lambda_k} \langle e_j, e_k \rangle = \begin{cases} 1 \text{ if } j = j. \\ 0 \text{ if } j \neq k. \end{cases}$$

Thus by our characterization of unitary operators ((a) $\iff$ (d)), S is unitary. Thus (a) holds. ■

7.4.3. QR Factorization

Definition (unitary matrix): An n by n matrix is called *unitary* if its columns form an orthonormal list in $\mathbb{F}^n$.

Proposition (characterization of unitary matrices): Suppose Q is an n by n matrix. Then the following are equivalent:

(a) Q is a unitary matrix.

(b) The rows of Q form an orthonormal list in $\mathbb{F}^n$.

(c) $\|Qv\| = \|v\|$ for every $v \in \mathbb{F}^n$.

(d) $Q^*Q = QQ^* = I$, the n by n matrix with 1's on the diagonals and 0's everywhere else.

Proof: Each condition is equivalent to the underlying operator of Q being unitary, which immediately implies the other conditions. ■

Theorem (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 only positive numbers on its diagonal, and

$$A = QR.$$

Proof: Let $v_1, ..., v_n$ be the columns of A. Apply the Gram-Schmidt procedure on the list to obtain an orthonormal basis $e_1, ..., e_n$ with $\operatorname{span}(v_1, ..., v_k) = \operatorname{span}(e_1, ..., e_k)$ for each k. Let R be the n by n matrix with $R_{j,k} = \langle v_k, e_j \rangle$. If $j > k$, then e_j is orthogonal to $\operatorname{span}(e_1, ..., e_k)$, and so e_j is orthogonal to v_k. Thus for $j > k$, $\langle v_k, e_j \rangle = 0$, so R is upper triangular.

Let Q be the unitary matrix whose columns are $e_1, ..., e_n$. The kth column of QR equals a linaer combination of the columns of Q with the coefficients for the linear combination coming from the kth column of R. Thus the kth column of QR equals

$$\langle v, e_1 \rangle e_1 + \cdots + \langle v_k, e_k \rangle e_k = v_k.$$

Thus $A = QR$.

The equations defining the Gram-Schmidt procedure show that each v_k equals a positive multiple of e_k plus a linear combination of $e_1, ..., e_{k-1}$. Thus each $\langle v_k, e_k \rangle$ is a positive number, so all the entries on the diagonal of R are positive.

Now we show Q and R are unique. Suppose $A = \hat{Q}\hat{R}$, where $\hat{Q}$ is unitary and $\hat{R}$ is upper triangular with positive diagonal entries. Let $q_1, ..., q_n$ be the columns of $\hat{Q}$. If we think of matrix multiplication as we did earlier in the proof, each v_k is a linear combination of $q_1, ..., q_k$, with coefficients coming from the kth column of $\hat{R}$. Since R is upper triangular, we have $\operatorname{span}(v_1, ..., v_k) = \operatorname{span}(q_1, ..., q_k)$ and $\langle v_k, q_k \rangle > 0$ ($v_k = \sum_{i=0}^{k} R_{i,k} q_i$, and $R_{k,k}$ is positive). By exercise 10 in section 6.2, we have $q_k = e_k$ for each k. Thus $\hat{Q} = Q$, whcih implies $\hat{R} = R$, which completes uniqueness. ■

7.4.4. Cholesky Factorization

Proposition (positive invertible operator): A self adjoint operator $T \in \mathcal{L}(V)$ is a positive invertible operator if and only if $\langle Tv, v \rangle > 0$ for every nonzero $v \in V$.

Proof: First suppose T is a positive invertible operator. Then $Tv \neq 0$ for nonzero v. Thus $\langle Tv, v \rangle \neq 0$, so $\langle Tv, v \rangle > 0$.

Now suppose $\langle Tv, v \rangle > 0$ for all nonzero v. Then $Tv \neq 0$, so T is invertible. ■

Definition (positive definite): A matrix $B \in \mathbb{F}^{n,n}$ is called *positive definite* if $B^* = B$ and

$$\langle Bx, x \rangle > 0$$

for every nonzero $x \in \mathbb{F}^n$.

Theorem (Cholesky factorization): Suppose B is a positive definite matrix. Then there exists a unique upper triangular matrix R with only positive numbers on its diagonal such that

$$B = R^* R.$$

Proof: Because B is positive definite, there exists an invertible square matrix A of the same size as B such that $B = A^* A$ (A is invertible since it cannot be 0 on nonzero v, as otherwise B would not be invertible). Since A is invertible, the underlying operator has range equal to V, which means the column rank is $\dim V$, which means it has linearly independent columns. Thus we can write $A = QR$ where Q is unitary and R is upper triangular with only positive numbers on the diagonal. Then $A^* = R^* Q^*$. Thus

$$B = A^* A = R^* Q^* Q R = R^* R,$$

as desired.

To prove uniqueness, suppose S is upper triangular with only positive diagonal entries such that $B = S^* S$. S is invertible because B is invertible as noted earlier. Thus $BS^{-1} = S^*$. Letting A be the matrix from the first paragraph, we have

$$\begin{aligned}(AS^{-1})^*(AS^{-1}) &= (S^*)^{-1} A^* A S^{-1} \\ &= (S^*)^{-1} B S^{-1} \\ &= (S^*)^{-1} S^* \\ &= I.\end{aligned}$$

Thus AS^{-1} is unitary, and so $A = (AS^{-1})S$ is a QR factorization of A. By the uniqueness of QR factorization, we have $S = R$. ■

7.4.5. Problems

Problem (Exercise 3):
(a) Show that the product of two unitary operators on V is a unitary operator.
(b) Show that the inverse of a unitary operator on V is a unitary operator.

Solution:
(a) Suppose S and T are unitary. Then $(ST)^* ST = T^* S^* ST = T^{-1} S^{-1} ST = I$. Similarly, $ST(ST)^* = I$, so ST is unitary.
(b) We have $S^* = S^{-1}$, and since S^* is unitary, S^{-1} must be unitary.

Problem (Exercise 4): Suppose $\mathbb{F} = \mathbb{C}$ and $A, B \in \mathcal{L}(V)$ are self adjoint. Show that $A = iB$ is unitary if and only if $AB = BA$ and $A^2 + B^2 = I$.

Solution: Suppose $A + iB$ is unitary. We have $(A + iB)^* = A^* - iB^* = A - iB$, so

$$\begin{aligned}(A + iB)(A - iB) = A^2 - iAB + iBA + B^2 = I \\ (A - iB)(A + iB) = A^2 - iBA + iAB + B^2 = I.\end{aligned}$$

Adding the two equations and dividing by 2 yields $A^2 + B^2 = I$ and subtracting and dividing by $2i$ yields $BA - AB = 0$.

Now suppose $AB = BA$ and $A^2 + B^2 = I$. The calculations above then show that $A + iB$ is unitary.

Problem (Exercise 5): Suppose $S \in \mathcal{L}(V)$. Prove that the following are equivalent.
(a) S is a self adjoint unitary operator.
(b) $S = 2P - I$ for some orthogonal projection P on V.
(c) There exists a subspace U of V such that $Su = u$ for every $u \in U$ and $Sw = -w$ for every $w \in U^\perp$.

Solution: Suppose (a) holds. We have some orthonormal eigenbasis. Note that since S is self adjoint, the eigenvalues are all real, and since S is unitary, eigenvalues have absolute value 1. Thus, the set of eigenvalues is contained in $\{1, -1\}$. Let $U = \operatorname{span}_{j \in A}(e_j)$, where $A = \{i : Se_i = e_i\}$. For $k \notin A$ we then must have $Se_k = -e_k$. For $i \in A$ we have $Se_i = e_i = (2P_U - I)e_i$, and for $j \notin A$ we have $Se_j = -e_j = (2P_U - I)e_j$, so $S = 2P_U - I$, as desired.

Suppose (b) holds. Let U be the subspace that P projects onto. Then clearly $Su = u$ for $u \in U$ and $Sw = -w$ for $w \in U^\perp$.

Suppose (c) holds. For $u \in U, w \in U^\perp$ we have $\|S(u + w)\| = \|u - w\| = \|u\| + \|w\| = \|u + w\|$, so S is unitary. For $u, u' \in U, w, w \in U^\perp$ we also have

$$\begin{aligned}\langle S(u+w), u'+w'\rangle &= \langle u - w, u' + w'\rangle = \langle u, u'\rangle + \langle u, w'\rangle - \langle w, u'\rangle - \langle w, w'\rangle = \langle u, u'\rangle - \langle w, w'\rangle \\ &= \langle u, u'\rangle - \langle u, w'\rangle + \langle w, u'\rangle - \langle w, w'\rangle = \langle u + w, u' - w'\rangle = \langle u + w, S(u' + w')\rangle.\end{aligned}$$

Thus, S is self adjoint, as desired.

Problem (Exercise 6): Suppose T_1, T_2 are both normal operators on $\mathbb{F}^3$ with $2, 5, 7$ eigenvalues. Prove there exists a unitary operator $S \in \mathcal{L}(\mathbb{F}^3)$ such that $T_1 = S^* T_2 S$.

Solution: Note that the eigenvalues are real, so both operators are self adjoint. Let e_1, e_2, e_3 be an orthonormal eigenbasis of T_1 and f_1, f_2, f_3 be an orthonormal eigenbasis of T_2, where e_i and f_i have the same eigenvalue. Let $S(e_i) = f_i$. We have

$$\left\langle S\left(\sum a_i e_i\right), S\left(\sum b_i e_i\right)\right\rangle = \left\langle \sum a_i f_i, \sum b_i f_i\right\rangle = \sum a_i \overline{b_i} = \left\langle \sum a_i e_i, \sum b_i e_i\right\rangle,$$

so S is unitary. Thus $S^* = S^{-1}$. Then we have

$$S^{-1} T_2 S e_i = S^* T_2 f_i = S^* \lambda_i f_i = \lambda_i e_i = T_1 e_i,$$

so we're done.

Remark: In general, when changing the basis of a normal/self-adjoint operator, the change of basis operator/matrix will be unitary.

Problem (Exercise 8): Prove or give a counterexample: if $S \in \mathcal{L}(V)$ and there exists an orthonormal basis $e_1, ..., e_n$ of V such that $\|Se_k\| = 1$ for each e_k, then S is a unitary operator.

Solution: The map $T(x, y) = \left(\frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y, \frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y\right)$ satsifies the property and is not unitary.

Problem (Exercise 10): Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$ is a self adjoint operator such that $\|Tv\| \leq \|v\|$ for all $v \in V$.

(a) Show that $I - T^2$ is a positive operator.
(b) Show that $T + i\sqrt{I - T^2}$ is a unitary operator.

Solution:
(a) We have $\langle (I - T^2)v, v\rangle = \langle v, v\rangle - \langle T^2 v, v\rangle = \|v\|^2 - \langle Tv, Tv\rangle \geq 0$, so $I - T^2$ is positive.
(b) We have $\left(T + i\sqrt{I - T^2}\right)\left(T - i\sqrt{I - T^2}\right) = T^2 - iT\sqrt{I - T^2} + iT\sqrt{I - T^2} + I - T^2 = I$, and we also have the commuted version, so $T + i\sqrt{I - T^2}$ is unitary.

Problem (Exercise 16): Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$ is self adjoint. Prove that $(T + iI)(T - iI)^{-1}$ is a unitary operator and 1 is not an eigenvalue of this operator.

Solution: Let $e_1, ..., e_n$ be an orthonormal eigenbasis of T. Note that since T is self adjoint, all the eigenvalue of T are real, so $T + iI$ is invertible. Let S be the operator. We have $(T - iI)^{-1}e_j = \frac{1}{\lambda_j - i}e_j$, so $Se_j = \frac{\lambda_j + i}{\lambda_j - i}e_j$. Note that $\|Se_j\| = 1$, so $Se_1, ..., Se_n$ is an orthonormal basis of V, which means S is unitary. Note that the eigenvalues of S are $\frac{\lambda_j + i}{\lambda_j - i}$, which can never be 1.

7.5. Singular Value decomposition

7.5.1. Singular Values

Proposition (properties of T^*T): Suppose $T \in \mathcal{L}(V, W)$. Then
(a) T^*T is a positive operator on V
(b) null $T^*T =$ null T
(c) range $T^*T =$ range T^*
(d) $\dim \operatorname{range} T = \dim \operatorname{range} T^* = \dim \operatorname{range} T^*T$

Proof:
(a) Clearly T^*T is self adjoint. We also have $\langle T^*Tv, v\rangle = \langle Tv, Tv\rangle = \|Tv\|^2 \geq 0$, so T^*T is positive.
(b) First suppose $v \in \text{null } T^*T$. Then

$$0 = \langle T^*Tv, v\rangle = \langle Tv, Tv\rangle = \|Tv\|^2,$$

so $Tv = 0 \Rightarrow v \in \text{null } T$. Clearly if $v \in \text{null } T$ then $T^*Tv = 0$, so the equality holds/
(c) Since T^*T is self adjoint, we have

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

(d) We have

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

The equalities then follow from (c).

■

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

Proposition (role of positive singular values): Supppose $T \in \mathcal{L}(V, W)$. Then
(a) T is injective $\Longleftrightarrow$ 0 is not a singular value of T.
(b) The number of positive singular values of T equals $\dim \text{range } T$.
(c) T is surjective $\Longleftrightarrow$ number of positive singular values of T equals $\dim W$.

Proof: T is injective if and only if null $T^*T = \{0\}$ (by the previous proposition), which happens if and only if 0 is not an eigenvalue of T^*T, which happens if and only if T is not a singular value of 0, so (a) is true.

By applying the spectral theorem to T^*T, we have that $\dim \operatorname{range} T$ equals the number of positive eigenvalues of T^*T, counting repetitions. Thus the previous proposition implies that $\dim \operatorname{range} T$ equals the number of positive singular values of T, so (b) holds.

(b) and the definition of surjectivity show that (c) holds. ■

Proposition: Suppose that $S \in \mathcal{L}(V, W)$. Then

$$S \text{ is an isomtery} \Longleftrightarrow \text{all singular values of } S \text{ equal } 1.$$

Proof:

$$\begin{aligned} S \text{ is an isometry} &\Longleftrightarrow S^*S = I \\ &\Longleftrightarrow \text{all eigenvalues of } S^*S \text{ are } 1 \\ &\Longleftrightarrow \text{all singular values of } S \text{ are } 1. \end{aligned}$$

■

7.5.2. SVD for Linear Maps and for Matrices

Theorem (singular value decomposition): Suppose $T \in \mathcal{L}(V, W)$ and the postiive 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 $v \in V$.

Proof: Let $s_1, \ldots, s_n$ be the singular values of T. Since T^*T is positive, by the spectral theorem we have an orthonormal basis $e_1, \ldots, e_n$ of V with

$$T^*Te_k = s_k^2 e_k.$$

Let $f_k = \frac{Te_k}{s_k}$. We have

$$\langle f_j, f_k \rangle = \frac{1}{s_j s_k} \langle Te_j, Te_k \rangle = \frac{1}{s_j s_k} \langle e_j, T^*Te_k \rangle = \frac{s_k}{s_j} \langle e_j, e_k \rangle = \begin{cases} 1 \text{ if } k = j, \\ 0 \text{ if } k \neq j. \end{cases}$$

Thus $f_1, \ldots, f_m$ is an orthonormal list in W.

Note that if $k > m$, then $s_k = 0 \Rightarrow T^*Te_k = 0 \Rightarrow Te_k = 0$.

For $v \in V$ we have

$$\begin{aligned} Tv &= T(\langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n) \\ &= \langle v, e_1 \rangle Te_1 + \cdots + \langle v, e_m \rangle Te_m \\ &= s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m, \end{aligned}$$

as desired. ■

Remark: If we consider $\mathcal{M}(T, (e_1, ..., e_{\dim V}), (f_1, ..., f_{\dim W}))$, where the bases come from SVD, then the matrix is a diagonal matrix (for non square matrices that means that only $A_{i,i}$ can be nonzero).

Proposition: Suppose $T \in \mathcal{L}(V, W)$ and the positive singular values of T are $s_1, ..., s_m$. Suppose $e_1, ..., e_m$ and $f_1, ..., f_m$ are orthonormal lists in V and 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 $v \in V$. then

$$T^* w = s_1 \langle w, f_1 \rangle e_1 + \cdots + s_m \langle w, f_m \rangle e_m$$

and

$$T^\dagger w = \frac{\langle w, f_1 \rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_m \rangle}{s_m} e_m$$

for every $w \in W$.

Proof: If $v \in V$ and $w \in V$ then

$$\begin{aligned} \langle Tv, w \rangle &= \langle s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m, w \rangle \\ &= s_1 \langle v, e_1 \rangle \langle f_1, w \rangle + \cdots + s_m \langle v, e_m \rangle \langle f_m, w \rangle \\ &= \langle v, s_1 \langle w, f_1 \rangle e_1 + \cdots + s_m \langle w, f_m \rangle e_m \rangle, \end{aligned}$$

which proves the first equation.

Now suppose $w \in W$ and let

$$v = \frac{\langle w, f_1 \rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_m \rangle}{s_m} e_m.$$

Applying T yields

$$\begin{aligned} Tv &= \frac{\langle w, f_1 \rangle}{s_1} Te_1 + \cdots + \frac{\langle w, f_m \rangle}{s_m} Te_m \\ &= \langle w, f_1 \rangle f_1 + \cdots + \langle w, f_m \rangle f_m \\ &= P_{\text{range } T} w, \end{aligned}$$

where the last line holds because $f_1, ..., f_m$ spans range T and thus is an orthonormal basis of range T. Extend $e_1, ..., e_m$ to an orthonormal basis $e_1, ..., e_n$. For $k > m$ we have $Tv = 0$ according to the definition of T in the problem. Since T is 0 only on the extension, null $T = \operatorname{span}(e_{m+1}, ..., e_n)$. Thus $v \in (\text{null } T)^\perp$. Thus we can restrict T to $(\text{null } T)^\perp$. Thus we have $\left(T|_{(\text{null } T)^\perp}\right) v = P_{\text{range } T} w$. By the definition of the pseudoinverse we have $T^\dagger w = v$, proving the second equation. ■

Theorem (matrix SVD): Suppose A is an M-by-n matrix of rank $m \geq 1$. Then there exist an M-by-m matrix B with orthonormal columns, an m-by-n diagonal matrix D with positive numbers on the diagonal, and an n-by-m matrix C with orthonormal columns such that

$$A = BDC^*.$$

Proof: Let $T : \mathbb{F}^n \to \mathbb{F}^M$ be the linear map whose matrix with respect to the standard basis equals A. Then $\dim \operatorname{range} T = m$. let

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \dots + s_m \langle v, e_m \rangle f_m$$

be the SVD of T. Let

$$\begin{aligned} B &= \text{the } M \text{ by } m \text{ matrix whose columns are } f_1, \dots, f_m \\ D &= \text{the } m \text{ by } m \text{ matrix whose diagonal entries are } s_1, \dots, s_m \\ C &= \text{the } n \text{ by } m \text{ matrix whose columns are } e_1, \dots, e_m. \end{aligned}$$

Let $u_1, \dots, u_m$ denote the standard basis of $\mathbb{F}^m$. If $k \in \{1, \dots, m\}$ Then

$$(AC - BD)u_k = Ae_k - B(s_k u_k) = s_k f_k - s_k f_k = 0,$$

so $AC = BD$.

Multiply both sides by C^* (the conjugate transpose of C) to get

$$ACC^* = BDC^*.$$

Note that the rows are the complex conjugates of $e_1, \dots, e_m$, so for $k \in \{1, \dots, m\}$ we have $C^* e_k = u_k$. This $CC^* e_k = e_k \Rightarrow ACC^* v = Av$ for all $v \in \operatorname{span}(e_1, \dots, e_m)$.

