1.1 definition: *complex numbers,* $\mathbf{C}$ $\mathbb{C}$

- A *complex number* is an ordered pair (a, b), where $a, b \in \mathbf{R}$, but we will write this as $a + bi$. $\mathbb{R}$
- The set of all complex numbers is denoted by $\mathbf{C}$:
$$\mathbf{C} = \{a + bi : a, b \in \mathbf{R}\}.$$
- *Addition* and *multiplication* on $\mathbf{C}$ are defined by
$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$
$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i;$$
here $a, b, c, d \in \mathbf{R}$.

$i^2 = -1$

$$(a+bi)(c+di) = a(c+di) + bi(c+di)$$
$$= ac + adi + bci + bdi^2$$
$$= (ac - bd) + (ad + bc)i$$

1.3 *properties of complex arithmetic*

commutativity
$\alpha + \beta = \beta + \alpha$ and $\alpha\beta = \beta\alpha$ for all $\alpha, \beta \in \mathbf{C}$.

associativity
$(\alpha + \beta) + \lambda = \alpha + (\beta + \lambda)$ and $(\alpha\beta)\lambda = \alpha(\beta\lambda)$ for all $\alpha, \beta, \lambda \in \mathbf{C}$.

identities
$\lambda + 0 = \lambda$ and $\lambda 1 = \lambda$ for all $\lambda \in \mathbf{C}$.

additive inverse
For every $\alpha \in \mathbf{C}$, there exists a unique $\beta \in \mathbf{C}$ such that $\alpha + \beta = 0$.

multiplicative inverse
For every $\alpha \in \mathbf{C}$ with $\alpha \neq 0$, there exists a unique $\beta \in \mathbf{C}$ such that $\alpha\beta = 1$.

distributive property
$\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$ for all $\lambda, \alpha, \beta \in \mathbf{C}$.

1.5 definition: $-\alpha$, *subtraction,* $1/\alpha$, *division*

Suppose $\alpha, \beta \in \mathbf{C}$.

- Let $-\alpha$ denote the additive inverse of α. Thus $-\alpha$ is the unique complex number such that
$$\alpha + (-\alpha) = 0.$$
- *Subtraction* on $\mathbf{C}$ is defined by
$$\beta - \alpha = \beta + (-\alpha).$$
- For $\alpha \neq 0$, let $1/\alpha$ and $\frac{1}{\alpha}$ denote the multiplicative inverse of α. Thus $1/\alpha$ is the unique complex number such that
$$\alpha(1/\alpha) = 1. \qquad \alpha^{-1}$$
- For $\alpha \neq 0$, *division* by α is defined by
$$\beta/\alpha = \beta(1/\alpha).$$

For $\alpha \in \mathbb{C}$ and m a positive integer, we define α^m to denote the product of α with itself m times:
$$\alpha^m = \underbrace{\alpha \cdots \alpha}_{m \text{ times}}.$$

Definition

Definition 1 (Field)
*A **field** is a set $\mathbb{F}$ equipped with two binary operations (usually written as addition and multiplication) such that:*

1. *Addition and multiplication are associative: $\forall a, b, c \in \mathbb{F}$, we have*

$$a + (b + c) = (a + b) + c$$
$$(ab)c = a(bc)$$

2. *Addition and multiplication commute: $\forall a, b, c \in \mathbb{F}$, we have*

$$a + b = b + a$$
$$ab = ba$$

3. *There exist distinct additive and multiplicative identities in $\mathbb{F}$:*

$$\exists 1 \in \mathbb{F} \text{ such that } 1a = a \quad \forall a \in \mathbb{F}$$
$$\exists 0 \in \mathbb{F} \text{ such that } a + 0 = a \quad \forall a \in \mathbb{F}$$

The distinctness requirement here means that $0 \neq 1$.

$\mathbb{F} = \{e\}$
$e + e = e$
$ee = e$

4. *Additive and multiplicative inverses exist in $\mathbb{F}$:*

$$\forall a \in \mathbb{F}, \exists (-a) \in \mathbb{F} \text{ such that } a + (-a) = 0$$
$$\forall a \neq 0 \in \mathbb{F}, \exists a^{-1} \in \mathbb{F} \text{ such that } aa^{-1} = 1$$

5. *Multiplication is distributive over addition:*

$$a(b + c) = ab + ac$$

Definition

Definition 2 ($\mathbb{F}^n$)
For n a positive integer, we define:

$$\mathbb{F}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{F}\}$$

Where the addition is component-wise.

Example

Example 3
Here is an example of addition in $\mathbb{F}^n$:

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

What about multiplication? Can we define $(x_1, \dots, x_n)(y_1, \dots, y_n) = (x_1y_1, \dots, x_ny_n)$?
Sure, but we lose a nice property!
Mainly, multiplicative inverses: Ex: $(0, x_2, \dots, x_n)$ does not have a multiplicative inverse! even when $x_i \neq 0$ for some $i \in \{2, \dots, n\}$
i.e. we have some $x \neq 0$ with no inverse wrt multiplication!

Notation

Notation 4
In $\mathbb{F}^n$, we let 0 denote the element $(0, 0, \ldots, 0) \in \mathbb{F}^n$, and for $x \in \mathbb{F}^n$, we let $(-x)$ denote the additive inverse of x–i.e., the element $(-x) \in \mathbb{F}^n$ such that $x + (-x) = 0$.

Definition

Definition 5 (Scaling)
For $\lambda \in \mathbb{F}$, define λx for $x \in \mathbb{F}^n$ to be:

$$\lambda x = (\lambda x_1, \ldots, \lambda x_n) \in \mathbb{F}^n$$

1.19 definition: *addition, scalar multiplication*

- An *addition* on a set V is a function that assigns an element $u + v \in V$ to each pair of elements $u, v \in V$.
- A *scalar multiplication* on a set V is a function that assigns an element $\lambda v \in V$ to each $\lambda \in \mathbf{F}$ and each $v \in V$.

1.20 definition: *vector space*

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

commutativity
$u + v = v + u$ for all $u, v \in V$.

associativity
$(u + v) + w = u + (v + w)$ and $(ab)v = a(bv)$ for all $u, v, w \in V$ and for all $a, b \in \mathbf{F}$.

additive identity
There exists an element $0 \in V$ such that $v + 0 = v$ for all $v \in V$.

additive inverse
For every $v \in V$, there exists $w \in V$ such that $v + w = 0$.

multiplicative identity
$1v = v$ for all $v \in V$.

different +'s!

distributive properties
$a(u + v) = au + av$ and $(a + b)v = av + bv$ for all $a, b \in \mathbf{F}$ and all $u, v \in V$.

1.21 definition: *vector, point*

Elements of a vector space are called *vectors* or *points*.

Example

$\mathbb{F}^n$ is a vector space for any field $\mathbb{F}$. If you take F^∞ to be all sequences of elements in $\mathbb{F}$, then so is F^∞.

Notation

For a set S and field $\mathbb{F}$, $\mathbb{F}^S$ denotes the set of all functions from S to $\mathbb{F}$.

Similar!

Example

If we define $(f + g)(x) = f(x) + g(x)$ and $(\lambda f)(x) = \lambda(f(x))$ for all $f, g \in \mathbb{F}^S$ and $\lambda \in \mathbb{F}$, then $\mathbb{F}^S$ is a vector space.

Notice: A function $f \in \mathbb{F}^S$ w/ $S = \{1, \dots, n\}$ is basically just a list $(x_1, \dots, x_n)$ ie. $f(i) = x_i$. Taking $S = \{1, \dots\}$ we get $\mathbb{F}^\infty$ too!

1.26 *unique additive identity*

A vector space has a unique additive identity.

Proof

Let $0, 0' \in V$ be two additive identities. Then,

$$0 + x = x \quad \text{and} \quad \boxed{0' + x = x}$$

for all $x \in V$. So,

$$\begin{aligned} \boxed{0'} &= 0 + 0' \\ &= \underline{0' + 0} \\ &= \underline{0}. \end{aligned}$$

commutativity!

1.27 *unique additive inverse*

Every element in a vector space has a unique additive inverse.

$\boxed{\text{Proof}}$ Let V be a vector space & let $v \in V$.

& w, w' are two additive inverses of v. Then:

$$
\begin{aligned}
w &= w + 0 \\
&= w + (v + w') \\
&= (w + v) + w' \\
&= 0 + w' \\
&= w'.
\end{aligned}
$$

Therefore, additive inverses are unique.

1.28 notation: $-v$, $w - v$

Let $v, w \in V$. Then

- $-v$ denotes the additive inverse of v;
- $w - v$ is defined to be $w + (-v)$.

1.30 *the number* 0 *times a vector*

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

$\boxed{\text{Proof}}$

$$0v = (0+0)v = 0v + 0v$$

So,
$$
\begin{aligned}
0 &= 0v + (-(0v)) \quad \leftarrow \text{Def'n of additive inverses} \\
&= (0v + 0v) + (-(0v)) \\
&= 0v + (0v + (-(0v))) \quad \leftarrow \text{Associativity of addition} \\
&= 0v + 0 \quad \leftarrow \text{Def'n of additive inverses} \\
&= 0v \quad \leftarrow \text{Def'n of additive identity.}
\end{aligned}
$$

1.31 *a number times the vector* 0

$a0 = 0$ for every $a \in \mathbf{F}$.

Proof

$$a0 = a(0+0)$$
$$= a0 + a0$$

Hence,
$$\begin{aligned} 0 &= a0 + (-(a0)) \\ &= (a0 + a0) + (-(a0)) \\ &= a0 + (a0 + (-(a0))) \\ &= a0 + 0 \\ &= a0. \end{aligned}$$

(Same reasons as before!)

1.32 *the number -1 times a vector*

$(-1)v = -v$ for every $v \in V$.

$\boxed{\text{Proof}}$ For $v \in V$,

$$\begin{aligned} v + (-1)v &= (1)v + (-1)v \\ &= (1 + (-1))v \\ &= 0v \\ &= 0 \end{aligned}$$

Hence, by the uniqueness of additive inverses,

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

1C Subspaces

$U \subseteq V \leftarrow$ subset
$U \leq V \leftarrow$ subspace.

we write $U \leq V$ to denote that U is a subspace.

1.33 definition: *subspace*

A subset U of V is called a *subspace* of V if U is also a vector space with the same additive identity, addition, and scalar multiplication as on V.

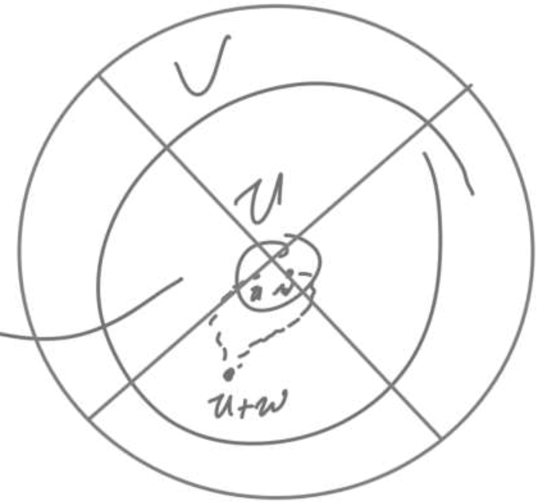
1.34 *conditions for a subspace*

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

1 **additive identity**
$0 \in U$.

2 **closed under addition**
$u, w \in U$ implies $u + w \in U$.

3 **closed under scalar multiplication**
$a \in \mathbf{F}$ and $u \in U$ implies $au \in U$.



$\boxed{\text{Proof (Sketch)}}$

If U is a subspace, then this holds by definition. In the reverse, notice that any of these being violated would prevent U from being a subspace. Moreover, they are precisely the 3 conditions defining a subspace.

1.36 definition: *sum of subspaces*

Suppose $V_1, \ldots, V_m$ are subspaces of V. The *sum* of $V_1, \ldots, V_m$, denoted by $V_1 + \cdots + V_m$, is the set of all possible sums of elements of $V_1, \ldots, V_m$. More precisely,

$$V_1 + \cdots + V_m = \{v_1 + \cdots + v_m : v_1 \in V_1, \ldots, v_m \in V_m\}.$$

$\boxed{\text{Ex}}$

$$U = \{(x,x,y,y) \in \mathbb{F}^4 : x,y \in \mathbb{F}\}$$

$$(x_1,x_1,y_1,y_1) + (x_2,x_2,x_2,y_2) = (x_1+x_2, x_1+x_2, y_1+x_2, y_1+y_2)$$

Same

$$W = \{(x,x,x,y) \in \mathbb{F}^4 : x,y \in \mathbb{F}\}$$

$$U + W = \{(x,x,y,z) \in \mathbb{F}^4 : x,y,z \in \mathbb{F}\}$$

1.40 *sum of subspaces is the smallest containing subspace*

Suppose $V_1, \ldots, V_m$ are subspaces of V. Then $V_1 + \cdots + V_m$ is the smallest subspace of V containing $V_1, \ldots, V_m$.

$\boxed{\text{Proof}}$ Clearly, $V_i \subseteq V_1 + \cdots + V_m$, since

$v_i = 0 + \cdots + 0 + v_i + 0 + \cdots 0 + 0 \in V_1 + \cdots + V_m$ for any $v_i \in V_i$.

Conversely, if $W \leq V$ such that $V_i \subseteq W$ for $i = 1, \ldots, m$,

then $V_1 + \cdots + V_m \subseteq W$ as well, as subspaces must contain all finite sums of their elements.

1.41 definition: *direct sum*, $\oplus$

Suppose $V_1, \ldots, V_m$ are subspaces of V.

- The sum $V_1 + \cdots + V_m$ is called a *direct sum* if each element of $V_1 + \cdots + V_m$ can be written in only one way as a sum $v_1 + \cdots + v_m$, where each $v_k \in V_k$.
- If $V_1 + \cdots + V_m$ is a direct sum, then $V_1 \oplus \cdots \oplus V_m$ denotes $V_1 + \cdots + V_m$, with the $\oplus$ notation serving as an indication that this is a direct sum.

$\boxed{\text{Ex}}$

Let $0 \neq U \subseteq V$. Then, $U+U$ is not direct!

If $w \in U+U$, then notice that $w \in U$, so $w = 0+w = w+0$

i.e. $w = u_1 + u_2 = u_1' + u_2'$ for $u_i' \neq u_i$.

1.45 *condition for a direct sum*

Suppose $V_1, \ldots, V_m$ are subspaces of V. Then $V_1 + \cdots + V_m$ is a direct sum if and only if the only way to write 0 as a sum $v_1 + \cdots + v_m$, where each $v_k \in V_k$, is by taking each v_k equal to 0.

$\boxed{\text{Proof}}$ If $V_1 + \cdots + V_m$ is a direct sum, then since $0 = 0 + \cdots + 0$, we get our desired result.

Now, suppose the only way to write $0 \in V_1 + \cdots + V_m$ is as $0 = 0 + \cdots + 0$, i.e., $0 = v_1 + \cdots + v_m$ implies $v_i = 0$ for $i = 1, \ldots, m$.

Now, if $w = v_1 + \cdots + v_m$ & $w = v_1' + \cdots + v_m'$ for some $w \in V_1 + \cdots V_m$.

Then, $0 = w - w = (v_1 + \cdots + v_m) - (v_1' + \cdots + v_m')$

$= (v_1 - v_1') + \cdots + (v_m - v_m')$

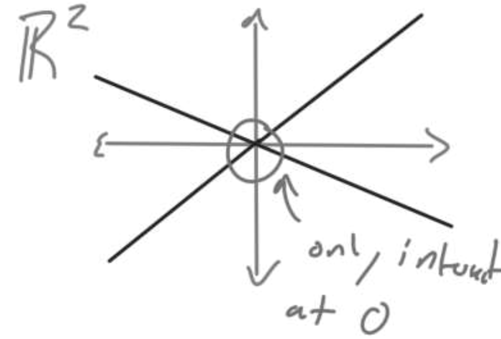
Since $v_i - v_i' \in V_i$ for $i = 1, \ldots, m$, we get $v_i - v_i' = 0$

Hence, $v_i = v_i'$ and there is only one way to write w as a sum of elements of V_i for $i = 1, \ldots, m$. That is, $V_1 + \cdots + V_m$ is direct.