If $v \in \operatorname{span}(e_1, \dots, e_m)^\perp$, then $Av = 0$ and $C^* v = 0$. This $ACC^* = A$, so we have $A = BDC^*$, as desired. ■

7.5.3. Problems

Problem (Exercise 1): Suppose $T \in \mathcal{L}(V, W)$. Show that $T = 0$ if and only if all singular values of T are 0.

Solution: Suppose $T = 0$. Then $T^*T = 0$, so all singular values of T are 0.

Now suppose all singular values of T are 0. This implies that $T^*T = 0$, since on an orthonormal eigenbasis, T^*T has 0 as its only eigenvalue. Since null $T^*T = V =$ null T, we have $T = 0$.

Problem (Exercise 2): Suppose $T \in \mathcal{L}(V, W)$ and $s > 0$. Prove that s is a singular value of T if and only if there exist nonzero vectors $v \in V$ and $w \in W$ such that

$$Tv = sw \text{ and } T^*w = sv.$$

Solution: Suppose s is a singular value, so $T^*Tv = s^2v$ for some v. Let $w = \frac{Tv}{s}$. Then w and v satisfy the equations.

Now suppose v, w satisfy the equations. Applying T^* to the first equations yields $T^*Tv = sT^*w = s^2v$, so s is a singular value of T.

Remark: This implies that s is a singular value of T if and only if it's also a singular value of T^*.

Problem (Exercise 3): Give an example of $T \in \mathcal{L}(\mathbb{C}^2)$ such that 0 is the only eigenvalue of T and the singular values of T are $5, 0$.

Solution: Let $T(x, y) = (5y, 0)$. Then $T^{*(x,y)} = (0, 5x)$, so $T^*T(x, y) = (25x, 0)$. This has eigenvalues $25, 0$, so the singulat values of T are $5, 0$.

Problem (Exercise 5): Suppose $T \in \mathcal{L}(\mathbb{C}^2)$ is defined by $T(x, y) = (-4y, x)$. Find the singular values of T.

Solution: Note that $T^*(x, y) = (y, -4x)$, so $T^*T(x, y) = (x, 16y)$. This has eigenvalues $16, 1$, so the singular values of T are $4, 1$.

Problem (Exercise 6): Find the singular values of the differentiation operator $D \in \mathcal{L}(\mathcal{P}_2(\mathbb{R}))$ defined by $Dp = p'$, where $\langle p, q \rangle = \int_{-1}^{1} pq$.

Solution: We have that $\sqrt{\frac{1}{2}}, \sqrt{\frac{3}{2}}x, \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right)$ is an orthonormal basis of $\mathcal{P}_2(\mathbb{R})$. The matrix of D with respect to this basis is

$$\begin{pmatrix} 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{15} \\ 0 & 0 & 0 \end{pmatrix}.$$

Since $\mathcal{M}(D)$ is with respect to the orthonormal basis, we have

$$\mathcal{M}(D^*) = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{3} & 0 & 0 \\ 0 & \sqrt{15} & 0 \end{pmatrix}.$$

Thus

$$D^*D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 15 \end{pmatrix}.$$

Thus the eigenvalues of D^*D are $0, 3, 15$, so the singular values of D are $0, \sqrt{3}, \sqrt{15}$.

Problem (Exercise 7): Suppose that $T \in \mathcal{L}(V)$ is self adjoint or that $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$ is normal. Let $\lambda_1, ..., \lambda_n$ be the eigenvalues of T, each included in the list as many times as the dimension of the corresponding eigenspace. Show that the singular values of T are $|\lambda_1|, ..., |\lambda_n|$, after these numbers have been sorted into decreasing order.

Solution: If T is self adjoint, then $T^*Te_i = T^2e_i = \lambda_i^2 e_i$, so $|\lambda_i|$ is a singular value.

If T is normal and $Te_i = \lambda_i e_i$, then $T^*e_i = \overline{\lambda_i}e_i$, so $T^*Te_i = |\lambda_i|^2 e_i$, which means $|\lambda_i|$ is a singular value of T.

Problem (Exercise 9): Suppose $T \in \mathcal{L}(V, W)$. Show that T and T^* have the same positive singular values.

Solution: See remark in exercise 2.

Problem (Exercise 10): Suppose $T \in \mathcal{L}(V, W)$ has singular values $s_1, ..., s_n$. Prove that if T is an invertible linear map, then T^{-1} has singular values

$$\frac{1}{s_n}, ..., \frac{1}{s_1}.$$

Solution: We have

$$Tv = \sum s_i \langle v, e_i \rangle f_i \text{ and } T^*w = \sum s_i \langle w, f_i \rangle e_i.$$

Thus $Te_i = s_i f_i \Rightarrow \frac{e_i}{s_i} = T^{-1} f_i$ and $T^* f_i = s_i e_i \Rightarrow \frac{f_i}{s_i} = \left(T^{-1}\right)^* e_i$. thus

$$\left(T^{-1}\right)^* T^{-1} f_i = \frac{1}{s_i^2} f_i,$$

so $\frac{1}{s_i}$ is a singular value of T^{-1}.

Problem (Exercise 11): Suppose that $T \in \mathcal{L}(V, W)$ and $v_1, ..., v_n$ is an orthonormal basis of V. Let $s_1, ..., s_n$ denote the singular values of T.
(a) Prove that $\|Tv_1\|^2 + \cdots + \|Tv_n\|^2 = s_1^2 + \cdots + s_n^2$.
(b) Prove that if $W = V$ and T is a positive operator, then

$$\langle Tv_1, v_1 \rangle + \cdots + \langle Tv_n, v_n \rangle = s_1 + \cdots + s_n.$$

Solution: First (a). Let $Tv = \sum_{i=1}^{m} s_i \langle v, e_i \rangle f_i$ be the SVD of T. Then

$$\begin{aligned}
\sum_{j=1}^{n} \langle Tv_j, Tv_j \rangle &= \sum_{j=1}^{n} \left\langle \sum_{i=1}^{m} s_i \langle v_j, e_i \rangle f_i, \sum_{i=1}^{m} s_i \langle v_j, e_i \rangle f_i \right\rangle \\
&= \sum_{j=1}^{n} \sum_{i=1}^{m} \left| s_i \langle v_j, e_i \rangle \right|^2 \\
&= \sum_{i=1}^{m} s_i^2 \sum_{j=1}^{n} \left| \langle v_j, e_i \rangle \right|^2 \\
&= \sum_{i=1}^{m} s_i^2 \|e_i\|^2 = \sum_{i=1}^{m} s_i^2 = \sum_{i=1}^{n} s_i^2.
\end{aligned}$$

Now (b). Note that $T^* T e_i = T^2 e_i = s_i^2 e_i$ for some orthonormal eigenbasis $e_1, ..., e_n$, so s_i is a singular value of T. Taking the square root of T^2 implies $Te_i = s_i e_i$. Thus $\sqrt{T}\sqrt{T} e_i = \sqrt{T}^* \sqrt{T} e_i = s_i e_i$, so $\sqrt{s_i}$ is a singular value of $\sqrt{T}$. Plugging in $\sqrt{T}$ and its singular values into (a) yields

$$\sum_{i=1}^{n} \left\langle \sqrt{T} v_i, \sqrt{T} v_i \right\rangle = \sum_{i=1}^{n} \langle Tv_i, v_i \rangle = \sum_{i=1}^{n} s_i.$$

Problem (Exercise 14): Suppose $T \in \mathcal{L}(V, W)$. Let s_n denote the smallest singular value of T. Prove that $s_n \|v\| \leq \|Tv\|$ for every $v \in V$.

Solution: If $s_n = 0$, then the inequality clearly holds, so we assume $s_n > 0$. Thus the $e_1, ..., e_n$ produced in the SVD of T is an orthonormal basis of V. We have $Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n$. Thus

$$\begin{aligned}
\|Tv\|^2 &= \sum |s_i \langle v, e_i \rangle|^2 \\
&= \sum s_i^2 |\langle v, e_i \rangle|^2 \\
&\geq s_n^2 \sum |\langle v, e_i \rangle|^2 = s_n^2 \|v\|^2,
\end{aligned}$$

as desired.

Remark: We also have $\|Tv\| \leq s_1 \|v\|$, where s_1 is the largest singular value, and this follows by basically the same argument.

Problem (Exercise 15): Suppose $T \in \mathcal{L}(V)$ and $s_1 \geq \ldots \geq s_n$ are the singular values of T. Prove that is λ is an eigenvalue of T, then $s_1 \geq |\lambda| \geq s_n$.

Solution: By the previous problem and remark, we have

$$s_n\|v\| \leq \|Tv\| \leq s_1\|v\|.$$

For eigenvectors $Tv = \lambda v$, we have

$$s_n\|v\| \leq \|\lambda v\| = |\lambda|\|v\| \leq s_1\|v\| \Rightarrow s_n \leq |\lambda| \leq s_1.$$

Problem (Exercise 16): Suppose $T \in \mathcal{L}(V, W)$. Prove that $(T^*)^\dagger = \left(T^\dagger\right)^*$.

Solution: We have

$$T^* w = \sum s_i \langle w, f_i \rangle e_i \text{ and } T^\dagger w = \sum \frac{\langle w, f_i \rangle}{s_i} e_i.$$

Applying $\dagger$ to the first equation and $*$ to the second yields

$$(T^*)^\dagger v = \sum \frac{1}{s_i} \langle v, e_i \rangle f_i = \left(T^\dagger\right)^*.$$

Problem (Exercise 17): Suppose $T \in \mathcal{L}(V)$. Prove that T is self adjoint if and only if $T^\dagger$ is self adjoint.

Solution: If T is self adjoint, then $\left(T^\dagger\right)^* = (T^*)^\dagger = T^\dagger$, so $T^\dagger$ is self adjoint. If $T^\dagger$ is self adjoint, then $(T^*)^\dagger = \left(T^\dagger\right)^* = T^\dagger$. Taking the pseudoinverse of both sides yields the desired result.

7.6. Consequences of Singular Value Decomposition

7.6.1. Norms of Linear Maps

Proposition (upper bound for $\|Tv\|$): Suppose $T \in \mathcal{L}(V, W)$. Let s_1 be the largest singular value of T. Then

$$\|Tv\| \leq s_1 \|v\|$$

for all $v \in V$.

Proof: Let $Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$ be the SVD of T. thus

$$\begin{aligned} \|Tv\|^2 &= s_1^2 |\langle v, e_1 \rangle|^2 + \cdots + s_m^2 |\langle v, e_m \rangle|^2 \\ &\leq s_1^2 \left(|\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_m \rangle|^2 \right) \\ &\leq s_1^2 \|v\|^2, \end{aligned}$$

where the last line follows from Bessel's inequality. Taking square roots yields the desired result. ■

Remark: This result shows that $\|Tv\| \leq s_1$ for all $\|v\| \leq 1$, and since $\|Te_1\| = \|s_1 f_1\| = s_1$, we have

$$\max\{\|Tv\| : v \in V \text{ and } \|v\| \leq 1\} = s_1,$$

so the following definition makes sense.

Definition (norm of a linear map): Suppose $T \in \mathcal{L}(V, W)$. Then the *norm* of T, denoted by $\|T\|$, is defined by

$$\|T\| = \max\{\|Tv\| : v \in V \text{ and } \|v\| \leq 1\}.$$

Proposition: Suppose $T \in \mathcal{L}(V, W)$. Then
(a) $\|T\| \geq 0$.
(b) $\|T\| = 0 \iff T = 0$.
(c) $\|\lambda T\| = |\lambda| \|T\|$ for all $\lambda \in \mathbb{F}$.
(d) $\|S + T\| \leq \|S\| + \|T\|$ for all $S \in \mathcal{L}(V, W)$.

Proof:
(a) Since $\|Tv\| \geq 0$, by definition $\|T\| \geq 0$.

(b) Suppose $\|T\| = 0$. Then $Tv = 0$ for all $\|v\| \leq 1$. We then have $Tu = \|u\| T\left(\frac{u}{\|u\|}\right) = 0$, since $\|u/\|u\|\| = 1$. Thus $T = 0$. Conversely, if $T = 0$, then $Tv = 0$, so $\|T\| = 0$.
(c) We have

$$\begin{aligned} \|\lambda T\| &= \max\{\|\lambda Tv\| : v \in V \text{ and } \|v\| \leq 1\} \\ &= |\lambda| \max\{\|Tv\| : v \in V \text{ and } \|v\| \leq 1\} \\ &= |\lambda| \|T\|. \end{aligned}$$

(d) Suppose $S \in \mathcal{L}(V, W)$. By definition of $\|S + T\|$, there exists $\|v\| \leq 1$ such that $\|S + T\| = \|(S + T)v\|$. We now have

$$\|S + T\| = \|(S + T)v\| = \|Sv + Tv\| \leq \|Sv\| + \|Tv\| \leq \|S\| + \|T\|,$$

as desired.

■

Proposition (alternative formulas for $\|T\|$): Suppoes $T \in \mathcal{L}(V, W)$. Then
(a) $\|T\| =$ the largest singular value of T.
(b) $\|T\| = \max\{\|Tv\| : v \in V \text{ and } \|v\| = 1\}$.
(c) $\|T\| =$ the smallest number c such that $\|Tv\| \leq c\|v\|$ for all $v \in V$.

Proof:
(a) See the earlier remark.
(b) Let $0 < \|v\| \leq 1$ and $u = \frac{v}{\|v\|}$. Then $\|u\| = 1$ and $\|Tu\| = \left\|T\left(\frac{v}{\|v\|}\right)\right\| = \frac{\|Tv\|}{\|v\|} \geq \|Tv\|$. Thus when finding the max of the set in the definition of the norm of a map, we can restrict to $\|v\| = 1$.
(c) For nonzero v, by definition we have

$$\left\|T\left(\frac{v}{\|v\|}\right)\right\| \leq \|T\| \Rightarrow \|Tv\| \leq \|T\|\|v\|.$$

Now suppose $c \geq 0$ and $\|Tv\| \leq c\|v\|$ for all $v \in V$. Then

$$\|Tv\| \leq c.$$

Taking the max of the left side over all $\|v\| \leq 1$ shows $\|T\| \leq c$, so $\|T\|$ is indeed the smallest c such that $\|Tv\| \leq c\|v\|$ for all $v \in V$.

■

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

$$\|Tv\| \leq \|T\|\|v\|$$

for all $v \in V$.

Proof: See proof of (c) in the previous result. ■

Proposition (norm of adjoint): Suppose $T \in \mathcal{L}(V, W)$. Then $\|T^*\| = \|T\|$.

Proof: For $w \in W$ we have

$$\|T^* w\|^2 = \langle T^* w, T^* w\rangle = \langle TT^* w, w\rangle \leq \|TT^* w\|\|w\| \leq \|T\|\|T^* w\|\|w\|.$$

Thus $\|T^* w\| \leq \|T\|\|w\|$. By (c) in the alternative formulas proposition, we have $\|T^*\| \leq \|T\|$. Replacing T with T^* then yields $\|T\| \leq \|T^*\|$, so $\|T\| = \|T^*\|$.

Remark: Alternatively, we can use the fact that a map and its adjoint have the same singular values.

■

7.6.2. Approximation by Linear Maps with Lower-Dimensional Range

Proposition: Suppose $T \in \mathcal{L}(V, W)$ and $s_1 \geq \cdots \geq s_m$ are the positive singular values of T. Suppose $1 \leq k < m$. then

$$\min\{\|T - S\| : S \in \mathcal{L}(V, W) \text{ and } \dim \operatorname{range} S \leq k\} = s_{k+1}.$$

Furthermore, if

$$Tv = s_1 \langle v, e_1\rangle f_1 + \cdots + s_m \langle v, e_m\rangle f_m$$

is a singular value decomposition of T and $T_k \in \mathcal{L}(V, W)$ is defined by

$$T_k v = s_1 \langle v, e_1\rangle f_1 + \cdots + s_k \langle v, e_k\rangle f_k$$

for each $v \in V$, then $\dim \operatorname{range} T_k = k$ and $\|T - T_k\| = s_{k+1}$.

Proof: If $v \in V$ Then

$$\begin{aligned}
\|(T - T_k)v\|^2 &= \left\|s_{k+1}\langle v, e_{k+1}\rangle f_{k+1} + \cdots + s_m \langle v, e_m\rangle f_m\right\|^2 \\
&= s_{k+1}^2 \left|\langle v, e_{k+1}\rangle\right|^2 + \cdots + s_m^2 |\langle v, e_m\rangle|^2 \\
&\leq s_{k+1}^2 \left(\left|\langle v, e_{k+1}\rangle\right|^2 + \cdots + |\langle v, e_m\rangle|^2\right) \\
&\leq s_{k+1}^2 \|v\|^2.
\end{aligned}$$

Thus $\|T - T_k\| \leq s_{k+1}$, and since $(T - T_k)e_{k+1} = s_{k+1} f_{k+1}$, we have $\|T - T_k\| = s_{k+1}$.

Suppose $S \in \mathcal{L}(V, W)$ and $\dim \operatorname{range} S \leq k$. Then $Se_1, \ldots, Se_{k+1}$ is linear dependent, so we have $a_1, \ldots, a_{k+1} \in \mathbb{F}$ not all 0 such that

$$a_1 S e_1 + \cdots + a_{k+1} S e_{k+1} = 0.$$

We also have $a_1 e_1 + \cdots + a_{k+1} e_{k+1} \neq 0$ since not all the a_i's are 0. We have

$$\begin{aligned}\left\|(T-S)(a_1e_1+\cdots+a_{k+1}e_{k+1})\right\|^2 &= \left\|T(a_1e_1+\cdots+a_{k+1}e_{k+1})\right\|^2 \\ &= \left\|s_1a_1f_1+\cdots+s_{k+1}a_{k+1}f_{k+1}\right\|^2 \\ &= s_1^2|a_1|^2+\cdots+s_{k+1}^2|a_{k+1}|^2 \\ &\geq s_{k+1}^2\left(|a_1|^2+\cdots+|a_{k+1}|^2\right) \\ &= s_{k+1}^2\left\|a_1e_1+\cdots+a_{k+1}e_{k+1}\right\|^2.\end{aligned}$$

Because $a_1e_1+\cdots+a_{k+1}e_{k+1}\neq 0$, the inequality above implies

$$\|T-S\|\geq s_{k+1}.$$

Thus $S=T_k$ minimizes $\|T-S\|$ amonf $S\in\mathcal{L}(V,W)$ with $\dim\operatorname{range} S\leq k$. ■

7.6.3. Polar Decomposition

Proposition (polar decomposition): Suppose $T\in\mathcal{L}(V)$. Then there exists a unitary operator $S\in\mathcal{L}(V)$ such that

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

Proof: Let

$$Tv=s_1\langle v,e_1\rangle f_1+\cdots+s_m\langle v,e_m\rangle f_m$$

be the SVD of T. Extened $e_1,\ldots,e_m$ and $f_1,\ldots,f_m$ to orthonormal bases of V (have length n).

Define $S\in\mathcal{L}(V)$ by

$$Sv=\langle v,e_1\rangle f_1+\cdots+\langle v,e_n\rangle f_n$$

for each $v\in V$. Then

$$\begin{aligned}\|Sv\|^2 &= \|\langle v,e_1\rangle f_1+\cdots+\langle v,e_n\rangle f_n\|^2 \\ &= |\langle v,e_1\rangle|^2+\cdots+|\langle v,e_n\rangle|^2 \\ &= \|v\|^2,\end{aligned}$$

so S is unitary.

Applying T^* to both sides of the very first equation and using the formula for the SVD of T^* yields

$$T^*Tv=s_1^2\langle v,e_1\rangle e_1+\cdots+s_m^2\langle v,e_m\rangle e_m$$

for every $v\in V$. thus

$$\sqrt{T^*T}v=s_1\langle v,e_1\rangle e_1+\cdots+s_m\langle v,e_m\rangle e_m.$$

Then we have

$$\begin{aligned} S\sqrt{T^*T}v &= S(s_1\langle v, e_1\rangle e_1 + \cdots + s_m\langle v, e_m\rangle e_m) \\ &= s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m \\ &= Tv, \end{aligned}$$

so we're done. ■

7.6.4. Operators Applies to Ellipsoids and Parallelepipeds

Definition (ball): The *ball* in V of radius 1 centered at 0, denote B, is defined by

$$B = \{v \in V : \|v\| < 1\}.$$

Definition (ellipsoid): Suppose $f_1, ..., f_n$ is an orthonormal basis of V and $s_1, ..., s_n$ are positive numbers. The *ellipsoid* $E(s_1 f_1, ..., s_n f_n)$ with *principal axes* $s_1 f_1, ..., s_n f_n$ is defined By

$$E(s_1 f_1, ..., s_n f_n) = \left\{ v \in V : \frac{|\langle v, f_1\rangle|^2}{s_1^2} + \cdots + \frac{|\langle v, f_n\rangle|^2}{s_n^2} < 1 \right\}.$$

Definition ($T(\Omega)$): For a function T defined on V and $\Omega \subseteq V$, define $T(\Omega)$ by

$$T(\Omega) = \{Tv : v \in \Omega\}.$$

Proposition (invertible operator takes ball to ellipsoid): Suppose $T \in \mathcal{L}(V)$ is invertible. Then T maps the ball B in V onto an ellipsoid in V.

Solution: Suppose

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

is the SVD of T. Since T is invertible, all the singular values are positive. If $v \in B$, then

$$\frac{|\langle Tv, f_1\rangle|^2}{s_1^2} + \cdots + \frac{|\langle Tv, f_n\rangle|^2}{s_n^2} = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2 = \|v\|^2 < 1,$$

so $Tv \in E(s_1 f_1, ..., s_n f_n)$, which implies $T(B) \subseteq E(s_1 f_1, ..., s_n f_n)$.

Now suppose $w \in E(s_1 f_1, ..., s_n f_n)$. let

$$v = \frac{\langle w, f_1\rangle}{s_1} e_1 + \cdots + \frac{\langle w, f_n\rangle}{s_n} e_n.$$

Since the norm of the left side is less than 1 (by definition of w being in the ellipsoid), we have $\|v\| < 1$, and we also have $Tv = \langle w, f_1 \rangle f_1 + \cdots + \langle w, f_n \rangle f_n = w$. Thus $E(s_1 f_1, \ldots, s_n f_n) \subseteq T(B)$.

Proposition (invertible operator takes ellipsoids to ellipsoids): Suppose $T \in \mathcal{L}(V)$ is invertible and E is an ellipsoid in V. Then $T(E)$ is an ellipsoid in V.

Proof: By definition there exist orthonormal basis $f_1, \ldots, f_n$ and positive numbers $s_1, \ldots, s_n$ such that $E = E(s_1 f_1, \ldots, s_n f_n)$. Define $S \in \mathcal{L}(V)$ by

$$S(a_1 f_1 + \cdots + a_n f_n) = a_1 s_1 f_1 + \cdots + a_n s_n f_n.$$

Note that S is invertible. Note also that $S(B) = E$, where B is the unit ball on V. Then

$$T(E) = T(S(B)) = (TS)(B).$$

Since T and S are invertible, TS is invertible, which means $(TS)(B) = T(E)$ is indeed an ellipsoid. ∎

Definition (parallelepiped): Suppose $v_1, \ldots, v_n$ is a basis of V. Let

$$P(v_1, \ldots, v_n) = \{a_1 v_1 + \cdots + a_n v_n : a_1, \ldots, a_n \in (0, 1)\}.$$

A *parallelepiped* is a set of the form $u + P(v_1, \ldots, v_n)$ for some $u \in V$. The vectors $v_1, \ldots, v_n$ are called the *edges* of this parallelepiped.

Proposition (invertible operator takes parallelepipeds to parallelepipeds): Suppose $u \in V$ and $v_1, \ldots, v_n$ is a basis of V. Suppose $T \in \mathcal{L}(V)$ is invertible. Then

$$T(u + P(v_1, \ldots, v_n)) = Tu + P(Tv_1, \ldots, Tv_n).$$

Proof: Because T is invertible, $Tv_1, \ldots, Tv_n$ is a basis of V. We then have

$$T(u + a_1 v_1 + \cdots + a_n v_n) = Tu + a_1 Tv_1 + \cdots + a_n Tv_n$$

for all $a_1, \ldots, a_n \in (0, 1)$. Thus $T(u + P(v_1, \ldots, v_n)) = Tu + P(Tv_1, \ldots, Tv_n)$. ∎

Definition (box): A *box* in V is a set of the form

$$u + P(r_1 e_1, \ldots, r_n e_n),$$

where $u \in V$, $r_1, \ldots, r_n$ are positive numbers, and $e_1, \ldots, e_n$ is an orthonormal basis of V.

Proposition (every invertible operator takes some boxes to boxes): Suppose $T \in \mathcal{L}(V)$ is invertible, and suppose T has SVD

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \dots + s_n \langle v, e_n \rangle f_n.$$

Then T maps the box $u + P(r_1 e_1, \dots, r_n e_n)$ onto the box $Tu + P(r_1 s_1 f_1, \cdots, r_n s_n f_n)$ for all positive numbers $r_1, \dots, r_n$ and all $u \in V$.

Proof: If $a_1, \dots, a_n \in (0, 1)$ and $r_1, \dots, r_n$ are positive numbers and $u \in V$, then

$$T(u + a_1 r_1 e_1 + \dots + a_n r_n e_n) = Tu + a_1 r_1 s_1 f_1 + \dots + a_n r_n s_n f_n.$$

Thus $T(u + P(r_1 e_1, \dots, r_n e_n)) = Tu + P(r_1 s_1 f_1, \cdots, r_n s_n f_n)$. ■

7.6.5. Volume via Singular values

Remark: A lot of the stuff in this subsection is nonrigorous, as it needs analysis to be defined properly.

Definition (volume of a box): Suppose $\mathbb{F} = \mathbb{R}$. If $u \in V$ and $r_1, \dots, r_n$ are positive numbers and $e_1, \dots, e_n$ is an orthonormal basis of V, then

$$\text{volume}(u + P(r_1 e_1, \dots, r_n e_n) = r_1 \times \dots \times r_n.$$

Definition (volume): Suppose $\mathbb{F} = \mathbb{R}$ and $\Omega \subseteq V$. Then the *volume* of Ω, denoted by volume Ω, is approximately the sum of the volumes of a collection of disjoint boxes that approximate Ω.

Proposition: Suppose $\mathbb{F} = \mathbb{R}$, $T \in \mathcal{L}(V)$ is invertible, and $\Omega \subseteq V$. Then

$$\text{volume } T(\Omega) = (\text{product of singular values of } T)(\text{volume } \Omega).$$

Proof: Suppose T has SVD

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \dots + s_n \langle v, e_n \rangle f_n.$$

Approximate Ω by boxes of the form $u + P(r_1 e_1, \dots, r_n e_n)$, which have volume $r_1 \times \dots \times r_n$. T maps each box $u + P(r_1 e_1, \dots, r_n e_n)$ to $Tu + P(r_1 s_1 f_1, \dots, r_n s_n f_n)$, which has volume $(r_1 \times \dots \times r_n)(s_1 \times \dots \times s_n)$. Since T changes the volume of the boxes that approximate Ω by $s_1 \times \dots \times s_n$, it changes the volume of Ω by that same factor. ■

7.6.6. Problems

Problem (Exercise 1): Prove that if $S, T \in \mathcal{L}(V, W)$, then $|\|S\| - \|T\|| \leq \|S - T\|$.

Solution: Follows from proof of reverse triangle inequality on normed spaces.

Problem (Exercise 2): Suppose that $T \in \mathcal{L}(V)$ is self adjoint or that $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$ is normal. Prove that

$$\|T\| = \max(|\lambda| : \lambda \text{ is an eigenvalue of } T).$$

Solution: Since under the conditions of the problem the singular values and the absolute value of the eigenvalues are equal, the max of the absolute values of the eigenvalues is equal to the max of the singular values of T.

Problem (Exercise 3): Suppose $T \in \mathcal{L}(V, W)$ and $v \in V$. Prove that

$$\|Tv\| = \|T\|\|v\| \Longleftrightarrow T^*Tv = \|T\|^2 v.$$

Solution: Suppose $T^*Tv = \|T\|^2 v$. Then

$$\langle T^*Tv, v\rangle = \|T\|^2 \langle v, v\rangle \Rightarrow \|Tv\|^2 = \|T\|\|v\|^2.$$

Now we prove the other direction. Letting $e_1, \ldots, e_m$ be the orthonormal list produced from the SVD of T, we calculate the norm of Tv and v explicitly to obtain

$$s_1^2|a_1|^2 + \cdots s_m^2|a_m|^2 = s_1^2|a_1|^2 + s_2^2|a_2|^2 + s_n^2|a_n|^2,$$

where we replaced $\|T\|$ with s_1, the largest singular value, and where a_i are the coefficients of v under the e_i basis. Rearranging yields

$$s_n^2|a_n|^2 + \cdots + s_{m+1}^2|a_{m+1}|^2 + (s_m^2 - s_1^2)|a_m|^2 + \cdots + (s_2^2 - s_1^2)|a_2|^2 = 0.$$

Note that each coefficient of an $|a_i|^2$ is positive, which means $a_2, \ldots, a_n = 0$. This $v = a_1 e_1$. Thus, under the SVD of T^*T, we have

$$T^*Tv = s_1^2 a_i e_i = s_1^2 v = \|T\|^2 v,$$

as desired.

Problem (Exercise 4): Suppose $T \in \mathcal{L}(V, W)$, $v \in V$, and $\|Tv\| = \|T\|\|v\|$. Prove that if $u \in V$ and $\langle u, v\rangle = 0$, then $\langle Tu, Tv\rangle$ = 0.

Solution: By the last exercise, we have $T^*Tv = \|T\|^2 v$. Thus

$$\langle Tu, Tv\rangle = \langle u, T^*Tv\rangle = \left\langle u, \|T\|^2 v\right\rangle = 0.$$

Problem (Exercise 5): Suppose U is a finite dimensional inner product space, $T \in \mathcal{L}(V, U)$, and $S \in \mathcal{L}(U, W)$. Prove that

$$\|ST\| \leq \|S\|\|T\|.$$

Solution: Pick the v such that $\|v\| = 1$ and $\|ST\| = \|STv\|$, which exists from our definitions of the map norm. Then we have

$$\|ST\| = \|STv\| \leq \|S\|\|Tv\| \leq \|S\|\|T\|\|v\| = \|S\|\|T\|.$$

Problem (Exercise 6): Prove or give a counterexample: If $S, T \in \mathcal{L}(V)$, then $\|ST\| = \|TS\|$.

Solution: Let $S(x, y) = (x, 0)$ and $T(x, y) = (0, x)$. Then $ST(x, y) = (0, x)$ and $TS(x, y) = (0, 0)$. Since ST is nonzero while TS is zero, clearly their norms aren't equal.

Problem (Exercise 7): Show that defining $d(S, T) = \|S - T\|$ for $S, T \in \mathcal{L}(V, W)$ makes d a metric on $\mathcal{L}(V, W)$.

Solution: Clearly $d(T, T) = 0$, and $d(S, T) \iff \|S - T\| = 0 \iff S - T = 0$. d is also symmetric and satisfies the triangle inequality.

Problem (Exercise 8):
(a) Prove that if $T \in \mathcal{L}(V)$ and $\|I - T\| < 1$, then T is invertible.
(b) Suppose that $S \in \mathcal{L}(V)$ is invertible. Prove that if $T \in \mathcal{L}(V)$ and $\|S - T\| < 1/\left\|S^{-1}\right\|$, then T is invertible.

Solution: Note that (a) is just (b) with $S = I$, so we prove (b) only. Suppose that T is not invertible. Then its SVD will have m positive singular values, where $m < \dim V$. Extend the basis from the SVD to a basis of V, and pick e_i such that $i > m$, since for these basis vectors, $Te_i = 0$. We have

$$\|S - T\| \geq \|(S - T)e_i\| = \|Se_i\| \geq s_n\|e_i\| = s_n = 1/\left\|S^{-1}\right\|.$$

Here s_n is the smallest singular values of S. The second inequality comes from the lower bound of $\|Sv\|$, and the last equality comes from the fact that S^{-1} has singular values $\frac{1}{s_n}, \ldots, \frac{1}{s_1}$, arranged from greatest to least. Thus we have a contradiction, so T must be invertible.

Problem (Exercise 17): Prove that if $u \in V$ and φ_u is the linear functional on V defined by the equation $\varphi_{u(v)} = \langle v, u \rangle$, then $\|\varphi_u\| = \|v\|$.

Solution: By CS, we have $\|\varphi_u(v)\| = |\langle v, u \rangle| \leq \|u\|\|v\|$ for all v, so $\|\varphi_u\| \leq \|u\|$. Then note that $\|\varphi_u(u)\| = \|u\|^2$, so we must have $\|\varphi_u\| = \|u\|$.

8. Operators on Complex Vector Spaces

8.1. Generalized Eigenvectors and Nilpotent Operators

8.1.1. Null Spcaes of Powers of an Operator

Proposition: Suppose $T \in \mathcal{L}(V)$. Then

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

Proof: Suppose $v \in \text{null } T^k$. Then $T^k v = 0 \Rightarrow T^{k+1} v = 0 \Rightarrow v \in \text{null } T^{k+1}$. ∎

Proposition: Suppose $T \in \mathcal{L}(V)$ and m is a nonnegative integer such that

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

Then

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

Proof: We want to show $\text{null } T^{m+k} = \text{null } T^{m+k+1}$. We already showed the inclusion in one direction (above proposition), so suppose $v \in \text{null } T^{m+k+1}$. then

$$T^{m+1}(T^k v) = T^{m+k+1} v = 0 \Rightarrow T^k v \in \text{null } T^{m+1} = \text{null } T^m,$$

which means that $T^{m+k} v = T^m (T^k v) = 0$, so $v \in \text{null } T^{m+k}$. ∎

Proposition: Suppose $T \in \mathcal{L}(V)$. Then

$$\text{null } T^{\dim V} = \text{null } T^{\dim V + 1} = \cdots.$$

Proof: We only need to prove that $\text{null } T^{\dim V} = \text{null } T^{\dim V+1}$. Suppose otherwise. then

$$\{0\} = \text{null } T^0 \subsetneq \text{null } T^1 \subsetneq \text{null } T^2 \cdots \subsetneq \text{null } T^{\dim V} \subsetneq \text{null } T^{\dim V+1}.$$

At each inclusion the dimension of the nullspace increases by at least 1, which means $\dim \text{null } T^{\dim V+1} \geq \dim V + 1$, which is clearly a contradiction. ∎

Proposition: Suppose $T \in \mathcal{L}(V)$. Then

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

Proof: Let $n = \dim V$. We first show

$$\text{null } T^n \cap \text{range } T^n = \{0\}.$$

Suppose u is in their intersection. Then $T^n u = 0$ and $T^n v = u$ for some v. Applying T^n to the second equations yields $T^{2n} v = 0 \Rightarrow v \in \text{null } T^{2n} \Rightarrow v \in \text{null } T^n \Rightarrow u = 0$, as desired.

Thus null T^n + range T^n is a direct sum, and since $\dim \text{null } T^n + \dim \text{range } T^n = \dim V$ by the fundamental theorem of linear maps, we have null $T^n \oplus$ range $T^n = V$, as desired. ■

8.1.2. Generalized Eigenvectors

Definition (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 is $v \neq 0$ and

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

for some positive integer k.

Proposition: Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then there is a basis of V consisting of generalized eigenvectors of T.

Proof: We induct on $n = \dim V$. Clearly the result holds when $n = 1$ since every map just scales the space.

Now suppose $n > 1$ and we;ve proved the result for every number up to n. Let λ be an eigenvalue of T. Then we have

$$V = \text{null}(T - \lambda I)^n \oplus \text{range } (T - \lambda I)^n.$$

If $\text{null}(T - \lambda I)^n = V$, then every vector is a generalized eigenvector, so we can assume $\text{null}(T - \lambda I)^n \neq V \Rightarrow \text{range}(T - \lambda I)^n \neq \{0\}$. Since λ is an eigenvalue, we have null $(T - \lambda I)^n \neq \{0\}$. Thus

$$0 < \dim \text{range } (T - \lambda I)^n < n.$$

We also have that range $(T - \lambda I)^n$ is invariant under T. Let $S = T|_{\text{range } (T - \lambda I)^n}$. By our induction hypothesis, there exists a basis of V made of generalized eigenvectors of S, which are clearly also generalized eigenvectors of T. Adjoining this basis with a basis of $\text{null}(T - \lambda I)^n$ yields our generalized eigenbasis/ ■

Proposition: Suppose $T \in \mathcal{L}(V)$. Then each generalized eigenvector of T corresponds to only one eigenvalue.

Proof: Suppose v is a generalized eigenvector of R corresponding to α and λ. Let m be the smallest positive integer such that $(T - \alpha I)^m v = 0$ and let $n = \dim V$. Then

$$\begin{aligned}0 &= (T-\lambda I)^n v \\ &= ((T-\alpha I)+(\alpha-\lambda I))^n v \\ &= \sum_{k=0}^{n}\binom{n}{k}(\alpha-\lambda)^{n-k}(T-\alpha I)^k v.\end{aligned}$$

Applying $(T-\alpha I)^{m-1}$ to both sides, everything on the right besides the first term of the sum disappears, since $k+m-1>m$ for those k. Thus we have

$$0=(\alpha-\lambda)^n(T-\alpha I)^{m-1}v.$$

Since $(T-\alpha I)^{m-1}v\neq 0$, we must have $\alpha=\lambda$, as desired. ■

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

Proof: Suppose otherwise. Then there exists a smllest m such that there exists a linearly dependent list $v_1,...,v_m$ of generalized eigenvectors corresponding to $\lambda_1,...,\lambda_m$. Thus these exist $a_1,...,a_m\in\mathbb{F}$, none of which are 0 by the minimality of m, such that

$$a_1v_1+\cdots+a_mv_m=0.$$

Let $n=\dim V$ and applying $(T-\lambda_m I)^n$ to both sides to get

$$\sum_{i=1}^{m-1}a_i(T-\lambda_m I)^n v_i=0.$$

For $k=1,...,m-1$, we have

$$(T-\lambda_m I)^n v_k\neq 0,$$

since otherwise v_k would be a generalized eigenvector corresponding to λ_m, which contradicts the previous proposition. We have

$$(T-\lambda_k I)^n((T-\lambda_m I)^n v_k)=(T-\lambda_m I)^n((T-\lambda_k I)^n v_k)=0,$$

so $(T-\lambda_m I)^n v_k$ is a generalized eigenvector of T corresponding to λ_k. Thus

$$(T-\lambda_m)^n v_1,...,(T-\lambda_m)^n v_{m-1}$$

is a linearly independent list of generalized eigenvectors corresponding to distinct eigenvalues (by the first equation), which contradicts the minimality of m. ■

8.1.3. Nilpotent Operators

Definition (nilpotent): An operator is called *nilpotent* if some power of it equals 0.

Proposition: Suppose $T \in \mathcal{L}(V)$ is nilpotent. Then $T^{\dim V} = 0$.