1.46 *direct sum of two subspaces*

Suppose U and W are subspaces of V. Then

$$U + W \text{ is a direct sum} \iff U \cap W = \{0\}.$$



Proof

Suppose $U+W$ is direct. If $v \in U \cap W$, then $0 = v + (-v)$ since $v \in U \cap W \Rightarrow v \in U$ & $v \in W$ (and, hence, $-v \in W$). So, by the previous result, $v = -v = 0$. Hence, $U \cap W = \{0\}$.

Now, suppose $\underline{U \cap W = \{0\}}$. Then, if $u \in U$ & $w \in W$ are such that $u + w = 0$, then $u = -w$. Hence, $u \in W$ as well (since W is a vector space). But, only 0 belongs to both W & U. So $w = 0$. Hence, $u + w = 0 \Rightarrow u = w = 0$ & by the previous result, $U+W$ is direct.

2A *Span and Linear Independence*

2.2 definition: *linear combination*

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

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

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

2.4 definition: *span*

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

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

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

2.6 *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.

$\boxed{\text{Proof}}$ Notice: If $v_1, \dots, v_m \in V$ is a list of vectors, then $0v_1 + \dots + 0v_m = 0$, so $0 \in \text{span}(v_1, \dots, v_m)$.

Moreover, if $v \in \text{Span}(v_1, \dots, v_m)$, then $v = a_1 v_1 + \dots + a_m v_m$ for some $a_1, \dots, a_m \in \mathbb{F}$ & for any $b \in \mathbb{F}$, we have

$$b v = b(a_1 v_1 + \dots + a_m v_m) = (ba_1)v_1 + \dots + (ba_m)v_m \in \text{Span}(v_1, \dots, v_m).$$

Similarly, if $v' = a_1' v_1 + \dots + a_m' v_m \in \text{Span}(v_1, \dots, v_m)$, then

$$\begin{aligned} v + v' &= (a_1 v_1 + \dots + a_m v_m) + (a_1' v_1 + \dots + a_m' v_m) \\ &= (a_1 v_1 + a_1' v_1) + \dots + (a_m v_m + a_m' v_m) \\ &= \underline{(a_1 + a_1')} v_1 + \dots + \underline{(a_m + a_m')} v_m \in \text{Span}(v_1, \dots, v_m) \end{aligned}$$

So, by a theorem we proved last time, $\text{Span}(v_1, \dots, v_m)$ is a subspace of V. Moreover, notice that $v_i = 0v_1 + \dots + 0v_{i-1} + (1)v_i + 0v_{i+1} + \dots + 0v_m$, so $v_i \in \text{Span}(v_1, \dots, v_m)$ for $i = 1, \dots, m$.

Span$(v_1, \dots, v_m)$ is a Subspace containing $v_1, \dots, v_m$

Now, $\oint$ $W \leq V$ & $v_1, \dots, v_m \in W$. Then, since linear combos of $v_1, \dots, v_m$ are certainly also contained in W, $\text{Span}(v_1, \dots, v_m) \subseteq W$ as well.

It is the smallest one.

2.7 definition: *spans*

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

2.9 definition: *finite-dimensional vector space*

A vector space is called *finite-dimensional* if some list of vectors in it spans the space.

Recall that by definition every list has finite length.

2.13 definition: *infinite-dimensional vector space*

A vector space is called *infinite-dimensional* if it is not finite-dimensional.

2.10 definition: *polynomial,* $\mathcal{P}(\mathbf{F})$

- A function $p\colon \mathbf{F} \to \mathbf{F}$ is called a *polynomial* with coefficients in $\mathbf{F}$ if there exist $a_0, \ldots, a_m \in \mathbf{F}$ such that

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

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

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

2.12 notation: $\mathcal{P}_m(\mathbf{F})$

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

$\boxed{\text{Ex}}$ $\mathcal{P}_m(\mathbb{F}) \subseteq \mathcal{P}(\mathbb{F})$

i.e. add togethr two polys of degree at most m, get back a poly. of degree at most m. Similarly, scale a poly. of degree at most m, get back a poly of the same degree or, if the scalar is 0, a lower degree poly.

Notice: $\mathcal{P}_m(\mathbb{F}) = \operatorname{span}(1, z, z^2, \ldots, z^m)$

So, $\mathcal{P}_m(\mathbb{F})$ is finite dimensional!

However, $\mathcal{P}(\mathbb{F})$ is not!

2.15 definition: *linearly independent*

- A list $v_1, \ldots, v_m$ of vectors in V is called *linearly independent* if the only choice of $a_1, \ldots, a_m \in \mathbf{F}$ that makes

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

is $a_1 = \cdots = a_m = 0$.

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

2.17 definition: *linearly dependent*

- A list of vectors in V is called *linearly dependent* if it is not linearly independent.
- In other words, a list $v_1, \ldots, v_m$ of vectors in V is linearly dependent if there exist $a_1, \ldots, a_m \in \mathbf{F}$, not all 0, such that $a_1 v_1 + \cdots + a_m v_m = 0$.

2.19 *linear dependence lemma*

Suppose $v_1, \ldots, v_m$ is a linearly dependent list in V. Then there exists $k \in \{1, 2, \ldots, m\}$ such that

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

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

$\boxed{\text{Proof}}$ Since $v_1, \ldots, v_m$ is lin. dep., there are scalars $a_1, \ldots, a_m \in \mathbb{F}$, not all zero, such that $a_1 v_1 + \ldots + a_m v_m = 0$. Let k be the largest # s.t. $a_k \neq 0$. Then,

$$0 = a_1 v_1 + \ldots + a_k v_k + a_{k+1} v_{k+1} + \ldots + a_m v_m$$
$$0 = a_1 v_1 + \ldots + a_k v_k + 0 + \ldots + 0$$

So, $-a_k v_k = a_1 v_1 + \ldots + a_{k-1} v_{k-1}$ and, since $a_k \neq 0$, we can multiply by $(a_k)^{-1}$ to get

$$v_k = \underline{(-a_k)^{-1} a_1 v_1 + \ldots + (-a_k)^{-1} a_{k-1} v_{k-1}}$$

Hence, $v_k \in \operatorname{Span}(v_1, \ldots, v_{k-1})$. To see the rest, notice that this lets us rewrite any scaling of v_k as a linear combo of $v_1, \ldots, v_{k-1}$.

2.22 *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.

$\boxed{\text{Proof}}$ Let $\mathrm{span}(w_1, \ldots, w_n) = V$, i.e. $w_1, \ldots, w_n$ is a spanning list. Let $u_1, \ldots, u_m$ be a linearly indep. list.

Since $w_1, \ldots, w_n$ spans V, so does $u_1, w_1, \ldots, w_n$.

However, since $u_1 \in V$, $u_1, w_1, \ldots, w_n$ is linearly dependent. So, by the lin. dep. lemma, $\exists K$ s.t. the list

$$\underbrace{u_1, w_1, \ldots, w_{K-1}, w_{K+1}, \ldots, w_n}_{\text{length } n} \text{ spans } V.$$

This is step 1. Here is step K for $K = 2, \ldots, m$:

Take your list from the previous step & notice that u_K is in its span. So, adding u_K in the K^{th} position in the list yields a lin. dep. list & we can use the lin. dep. lemma to remove an element. Since the start of this list is all u_i's & $u_1, \ldots, u_m$ is lin. indep., we get that the removed element must be a w_i.

Then, at the end of this process, we get a list that looks like $\underbrace{u_1, \ldots, u_m}, w_{t_1}, w_{t_2}, \ldots, w_{t_{n-m}}$ (where there are no w's in the case where $n = m$) that spans V.

So, the length of such a list must be at least as long as the length of the lin. indep. list $u_1, \ldots, u_m$ (i.e. m).

2.25 *finite-dimensional subspaces*

Every subspace of a finite-dimensional vector space is finite-dimensional.

$\boxed{\text{Proof}}$ Let V be a fin. dim. V.S. . Then,

Let $U \leq V$.