Proof: Because T is nilpotent, there exists a positive integer k such that $T^k = 0$. Thus $\operatorname{null} T^k = V$, which implies $\operatorname{null} T^{\dim V} = V \Rightarrow T^{\dim V} = 0$. ■

Proposition: Suppose $T \in \mathcal{L}(V)$.
(a) If T is nilpotent, then 0 is an eigenvalue of T and T has no other eigenvalues.
(b) If $\mathbb{F} = \mathbb{C}$ and 0 is the only eigenvalue of T, then T is nilpotent.

Proof:

(a) Suppose T is nilpotent. Then $T^m = 0$ for some m, which implies T is not injective. Thus 0 is an eigenvalue of T. Now suppose some nonzero λ is an eigenvalue of T. Then for its corresponding eigenvector, we have $T^m v = \lambda^m v \neq 0$, which is a contradiction. Thus, 0 is the only eigenvalue of T.

(b) If $\mathbb{F} = \mathbb{C}$ and 0 is the only eigenvalue, then z^m is the minimal polynomial of T for some m, which implies $T^m = 0$, as desired.

■

Proposition: Suppose $T \in \mathcal{L}(V)$. Then the following are equivalent.
(a) T is nilpotent.
(b) The minimal polynomial of T is z^m for some positive integer m.
(c) There is a basis of V with respect to which the matrix of T has the form

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

where all entries on and below the diagonal equal 0.

Proof: If (a) holds, then $T^n = 0$ for some n, which means the minimal polynomial is T^m for some $m \leq n$, so (b) holds.

If (b) holds, then 0 is the only eigenvalue of T. It also implies that there exists a basis of V for which T is upper triangular. Both of these imply that all the entries on the diagonal are 0, so (c) holds.

If (c) holds, then $T^{\dim V} = 0$, so T is nilpotent. ■

8.1.4. Problems

Problem (Exercise 1): Suppose $T \in \mathcal{L}(V)$. Prove that if $\dim \operatorname{null} T^4 = 8$ and $\dim \operatorname{null} T^6 = 9$, then $\dim \operatorname{null} T^m = 9$ for all integers $m \geq 5$.

Solution: Note that $\dim \operatorname{null} T^5$ can only be 8 or 9. If it's 8, then that would imply that $\dim \operatorname{null} T^6 = 8$, which is impossible, so it must be 9, which then implies the desired result.

Problem (Exercise 2): Suppose $T \in \mathcal{L}(V)$, m is a positive integer, $v \in V$, and $T^{m-1}v \neq 0$, but $T^m v = 0$. Prove that $v, Tv, T^2 v, \ldots, T^{m-1}v$ are linearly independent.

Solution: Suppose FTSOC they are linearly dependent. Then there is some smallest subset of them $T^{i_1}v, \ldots, T^{i_k}v$, where $\{i_1, \ldots, i_k\} \subseteq \{0, \ldots, m-1\}$ and $\sum_{j=1}^{k} a_j T^{i_j} v = 0$ where none of the a_j are 0. Suppose WLOG that the $i_n < i_{n+1}$. Applying T^{m-i_k} to both sides of the sum yields

$$\sum_{j=1}^{k} a_j T^{i_j+m-i_k} v = 0.$$

Note that last term in the sum is 0 since $T^m v = 0$, while none of the other terms aren't. This yields a smaller subset that is linearly dependent with nonzero coefficients, a contradiction.

Problem (Exercise 3): Suppose $T \in \mathcal{L}(V)$. Prove that

$$V = \operatorname{null} T \oplus \operatorname{range} T \Longleftrightarrow \operatorname{null} T^2 = \operatorname{null} T.$$

Solution: First suppose $V = \operatorname{null} T \oplus \operatorname{range} T$. Clearly we have $\operatorname{null} T \subseteq \operatorname{null} T^2$. Now suppose $v \in \operatorname{null} T^2$. Then $T(Tv) = 0$, which implies $Tv \in \operatorname{null} T$, but we also have $Tv \in \operatorname{range} T$, so $Tv = 0$, which implies $v \in \operatorname{null} T$, as desired.

Now suppose $\operatorname{null} T^2 = \operatorname{null} T$. We just need to prove $\operatorname{null} T \oplus \operatorname{range} T = \{0\}$, and then the result follows from the fundamental theorem of linear maps. If $v \in \operatorname{null} T \cap \operatorname{range} T$, then $Tv = 0$ and $v = Tu$ for some u. Applying T to the last equation yields $0 = Tv = T^2 u$, which implies $u \in \operatorname{null} T^2 = \operatorname{null} T$, which implies $v = Tu = 0$, as desired.

Problem (Exercise 13): Suppose $S, T \in \mathcal{L}(V)$ and ST is nilpotent. Prove that TS is nilpotent.

Solution: Suppose $(ST)^m = 0$. We have $(TS)^{m+1} = T(ST)^m S = 0$, as desired.

Problem (Exercise 23): Give an example of an operator T ona finite dimensional real vector space such that 0 is the only eigenvalue of T but T is not nilpotent.

Solution: Let $T(x, y, z) = (0, -y + z, -y)$. The minimal polynomial of T is $z(z^2 + z + 1)$. 0 is the only eigenvalue of T, but T is not nilpotent, since $T^3 \neq 0$.

8.2. Generalized Eigenspace Decomposition

8.2.1. Generalized Eigenspaces

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

$$G(\lambda, T) = \left\{v \in V : (T - \lambda I)^k v = 0 \text{ for some positive integer } k\right\}.$$

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

Proof: If $v \in \operatorname{null}(T - \lambda I)^{\dim V}$, then clearly $v \in G(\lambda, T)$, so $\operatorname{null}(T - \lambda I)^k \subseteq G(\lambda, T)$. If $v \in G(\lambda, T)$, then $v \in \operatorname{null}(T - \lambda I)^k$ for some k, which implies $v \in \operatorname{null}(T - \lambda I)^k$. Thus $G(\lambda, T) \subseteq \operatorname{null}(T - \lambda I)^{\dim V}$. ■

Theorem (generalized eigenspace decomposition): Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. Then
(a) $G(\lambda_k, T)$ is invariant under T for each k.
(b) $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent for each k.
(c) $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$.

Proof:
(a) Since $G(\lambda_k, T) = \operatorname{null}\ (T - \lambda_k I)^{\dim V}$, it's clearly invariant under T.
(b) If $v \in G(\lambda_k, T)$, then $(T - \lambda_k I)^{\dim V} v = 0$, which implies $\left((T - \lambda_k I)|_{G(\lambda_k, T)}\right)^{\dim V} = 0$, as desired.
(c) Suppose

$$v_1 + \cdots + v_m = 0,$$

where $v_k \in G(\lambda_k, T)$. Because generalized eigenvectors corresponding to distinct eigenvalues are linearly independent, each v_k must be 0, which implies that $V = G(\lambda_1, T) + \cdots + G(\lambda_m, T)$ is a direct sum. Since each vector in V can be written as a sum of generalized eigenvectors (generalized eigenbasis), we have $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$, as desired.

■

8.2.2. Multiplicity of an Eigenvalue

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

Proposition: Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then the sum of the multiplicities of all eigenvalues of T equals $\dim V$.

Proof: Follows from the generalized eigenspace decomposition. ■

Definition (characteristic polynomial): Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Let $\lambda_1, ..., \lambda_m$ denote the distinct eigenvalues of T, with multiplicaties $d_1, ..., d_m$. The polynomial

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

is called the *characteristic polynomial* of T.

Proposition: Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then the characteristic polynomial of T has degree $\dim V$, and the zeros of the characteristic polynomial of T are the eigenvalues of T.

Proof: (a) follows from the previous proposition, and (b) follows by definition. ■

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

Proof: Let $\lambda_1, ..., \lambda_m$ be the distinct eigenvalues of T, and let $d_k = \dim G(\lambda_k, T)$. For each k, we know $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent, so

$$(T - \lambda_k I)^{d_k}|_{G(\lambda_k, T)} = 0.$$

By the generalized eigenspace decomposition, we only need to show $q(T)|_{G(\lambda_k, T)} = 0$. However, this is easy, as the factors of the polynomial commute, and so pushing that factor involving k to the end yields 0. ■

Corollary: Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then the characteristic polynomial of T is a polynomial mutiple of the minimal polynomial of T.

Proof: Follows since $q(T) = 0$. ■

Proposition: Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Suppose $v_1, ..., v_n$ is a basis of V such that $\mathcal{M}(T, (v_1, ..., v_n))$ is upper triangular. Then the number of times that each eigenvalue of T appears on the diagonal of $\mathcal{M}(T, (v_1, ..., v_n))$ equals the multiplicity of λ as an eigenvalue.

Proof: Let $A = \mathcal{M}(T, (v_1, ..., v_n))$. Let $\lambda_1, ..., \lambda_n$ be the entries on the diagonal of A. For each k we have

$$Tv_k = u_k + \lambda_k v_k,$$

where $u_k \in \operatorname{span}(v_1, ..., v_{k-1})$. Thus is $\lambda_k \neq 0$, then Tv_k is not a linear combination of $Tv_1, ..., Tv_{k-1}$. The converse of the linear dependence lemma then implies that the list of Tv_k with $\lambda_k \neq 0$ is linearly independent.

Let d denote the number of indices such that $\lambda_k = 0$. From the last paragraph, we have

$$\dim \operatorname{range} T \geq n - d \Rightarrow \dim \operatorname{null} T \leq d.$$

Since $\mathcal{M}(T^n, (v_1, ..., v_n)) = A^n$ has diagonal entries λ_k^n, the number of times 0 appears on the diagonal of A^n is d, so using the last inequality we have

$$\dim \operatorname{null} T^n \leq d.$$

For an eigenvalue λ of T, let m_λ denote its multiplicity, and let d_λ denote the number of times it appears on the diagonal of A. Replacing T with $T - \lambda I$ is the previous equation yields

$$m_\lambda \leq d_\lambda.$$

Summing over all λ, we get n on both sides, which implies the inequality is actually in equality, so we're done. ■

8.2.3. Block Diagonal Matrices

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

Proposition: Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Let $\lambda_1, ..., \lambda_m$ be the distinct eigenvalues of T, with multiplicities $d_1, ..., d_m$. 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_k is a d_k by d_k upper triangular matrix of the form

$$A_k = \begin{pmatrix} \lambda_k & & * \\ & \ddots & \\ 0 & & \lambda_k \end{pmatrix}.$$

Proof: We know each $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent. For each of these maps, choose a matrix such that $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is upper triangular, and thus has all 0s on its diagonal. Then, with respect to this basis, we know that $T|_{G(\lambda_k, T)} = (T - \lambda_k I)|_{G(\lambda_k, T)} + \lambda_k I|_{G(\lambda_k, T)}$, which gives us the desired form for A_k. Then, by the generalized eigenspace decomposition, we can put these bases together and obtain the desired matrix. ■

8.2.4. Problems

Problem (Exercise 1): Define $T \in \mathcal{L}(\mathbb{C}^2)$ by $T(w, z) = (-z, w)$. Find the generalized eigenspaces corresponding to the distinct eigenvalues of T.

Solution: We can compute the eigenvalues are $\pm i$, both with 1 eigenvector. Since the the sum of the dimensions of the eigenspaces is equal the sum of $\dim V$, the eigenspaces are equal to the generalized eigenspaces (we can also compute $(T - \lambda I)^2$ and see that they're both actually equal to $T - \lambda I$).

Problem (Exercise 2): Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that $G(\lambda, T) = G\left(\frac{1}{\lambda}, T^{-1}\right)$ for every $\lambda \in \mathbb{F}$ with $\lambda \neq 0$.

Solution: Suppose $v \in G(\lambda, T)$, meaning $(T - \lambda I)^k v = 0$ for some k. Then we have

$$(T - \lambda I)^k v = 0 \Rightarrow \lambda^k (-1)^k T^k \left(T^{-1} - \frac{1}{\lambda} I\right)^k v = 0.$$

Since T is invertible, 0 is not an eigenvalue, so $T^k v \neq 0$ for any v. Thus we must have $\left(T^{-1} - \frac{1}{\lambda} I\right)^k v = 0$, which implies $V \in G\left(\frac{1}{\lambda}, T^{-1}\right)$. Thus $G(\lambda, T) \subseteq G\left(\frac{1}{\lambda}, T^{-1}\right)$. The other direction follows similarly, so the sets are equal.

Problem (Exercise 3): Suppose $T \in \mathcal{L}(V)$ and $S \in \mathcal{L}(V)$ is invertible. Prove that T and $S^{-1}TS$ have the same eigenvalues with the same multiplicities.

Solution: Suppose $v \in G(\lambda, T)$. Then $(T - \lambda I)^k v = 0$ for some k. We can expand this to get

$$\sum_{i=0}^{k} \binom{k}{i} T^i (-\lambda)^{k-i} v = 0.$$

Note that $\left(S^{-1}TS\right)^j = S^{-1}T^jS$. Thus we can write $T^i v = S^{-1}T^i S u = \left(S^{-1}TS\right)^i u$ for some u. Thus

$$0 = \sum_{i=0}^{k} \binom{k}{i} T^i (-\lambda)^{k-i} v = \sum_{i=0}^{k} \binom{k}{i} \left(S^{-1}TS\right)^i (-\lambda)^{k-i} u = \left(S^{-1}TS - \lambda I\right)^k u = 0.$$

Thus λ is an eigenvalue of $S^{-1}TS$. Consider $n = \dim G(\lambda, T)$ linearly independent vectors in $G(\lambda, T)$. Note that in the equation above, $Su = v \Rightarrow S^{-1}v = u$. Thus, since S is invertible, the u's are linearly independent, and this $\dim G(\lambda, T) \leq \dim G(\lambda, S^{-1}TS)$. The other direction follows similarly.

Problem (Exercise 4): Suppose $\dim V \geq 2$ and $T \in \mathcal{L}(V)$ is such that null $T^{\dim V - 2} \neq$ null $T^{\dim V - 1}$. Prove that T has at most two distinct eigenvalues.

Solution: Suppose for the sake of contradiction that T has at least 3 distinct eigenvalues. Note that T cannot be injective, otherwise the given condition is false, so 0 is an eigenvalue. Thus, we have $\dim \operatorname{null} T^n \leq n - 2$. However, since the condition on T implies that null T^k keeps growing untill at least $k = n - 1$, we have $n - 1 \leq \dim \operatorname{null} T^{n-1} < \dim \operatorname{null} T^n \leq n - 2$, contradiction.

Problem (Exercise 6): Suppose $T \in \mathcal{L}(V)$ is an eigenvalue of T. Prove that the exponent of $(z - \lambda)$ in the factorization of the minimal polynomial of T is the smallest m such that $(T - \lambda I)^m |_{G(\lambda, T)} = 0$.

Solution: Note that the minimal polynomial of $T|_{G(\lambda,T)}$ is $p(z) = (z - \lambda)^m$. Thus, the minimal polynomial of T is of the form $(z - \lambda)^{m+k} q(z)$. We can show $k = 0$ by noticing the portion of a vector v contained in $G(\lambda, T)$ vanished under $(T - \lambda I)^m$, and the portion not in $G(\lambda, T)$ must vanish in $q(T)$, since null $(T - \lambda I)^m$ is constant for $n \geq m$.

Problem (Exercise 10): Suppose V is a complex inner product space, $e_1, \ldots, e_n$ is an orthonormal basis of T, and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of T, each included as many times as its multiplicity. Prove that $|\lambda_1|^2 + \cdots + |\lambda_n|^2 \leq \|Te_1\|^2 + \cdots + \|Te_n\|^2$.

Solution: We know from a previous exercise that the right side is the same regardless of the orthonormal basis chosen. Thus, we choose a basis such that the matrix of T with respect to that basis is upper triangular. Then the conclusion clearly follows: for each i, we have $\|Te_i\|^2 = |A_{i,i}|^2 + |A_{i-1,i}|^2 + \dots + |A_{1,i}|^2 = |\lambda_i|^2 + |A_{i-1,i}|^2 + \dots + |A_{1,i}|^2 \geq |\lambda_i|^2$.

Remark: While solving, I also showed that

$$(T - \lambda I)^k v = 0 \Longleftrightarrow \left(T^* - \overline{\lambda} I\right)^k v = 0,$$

and this can be shown pretty easily using inner products.

8.3. Consequences of Gernalized Eigenspace Decomposition

8.3.1. Square Roots of Operators

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

Proof: Consider the Taylor series for $\sqrt{1+x}$, and plug in T for x. Since T is nilpotent, at some point the series will stop, and there will be m terms. We label the coefficients of these terms a_i. Squaring the left over series, we equate each coefficient to the coefficient of its corresponding operator in $I + T$ (1 is its the coefficient of I or T and 0 otherwise). Then we solve for each a_i and get the square root. ■

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

Proof: Let $\lambda_1, ..., \lambda_m$ be the distinct eigenvalues of T. For each k we have let $T_k = T|_{G(\lambda_k T)} - \lambda_k I$. Note that T_k is nilpotent. Since T is invertible, none of the λ's are 0, so we can write

$$T|_{G(\lambda_k,T)} = \lambda_k \left(I + \frac{T_k}{\lambda_k}\right).$$

Because $\frac{T_k}{\lambda_k}$ is nilpotent, $I + \frac{T_k}{\lambda_k}$ has a square roto, and multiplying it by $\sqrt{\lambda_k}$ gives us a square root of $T|_{G(\lambda_k,T)}$, which we denote R_k.

By the generalized eigenspace decomposition, we have write any $v \in V$ as

$$v = u_1 + \cdots + u_m,$$

where $u_j \in G(\lambda_j, T)$. Define $R \in \mathcal{L}(V)$ by

$$Rv = R_1 u_1 + \cdots + R_m u_m.$$

Since the R_j's are invariant on their generalized eigenspaces, Rv is also uniquely written as vectors from each generalized eigenspace. Thus we can apply R again and get

$$R^2 v = R_1^2 u_1 + \cdots + R_m^2 u_m = T u_1 + \cdots + T u_m = T v,$$

so R is a square of T, as desired. ■

8.3.2. Jordan Form

Definition (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

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

in which each A_k is an upper triangular matrix of the form

$$A_k = \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{pmatrix}.$$

Proposition (every nilpotent operator has a Jordan basis): Suppose $T \in \mathcal{L}(V)$ is nilpotent. Then there is a basis of V that is a Jordan basis for T.

Proof: We induct on $\dim V$. Clearly it holds for $\dim V = 1$, since the only nilpotent operator is 0. Now suppose it holds for all k less than $\dim V$.

Let m be the smallest positive integer such that $T^m = 0$. Thus there exists $u \in V$ such taht $T^{m-1}u \neq 0$. Let

$$U = \operatorname{span}(u, Tu, ..., T^{m-1}u).$$

The list in the above span is linearly independent by exercise 2 in 8.1. If $U = V$, then writing the list in reverse order gives a Jordan basis, so we assume $U \neq V$.

Note U is invariant under T, so by our induction hypothesis, there exists a Jordan basis for $T|_U$. Our goal is to find a subspace W such that $V = U \oplus W$ and such that W is invariant under T, as then there will be a Jordan basis of W by our induction hypothesis, and putting the bases for U and W together gives us the desired result.

Let $\varphi \in V'$ such that $\varphi(T^{m-1}u) \neq 0$. Let

$$W = \{v \in V : \varphi(T^k v) = 0 \text{ for each } k = 0, ..., m-1\}.$$

W is a subspace of V invariant under T. We show that $U \oplus W = V$.

Suppose $0 \neq v \in U \cap W$. Because $v \in U$, there exist $c_0, ..., c_{m-1} \in \mathbb{F}$ such that

$$v = c_0 u + c_1 Tu + \cdots + c_{m-1} T^{m-1} u.$$

Let j be the smallest index such taht $c_j \neq 0$. Apply T^{m-j-1} to both sides to get

$$T^{m-j-1} v = c_j T^{m-1} u,$$

where the terms after disappeared because $T^m = 0$, and the terms before disappeared because their coefficients were 0. Apply φ to both sides to get

$$\varphi(T^{m-j-1} v) = c_j \varphi(T^{m-1} u) \neq 0.$$

Thus $v \notin W$, so $U \cap W = \{0\}$. Thus $U + W$ is a direct sum.

Now define $S : V \to \mathbb{F}^m$ By

$$Sv = (\varphi(v), \varphi(Tv), ..., \varphi(T^{m-1}v)).$$

Thus null $S = W$. We have

$$\dim W = \dim \operatorname{null} S = \dim V - \dim \operatorname{range} S \geq \dim V - m,$$

where the last inequality comes from the fact that S maps into $\mathbb{F}^m$, so its range is at most m. Then we have

$$\dim(U \oplus W) = \dim U + \dim W \geq m + (\dim V - m) = \dim V.$$

However, since $U \oplus W$ is a subspace of V, we also have $\dim V \geq \dim(U \oplus W)$. Thus, we have $V = U \oplus W$, as desired. ■

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

Proof: Let $\lambda_1, ..., \lambda_m$ be the distinct eigenvalues of T. By the generalized eigenspace decomposition, we have

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

where $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent. Thus the previous result implies there exists a Jordan basis for $(T - \lambda_k I)|_{G(\lambda_k, T)}$, which is then also a Jordan basis for $T|_{G(\lambda_k, T)}$. Putting these bases together yields a basis of V that is a Jordan basis of T. ■

8.3.3. Problems

Problem (Exercise 6): Find a basis of $\mathcal{P}_4(\mathbb{R})$ that's a Jordan basis for the differentiation operator.

Solution:

$$6, 6x, 3x^2, x^3$$

8.4. Trace

Definition (trace of a matrix): Suppose A is a square matrix with entries in $\mathbb{F}$. The *trace* of A, denoted by $\operatorname{tr} A$, is defined to be the sum of the diagonal entries of A.

Proposition: Suppose A is an m by n matrix and B is an n by , matrix. Then

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

Proof:

$$\begin{aligned}
\operatorname{tr}(AB) &= \sum_{j=1}^{m} \sum_{k=1}^{n} A_{j,k} B_{k,j} \\
&= \sum_{k=1}^{n} \sum_{j=1}^{m} B_{k,j} A_{j,k} \\
&= \sum_{k=1}^{n} \left(k^{\text{th}} \text{ term on the diagonal of the } n \text{ by } n \text{ matrix } BA\right) \\
&= \operatorname{tr}(BA).
\end{aligned}$$

■

Proposition: Supose $T \in \mathcal{L}(V)$. Suppose $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ are bases of V. Then

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

Proof: Let A be the first matrix and B be the second. By change of basis there's some C such that $A = C^{-1}BC$. Thus

$$\begin{aligned}
\operatorname{tr} A &= \operatorname{tr}\left((C^{-1}B)C\right) \\
&= \operatorname{tr}\left(C(C^{-1}B)\right) \\
&= \operatorname{tr}\left((CC^{-1})B\right) \\
&= \operatorname{tr}(B),
\end{aligned}$$

where the second line comes from the previous proposition. ■

Definition (trace of an operator): Suppose $T \in \mathcal{L}(V)$. The *trace* of T, denoted $\operatorname{tr} T$, is defined by

$$\operatorname{tr} T = \operatorname{tr} \mathcal{M}(T, (v_1, \ldots, v_n)),$$

where $v_1, \ldots, v_n$ is any basis of V.

Proposition: Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then $\operatorname{tr} T$ equals the sum of the eigenvalues of T, with each eigenvalue included as many times as its multiplicity.

Proof: There's a basis of V such that the matrix of T is upper triangular with the eigenvalues included as many times as their multiplicities on the diagonal. Thus, by the definition of trace, the result follows. ∎

Proposition: Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then $\operatorname{tr} T$ equals the negative coefficient of z^{n-1} in the characteristic polynomial of T.

Proof: Follows from Vieta's on the characteristic polynomial. ∎

Proposition: Suppose V is an inner product space, $T \in \mathcal{L}(V)$, and $e_1, \ldots, e_n$ is an orthonormal basis of V. Then

$$\operatorname{tr} T = \langle Te_1, e_1 \rangle + \cdots + \langle Te_n, e_n \rangle.$$

Proof: Follows from the fact that the the ith diagonal entry is equal to $\langle Te_i, e_i \rangle$. ∎

Proposition (trace is linear): The function $\operatorname{tr} : \mathcal{L}(V) \to \mathbb{F}$ is a linear functional on $\mathcal{L}(V)$ such that

$$\operatorname{tr}(ST) = \operatorname{tr}(TS)$$

for all $S, T \in \mathcal{L}(V)$.

Proof: Choose a basis of V. It will be used in the duration of this proof. Suppose $S, T \in \mathcal{L}(V)$.

If $\lambda \in \mathbb{F}$, then

$$\operatorname{tr}(\lambda T) = \operatorname{tr} \mathcal{M}(\lambda T) = \operatorname{tr}(\lambda \mathcal{M}(T)) = \lambda \operatorname{tr} \mathcal{M}(T) - \lambda \operatorname{tr} T.$$

We also have

$$\operatorname{tr}(S + T) = \operatorname{tr} \mathcal{M}(S + T) = \operatorname{tr}(\mathcal{M}(S) + \mathcal{M}(T)) = \operatorname{tr} \mathcal{M}(S) + \operatorname{tr} \mathcal{M}(T) = \operatorname{tr} S + \operatorname{tr} T.$$

Thus tr is a linear functional.

We also have

$$\operatorname{tr}(ST) = \operatorname{tr} \mathcal{M}(ST) = \operatorname{tr}(\mathcal{M}(S)\mathcal{M}(T)) = \operatorname{tr}(\mathcal{M}(T)\mathcal{M}(S)) = \operatorname{tr} \mathcal{M}(TS) = \operatorname{tr}(TS),$$

where the third equality comes from trace commuting on matrices. ∎

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

Proof: We have

$$\operatorname{tr}(ST - TS) = \operatorname{tr}(ST) - \operatorname{tr}(TS) = 0,$$

but $\operatorname{tr}(I) = \dim V$, so the operators can't be equal. ■

8.4.1. Problems

Problem (Exercise 1): Suppose V is an inner product space and $v, w \in V$. Define an operator $T \in \mathcal{L}(V)$ by $Tu = \langle u, v \rangle w$. Find a formula for $\operatorname{tr} T$.

Solution: Let $e_1, ..., e_n$ be an orthonormal basis of V. Let $v = \sum a_i e_i$ and $w = \sum b_i e_i$. Letting $A = \mathcal{M}(T, (e_1, ..., e_n))$, we have

$$A_{j,j} = \langle \langle e_j, v \rangle w, e_j \rangle = \overline{\langle e_j, v \rangle} \langle w, e_j \rangle.$$

Summing over all j yields $\sum \overline{a_j} b_j = \langle w, v \rangle = \operatorname{tr} T$.

Problem (Exercise 4): Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Prove that

$$\operatorname{tr} T^* = \overline{\operatorname{tr} T}.$$

Solution: Follows since the matrix of T^* is the conjugate transpose of the matrix of T.

Problem (Exercise 5): Suppose V is an innner product space and $T \in \mathcal{L}(V)$ is a positive operator with $\operatorname{tr} T = 0$. Prove that $T = 0$.

Solution: Since T is positive, all its eigenvalues are positive real numbers, and thus they must all be 0 in order for $\operatorname{tr} T = 0$. Since all the eigenvalues are 0 and since by the spectral theorem T is diagonalizable, T must be the zero map.

Problem (Exercise 12): Suppose V and W are finite dimensional inner product spaces.

(a) Prove that $\langle S, T \rangle = \operatorname{tr}(T^* S)$ defines an inner product space on $\mathcal{L}(V, W)$.
(b) Suppose $e_1, ..., e_n$ is an orthonormal basis of V and $f_1, ..., f_n$ is an orthonormal basis of W. Show that the inner product on $\mathcal{L}(V, W)$ from (a) is the same as the standard inner product on $\mathbb{F}^{mn}$, where we identify each element of $\mathcal{L}(V, W)$ with its matrix from $\mathbb{F}^{mn}$ with the bases mentioned above.

Solution: Clearly the additivity and scalar multiplcation work like they should on an inner product. Note that $\langle T, T\rangle = \mathrm{tr}(T^* T) \geq 0$, and as we showed in an earlier exercise, is only 0 when $T = 0$. Thus the function is an inner product.

Note that $T^*_{i,j} = \overline{T_{j,i}}$. Thus

$$T^* S_{k,k} = \sum_{j=1}^{n} \overline{T_{j,k}} S_{j,k} \Rightarrow \mathrm{tr}(T^* S) = \sum_{k=1}^{n} \sum_{j=1}^{n} \overline{T_{j,k}} S_{j,k},$$

which is the same formula as the standard inner product.

Problem (Exercise 13): Find $S, T \in \mathcal{L}(\mathcal{P}(\mathbb{F}))$ such that $ST - TS = I$.

Solution: Let D be the differentiation operator and let T be the operator that multiplies the input by x. We have

$$(DT - TD)p = p + xp' - xp' = p = I(p).$$

9. Multilinear Algebra and Determinants

9.1. Bilinear Forms and Quadratic Forms

9.1.1. Bilinear Forms

Definition (bilinear form): A *bilinear form* on V is a function $\beta : V \times V \to \mathbb{F}$ such that

$$v \mapsto \beta(v, u) \quad \text{and } v \mapsto \beta(u, v)$$

are both linear functionals on V for every $u \in V$.

Definition ($V^{(2)}$): The set of bilinear forms on V is denoted by $V^{(2)}$.

Definition (matrix of a bilinear form): Suppose β is a bilinear form on V and $e_1, ..., e_n$ is a basis of V. The matrix of β with respect to this basis is the n by n matrix $\mathcal{M}(\beta)$ whose j, k entry is

$$\mathcal{M}(\beta)_{j,k} = \beta(e_j, e_k).$$

Proposition: Suppose $e_1, ..., e_n$ is a basis of V. Then the map $\beta \mapsto \mathcal{M}(\beta)$ is an isomorphism of $V^{(2)}$ onto $\mathbb{F}^{n,n}$.

Proof: The map $\beta \mapsto \mathcal{M}(\beta)$ is clearly a linear map from $V^{(2)}$ to $\mathbb{F}^n$. For $A \in \mathbb{F}^{n,n}$, define a bilinear form β_A on V by

$$\beta_A(x_1 e_1 + \dots + x_n e_n, y_1 e_1 + \dots + y_n e_n) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j y_k$$

for $x_1, ..., x_n, y_n, ..., y_n \in \mathbb{F}$.

Note that for any $\beta \in V^{(2)}$ and for any $u, v \in V$ where $u = \sum_{i=1}^{n} x_i e_i$ and $v = \sum_{i=1}^{n} y_i e_i$, we have

$$\begin{aligned}\beta_{\mathcal{M}(\beta)}(u,v) &= \sum_{k=1}^{n}\sum_{j=1}^{n}\beta(e_j,e_k)x_j y_k \\ &= \sum_{k=1}^{n} y_k \sum_{j=1}^{n}\beta(x_j e_j, e_k) \\ &= \sum_{k=1}^{n} y_k \beta\left(\sum_{j=1}^{n} x_j e_j, e_k\right) \\ &= \beta\left(\sum_{j=1}^{n} x_j e_j, \sum_{k=1}^{n} y_k e_k\right).\end{aligned}$$

Thus $\beta_{\mathcal{M}(\beta)} = \beta$. For any $A \in \mathbb{F}^{n,n}$, we also have

$$\mathcal{M}(\beta_A)_{j,k} = \beta_A(e_j,e_k) = A_{j,k},$$

so $\mathcal{M}(\beta_A) = A$. Thus, the two maps defined are inverses, and therefore isomorphisms. ■

Corollary: $\dim V^{(2)} = (\dim V)^2$.

Proof: $V^{(2)}$ and $\mathbb{F}^{n,n}$ are isomorphisms by the previous result, and $\dim \mathbb{F}^{n,n} = (\dim V)^2$. ■

Proposition: Suppose β is a bilinear form on V and $T \in \mathcal{L}(V)$. Define bilinear forms α and ρ on V by

$$\alpha(u,v) = \beta(u,Tv) \quad \text{and} \quad \rho(u,v) = \beta(Tu,v).$$

Let $e_1, ..., e_n$ be a basis of V. Then

$$\mathcal{M}(\alpha) = \mathcal{M}(\beta)\mathcal{M}(T) \quad \text{and} \quad \mathcal{M}(\rho) = \mathcal{M}(T)^t\mathcal{M}(\beta).$$

Proof: We have

$$\begin{aligned}\mathcal{M}(\alpha)_{j,k} &= \alpha(e_j,e_k) \\ &= \beta(e_j, Te_k) \\ &= \beta\left(e_j, \sum_{r=1}^{n}\mathcal{M}(T)_{r,k}e_r\right) \\ &= \sum_{r=1}^{n}\beta(e_j,e_r)\mathcal{M}(T)_{r,k} \\ &= (\mathcal{M}(\beta)\mathcal{M}(T))_{j,k}.\end{aligned}$$

The proof of the second assertion follows similarly. ■

Theorem (change of basis for bilinear forms): Suppose $\beta \in V^{(2)}$. Suppose $e_1, ..., e_n$ and $f_1, ..., f_n$ are bases of V. Let

$$A = \mathcal{M}(\beta, (e_1, ..., e_n)) \quad \text{and } B = \mathcal{M}(\beta, (f_1, ..., f_n))$$

and $C = \mathcal{M}(I, (e_1, ..., e_n), (f_1, ..., f_n))$. Then

$$A = C^t BC.$$

Proof: Define $T \in \mathcal{L}(V)$ by $Tf_k = e_k$. Then $\mathcal{M}(T, (f_1, ..., f_m)) = C$.

Define $\alpha, \rho \in V^{(2)}$ by

$$\alpha(u, v) = \beta(u, Tv) \quad \text{and } \rho(u, v) = \alpha(Tu, v) = \beta(Tu, Tv).$$

Then $\beta(e_j, e_k) = \beta(Tf_j, Tf_k) = \rho(f_j, f_k)$. Thus

$$\begin{aligned} A &= \mathcal{M}(\rho, (f_1, ..., f_n)) \\ &= C^t \mathcal{M}(\alpha, (f_1, ..., f_n)) \\ &= C^t BC, \end{aligned}$$

where the last two equalities hold by the previous result. ∎

9.1.2. Symmetric and Alternating Bilinear Forms

Definition (symmetric bilinear form): A bilinear form $\rho \in V^{(2)}$ is called *symmetric* if

$$\rho(u, w) = \rho(w, u)$$

for all $u, w \in V$. The set of symmetric bilinear forms on V is denoted by $V^{(2)}_{\text{sym}}$.

Definition (symmetric matrix): A square matrix A is called *symmetric* if it equals its transpose.

Proposition: Suppose $\rho \in V^{(2)}$. Then the following are equivalent:
(a) ρ is a symmetric bilinear form on V.
(b) $\mathcal{M}(\rho, (e_1, ..., e_n))$ is a symmetric matrix for every basis $e_1, ..., e_n$ of V.
(c) $\mathcal{M}(\rho, (e_1, ..., e_n))$ is a symmetric matrix for some basis $e_1, ..., e_n$ of V.
(d) $\mathcal{M}(\rho, (e_1, ..., e_n))$ is a diagonal matrix for every basis $e_1, ..., e_n$ of V.

Proof: First suppose (a) holds. Suppose $e_1, ..., e_n$ is a basis of V. Then $\mathcal{M}(\rho)_{j,k} = \rho(e_j, e_k) = \rho(e_k, e_j) = \mathcal{M}(\rho)_{k,j}$, so $\mathcal{M}(\rho)$ is symmetric.

Clearly (b) implies (c).

Suppose (c) holds and $e_1, ..., e_n$ is a basis of V such that $\mathcal{M}(\rho(e_1, ..., e_n))$ is symmetric. Let $u = a_1 e_1 + \cdots + a_n e_n$ and $v = b_1 e_1 + \cdots b_n e_n$. Then

$$\begin{aligned}\rho(u, w) &= \rho\left(\sum_{j=1}^{n} a_j e_j, \sum_{k=1}^{n} b_k e_k\right) \\ &= \sum_{j=1}^{n}\sum_{k=1}^{n} a_j b_k \rho(e_j, e_k) \\ &= \sum_{j=1}^{n}\sum_{k=1}^{n} a_j b_k \rho(e_k, e_j) \\ &= \rho\left(\sum_{k=1}^{n} b_k e_k, \sum_{j=1}^{n} a_j e_j\right) = \rho(w, u),\end{aligned}$$

where the third line holds since $\mathcal{M}(\rho)$ is symmetric. Thus, ρ is a symmetric bilinear form.

Note that (d) clearly implies (c), so we can just show that (a) implies (d). We show this via induction on $n = \dim V$. If $n = 1$, then the statement is clearly true. Now suppose it holds for all numbers less than n.

If $\rho = 0$, then the matrix is the 0 matrix, which is diagonal, so we can assume $\rho \neq 0$. Let $u, w \in V$ such that $\rho(u, w) \neq 0$. Then

$$\begin{gathered}\rho(u + w, u + w) = \rho(u, u) + \rho(u, w) + \rho(w, u) + \rho(w, w) \Rightarrow \\ 2\rho(u, w) = \rho(u + w, u + w) - \rho(u, u) - \rho(w, w).\end{gathered}$$

Since the right side is nonzero, at least one of the values on the left must be nonzero, so there exists $v \in V$ such that $\rho(v, v) \neq 0$.

Let $U = \{u \in V : \rho(u, v) = 0\}$. U is the null space of the linear functional $u \mapsto \rho(u, v)$, which is nonzero since $v \notin U$. Thus $\dim U = n - 1$. By our induction hypothesis, there exists a basis $e_1, ..., e_{n-1}$ of U such taht $\rho|_{U \times U}$ has a diagonal matrix with respect to this basis.

Since $v \notin U$, the list $e_1, ..., e_{n-1}, v$ is a basis of V. By construction, $\rho(e_k, v) = 0$, and since ρ is symmetric, $\rho(v, e_k) = 0$. Thus, with respect to $e_1, ..., e_{n-1}, v$, the matrix of ρ is diagonal. ■

Proposition: Suppose V is a real inner product space and ρ is a symmetric bilinear form on V. Then ρ has a diagonal matrix with respect to some orthonormal basis of V.

Proof: Let $f_1, ..., f_n$ be an orthonormal basis of V, and let $B = \mathcal{M}(\rho, (f_1, ..., f_n))$. Then B is symmetric by the previous result. Let $T \in \mathcal{L}(V)$ such that $\mathcal{M}(T, (f_1, ..., f_n)) = B$. Thus T is self adjoint.

By the real spectral theorem, T has a diagonal matrix with respect to the orthonormal basis $e_1, ..., e_n$. Let $C = \mathcal{M}(I, (e_1, ..., e_n), (f_1, ..., f_n))$. By the change of basis formula, $C^{-1}BC$ is the matrix of T with respect to $e_1, ..., e_n$, so $C^{-1}BC$ is a diagonal matrix. Note that since I is a unitary operator, $\mathcal{M}(C)$ is a unitary matrix. Thus we have

$$\mathcal{M}(\rho, (e_1, ..., e_n)) = C^t BC = C^{-1}BC,$$

where the first equality holds by the change of basis formula for bilinear forms. ■

Definition (alternating bilinear form): A bilinear form $\alpha \in V^{(2)}$ is called *alternating* if

$$\alpha(v, v) = 0$$

for all $v \in V$. The set of alternating bilinear forms on V is denoted by $V^{(2)}_{\text{alt}}$.

Proposition: A bilinear form α on V is alternating if and only if

$$\alpha(u, w) = -\alpha(w, u)$$

for all $u, w \in V$.