Step 1: If $U = \{0\}$, we are done since $U = \operatorname{span}(\,)$.
If not, let $u_1 \in U$ be non-zero.

Step j : Take $\operatorname{Span}(u_1, \ldots, u_{j-1})$. If this is equal to U,
for $j \in \{2, \ldots, n\}$ we are done. If not, take some $u_j \notin \operatorname{Span}(u_1, \ldots, u_{j-1})$
& repeat now with $\operatorname{Span}(u_1, \ldots, u_j)$ as the
next step.

Now, since $u_j \notin \operatorname{Span}(u_1, \ldots, u_{j-1})$, the list $u_1, \ldots, u_j$ is lin. indep. By the previous result, the length of every lin. indep. list in V is less than or equal to the length of any list that spans V. Since V is finite dim., there is a finite length list that spans U, so there is an upper bound on how long a lin. indep. list in U can be.

This applies to the u_i's — the process has to eventually stop at some k that is less than or equal to the length of any choice of a finite spanning list for U.
Hence, we get $U = \operatorname{Span}(u_1, \ldots, u_k)$ & U is finite dim.

2B Bases

tells us when the span doesnt have any extraneous info

2.26 definition: *basis*

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

How to make subspaces from lists of vectors

$\boxed{\text{Ex}}$ $v_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $v_2 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

Then, if $v = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$, notice that $a_1 v_1 + a_2 v_2 = \begin{bmatrix} 0 \\ a_1 \end{bmatrix} + \begin{bmatrix} 2a_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2a_2 \\ a_1 \end{bmatrix}$

& $v = a_1 v_1 + a_2 v_2$ has the unique solution $a_2 = \frac{x}{2}$, $a_1 = y$.

So, v_1, v_2 is a basis of $\mathbb{R}^2$.

2.28 *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

2.29 $$v = a_1 v_1 + \cdots + a_n v_n,$$

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

$\boxed{\text{Proof}}$ & $v_1, \ldots, v_n$ is a basis of V.

Let $v \in V$. Since $v_1, \ldots, v_n$ spans V, it follows that there are scalars $a_1, \ldots, a_n \in \mathbb{F}$ s.t. $v = a_1 v_1 + \cdots + a_n v_n$. Now, & that $c_1, \ldots, c_n \in \mathbb{F}$ also satisfy $v = c_1 v_1 + \cdots + c_n v_n$. Then,

$$0 = v - v = (a_1 v_1 + \cdots + a_n v_n) - (c_1 v_1 + \cdots + c_n v_n)$$
$$= (a_1 - c_1) v_1 + \cdots + (a_n - c_n) v_n$$

But, this means that $a_i - c_i = 0$ for $i = 1, \ldots, n$.

That is, $a_i = c_i$.

In the reverse, & $v = a_1 v_1 + \cdots + a_n v_n$ has a unique solution $a_1, \ldots, a_n \in \mathbb{F}$ for $\forall v \in V$. This certainly implies $v_1, \ldots, v_n$ spans V, since it says every $v \in V$ is a linear combo of $v_1, \ldots, v_n$. For lin. indep., notice that this implies in particular that the solution $a_1 = \cdots = a_n = 0$ to $0 = a_1 v_1 + \cdots + a_n v_n$ is unique. So, $v_1, \ldots, v_n$ is also lin. indep. Hence, it is a basis.

2.30 *every spanning list contains a basis*

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

proof Let $v_1, \ldots, v_n$ be a spanning list for V.

Step 1: If $v_1 = 0$, delete v_1. If not, do nothing.

Step k: If v_k is in $\operatorname{span}(v_1, \ldots, v_{k-1})$, delete v_k.
If not, do nothing.

This process eventually stops: Either we discard nothing & have only the list $v_1, \ldots, v_n$ which spans V, or we have some new list $w_1, \ldots, w_m$ w/ $w_i \in \{v_1, \ldots, v_n\}$ & $w_i = v_j \Rightarrow w_t \in \{v_1, \ldots, v_j\}$ for $t < i$ (i.e. still in order, just missing a few) and, by the lin. dep. lemma, this list is independent since, if it were dependent, we could choose some w_k s.t. $w_k \in \operatorname{Span}(w_1, \ldots, w_{k-1})$ & by construction this is not the case. Hence, we are done!

2.31 *basis of finite-dimensional vector space*

Every finite-dimensional vector space has a basis.

$\boxed{\text{Proof}}$ If V is finite dim., it has a spanning list of vectors. By the previous, this can be reduced to a basis.

2.32 *every linearly independent list extends to a basis*

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

$\boxed{\text{Proof}}$ Let $\underline{u_1, \ldots, u_m}$ be a lin. indep. list of vectors in V, a finite dimensional vector space. Since V is finite dim, there exists a spanning list for V, call it $\underline{w_1, \ldots, w_n}$. Form the list $\underline{u_1, \ldots, u_m, w_1, \ldots, w_n}$. This certainly spans V and can be reduced to a basis using the process from the proof of the result before last. This process deletes none of the u_i, since $u_i \notin \operatorname{Span}(u_1, \ldots, u_{i-1})$ for any $i \in 1, \ldots, m$ as $u_1, \ldots, u_m$ is linearly indep. Hence, the obtained basis "starts w/" the u_i's. So, we are done!

2.33 *every subspace of V is part of a direct sum equal to V*

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

$\boxed{\text{Proof}}$ $\ni$ V is finite dim. Let $U \leq V$. Then, U is also finite dim. Hence, there is a basis $u_1, \dots, u_m$ of U. Notice that $u_1, \dots, u_m$ is linearly indep. in V, so we can extend it to a basis $u_1, \dots, u_m, \underline{w_1, \dots, w_n}$ of V, by the previous result.

Let $W = \operatorname{span}(w_1, \dots, w_n)$. Then, since $u_1, \dots, u_m, w_1, \dots, w_n$ is a basis for V, every $v \in V$ can be written as

$$v = a_1u_1 + \dots + a_mu_m + b_1w_1 + \dots + b_nw_n$$

for some $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbb{F}$. To get directness, recall that we need only show that $V = U + W$ & $U \cap W = \{0\}$. Since $a_1u_1 + \dots + a_mu_m \in U$ for any $a_1, \dots, a_m \in \mathbb{F}$ & $b_1w_1 + \dots + b_nw_n \in W$ for any $b_1, \dots, b_n \in \mathbb{F}$, the fact here tells us we can write any $v \in V$ as $v = u + w$ for some $u \in U$ & $w \in W$. Hence, $\underline{V = U + W}$. Now, since $u_1, \dots, u_m, w_1, \dots, w_n$ is a basis, it is linearly indep. & the only way to write $0 = a_1u_1 + \dots + a_mu_m + b_1w_1 + \dots + b_nw_n$ is w/ $a_1 = \dots = a_m = b_1 = \dots = b_n = 0$.
So, $\ni$ $\underline{v \in U \cap W}$. Then, $v = a_1u_1 + \dots + a_mu_m$ for some $a_1, \dots, a_m \in \mathbb{F}$ & $v = b_1w_1 + \dots + b_nw_n$ for some $b_1, \dots, b_n \in \mathbb{F}$. Hence, $0 = v - v = a_1u_1 + \dots + a_mu_m - b_1w_1 - \dots - b_nw_n$ & we get $a_1 = \dots = a_m = -b_1 = \dots = -b_n = 0$ & $\underline{v = 0}$. So, $V = U \oplus W$.

2C Dimension

Recall:

2.22 *length of linearly independent list ≤ 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.

Since a basis is a lin. indep. spanning list, this applies.

2.34 *basis length does not depend on basis*

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

$\boxed{\text{Proof}}$ Let $v_1, \ldots, v_n$ and $w_1, \ldots, w_m$ be two bases for V. By the result above, $n \leq m$ since $v_1, \ldots, v_n$ is lin indep. & $w_1, \ldots, w_m$ spans V. Now, notice that $w_1, \ldots, w_m$ is also lin. indep. & $v_1, \ldots, v_n$ is also a spanning list. Hence, we also have $m \leq n$. Therefore, $m = n$.

This only works b/c every basis has the same length!

2.35 definition: *dimension*, $\dim V$

- The *dimension* of a finite-dimensional vector space is the length of any basis of the vector space.
- The dimension of a finite-dimensional vector space V is denoted by $\dim V$.

i.e. $\dim : \text{finite dim Vector spaces} \to \mathbb{N}$

where $\mathbb{N} = \{0, 1, 2, \ldots\}$

2.37 *dimension of a subspace*

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

$\boxed{\text{Proof}}$ Let $U \leq V$, where V is finite dim.
Let $u_1, \dots, u_n$ be a basis for U (i.e. $\dim U = n$)
& let $v_1, \dots, v_m$ be a basis for V. Then, notice
that $u_1, \dots, u_n$ is a linearly indep. list in V as well, and
$v_1, \dots, v_m$ is a spanning list in V. By the result
we recalled at the beginning of lecture, $n \leq m$.
That is, $\dim U \leq \dim V$.

2.38 *linearly independent list of the right length is a basis*

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

$\boxed{\text{Proof}}$ Suppose V is finite dim. & $v_1, \dots, v_n$ is a lin. indep.
list in V. Suppose further that $\dim V = n$. Recall that
any lin. indep. list in a finite dim. V.S. can be extended
to a basis. But, since the length of every basis is
the same in a finite dim V.S., this extention must
be trivial (i.e. it must not add anything!). Hence, $v_1, \dots, v_n$
is already a basis whenever it is linearly indep. & $\dim V = n$.

2.39 *subspace of full dimension equals the whole space*

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

$\boxed{\text{Proof}}$ Let V be finite dim & $U \leq V$ & s.t. $\dim U = \dim V$. Let $u_1, \ldots, u_n$ be a basis for U. Then, $u_1, \ldots, u_n$ is a linearly indep. list in V & its length is $\dim V$ since $\dim U = \dim V$. But, by the previous result, any lin. indep. list of length $\dim V$ is a basis. So, $u_1, \ldots, u_n$ is a basis for V as well; i.e., $U = \text{Span}(u_1, \ldots, u_n) = V$.

2.42 *spanning list of the right length is a basis*

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

$\boxed{\text{Proof}}$ S.p. V is finite dim. & let $v_1, \ldots, v_n$ be a spanning list of length $\dim V = n$. Since every spanning list in a finite dim. V.S. can be reduced to a basis, $v_1, \ldots, v_n$ can as well. But, since every basis has the same length, this reduction must be trivial (i.e., we remove nothing!). So, $v_1, \ldots, v_n$ is already a basis for V.

2.43 *dimension of a sum*

If V_1 and V_2 are subspaces of a finite-dimensional vector space, then

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

$\boxed{\text{Proof}}$ let $V_1, V_2 \subseteq V$, a finite dim. V.S..

Let $v_1, \dots, v_m$ be a basis for $V_1 \cap V_2$ so that $\dim V_1 \cap V_2 = m$. Since $v_1, \dots, v_m$ is a linearly indp. spanning list in $V_1 \cap V_2$, it is a lin. indp list in both V_1 & V_2, so we can extend it in two ways:

$v_1, \dots, v_m, u_1, \dots, u_j$ a basis of V_1

$v_1, \dots, v_m, w_1, \dots, w_k$ a basis of V_2

Then, we are saying that $\dim V_1 = m+j$ & $\dim V_2 = m+k$

We will show that $B = v_1, \dots, v_m, u_1, \dots, u_j, w_1, \dots, w_k$ is a basis for $V_1 + V_2$, since this says that:

$$\begin{aligned} \dim V_1 + V_2 = m+j+k &= (m+j) + (m+k) - m \\ &= \dim V_1 + \dim V_2 - \dim V_1 \cap V_2. \end{aligned}$$

Notice that B is contained in $V_1 \cup V_2$ & is therefore contained in $V_1 + V_2$. The span of B certainly contains both V_1 & V_2, so it is equal to $V_1 + V_2$ (i.e. $\operatorname{span} B \subseteq V_1 + V_2$ & $V_1 + V_2 \subseteq \operatorname{span} B$, so $\operatorname{span} B = V_1 + V_2$). Hence, B spans $V_1 + V_2$.

Now, we need only show linear indp.

Proof (continued)

To show lin. independence, suppose that

$$a_1v_1 + \dots + a_mv_m + b_1u_1 + \dots + b_ju_j + c_1w_1 + \dots + c_kw_k = 0$$

for some $a_1, \dots, a_m, b_1, \dots, b_j, c_1, \dots, c_k \in \mathbb{F}$.

Notice that this can be rewritten as $c_1w_1 + \dots + c_kw_k = -(a_1v_1 + \dots + a_mv_m + b_1u_1 + \dots + b_ju_j)$

Hence, $c_1w_1 + \dots + c_kw_k \in V_1$, since $v_1, \dots, v_m, u_1, \dots, u_j$ is a basis for V_1.

All of $w_1, \dots, w_k$ belong to V_2 since $v_1, \dots, v_m, w_1, \dots, w_k$ is a basis for V_2.

So, $c_1w_1 + \dots + c_kw_k \in V_1 \cap V_2$. Since $v_1, \dots, v_m$ is a basis for $V_1 \cap V_2$,

$c_1w_1 + \dots + c_kw_k = d_1v_1 + \dots + d_mv_m$ for some $d_1, \dots, d_m \in \mathbb{F}$. However, $v_1, \dots, v_m, w_1, \dots, w_k$ is lin. indep. since it is a basis for V_2, so this implies that

$c_1 = \dots = c_k = d_1 = \dots = d_m = 0$, since ✓ implies $c_1w_1 + \dots + c_kw_k - d_1v_1 - \dots - d_mv_m = 0$.

So, we can rewrite ✓ as $a_1v_1 + \dots + a_mv_m + b_1u_1 + \dots + b_ju_j = 0$. Now, repeat for the list $v_1, \dots, v_m, u_1, \dots, u_j$, which is also lin. indep. since it is a basis for V_1. Hence, $a_1 = \dots = a_m = b_1 = \dots = b_j = 0$. So, B is lin. indep. and, since it spans $V_1 + V_2$, it is a basis of $V_1 + V_2$.

sets	**vector spaces**
S is a finite set	V is a finite-dimensional vector space
$\#S$	$\dim V$
for subsets S_1, S_2 of S, the union $S_1 \cup S_2$ is the smallest subset of S containing S_1 and S_2	for subspaces V_1, V_2 of V, the sum $V_1 + V_2$ is the smallest subspace of V containing V_1 and V_2
$\#(S_1 \cup S_2) = \#S_1 + \#S_2 - \underline{\#(S_1 \cap S_2)}$	$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \underline{\dim(V_1 \cap V_2)}$
$\#(S_1 \cup S_2) = \#S_1 + \#S_2 \iff \underline{S_1 \cap S_2} = \underline{\emptyset}$	$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 \iff V_1 \cap V_2 = \{0\}$
$S_1 \cup \dots \cup S_m$ is a disjoint union $\iff \#(S_1 \cup \dots \cup S_m) = \#S_1 + \dots + \#S_m$	$V_1 + \dots + V_m$ is a direct sum $\iff \dim(V_1 + \dots + V_m) = \dim V_1 + \dots + \dim V_m$

Basically, the same thing!
This is the power of dimension!

3A *Vector Space of Linear Maps*

vector spaces over the same field

3.1 definition: *linear map*

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

additivity

W's +

$T(u + v) = Tu + Tv$ for all $u, v \in V$.

V's + — Different +'s!

homogeneity

$T(\lambda v) = \lambda(Tv)$ for all $\lambda \in \mathbf{F}$ and all $v \in V$.

Different scalings!

3.2 notation: $\mathcal{L}(V, W)$, $\mathcal{L}(V)$

- The set of linear maps from V to W is denoted by $\mathcal{L}(V, W)$.
- The set of linear maps from V to V is denoted by $\mathcal{L}(V)$. In other words, $\mathcal{L}(V) = \mathcal{L}(V, V)$.