Proof: If α is alternating, then

$$\begin{aligned} 0 &= \alpha(u + w, u + w) \\ &= \alpha(u, u) + \alpha(u, w) + \alpha(w, u) + \alpha(w, w) \\ &= \alpha(u, w) + \alpha(w, u) \Rightarrow \alpha(u, w) = -\alpha(w, u). \end{aligned}$$

If $\alpha(u, w) = -\alpha(w, u)$ for all $u, w \in V$, then taking $u, w = v$ yields $\alpha(v, v) = -\alpha(v, v)$, which implies $\alpha(v, v) = 0$ for all $v \in V$. ■

Proposition:

$$V^{(2)} = V^{(2)}_{\text{sym}} \oplus V^{(2)}_{\text{alt}}.$$

Proof: First we show $V^{(2)} = V^{(2)}_{\text{sym}} + V^{(2)}_{\text{alt}}$. For $\beta \in V^{(2)}$, define

$$\rho(u, w) = \frac{\beta(u, w) + \beta(w, u)}{2} \quad \text{and} \quad \alpha(u, w) = \frac{\beta(u, w) - \beta(w, u)}{2}.$$

Then $\rho \in V^{(2)}_{\text{sym}}$, $\alpha \in V^{(2)}_{\text{alt}}$, and $\beta = \rho + \alpha$. Thus $V^{(2)} = V^{(2)}_{\text{sym}} + V^{(2)}_{\text{alt}}$.

Now suppose $\beta \in V^{(2)}_{\text{sym}} \cap V^{(2)}_{\text{alt}}$. Then

$$\beta(u, w) = \beta(w, u) = \beta(u, w)$$

for all $u, w \in V$, which implies $\beta = 0$. ■

9.1.3. Quadratic Forms

Definition (quadratic form): For a bilinear form β on V, define $q_\beta : V \to \mathbb{F}$ by $q_\beta(v) = q(v, v)$. A function $q : V \to \mathbb{F}$ is called a *quadratic form* on V is there exists a bilinear form β on B such that $q = q_\beta$.

Proposition (quadratic forms on $\mathbb{F}^n$): Suppose q is a function from $\mathbb{F}^n$ to $\mathbb{F}$. Then q is a quadratic form on $\mathbb{F}^n$ if and only if there exist numbers $A_{j,k} \in \mathbb{F}$ for $j, k \in \{1, ..., n\}$ such that

$$q(x_1, ..., x_n) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j x_k$$

for all $(x_1, ..., x_n) \in \mathbb{F}^n$.

Proof: First suppose q is a quadratic form on $\mathbb{F}^n$. Thus there exists a bilinear form β on $\mathbb{F}^n$ such that $q = q_\beta$. Let A be the matrix of β with respect to the standard basis of $\mathbb{F}^n$. Then for all $(x_1, ..., x_n) \in \mathbb{F}^n$, we have the desired equation

$$q(x_1, ..., x_n) = \beta((x_1, ..., x_n), (x_1, ..., x_n)) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j x_k.$$

Now suppose $A_{j,k} \in \mathbb{F}$ exists such that

$$q(x_1, ..., x_n) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j x_k$$

for all $(x_1, ..., x_n) \in \mathbb{F}^n$. Define a bilinear form β on $\mathbb{F}^n$ by

$$\beta((x_1, ..., x_n), (x_1, ..., x_n)) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k} x_j x_k.$$

Then $q = q_\beta$, as desired. ■

Proposition (characterization of quadratic forms): Suppose $q : V \to \mathbb{F}$ is a function. The following are eqivalent:

(a) q is a quadratic form.
(b) There exists a unique symmetric bilinear form ρ on V such that $q = q_\rho$.
(c) $q(\lambda v) = \lambda^2 q(v)$ for all $\lambda \in \mathbb{F}$ and all $v \in V$, and the function $(u, w) \mapsto q(u + w) - q(u) - q(w)$ is a symmetric bilinear form on V.
(d) $q(2v) = 4q(v)$ for all $v \in V$, and the function $(u, w) \mapsto q(u + w) - q(u) - q(w)$ is a symmetric bilinear form on V.

Proof: First suppose (a) holds. Thus there exists a bilinear form β such that $q = q_\beta$. We have a ρ symmetric and α alternating such that $\beta = \rho + \alpha$. Thus

$$q = q_\beta = q_\rho + q_\alpha = q_\rho.$$

If ρ' is another symmetric bilinear form such that $q_{\rho'} = q$, then $q_{\rho' - \rho} = 0$, which implies $\rho' - \rho \in V^{(2)}_{\text{sym}} \cap V^{(2)}_{\text{alt}} = \{0\}$, so $\rho' = \rho$. Thus, (b) holds.

Now suppose (b) holds. Let $\rho \in V^{(2)}$ such that $q = q_\rho$. Then

$$q(\lambda v) = \rho(\lambda v, \lambda v) = \lambda^2 \rho(v, v) = \lambda^2 q(v).$$

Also,

$$q(u+w) - q(u) - q(w) = \rho(u+w, u+w) - \rho(u,u) - \rho(w,w) = 2\rho(u,w).$$

Thus, the function in the statment is 2ρ, which is a symmetric bilinear form. Thus (d) holds.

Clearly (c) implies (d). Now suppose (d) holds. Let ρ be the symmetric bilinear form defined by

$$\rho(u,w) = \frac{q(u+w) - q(u) - q(w)}{2}.$$

Then

$$\rho(v,v) = \frac{q(2v) - q(v) - q(v)}{2} = \frac{4q(v) - 2q(v)}{2} = q(v).$$

Thus $q = q_\beta$, so (a) holds. ∎

Proposition (diagonalization of quadratic form): Suppose q is a quadratic form on V.

(a) There exists a basis $e_1, ..., e_n$ of V and $\lambda_1, ..., \lambda_n \in \mathbb{F}$ such that

$$q(x_1 e_1 + \dots + x_n e_n) = \lambda_1 x_1^2 + \dots + \lambda_n x_n^2$$

for all $x_1, ..., x_n \in \mathbb{F}$.

(b) If $\mathbb{F} = \mathbb{R}$ and V is an inner product space, then the basis in (a) can be chosen to be an orthonormal basis.

Proof:

(a) There exists a symmetric bilinear form ρ on V such that $q = q_\rho$. Then there exists a basis $e_1, ..., e_n$ of V such that $\mathcal{M}(\rho, (e_1, ..., e_n))$ is a diagonal matrix. Let $\lambda_1, ..., \lambda_n$ be the diagonal entries. Thus

$$\rho(e_j, e_k) = \begin{cases} \lambda_j & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases}.$$

Then we have

$$\begin{aligned} q(x_1 e_1 + \dots + x_n e_n) &= \rho(x_1 e_1 + \dots + x_n e_n, x_1 e_1 + \dots + x_n e_n) \\ &= \sum_{k=1}^{n} \sum_{j=1}^{n} x_j x_k \rho(e_j, e_k) \\ &= \lambda_1 x_1^2 + \dots + \lambda_n x_n^2, \end{aligned}$$

as desired.

(b) From the previous subsection, we know we can choose the basis for ρ to be orthonormal.

∎

9.1.4. Problems

Problem (Exercise 1): Prove that if β is a bilinear form on $\mathbb{F}$, there exists $c \in \mathbb{F}$ such that

$$\beta(x, y) = cxy$$

for all $x, y \in \mathbb{F}$.

Proof: Note that

$$\beta(x, y) = \beta(1, y)x = \beta(1, 1)xy,$$

so $c = \beta(1, 1)$. ■

Problem (Exercise 3): Suppose $\beta : V \times V \to \mathbb{F}$ is a bilinear form on V and also is a linear functional on $V \times V$. Prove that $\beta = 0$.

Solution: Since $v \mapsto \beta(u, v)$ is a linear functional for fixed u, $\beta(u, 0) = 0$ for all $u \in V$. Similarly, $\beta(0, v) = 0$ for all $v \in V$. Adding these together yields $\beta(u, v) = 0$ for all $u, v \in V$, which implies $\beta = 0$.

Problem (Exercise 4): Suppose V is a real inner product space and β is a bilinear form on V. Show that there exists a unique operator $T \in \mathcal{L}(V)$ such that

$$\beta(u, v) = \langle u, Tv \rangle$$

for all $u, v \in V$.

Solution: Let $e_1, ..., e_n$ be an orthonormal basis of V. Define $T \in \mathcal{L}(V)$ by

$$Te_j = \sum_{i=1}^{n} \beta(e_i, e_j)e_i.$$

Let

$$u = \sum_{j=1}^{n} a_j e_j \quad \text{and} \quad v = \sum_{j=1}^{n} b_j e_j.$$

We have

$$\begin{aligned}
Tv &= \sum_{j=1}^{n} b_j \sum_{i=1}^{n} \beta(e_i, e_j) e_i \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} b_j \beta(e_i, e_j) e_i \\
&= \sum_{i=1}^{n} e_i \sum_{j=1}^{n} \beta(e_i, b_j e_j) \\
&= \sum_{i=1}^{n} \beta(e_i, v) e_i.
\end{aligned}$$

Thus we have

$$\langle u, Tv \rangle = \sum_{i=1}^{n} \beta(e_i, v) a_i = \sum_{i=1}^{n} \beta(a_i e_i, v) = \beta(u, v).$$

Now suppose T' also works. Then

$$0 = \beta(u, v) - \beta(u, v) = \langle u, (T - T')v \rangle.$$

Note that for any v, we must have $(T - T')v \in V^{\perp} = \{0\}$, since the above equation holds for all $u \in V$. Thus, $T - T' = 0$, so T is unique.

Problem (Exercise 9): Suppose that n is a positive integer and $V = \{p \in \mathcal{P}_n(\mathbb{R}) : p(0) = p(1)\}$. Define $\alpha : V \times V \to \mathbb{R}$ by

$$\alpha(p, q) = \int_0^1 pq'.$$

Show that α is an alternating bilinear form on V.

Solution:

$$\alpha(p, p) = \int_0^1 pp' = p^2 \mid_0^1 - \int_0^1 pp' \Rightarrow 2\int_0^1 pp' = p^2(1) - p^2(0) = 0 \Rightarrow \int_0^1 pp' = 0.$$

Problem (Exercise 10): Suppose that n is a positive integer and

$$V = \{p \in \mathcal{P}_n(\mathbb{R}) : p(0) = p(1) \text{ and } p'(0) = p'(1)\}.$$

Define $\rho : V \times V \to \mathbb{R}$ by

$$\rho(p, q) = \int_0^1 pq''.$$

Show that ρ is a symmetric bilinear form on V.

Proof: Note that $\dim V = n - 1$, since we can choose the coefficients of $x^3, x^4, ..., x^n$, and then x and x^2 are fixed, and then we can choose whatever constant term. Thus we have the following basis for V:

$$1, x^3 - \frac{3}{2}x^2 + \left(\frac{3}{2} - 1\right)x, ..., x^n - \frac{n}{2}x^2 + \left(\frac{n}{2} - 1\right)x.$$

Let this basis be $e_0, e_3, ..., e_n$. We have $\rho(e_0, e_k) = \rho(e_k, e_0) = 0$ for any k. We also have that $\rho(e_i, e_j)$ is symmetric in i and j, so $\rho(e_j, e_i) = \rho(e_i, e_j)$. Thus the matrix of ρ is symmetric, so ρ is a symmetric bilinear form. ■

9.2. Alternating Multilinear Forms

9.2.1. Multilinear Forms

Definition (V^m): For m a positive integer, define V^m by

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

Definition (m-linear form): For a positive integer m, an *m-linear form* on V is a function $\beta : V^m \to \mathbb{F}$ that is linear in each slot when the other slots are fixed. This means for each $k \in \{1, ..., m\}$ and all $u_1, ..., u_m \in V$, the function

$$v \mapsto \beta(u_1, ..., u_{k-1}, v, u_{k+1}, ..., u_m)$$

is a linear functional on $\mathbb{F}$. The set of m-linear forms on V is denoted by $V^{(2)}$.

A function β is called a *multilinear form* if it is an m-linear form on V for some positive integer m.

Definition (alternating forms): An m-linear form α on V is called *alternating* if $\alpha(v_1, ..., v_m) = 0$ whenever $v_1, ..., v_m$ is a list of vectors in V with $v_j = v_k$ and $j \neq k$. $V_{\text{alt}}^{(m)}$ denotes the set of all alternating m-linear forms on V.

Proposition: Suppose m is a positive integer and α is an alternating m linear form on V. If $v_1, ..., v_m$ is a linearly dependent list in V, then

$$\alpha(v_1, ..., v_m) = 0.$$

Proof: By the linear dependence lemma, there exists k such that v_k is a linear combination of $v_1, ..., v_{k-1}$. Thus

$$\begin{aligned} \alpha(v_1, ..., v_m) &= \alpha\left(v_1, ..., v_{k-1}, \sum_{j=1}^{k-1} a_j v_j, v_{k+1}, ..., v_m\right) \\ &= \sum_{j=1}^{k-1} a_j \alpha(v_1, ..., v_{k-1}, v_j, v_{k+1}, ..., v_m) = 0. \end{aligned}$$

∎

Proposition: Suppose $m > \dim V$. Then 0 is the only alternating m linear form on V.

Proof: Since $m > \dim V$, any list $v_1, ..., v_m$ in V is linearly dependent. Thus the previous proposition tells us that $\alpha(v_1, ..., v_m) = 0$ for any list, which implies $\alpha = 0$. ■

9.2.2. Alternating Multilinear Forms and Permutations

Proposition: Suppose α is an alternating m linear form on V, and $v_1, ..., v_m$ is a list of vectors in V. Then swapping the vectors in any two slots of $\alpha(v_1, ..., v_m)$ changes the value of α by -1.

Proof: We have

$$\begin{aligned} 0 &= \alpha(v_1 + v_2, v_1 + v_2, v_3, ..., v_m) \\ &= \alpha(v_1, v_2, v_3, ..., v_m) + \alpha(v_2, v_1, v_3, ..., v_m) + \alpha(v_1, v_1, v_3, ..., v_m) + \alpha(v_2, v_2, v_3, ..., v_m) \\ &= \alpha(v_1, v_2, v_3, ..., v_m) + \alpha(v_2, v_1, v_3, ..., v_m) \Rightarrow -\alpha(v_1, v_2, v_3, ..., v_m) = \alpha(v_2, v_1, v_3, ..., v_m). \end{aligned}$$

The same idea shows the result for any other two slots. ■

Definition (sign of a permutation): The *sign* of a permutation $(j_1, ..., j_m)$ is defined by

$$\operatorname{sgn}(j_1, ..., j_m) = (-1)^N,$$

where N is the number of pairs of integer (k, ℓ) with $1 \leq k < \ell \leq m$ such that k appears after ℓ in the list $(j_1, ..., j_m)$.

Proposition (swapping two entries in a permutation): Swppaing two entries in a permutation multiplies the sign of the permutation by -1.

Proof: Just consider the orders of the swapped elements with the elements between them. You get a contribution of ± 1 to N from swapping the two elements, and then for each element between them, you get $\pm 2, 0$ to N. Thus N changes parity, so the sign of the permutation changes. ■

Proposition: Suppose $\alpha \in V_{\text{alt}}^{(m)}$. Then

$$\alpha\left(v_{j_1}, ..., v_{j_m}\right) = (\operatorname{sgn}(j_1, ..., j_m))\alpha(v_1, ..., v_m)$$

for every list $v_1, ..., v_m$ of vectors in V and all $(j_1, ..., j_m) \in \operatorname{perm} m$.

Proof: We can get from $(j_1, ..., j_m)$ to $(1, ..., m)$ by swapping elements. Each swap changes the sign of α and the sign of the remaining permutation by -1. After we're done swapping, we're left with $(1, ..., m)$, which has sign 1. Thus the value of α changed signs an even number times if $\operatorname{sgn}(j_1, ..., j_m) = 1$ and an odd number of times if $\operatorname{sgn}(j_1, ..., j_m) = -1$, which yields the desired result. ■

Proposition (formula for $\dim V$ linear alternating forms on V): Let $n = \dim V$. Suppose $e_1, ..., e_n$ is a basis of V and $v_1, ..., v_n \in V$. For each $k \in \{1, ..., n\}$, let $b_{1,k}, ..., b_{n,k} \in \mathbb{F}$ be such that

$$v_k = \sum_{j=1}^{n} b_{j,k} e_j.$$

Then

$$\alpha(v_1, ..., v_n) = \alpha(e_1, ..., e_n) \sum_{(j_1,...,j_n) \in \text{ perm } n} (\operatorname{sgn}(j_1, ..., j_n)) b_{j_1,1} \cdots b_{j_n,n}$$

for every alternating n-linear form α on V.

Proof:

$$\begin{aligned}
\alpha(v_1, ..., v_n) &= \alpha\left(\sum_{j_1=1}^{n} b_{j_1,1} e_{j_1}, ..., \sum_{j_n=1}^{n} b_{j_n,n} e_{j_n}\right) \\
&= \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} b_{j_1,1} \cdots b_{j_n,n} \alpha\left(e_{j_1}, ..., e_{j_n}\right) \\
&= \sum_{(j_1,...,j_n) \in \text{ perm } n} b_{j_1,1} \cdots b_{j_n,n} \alpha\left(e_{j_1}, ..., e_{j_n}\right) \\
&= \alpha(e_1, ..., e_n) \sum_{(j_1,...,j_n) \in \text{ perm } n} (\operatorname{sgn}(j_1, ..., j_n)) b_{j_1,1} \cdots b_{j_n,n}.
\end{aligned}$$

The third line holds since any tuple with two entries that are equal will be 0 on α, and the last line holds from the previous proposition. ■

Proposition:

$$\dim V_{\text{alt}}^{(\dim V)} = 1.$$

Proof: Let $n = \dim V$. Suppose $\alpha, \alpha' \in V_{\text{alt}}^{(\dim V)}$ with $\alpha \neq 0$. Let $e_1, ..., e_n$ be such that $\alpha(e_1, ..., e_n) \neq 0$. Then there exists $c \in \mathbb{F}$ such that

$$\alpha'(e_1, ..., e_n) = c\alpha(e_1, ..., e_n).$$

Since $\alpha(e_1, ..., e_n) \neq 0$, the converse of the result about linearly dependent lists shows that $e_1, ..., e_n$ is linearly independent, and thus is a basis of V.

Let $v_1, ..., v_n \in V$, and use the same notation as in the last result to denote the coefficients of the v's. Then

$$\begin{aligned}\alpha'(v_1, \ldots, v_n) &= \alpha'(e_1, \ldots, e_n) \sum_{(j_1, \ldots, j_n) \in \text{ perm } n} (\operatorname{sgn}(j_1, \ldots, j_n)) b_{j_1,1} \cdots b_{j_n,n} \\ &= c\alpha(e_1, \ldots, e_n) \sum_{(j_1, \ldots, j_n) \in \text{ perm } n} (\operatorname{sgn}(j_1, \ldots, j_n)) b_{j_1,1} \cdots b_{j_n,n} \\ &= c\alpha(v_1, \ldots, v_n).\end{aligned}$$

Thus the list α, α' is linearly dependent, so $\dim V_{\text{alt}}^{(\dim V)} \leq 1$. Now we just need to show there exists a nonzero alternating $\dim V$ linear form on V to finish.

Let $e_1, \ldots, e_n$ be any basis of V, and let $\varphi_1, \ldots, \varphi_n \in V'$ be its dual basis. Then for $v_1, \ldots, v_n \in V$, define

$$\alpha(v_1, \ldots, v_n) = \sum_{(j_1, \ldots, j_n) \in \text{ perm} \dim n} (\operatorname{sgn}(j_1, \ldots, j_n)) \varphi_{j_1}(v_1) \cdots \varphi_{j_n}(v_n).$$

It's easy to see α is an n linear form.

Suppose $v_1 = v_2$. Then for each permutation $(j_1, j_2, \ldots, j_n)$, the permutation $(j_2, j_1, \ldots, j_n)$ has opposite sign, so the contributions of these permutations in the sum in the definition of α will cancel out. Thus, $\alpha(v_1, v_1, v_3, \ldots, v_n) = 0$. This similarly holds for any two positions on α, so α is alternating.

Now consider $\alpha(e_1, \ldots, e_n)$. Since the φ's are a dual basis, the only term in the sum that isn't 0 is when $(j_1, \ldots, j_n) = (1, \ldots, n)$, and that term contributes 1 to the sum. Thus, $\alpha(e_1, \ldots, e_n) = 1$, so α is nonzero. ■

Proposition: Let $n = \dim V$. Suppose α is a nonzero alternating n-linear form on V and $e_1, \ldots, e_n$ is a list of vectors in V. Then

$$\alpha(e_1, \ldots, e_n) \neq 0$$

if and only if $e_1, \ldots, e_n$ is linearly independent.

Proof: The forward direction follows from the converse of the linearly dependent proposition on alternating forms. For the other direction, suppose $e_1, \ldots, e_n$ is linearly dependent, and is thus a basis of V. Because α is nonzero, there exists $v_1, \ldots, v_n \in V$ such that $\alpha(v_1, \ldots, v_n) \neq 0$. Thus, by the formula for alternating linear forms, $\alpha(e_1, \ldots, e_n) \neq 0$. ■

9.2.3. Problems

Problem (Exercise 1): Show that $\dim V^{(m)} = (\dim V)^m$.

Solution: $n = \dim V$. Let $\mathcal{M}(\alpha)$ be an m-dimensional array of size n such that

$$\mathcal{M}(\alpha)_{i_1, \ldots, i_m} = \alpha\left(e_{i_1}, \ldots, e_{i_m}\right),$$

where $\alpha \in V^{(m)}$ and $1 \leq i_k \leq n$. Then $\alpha \mapsto \mathcal{M}(\alpha)$ is a linear map.

For an m-dimensional array of size n, define $\alpha_A \in V^{(m)}$ such that

$$\alpha_A = \left(\sum_{i_1=1}^{n} x_{i_1,1} e_{i_1}, \ldots, \sum_{i_m=1}^{n} x_{i_m,m} e_{i_m}\right) = \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} A_{i_1,\ldots,i_m} x_{i_1,1} \cdots x_{i_m,m},$$

where $x_{i_j,k} \in \mathbb{F}$. The maps $\alpha \mapsto \mathcal{M}(\alpha)$ and $A \mapsto \alpha(A)$ are inverses of each, which can be proven in the same way as be proved to show that $\mathbb{F}^{n,n}$ and $V^{(2)}$ are isomorphisms. Since it's clear that the vector space of all m-dimensionals arrays of size n has dimension n^m, we have $\dim V^{(n)} = n^m = (\dim V)^m$.

9.3. Determinants

9.3.1. Defining the Determinant

Definition (α_T): Suppose $T \in \mathcal{L}(V)$. For $\alpha \in V_{\text{alt}}^{(m)}$, define $\alpha_T \in V_{\text{alt}}^{(m)}$ by

$$\alpha_T(v_1, ..., v_m) = \alpha(Tv_1, ..., Tv_m)$$

for every list $v_1, ..., v_m$ of vectors in V.

Note that $\alpha_T \in V_{\text{alt}}^{(m)}$.

Definition (determinant of an operator): Suppose $T \in \mathcal{L}(V)$. The *determinant* of T, denoted by $\det T$, is defined to be the unique number in $\mathbb{F}$ such that

$$\alpha_T = (\det T)\alpha$$

for all $\alpha \in V_{\text{alt}}^{(m)}$.

Definition (determinant of a matrix): Suppose A is an n by n matrix with entries in $\mathbb{F}$. Let $T \in \mathcal{L}(\mathbb{F}^n)$ be the operator whose matrix with respect to the standard basis of $\mathbb{F}^n$ equals A. The *determinant* of A, denoted $\det A$, is defined by $\det A = \det T$.

The notation $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ denotes an n by n matrix, where v_i is an n by 1 column vector.

Proposition: The map that takes a list $v_1, ..., v_n$ of vectors to $\det\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is an alternating n linear form on $\mathbb{F}^n$.

Proof: Let $e_1, ..., e_n$ be the standard basis of $\mathbb{F}^n$. Let $T \in \mathcal{L}(\mathbb{F}^n)$ such that $Te_k = v_k$. Thus the matrix of T with respect to $e_1, ..., e_n$ is $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$. Thus $\det\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} = \det T$ by definition.

Let α be an alternating n-linear form on $\mathbb{F}^n$ such that $\alpha(e_1, ..., e_n) = 1$. Then

$$\begin{aligned} \det\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} &= \det T \\ &= (\det T)\alpha(e_1, ..., e_n) \\ &= \alpha(Te_1, ..., Te_n) \\ &= \alpha(v_1, ..., v_n). \end{aligned}$$

Thus the map that takes $v_1, ..., v_n$ to $\det\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is the alternating n linear form on $\mathbb{F}^n$. ∎

Proposition (Leibniz formula for determinants): Suppose A is an n by n square matrix. Then

$$\det A = \sum_{(j_1,...,j_n)\in \text{ perm } n} (\operatorname{sgn}(j_1, ..., j_n)) A_{j_1,1} \cdots A_{j_n,n}.$$

Proof: Apply formula for alternating linear forms with $V = \mathbb{F}^n$, $e_1, ..., e_n$ as the standard basis of $\mathbb{F}^n$, and α as the alternating n linear form that takes $v_1, ..., v_n$ to $\det(\, v_1 \cdots v_n \,)$. If each v_k is the kth column of A, then each $b_{j,k}$ in the formula equals $A_{j,k}$. Finally,

$$\alpha(e_1, ..., e_n) = \det(\, e_1 \cdots e_n \,) = \det I = 1.$$

Thus we have our desired formula. ■

Proposition (determinant of upper triangular matrix): Suppose A is an upper triangular matrix with $\lambda_1, ..., \lambda_n$ on the diagonal. Then $\det A = \lambda_1 \cdots \lambda_n$.

Proof: The identity permutation is the only one that produced a nonzero term in the formula above, and thus the determinant is the product of the diagonal entries. ■

9.3.2. Properties of Determinants

Proposition (multiplicative):
(a) Suppose $S, T \in \mathcal{L}(V)$. Then $\det(ST) = (\det S)(\det T)$.
(b) Suppose A and B are square matrices of the same size. Then

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

Proof:
(a) Suppose $\alpha \in V_{\text{alt}}^{(\dim V)}$ and $v_1, ..., v_n \in V$. Then

$$\begin{aligned} \alpha_{ST}(v_1, ..., v_n) &= \alpha(STv_1, ..., STv_n) \\ &= (\det S)\alpha(Tv_1, ..., Tv_n) \\ &= (\det S)(\det T)\alpha(v_1, ..., v_n). \end{aligned}$$

Thus $\det ST = (\det S)(\det T)$.
(b) Let $S, T \in \mathcal{L}(\mathbb{F}^n)$ be such that $\mathcal{M}(S) = A$ and $\mathcal{M}(T) = B$ with respect to the standard basis. Then

$$\det(AB) = \det(ST) = (\det S)(\det T) = (\det A)(\det B).$$

■

Proposition: An operator $T \in \mathcal{L}(V)$ is invertible if and only if $\det T \neq 0$. Furthermore, if T is invertible, then $\det(T^{-1}) = \frac{1}{\det T}$.

Proof: If T is invertible, then $TT^{-1} = I \Rightarrow \det(TT^{-1}) = \det I \Rightarrow \det(T)\det(T^{-1}) = 1$. Thus $\det T$ is nonzero and $\det T^{-1} = \frac{1}{\det T}$.

Now suppose $\det T \neq 0$ and $0 \neq v \in V$. Let $v, e_2, ..., e_n$ be a basis of V and let $0 \neq \alpha \in V_{\text{alt}}^{(\dim V)}$. Then $\alpha(v, e_2, ..., e_n) \neq 0$. Thus

$$\alpha(Tv, Tv_2, ..., Tv_n) = (\det T)\alpha(v, e_2, ..., e_n) \neq 0,$$

which means $Tv \neq 0$. Thus T is invertible. ■

Proposition: Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. Then λ is an eigenvalue of T if and only if $\det(\lambda I - T) = 0$.

Proof: Is λ is an eigenvalue, then $\lambda I - T$ is not injective, which means it's not invertible, which implies $\det(\lambda I - T) = 0$. If $\det(\lambda I - T) = 0$, then $\lambda I - T$ is not injective, which implies the existence of an eigenvector for λ. ■

Proposition (determinant is similarity invariant): Suppose $T \in \mathcal{L}(V)$ and $S : W \to V$ is an invertible linear map. Then

$$\det(S^{-1}TS) = \det T.$$

Proof: Let $n = \dim V = \dim W$, and suppose $\tau \in W_{\text{alt}}^{(n)}$. Define $\alpha \in V_{\text{alt}}^{(n)}$ by

$$\alpha(v_1, ..., v_n) = \tau(S^{-1}v_1, ..., S^{-1}v_n).$$

Suppose $w_1, ..., w_n \in W$. Then

$$\begin{aligned}
\tau_{S^{-1}TS}(w_1, ..., w_n) &= \tau(S^{-1}TSw_1, ..., S^{-1}TSw_n) \\
&= \alpha(TSw_1, ..., TSw_n) \\
&= \alpha_T(Sw_1, ..., Sw_n) \\
&= \det(T)\alpha(Sw_1, ..., Sw_n) \\
&= (\det T)\tau(w_1, ..., w_n).
\end{aligned}$$

Thus $\det(S^{-1}TS) = \det T$. ■

Proposition: Suppose $T \in \mathcal{L}(V)$ and $e_1, ..., e_n$ is a basis of V. Then

$$\det T = \det \mathcal{M}(T, (e_1, ..., e_n)).$$

Proof: Let $f_1, ..., f_n$ be the standard basis of $\mathbb{F}^n$. Let $S : \mathbb{F}^n \to V$ be the linear map such that $S f_k = e_k$ (change of basis matrix). Thus $\mathcal{M}(S, (f_1, ..., f_n), (e_1, ..., e_n))$ and $\mathcal{M}(S^{-1}, (e_1, ..., e_n), (f_1, ..., f_n))$ both equal the identity matrix. By the change of basis formula,

$$\mathcal{M}(S^{-1}TS, (f_1, ..., f_n)) = \mathcal{M}(T, (e_1, ..., e_n)).$$

Thus

$$\begin{aligned}\det T &= \det(S^{-1}TS) \\ &= \det \mathcal{M}(S^{-1}TS, (f_1, ..., f_n)) \\ &= \det \mathcal{M}(T, (e_1, ..., e_n)).\end{aligned}$$

■

Proposition: Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then $\det T$ equals the product of the eigenvalues of T, with each eigenvalue included as many times as its multiplicity.

Proof: Follows since there exists an upper triangular matrix for T with eigenvalues repeated with mutliplicity on the diagonal. ■

Proposition (determinant of transpose, dual, adjoint):
(a) Suppose A is a square matrix. Then $\det A^t = \det A$.
(b) Suppose $T \in \mathcal{L}(V)$. Then $\det T' = \det T$.
(c) Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Then $\det(T^*) = \overline{\det T}$.

Proof:

(a) Define $\alpha : (\mathbb{F}^n)^n \to \mathbb{F}$ by $\alpha((v_1 \cdots v_n)) = \det\left((v_1 \cdots v_n)^t\right)$ for all $v_1, ..., v_n \in \mathbb{F}^n$. By the Leibniz formula, this is an n-linear form (when multiplying by a constant, each term in the sum will grow by that constant. When adding two elements, each term in the sum will distribute out to the original two elements).

   Suppose $v_j = v_k$ for distinct j, k. Then $(v_1 \cdots v_n)^t B$ cannot be the identity for some n-by-n matrix B, as row j and k of $(v_1 \cdots v_n)^t B$ are equal. Thus $(v_1 \cdots v_n)^t$ us not invertible, which implies $\alpha((v_1 \cdots v_n)) = \det\left((v_1 \cdots v_n)^t\right) = 0$. Thus α is alternating as well (we could also calculate the rank of $(v_1 \cdots v_n)^t$ to show it's non invertible).

   Finally, note that α applied to the standard basis of $\mathbb{F}^n$ equals 1. Because the space of n linear forms on $\mathbb{F}^n$ has dimension 1, α must be the determinant function.

(b) Follows since the matrix of T' is the transpose of the matrix of T.

(c) With respect to an orthonormal basis, the matrix of T^* is the conjugate transpose of the matrix of T. Thus using the Leibniz and (a) yields the desired result.

■

Corollary: The following are all methods to help evaluate determinants. Each follow from the fact that the determinant on rows or columns is an alternating linear form.

(a) If either two columns or two rows of a square matrix are equal, then the determinant of the matrix equals 0.
(b) Suppose A is a square matrix and B is the matrix obtained from A by swapping either two columns or two rows. Then $\det A = -\det B$.
(c) If one column or row of a square matrix is multiplied by a scalar, then the value of the determinant is multiplied by the same scalar.
(d) If a scalar multiple of one row/column of a square matrix is added to another row/column, then the value of the determinant is unchanged.

Proposition: Suppose V is an inner product space and $S \in \mathcal{L}(V)$ is unitary. Then $|\det S| = 1$.

Proof: Since S is unitary, we have $I = S^* S$. Thus

$$1 = \det I = \det(S^*)(\det S) = \overline{\det S}(\det S) = |\det S|^2 \Rightarrow 1 = |\det S|.$$

■

Proposition: Suppose V is an inner product space and $T \in \mathcal{L}(V)$ is a positive operator. Then $\det T \geq 0$.

Proof: By the spectral, we pick an orthonormal eigenbasis of V. Note that since T is positive, every eigenvalue is positive, which implies that the product of them all, which implies the result.

Remark: Note we haven't proved that the product of eigenvalues is the determinant on real vector spaces, since on real vectors spaces we can't gurantee an upper triangular matrix. However, since the operator is positive, this isn't an issue.

■

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

$$|\det T| = \sqrt{\det(T^* T)} = \text{product of singular values of } T.$$

Proof: We have

$$|\det T|^2 = \overline{\det T}(\det T) = (\det(T^*))(\det T) = \det(T^* T).$$

Taking square roots yields the desired result.

Let $s_1, ..., s_n$ be the singular values of T. Then $s_1^2, ..., s_n^2$ is the list of eigenvalues of T^*T. Thus we have $\det(T^*T) = (s_1 \cdots s_n)^2$. Taking square roots yields the desired result. ■

Corollary: Suppose $T \in \mathcal{L}(\mathbb{R}^n)$ and $\Omega \subseteq \mathbb{R}^n$. Then

$$\text{volume } T(\Omega) = |\det T|(\text{volume } \Omega).$$

Proof: Follows since the volume scaling caused by a map is equal to the product of the singular values. ■

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

$$\det(zI - T) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}.$$

Proof: There exists a basis of V for which T is upper triangular with eigenvalues repeated by multiplicity on the diagonal. Thus the matrix of $zI - T$ with respect to that matrix has $z - \lambda_k$ appear on the diagonal d_k times, which implies the result. ■

Remark: The definition and theorem that follow now hold on real vector spaces as well as complex ones.

Definition (characteristic polynomial): Suppose $T \in \mathcal{L}(V)$. The polynomial defined by

$$z \mapsto \det(zI - T)$$

is called the *characteristic polynomial* of T.

Theorem (Cayley-Hamilton): If q is the characteristic polynomial of $T \in \mathcal{L}(V)$, then $q(T) = 0$.

Proof: We already proved the complex case, so suppose $\mathbb{F} = \mathbb{R}$. Pick a basis of V and let A be the matrix of T with respect to this basis. Let S be the operator on $\mathbb{C}^{\dim V}$ such that the matrix of S with respect to the standard basis of $\mathbb{C}^{\dim V}$ is A. For all $z \in \mathbb{R}$, we have

$$q(z) = \det(zI - T) = \det(zI - A) = \det(zI - S).$$

Since q agrees with $\det(zI - S)$ on all reals, q must be the characteristic polynomial of S. Then Cayley-Hamilton on $\mathbb{C}$ implies $0 = q(S) = q(A) = q(T)$. ■

Proposition: Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then the characteristic polynomial of T can be written as

$$z^n - (\operatorname{tr} T)z^{n-1} + \cdots + (-1)^n(\det T).$$

Proof: To get the constant term, just plug in 0, which yields $\det(-T) = (-1)^n \det T$.

Now fix a basis of T, and let A be the matrix of T with respect to that basis. The matrix of $zI - T$ with respect to that basis is $zI - A$. The term coming from the identity permutation is the Leibniz formula is

$$(z - A_{1,1})\cdots(z - A_{n,n}).$$

The coefficient of z^n is 1, and the coefficient of z^{n-1} is $-(A_{1,1} + \cdots + A_{n,n}) = -\operatorname{tr} T$. The other terms in the Leibniz formula come from other non identity permutations, which can contain only at most $n - 2$ factors from the diagonal, so nothing else contributes to the coefficient of z^n and z^{n-1}. ■

Proposition (Hadamard's inequality): Suppose A is an n-by-n matrix. Let $v_1, \ldots, v_n$ denote the columns of A. Then

$$|\det A| \leq \prod_{k=1}^{n} \|v_k\|.$$

Proof: If A is not invertible, then $\det A = 0$ and the inequality holds, so suppose A is invertible. Thus the rank of A is n, and so the columns must be linearly independent, which implies unique unitary Q and upper triangular R with only positive entries on the diagonal such that $A = QR$. We have

$$\begin{aligned}
|\det A| &= |\det Q||\det R| \\
&= |\det R| \\
&= \prod_{k=1}^{n} R_{k,k} \\
&\leq \prod_{k=1}^{n} \|R_{\cdot,k}\| \\
&= \prod_{k=1}^{n} \|QR_{\cdot,k}\| \\
&= \prod_{k=1}^{n} \|v_k\|,
\end{aligned}$$

where the 5th line holds becase Q is unitary. ■

Proposition (Vandermonde matrix): Suppose $n > 1$ and $\beta_1, ..., \beta_n \in \mathbb{F}$. Then

$$\det \begin{pmatrix} 1 & \beta_1 & \beta_1^2 & \cdots & b_1^{n-1} \\ 1 & \beta_2 & \beta_2^2 & \cdots & b_2^{n-1} \\ & & \ddots & & \\ 1 & \beta_n & \beta_n^2 & \cdots & b_n^{n-1} \end{pmatrix} = \prod_{1 \leq j < k \leq n} (\beta_k - \beta_j).$$

Proof: Let $1, z, ..., z^{n-1}$ be the standard basis of $\mathcal{P}_{n-1}(\mathbb{F})$ and let $e_1, ..., e_n$ be the standard basis of $\mathbb{F}^n$. Define $S : \mathcal{P}_{n-1}(\mathbb{F}) \to \mathbb{F}^n$ by

$$Sp = (p(\beta_1), ..., p(\beta_n)).$$

Let A denote the Vandermonde matrix. Note that $A = \mathcal{M}(S, (1, z, ..., z^{n-1}), (e_1, ..., e_n))$. Let $T : \mathcal{P}_{n-1}(\mathbb{F}) \to \mathcal{P}_{n-1}(\mathbb{F})$ be such that $T(1) = 1$ and

$$T(z^k) = (z - \beta_1)\cdots(z - \beta_k).$$

Let $B = \mathcal{M}(T, (1, z, ..., z^{n-1}))$. Then B is upper triangular with 1's on the diagonal, so $\det B = 1$.