$\boxed{\text{Ex}}$

$T: V \to W$ — zero in

- Let $Tv = 0$ for all $v \in V$ i.e. T is the zero map, which is usually also denoted by 0.
  Then, $T(v+w) = 0 = 0 + 0 = Tv + Tw$.
  Similarly, $T(\lambda v) = 0 = \lambda 0 = \lambda(Tv)$.
  So, the zero map is a linear map.
- Let $I: V \to V$ be defined by $Iv = v$ for $\forall v \in V$.
  Then, $I(v+w) = v+w = Iv + Iw$ & $I(\lambda v) = \lambda v = \lambda(Iv)$.
  So, this map (which is called the identity map) is linear too.
- Define $D \in \mathcal{L}(\mathcal{P}(\mathbb{R}))$ by $Dp = p'$. Then, since the derivation is linear, so is D (i.e. $D(p+q) = p' + q'$ & $D(\lambda p) = \lambda p'$)
- Integration also is linear: Define, for example, $T \in \mathcal{L}(\mathcal{P}(\mathbb{R}), \mathbb{R})$ by $Tp = \int_0^1 p(x)\,dx$. Then, $T(p+q) = \int_0^1 p(x) + q(x)\,dx = \int_0^1 p(x)\,dx + \int_0^1 q(x)\,dx = Tp + Tq$
  & $T(\lambda p) = \int_0^1 \lambda p(x)\,dx = \lambda \int_0^1 p(x)\,dx = \lambda Tp$

3.4 *linear map lemma*

V finite-dim

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

!

$$\boxed{Tv_k = w_k}$$

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

$\boxed{\text{Proof}}$ Let V be a finite-dim. V.S. & $\S$ $v_1, \ldots, v_n$ is a basis for V. Let W be a vector space & let $w_1, \ldots, w_n \in W$. Define $T: V \to W$

by $\underline{Tv = T(a_1v_1 + \cdots + a_nv_n) = a_1w_1 + \cdots + a_nw_n}$

(which is possible since $v_1, \ldots, v_n$ is a basis & therefore every $v \in V$ can be written uniquely as $v = a_1v_1 + \cdots + a_nv_n$ for some $a_1, \ldots, a_n \in \mathbb{F}$).

Taking $a_k = 1$ & $a_i = 0$ for $i \neq k$, we get $Tv_k = w_k$.

To see $T \in \mathcal{L}(V, W)$, note that

$$\begin{aligned}
T(v+w) &= T(\underbrace{(a_1v_1 + \cdots + a_nv_n)}_{v} + \underbrace{(b_1v_1 + \cdots + b_nv_n)}_{w}) \\
&= T((a_1+b_1)v_1 + \cdots + (a_n+b_n)v_n) \\
&= (a_1+b_1)w_1 + \cdots + (a_n+b_n)w_n \\
&= (a_1w_1 + \cdots + a_nw_n) + (b_1w_1 + \cdots + b_nw_n) \\
&= T(a_1v_1 + \cdots + a_nv_n) + T(b_1v_1 + \cdots + b_nv_n) \\
&= Tv + Tw
\end{aligned}$$

$$\begin{aligned}
T(\lambda v) &= T(\lambda(a_1v_1 + \cdots + a_nv_n)) \\
&= T((\lambda a_1)v_1 + \cdots + (\lambda a_n)v_n) \\
&= (\lambda a_1)w_1 + \cdots + (\lambda a_n)w_n \\
&= \lambda(a_1w_1 + \cdots + a_nw_n) \\
&= \lambda T(a_1v_1 + \cdots + a_nv_n) \\
&= \lambda Tv
\end{aligned}$$

So, T is linear. For uniqueness, $\S$ $T' \in \mathcal{L}(V, W)$ satisfies $T'v_k = w_k$.

Then, $T'(v) = T'(a_1v_1 + \cdots + a_nv_n) = a_1(T'v_1) + \cdots + a_n(T'v_n) = a_1w_1 + \cdots + a_nw_n = T(v)$.

So, $T' = T$ since they agree at every $v \in V$.

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

3.5 definition: *addition and scalar multiplication on $\mathcal{L}(V, W)$*

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

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

for all $v \in V$.

linearity:

$$\begin{aligned}(S+T)(v+w) &= S(v+w)+T(v+w)\\ &= Sv + Sw + Tv + Tw\\ &= Sv + Tv + Sw + Tw\\ &= (S+T)(v) + (S+T)(w)\end{aligned}$$

$$\begin{aligned}(S+T)(\lambda v) &= S(\lambda v) + T(\lambda v)\\ &= \lambda(Sv) + \lambda(Tv)\\ &= \lambda(Sv + Tv)\\ &= \lambda(S+T)(v)\end{aligned}$$

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

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

Proof Routine – leave as an exercise!

3.7 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 $u \in U$.

$$\begin{tikzcd} U \arrow[r, "T"] \arrow[rr, bend right, "ST"'] & V \arrow[r, "S"] & W \end{tikzcd}$$

3.8 *algebraic properties of products of linear maps*

associativity

$(T_1T_2)T_3 = T_1(T_2T_3)$ whenever T_1, T_2, and T_3 are linear maps such that the products make sense (meaning T_3 maps into the domain of T_2, and T_2 maps into the domain of T_1).

$$U \xrightarrow{T_1} V \xrightarrow{T_2} W \xrightarrow{T_3} Y$$

identity

$TI = IT = T$ whenever $T \in \mathcal{L}(V, W)$; here the first I is the identity operator on V, and the second I is the identity operator on W.

distributive properties

$(S_1 + S_2)T = S_1T + S_2T$ and $S(T_1 + T_2) = ST_1 + ST_2$ whenever $T, T_1, T_2 \in \mathcal{L}(U, V)$ and $S, S_1, S_2 \in \mathcal{L}(V, W)$.

Proof exercise – routine

3.10 *linear maps take* 0 *to* 0

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

Proof

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

$T(0) = T(0+0)$ (since 0 is the additive id. of V)

$T(0) = T(0) + T(0)$ (by linearity of T)

$-T(0) + T(0) = -T(0) + (T(0) + T(0))$ (add $-T(0)$ by existence of additive inverses in W)

$0 = 0 + T(0)$ (def'n of additive inverses & associativity of $+$ in W)

$0 = T(0)$ (def'n of additive identity)

3B Null Spaces and Ranges

Null Space and Injectivity

3.11 definition: *null space*, null T

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

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

3.13 *the null space is a subspace*

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

$\boxed{\text{Proof}}$ Since T is linear, $T(0)=0$ (we proved this last time). So, $0 \in \text{null}\,T$. Now, if $u, v \in \text{null}\,T$, then note that, by linearity, $T(u+v) = Tu + Tv = 0+0 = 0$. So, $u+v \in \text{null}\,T$ as well. Finally, if $v \in \text{null}\,T$ $\ni$ $\lambda \in \mathbb{F}$, then by linearity, we have $T(\lambda v) = \lambda Tv = \lambda 0 = 0$. So, null T is a subspace.

3.14 definition: *injective*

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

3.15 *injectivity* $\Longleftrightarrow$ *null space equals* $\{0\}$

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

$\boxed{\text{Proof}}$ First, $\text{\$}$ $T \in \mathcal{L}(V,W)$ is injective. Then, since $T0=0$, null $T = \{0\}$ since anything else would imply $Tv = 0 = T0$ for some $v \neq 0$, contradicting injectivity.

Now, $\text{\$}$ $T \in \mathcal{L}(V,W)$ $\ni$ null $T = \{0\}$. $\text{\$}$ $u, v \in V$ $\ni$ $Tu = Tv$. Then, by linearity, we have $0 = Tu - Tv = T(u-v)$. So, $u - v \in \text{null}\,T$. But, since null $T = \{0\}$, this says $u - v = 0$, i.e. $u = v$. So, T is injective.