Let $C = \mathcal{M}(ST, (1, z, ..., z^{n-1}), (e_1, ..., e_n))$. Thus $C = AB$, which implies $\det A = (\det A)(\det B) = \det C$. By the definition of the maps, we can see that C is

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & \beta_1 - \beta_2 & 0 & \cdots & 0 \\ 1 & \beta_3 - \beta_1 & (\beta_3 - \beta_1)(\beta_3 - \beta_2) & \cdots & 0 \\ & & & \ddots & \\ 1 & \beta_n - \beta_1 & (\beta_n - \beta_1)(\beta_n - \beta_2) & \cdots & (\beta_n - \beta_1)(\beta_n - \beta_2)\cdots(\beta_n - \beta_{n-1}) \end{pmatrix},$$

Since $\det C^t = \det C = \det A$, we have

$$\det A = \prod_{1 \leq j < k \leq n} (\beta_k - \beta_j),$$

as desired.

■

9.3.3. Problems

Problem (Exercise 3): Suppose T is nilpotent. Prove that $\det(T + I) = 1$.

Solution: Conside the matrix of $T + I$ with respect to a Jordan basis. Then the matrix of $T + I$ with respect to that basis is upper triangular with 1's on the diagonal, so the result follows.

Problem (Exercise 4): Suppose $S \in \mathcal{L}(V)$. Prove that S is unitary if and only if $|\det S| = \|S\| = 1$.

Solution: If S is unitary then the result follows by properties of unitary operators.

Let $s_1, ..., s_n$ be the singular values of S from smallest to largest. If the equation holds, then $s_n = 1$, but we also have $s_1 \cdots s_n = s_n \Rightarrow s_1 \cdots s_{n-1} = 1$, which implies all the singular values of S are 1, which means S is unitary.

Problem (Exercise 5): Suppose A is a block upper-triangular matrix

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

where each A_k along the diagonal is a square matrix. Prove that

$$\det A = (\det A_1) \cdots (\det A_m).$$

Solution: Note that we can row reduce each square matrix into an upper triangular matrix without messing with the rows in the other matrices. Thus $\det A$ is the product of all the elements along the diagonal, but the product of the elements along the diagonal of A_k equals $\det A_k$, so the result follows.

Problem (Exercise 12): Suppose $S, T \in \mathcal{L}(V)$ and S is invertible. Define $p : \mathbb{F} \to \mathbb{F}$ by

$$p(z) = \det(zS - T).$$

Prove that p is a polynomial of degree $\dim V$ and that the coefficient of $z^{\dim V}$ in this polynomial is $\det S$.

Solution: We can write

$$p(z) = \det\left(S\left(zI - S^{-1}T\right)\right) = (\det S) \det\left(zI - S^{-1}T\right).$$

The second term is a degree $\dim V$ polynomial with leading coefficient 1, so the result follows.

Problem (Exercise 14): Suppose V is an inner product space and T is a positive operator on V. Prove that

$$\det \sqrt{T} = \sqrt{\det T}.$$

Solution: Pick an orthonormal eigenbasis of T. With respect to that basis, the matrix of T is diagonal only nonngeative numbers on the diagonal. The matrix of $\sqrt{T}$ with respect to that basis is diagonal and the entries are just the square roots of the entries of the matrix of T. Thus the result follows easily.

Problem (Exercise 15): Suppse V is an inner product space and $T \in \mathcal{L}(V)$. Use polar decomposition to give a proof that

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

Solution: By polar decomposition and the last exercise we have

$$\det T = (\det S)\left(\sqrt{\det(T^* T)}\right).$$

Taking absolute values yields

$$|\det T| = \left|\sqrt{\det(T^* T)}\right| = \sqrt{\det(T^* T)}.$$

Problem (Exercise 16): Suppose $T \in \mathcal{L}(V)$. Define $g : \mathbb{F} \to \mathbb{F}$ by $g(x) = \det(I + xT)$. Show that $g'(0) = \operatorname{tr} T$.

Solution: Let $n = \dim V$. Note that

$$g\left(-\frac{1}{x}\right) = \frac{1}{x^n}\det(xI - T) = \frac{1}{x^n}p(x),$$

where $p(x)$ is the characteristic polynomial of T. Taking the derivative of both sides to get

$$\frac{1}{x^2}g'\left(-\frac{1}{x}\right) = -\frac{n}{x^{n+1}}p(x) + \frac{1}{x^n}p'(x) \Rightarrow g'\left(-\frac{1}{x}\right) = -\frac{n}{x^{n-1}}p(x) + \frac{1}{x^{n-2}}p'(x).$$

We can write the right side as

$$-\frac{n}{x^{n-1}}(x^n - (\operatorname{tr} T)x^{n-1} + \cdots) + \frac{1}{x^{n-2}}(nx^{n-1} - (n-1)(\operatorname{tr} T)x^{n-2} + \cdots) =$$
$$-nx + n(\operatorname{tr} T) + nx - (n-1)(\operatorname{tr} T) + \cdots = \operatorname{tr} T + \cdots.$$

The other terms not written out are all of the form $\frac{c}{x^k}$, for some $c \in \mathbb{F}$ and positive integer k. Thus, taking the limit to infinity on both sides yields

$$\lim_{x \to \infty} g'\left(-\frac{1}{x}\right) = g'(0) = \operatorname{tr} T.$$

Problem (Exercise 17): Suppose a, b, c are positive numbers. Find the volume of the ellipsoid

$$\left\{(x, y, z) \in \mathbb{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} < 1\right\}.$$

Solution: Let Ω be the unit ball in $\mathbb{R}^3$, which has volume $\frac{4}{3}\pi$. Define $T \in \mathcal{L}(\mathbb{R}^3)$ by $T(x, y, z) = (ax, by, cz)$. Then $T(\Omega)$ equals the ellipse defined in the problem. The determinant of T is abc, so the volume of the ellipse is $\frac{4}{3}\pi abc$.

Problem (Exercise 18): Suppose A is an invertible square matrix. Prove that Hadamard's inequality is an equality if and only if the columns of A form an orthogonal list.

Proof: First suppose the inequality is an equality. Then we have

$$\prod_{k=1}^{n} R_{k,k}^2 = \prod_{k=1}^{n} \left\|R_{\cdot,k}\right\|^2.$$

Note that $R_{k,k}^2 \leq R_{k,k}^2 + \left|R_{k-1,k}\right|^2 + \cdots + \left|R_{1,k}\right|^2 = \left\|R_{\cdot,k}\right\|^2$ for each k. Since the product of this inequality over all k is an equality, then the inequality must be an equality as well, which implies $R_{j,k} = 0$ for $j \neq k$, which implies R is a diagonal matrix. Next note that right multiplying by a diagonal matrix simply scales the columns of the other matrix by the diagonal entries. Since Q is unitary, its columns form an orthogonal list, which means the columns of $A = QR$ form an orthogonal list.

Now suppose the columns of A form an orthogonal list. By the definition of R is the QR factorization, R will be diagonal, and thus $R_{k,k} = \left\|R_{\cdot,k}\right\|$. ■

Problem (Exercise 20): Suppose A is an n-by-n matrix, and suppose c is such that $\left|A_{j,k}\right| \leq c$. Prove that

$$|\det A| \leq c^n n^{\frac{n}{2}}.$$

Solution: Note that the norm of each column is at most $c\sqrt{n}$, so the result follows by Hadamard's inequality.

9.4. Tensor Products

9.4.1. Tensor Product of Two Vector Spaces

Definition (bilinear form on $V \times W$): A *bilinear functional* on $V \times W$ is a function $\beta : V \times W \to \mathbb{F}$ such that $v \mapsto \beta(v, w)$ is a linear functional on V for each $w \in W$ and $w \mapsto \beta(v, w)$ is a linear functional on W for each $v \in V$. The vector space of bilinear functionals on $V \times W$ is denoted $\mathcal{B}(V, W)$.

Proposition:

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

Proof: Essentially the same proof we used to show that $\dim V^{(2)} = (\dim V)^2$, except we now map to a $\dim V$ by $\dim W$ matrix. ∎

Definition (tensor product): The *tensor product* $V \otimes W$ is defined to be $\mathcal{B}(V', W')$. For $v \in V$ and $w \in W$, the tensor product $v \otimes w$ is the element of $V \otimes W$ defined by

$$(v \otimes w)(\varphi, \tau) = \varphi(v)\tau(w)$$

for all $(\varphi, \tau) \in V' \times W'$.

Corollary:

$$\dim(V \otimes W) = (\dim V)(\dim W)$$

Proof:

$$\dim(V \otimes W) = \dim \mathcal{B}(V', W') = (\dim V')(\dim W') = (\dim V)(\dim W).$$

∎

Proposition (bilinearity of tensor product): Suppose $v_1, v_2, v \in V$ and $w_1, w_2, w \in W$ and $\lambda \in \mathbb{F}$. Then

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w \quad \text{and } v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$$

and

$$\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w).$$

Proof: Suppose $(\varphi, \tau) \in V' \times W'$. Then

$$\begin{aligned} ((v_1 + v_2) \otimes w)(\varphi, \tau) &= \varphi(v_1 + v_2)\tau(w) \\ &= \varphi(v_1)\tau(w) + \varphi(v_2)\tau(w) \\ &= (v_1 \otimes w)(\varphi, \tau) + (v_2 \otimes w)(\varphi, \tau) \\ &= (v_1 \otimes w + v_2 \otimes w)(\varphi, \tau). \end{aligned}$$

Thus $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$. The other two equations follows similarly. ■

Proposition (basis of $V \otimes W$): Suppose $e_1, ..., e_m$ is a list of vectors in V and $f_1, ..., f_n$ is a list of vectors in W.

(a) If $e_1, ..., e_m$ and $f_1, ..., f_n$ are both linearly independent lists, then $\{e_j \otimes f_k\}_{j=1,...,m;k=1,...,n}$ is a linearly independent list in $V \otimes W$.

If $e_1, ..., e_m$ and $f_1, ..., f_n$ are both bases of their respective vector spaces, then $\{e_j \otimes f_k\}_{j=1,...,m;k=1,...,n}$ is a basis of $V \otimes W$.

Proof: To prove (a) pick the "dual basis" for each list in V' and W' (i.e. $\varphi_i(v_j) = \delta_{ij}$ where $\varphi_i \in V'$ and v_j is in V). Let $\varphi_1, ..., \varphi_m$ be this "dual basis" for V and let $\tau_1, ..., \tau_n$ be this "dual basis" for W. Now suppose

$$\sum_{k=1}^{n} \sum_{j=1}^{m} a_{j,k}(e_j \otimes f_k) = 0.$$

Note that $(e_j \otimes f_k)(\varphi_M, \tau_N)$ equals 1 if $j = M, k = N$ and 0 otherwise. Thus, applying the equation above to (φ_m, τ_N) shows $a_{M,N} = 0$, and thus the list is linearly independent. (b) follows directly since the list in $V \otimes W$ has length $(\dim V)(\dim W) = \dim(V \otimes W)$. ■

Definition (bilinear map): A *bilinear map* from $V \times W$ to a vector space U is a function $\Gamma : V \times W \to U$ such that $v \mapsto \Gamma(v, w)$ is a linear map from V to U for each $w \in W$ and $w \mapsto \Gamma(v, w)$ is a linear map from W to U for each $v \in V$.

Proposition (converting bilinear maps to linear maps): Suppose U is a vector space.

(a) Suppose $\Gamma : V \times W \to U$ is a bilinear map. Then there exists a unique linear map $\hat{\Gamma} : V \otimes W \to U$ such that

$$\hat{\Gamma}(v \otimes w) = \Gamma(v, w)$$

for all $(v, w) \in V \times W$.

(b) Suppose $T : V \otimes W \to U$ is a linear map. Then there exists a unique bilinear map $T^{\#} : V \times W \to U$ such that

$$T^{\#}(v, w) = T(v \otimes w)$$

for all $(v, w) \in V \times W$.

Proof: Let $e_1, ..., e_m$ be a basis of V and let $f_1, ..., f_n$ be a basis of W. Then we define the unique map $\hat{\Gamma} : V \otimes W \to U$ by

$$\hat{\Gamma}(e_j \otimes f_k) = \Gamma(e_j, f_k).$$

Now suppose $(v, w) \in V \times W$. Thus we can write $v = \sum a_i e_i$ and $w = \sum b_j f_j$. Then

$$\begin{aligned}\hat{\Gamma}(v \otimes w) &= \hat{\Gamma}\left(\sum_{k=1}^{n}\sum_{j=1}^{m} a_j b_k (e_j \otimes f_k)\right) \\ &= \sum_{k=1}^{n}\sum_{j=1}^{m} a_j b_k \hat{\Gamma}(e_j \otimes f_k) \\ &= \sum_{k=1}^{n}\sum_{j=1}^{m} a_j b_k \Gamma(e_j, f_k) = \Gamma(v, w).\end{aligned}$$

This map is unique because we must have $\hat{\Gamma}(e_j \otimes f_k) = \Gamma(e_j, f_k)$.

For (b), define $T^{\#} : V \times W \to U$ by $T^{\#}(v, w) = T(v \otimes W)$ for all $(v, w) \in V \times W$. It's easy to see that $T^{\#}$ is bilinear, and since again we must have $T^{\#}(e_j, f_k) = T(e_j \otimes f_k)$, the map is unique. ■

9.4.2. Tensor Product of Inner Product Spaces

Proposition: Suppose V and W are inner product spaces. Then there is a unique inner product on $V \otimes W$ such that

$$\langle v \otimes w, u \otimes x \rangle = \langle v, u \rangle \langle w, x \rangle$$

for all $v, u \in V$ and $w, x \in W$.

Proof: Suppose $e_1, ..., e_m$ is an orthonormal basis of V and $f_1, ..., f_n$ is an orthonormal basis of W. Define an inner product on $V \otimes W$ by

$$\left\langle \sum_{k=1}^{n}\sum_{j=1}^{m} b_{j,k} e_j \otimes f_k, \sum_{k=1}^{n}\sum_{j=1}^{m} c_{j,k} e_j \otimes f_k \right\rangle = \sum_{k=1}^{n}\sum_{j=1}^{m} b_{j,k}\overline{c_{j,k}}.$$

We can easily verify that this is an inner product. Write $v = v_1 e_1 + \cdots + v_m e_m$ and similarly for the other vectors. Then

$$\begin{aligned}
\langle v \otimes w, u \otimes x \rangle &= \left\langle \sum_{j=1}^{m} v_j e_j \otimes \sum_{k=1}^{n} w_k f_k, \sum_{j=1}^{m} u_j e_j \otimes \sum_{k=1}^{n} x_k f_k \right\rangle \\
&= \left\langle \sum_{k=1}^{n}\sum_{j=1}^{m} v_j w_k e_j \otimes f_k, \sum_{k=1}^{n}\sum_{j=1}^{m} u_j x_k e_j \otimes f_k \right\rangle \\
&= \sum_{k=1}^{n}\sum_{j=1}^{m} v_j \overline{u_j} w_k \overline{x_k} \\
&= \left(\sum_{j=1}^{m} v_j \overline{u_j}\right)\left(\sum_{k=1}^{n} w_k \overline{x_k}\right) \\
&= \langle v, u\rangle \langle w, x\rangle.
\end{aligned}$$

There is only one inner product on $V \otimes W$ such that $\langle v \otimes w, u \otimes x \rangle = \langle v, u\rangle\langle w, x\rangle$ because every element of $V \otimes W$ can be written as a linear combination of elements of the form $v \otimes w$.

Remark: Not gonna lie, don't really understand that last part, so I'll probably come back later and add extra detail once I understand it.

■

Definition: Suppose V and W are inner product spaces. Then inner product on $V \otimes W$ is the unique function $\langle \cdot, \cdot \rangle$ from $(V \otimes W) \times (V \otimes W)$ to $\mathbb{F}$ such that

$$\langle v \otimes w, u \otimes x \rangle = \langle v, u\rangle\langle w, x\rangle$$

for all $v, u \in V$ and $w, x \in W$.

Proposition (orthonormal basis of $V \otimes W$): Suppose V and W are inner product spaces, and $e_1, \ldots, e_m$ is an orthonormal basis of V and $f_1, \ldots, f_n$ is an orthonormal basis of W. Then

$$\{e_j \otimes f_k\}_{j=1,\ldots,m; k=1,\ldots,n}$$

is an orthonormal basis of $V \otimes W$.

Proof: We already know it's a basis of $V \otimes W$, so we only need to prove orthonormality. However, this follows pretty easily, since

$$\langle e_j \otimes f_k, e_M \otimes f_n \rangle = \langle e_j, e_M\rangle\langle f_k, f_N\rangle = \delta_{jM}\delta_{kN}.$$

■

9.4.3. Tensor Products of Multiple Vector Spaces

Remark: Every proposition below will be stated without proof, since the proofs are essentially identical to the ones for the tensor product of two vector spaces. I'm also tired of writing indexes and subindexes and all that.

Definition (m-linear functional): An *m-linear functional* on $V_1 \times \cdots \times V_m$ is a function $\beta : V_1 \times \cdots \times V_m \to \mathbb{F}$ that is a linear functional in each slot when the other slots are held fixed. The vector space of m-linear functionals on $V_1 \times \cdots \times V_m$ is denoted by $\mathcal{B}(V_1, \ldots, V_m)$.

Proposition:

$$\dim \mathcal{B}(V_1 \times \cdots \times V_m) = (\dim V_1) \times \cdots \times (\dim V_m)$$

Definition (tensor product): The *tensor product* $V_1 \otimes \cdots \otimes V_m$ is defined to be $\mathcal{B}(V_1', \ldots, V_m')$. For $v_1 \in V_1, \ldots, v_m \in V_m$, the *tensor product* $v_1 \otimes \cdots \otimes v_m$ is the element of $V_1 \otimes \cdots \otimes V_m$ defined by

$$(v_1 \otimes \cdots \otimes v_m)(\varphi_1, \ldots, \varphi_m) = \varphi_1(v_1) \cdots \varphi_m(v_m)$$

for all $(\varphi_1, \ldots, \varphi_m) \in V_1' \times \cdots \times V_m'$.

Corollary:

$$\dim(V_1 \otimes \cdots \otimes V_m) = (\dim V_1) \cdots (\dim V_m).$$

Proposition (basis of $V_1 \otimes \cdots \otimes V_m$): Suppose $\dim V_k = n_k$ and $e_1^k, \ldots, e_{n_k}^k$ is a basis of V_k. Then

$$\left\{ e_{j_1}^1 \otimes \otimes \cdots \otimes e_{j_m}^m \right\}_{j_1 : 1, \ldots, n_1, \ldots, j_m = 1, \ldots, n_m}$$

is a basis of $V_1 \otimes \cdots \otimes V_m$.

Definition (m-linear map): An *m-linear map* from $V_1 \times \cdots \times V_m$ to a vector space U is a function $\Gamma : V_1 \times \cdots \times V_m$ that is a linear map in each slot when the other slots are held fixed.

Proposition: Suppose U is a vector space.

(a) Suppose $\Gamma : V_1 \times \cdots \times V_m \to U$ is an m-linear map. Then there exists a unique linear map $\hat{\Gamma} : V_1 \otimes \cdots \otimes V_m \to U$ such that

$$\hat{\Gamma}(v_1 \otimes \cdots \otimes v_m) = \Gamma(v_1, \ldots, v_m)$$

for all $(v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m$.

(b) Suppose $T : V_1 \otimes \cdots \otimes V_m \to U$ is a linear map. Then there exists a unique m-linear map $T^\# : V_1 \times \cdots \times V_m \to U$ such that

$$T^\#(v_1, \ldots, v_m) = T(v_1 \otimes \cdots \otimes v_m)$$

for all $(v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m$.

9.4.4. Problems