Range and Surjectivity

3.16 definition: *range*

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

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

3.18 *the range is a subspace*

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

$\boxed{\text{Proof}}$ Let $T \in \mathcal{L}(V,W)$. First, note that $0 = T0$, so $0 \in \operatorname{range} T$. Now, $\S$ $w_1, w_2 \in \operatorname{range} T$. Then, $\exists v_1, v_2 \in V$ st. $Tv_1 = w_1$ \$ $Tv_2 = w_2$. Then, by linearity, $T(v_1 + v_2) = Tv_1 + Tv_2 = w_1 + w_2$. So, $w_1 + w_2 \in \operatorname{range} T$ as well. Finally, $\S$ $w \in \operatorname{range} T$ \$ $\lambda \in \mathbb{F}$. Then, $\exists v \in V$ st. $Tv = w$ and, by linearity, we have $T(\lambda v) = \lambda Tv = \lambda w$. So, $\lambda w \in \operatorname{range} T$ as well.

3.19 definition: *surjective*

A function $T \colon V \to W$ is called *surjective* if its range equals W.

3.20 example: *surjectivity depends on the target space*

The differentiation map $D \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}))$ defined by $Dp = p'$ is not surjective, because the polynomial x^5 is not in the range of D. However, the differentiation map $S \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}), \mathcal{P}_4(\mathbf{R}))$ defined by $Sp = p'$ is surjective, because its range equals $\mathcal{P}_4(\mathbf{R})$, which is the vector space into which S maps.

Fundamental Theorem of Linear Maps

3.21 *fundamental theorem of linear maps*

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

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

$\boxed{\text{Proof}}$ Since $\operatorname{null} T \leq V$, it is finite dimensional since V is. Let $u_1, \ldots, u_m$ be a basis of null T. Since this is a lin. indep. list in V, it can be extended to a basis of V. Let $u_1, \ldots, u_m, v_1, \ldots, v_n$ be this extension. Then, $\dim V = n + m = n + \dim \operatorname{null} T$. We will show that $\dim \operatorname{range} T = n$ by showing that $Tv_1, \ldots, Tv_n$ is a basis for range T. Let $v \in V$. Since $u_1, \ldots, u_m, v_1, \ldots, v_n$ is a basis of V, $\exists\, a_1, \ldots, a_m, b_1, \ldots, b_n \in \mathbb{F}$ s.t. $v = a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_n v_n$. Then, by linearity, we have $Tv = a_1 Tu_1 + \cdots + a_m Tu_m + b_1 Tv_1 + \cdots + b_n Tv_n$. Since $u_1, \ldots, u_m \in \operatorname{null} T$, $Tu_i = 0$ for $i = 1, \ldots, m$. So, $\underbrace{Tv}_{\in \operatorname{range} T} = b_1 Tv_1 + \cdots + b_n Tv_n$. Hence, $Tv_1, \ldots, Tv_n$ spans range T. Now, we show $Tv_1, \ldots, Tv_n$ is lin. indep.: Let $c_1, \ldots, c_n \in \mathbb{F}$ be s.t. $c_1 Tv_1 + \cdots + c_n Tv_n = 0$. Then, by linearity, this says $T(c_1 v_1 + \cdots + c_n v_n) = 0$. So, $c_1 v_1 + \cdots + c_n v_n \in \operatorname{null} T$, and $\exists\, d_1, \ldots, d_m \in \mathbb{F}$ s.t.

$c_1 v_1 + \cdots + c_n v_n = d_1 u_1 + \cdots + d_m u_m$ (since $u_1, \ldots, u_m$ is a basis for null T).

But all the c's & d's must then be 0 since $u_1, \ldots, u_m, v_1, \ldots, v_n$ is a basis for V & is lin. indep. So, $Tv_1, \ldots, Tv_n$ is a lin. indep. list of vectors that spans range T, i.e., it is a basis of range T.

3.22 *linear map to a lower-dimensional space is not injective*

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

$\boxed{\text{Proof}}$ Let $T \in \mathcal{L}(V, W)$. Then,

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

(Since $\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$ by the FToLM.)

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

$$\geq \dim V - \dim W$$

$$> 0$$

So, $\dim \operatorname{null} T \neq 0$ & therefore $\operatorname{null} T \neq \{0\}$.

So, T is not injective by a previous result.

3.24 *linear map to a higher-dimensional space is not surjective*

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

$\boxed{\text{Proof}}$ Let $T \in \mathcal{L}(V, W)$. Then, by the FToLM, we have

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

$$\leq \dim V$$

$$< \dim W$$

So, $\operatorname{range} T \neq W$ & T cannot be surjective.

Fix n, m positive integers & let $A_{jk} \in \mathbb{F}$ for $j = 1, \dots, m$ & $k = 1, \dots, n$. Consider the system

$$\left.\begin{aligned} A_{11}x_1 + \dots + A_{1n}x_n &= 0 \\ &\vdots \\ A_{m1}x_1 + \dots + A_{mn}x_n &= 0 \end{aligned}\right] \text{rhs}$$

Clearly, $x_1 = \dots = x_n = 0$ is a solution. Does any other solution exist?

Define $T : \mathbb{F}^n \to \mathbb{F}^m$ by

$$T(x_1, \dots, x_n) = (A_{11}x_1 + \dots + A_{1n}x_n, \dots, A_{m1}x_1 + \dots + A_{mn}x_n)$$

3.26 *homogeneous system of linear equations*

A homogeneous system of linear equations with more variables than equations has nonzero solutions.

$\boxed{\text{Proof}}$ The map T above is linear, i.e. $T \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ (exercise: check this!). So, if $n > m$, we have $n = \dim \mathbb{F}^n > \dim \mathbb{F}^m = m$ & T cannot be injective, i.e. $\operatorname{null} T \neq \{0\}$. So, nontrivial solutions exist.

3.28 *inhomogeneous system of linear equations*

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

$\boxed{\text{Proof}}$ The inhomogeneous systems w/ solutions are those ones whose rhs lies in $\operatorname{range} T$. But, if $m > n$, this cannot be possible for every $w \in \mathbb{F}^m$ since T cannot be surjective.

3C Matrices

Representing a Linear Map by a Matrix

3.29 definition: *matrix*, $A_{j,k}$

Suppose m and n are nonnegative integers. An m-by-n *matrix* A is a rectangular array of elements of $\mathbf{F}$ with m rows and n columns:

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

The notation $A_{j,k}$ denotes the entry in row j, column k of A.

3.31 definition: *matrix of a linear map*, $\mathcal{M}(T)$

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

in W

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

lin. combo. of a basis of W

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

$\boxed{\text{Idea}}$

$$\mathcal{M}(T) = \begin{array}{c c} & \begin{array}{ccccc} v_1 & \cdots & v_k & \cdots & v_n \end{array} \\ \begin{array}{c} w_1 \\ \vdots \\ w_m \end{array} & \left(\begin{array}{ccccc} & & A_{1,k} & & \\ & & \vdots & & \\ & & A_{m,k} & & \end{array}\right) \end{array}.$$

The k^{th} column of $\mathcal{M}(T)$ consists of the scalars needed to write Tv_k as a linear combination of $w_1, \ldots, w_m$:

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

$\boxed{\text{Ex}}$

Suppose $T \in \mathcal{L}(\mathbf{F}^2, \mathbf{F}^3)$ is defined by

$$T(x, y) = (x + 3y, 2x + 5y, 7x + 9y).$$

With respect to the basis $(1,0),(0,1)$ of $\mathbb{F}^2$
and the basis $(1,0,0),(0,1,0),(0,0,1)$ of $\mathbb{F}^3$,

$M(T) = \begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{bmatrix}$ Since $T(1,0) = (1+3(0), 2(1)+5(0), 7(1)+9(0))$

$= (1, 2, 7) = 1(1,0,0) + 2(0,1,0) + 7(0,0,1)$

$T(0,1) = (3, 5, 9)$

$= 3(1,0,0) + 5(0,1,0) + 9(0,0,1)$

$\boxed{\text{Ex}}$

Suppose $D \in \mathcal{L}(\mathcal{P}_3(\mathbf{R}), \mathcal{P}_2(\mathbf{R}))$ is the differentiation map defined by $Dp = p'$.

With respect to the basis $1, x, x^2, x^3$ of $\mathcal{P}_3(\mathbb{R})$
and the basis $1, x, x^2$ of $\mathcal{P}_2(\mathbb{R})$, we have

$M(T) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$ since $D1 = 0 = 0(1) + 0(x) + 0(x^2)$

$Dx = 1 = (1)(1) + 0(x) + 0(x^2)$

$Dx^2 = 2x = 0(1) + 2(x) + 0(x^2)$

$Dx^3 = 3x^2 = 0(1) + 0(x) + 3(x^2)$

Addition and Scalar Multiplication of Matrices

3.34 definition: *matrix addition*

The *sum of two matrices of the same size* is the matrix obtained by adding corresponding entries in the matrices:

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

$$\mathcal{M}(T)v_k = \mathcal{M}(T)_{\cdot,k}\begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$
$$\mathcal{M}(S)v_k = \mathcal{M}(S)_{\cdot,k}\begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$
$$[\mathcal{M}(S)+\mathcal{M}(T)]v_k = (\mathcal{M}(S)_{\cdot,k}+\mathcal{M}(T)_{\cdot,k})\begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$
$$\shortparallel$$
$$\mathcal{M}(S+T)v_k$$

3.35 *matrix of the sum of linear maps*

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

$\boxed{\text{Proof}}$ Since $\mathcal{M}(S)$ & $\mathcal{M}(T)$ tell you the coefficients needs to write the image of a basis element in V as a liner combo of a basis of W, add together these coefficients and you, by linearity, get a liner combo that is equal to the sum of the map's image of that basis element.

3.36 definition: *scalar multiplication of a matrix*

The product of a scalar and a matrix is the matrix obtained by multiplying each entry in the matrix by the scalar:

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

3.38 *the matrix of a scalar times a linear map*

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

$\boxed{\text{Proof}}$ exercise.

3.39 notation: $\mathbf{F}^{m,n}$

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

3.40 $\dim \mathbf{F}^{m,n} = mn$

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

$\boxed{\text{Proof}}$ That $\mathbb{F}^{m,n}$ is a vector space will be left as an exercise (idea: write a matrix as a long list. Then, everything works since $\mathbb{F}^k$ is a vector space for all k a positive integer). To show $\dim \mathbb{F}^{m,n} = mn$, notice that the list of matrices $B_{1,1}, \ldots, B_{1,n}, \ldots, B_{m,1}, \ldots, B_{m,n}$, where B_{ij} is the matrix w/ zeros everywhere except the ij^{th} entry, is a basis for $\mathbb{F}^{m,n}$. So, since this list has length mn, $\dim \mathbb{F}^{m,n} = mn$.

Matrix Multiplication

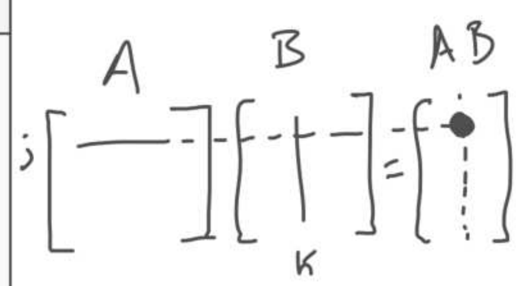
Consider linear maps $T : U \to V$ and $S : V \to W$. The composition ST is a linear map from U to W. Does $\mathcal{M}(ST)$ equal $\mathcal{M}(S)\mathcal{M}(T)$? This question does not yet make sense because we have not defined the product of two matrices. We will choose a definition of matrix multiplication that forces this question to have a positive answer. Let's see how to do this.

3.41 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 the equation

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

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.



$\boxed{\text{Ex}}$ $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 6 & 5 & 4 & 3 \\ 2 & 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 10 & 7 & 4 & 1 \\ 26 & 19 & 12 & 5 \\ 42 & 31 & 20 & 9 \end{bmatrix}$

3.43 *matrix of product of linear maps*

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

$\boxed{\text{proof}}$ Let $T \in \mathcal{L}(U,V)$, $S \in \mathcal{L}(V,W)$. Let $v_1, \dots, v_n$ be a basis of V, $w_1, \dots, w_m$ be a basis of W, $u_1, \dots, u_p$ be a basis for U.

Write $A = \mathcal{M}(S)$ & $B = \mathcal{M}(T)$. Now, for $k = 1, \dots, p$,

$$\begin{aligned}
\underbrace{(ST)u_k} &= S(Tu_k) \\
&= S\left(\sum_{r=1}^{n} B_{r,k} v_r\right) \\
&= \sum_{r=1}^{n} B_{r,k} S(v_r) \quad \leftarrow \text{linearity} \\
&= \sum_{r=1}^{n} B_{r,k}\left(\sum_{j=1}^{m} A_{j,r} w_j\right) \\
&= \sum_{r=1}^{n}\sum_{j=1}^{m}\left(A_{j,r}B_{r,k}\, w_j\right) \leftarrow \text{distributivity} \\
&= \sum_{j=1}^{n}\left(\sum_{r=1}^{m} A_{j,r}B_{r,k}\right) w_j \leftarrow \text{commutativity}
\end{aligned}$$

But this says that the k^{th} column of $\underbrace{\mathcal{M}(ST)}$ has as its j^{th} entry $\sum_{r=1}^{m} A_{j,r}B_{r,k}$. But, this is just what we get from looking at the j^{th} entry of the k^{th} column of $\underbrace{\mathcal{M}(S)\mathcal{M}(T)}$. So, $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.

3.44 notation: $A_{j,\cdot}$, $A_{\cdot,k}$

Suppose A is an m-by-n matrix.

- If $1 \le j \le m$, then $A_{j,\cdot}$ denotes the 1-by-n matrix consisting of row j of A.
- If $1 \le k \le n$, then $A_{\cdot,k}$ denotes the m-by-1 matrix consisting of column k of A.

3.46 *entry of matrix product equals row times column*

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

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

if $1 \le j \le m$ and $1 \le k \le p$. In other words, the entry in row j, column k, of AB equals (row j of A) times (column k of B).

$\boxed{\text{Proof}}$ By definition, $(AB)_{j,k} = A_{j,1}B_{1,k} + \cdots + A_{j,n}B_{n,k}$

$$= \underbrace{\begin{bmatrix} A_{j,1} & \cdots & A_{j,n} \end{bmatrix}}_{} \underbrace{\begin{bmatrix} B_{1,k} \\ \vdots \\ B_{n,k} \end{bmatrix}}_{}$$

$$= A_{j,\cdot}\, B_{\cdot,k}$$

3.48 *column of matrix product equals matrix times column*

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

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

if $1 \le k \le p$. In other words, column k of AB equals A times column k of B.

$\boxed{\text{Proof}}$ By the previous result, $(AB)_{j,k} = A_{j,\cdot}\, B_{\cdot,k}$.

So, $(AB)_{\cdot,k} = A_{\cdot,\cdot}\, B_{\cdot,k} = AB_{\cdot,k}$.

3.50 *linear combination of columns*

Suppose A is an m-by-n matrix and $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ is an n-by-1 matrix. Then

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

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

$\boxed{\text{Proof}}$ By the previous result, $(AB)_{\cdot,k} = AB_{\cdot,k}$.

But, if $B = b$, we have $Ab = (Ab)_{\cdot,1} = Ab_{\cdot,1} =$

This says $\begin{bmatrix} A_{\cdot,1} & \cdots & A_{\cdot,n} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = b_1 A_{\cdot,1} + \cdots + b_n A_{\cdot,n}$.

3.51 *matrix multiplication as linear combinations of columns*

Suppose C is an m-by-c matrix and R is a c-by-n matrix.

(a) 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.

(b) If $j \in \{1, ..., m\}$, 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.

Same logic

$\boxed{\text{Proof}}$ Fix $k \in \{1, \dots, n\}$. Then, column k of CR i.e. $(CR)_{\cdot,k}$, is equal to a linear combo of the columns of C w/ coefficients coming from $R_{\cdot,k}$ (i.e. the k^{th} column of R) by the previous result: $(CR)_{\cdot,k} = C(R_{\cdot,k}) = \begin{bmatrix} C_{\cdot,1} & \cdots & C_{\cdot,n} \end{bmatrix} \begin{bmatrix} R_{1,k} \\ \vdots \\ R_{n,k} \end{bmatrix} = R_{1k} C_{\cdot,1} + \cdots + R_{nk} C_{\cdot,n}$

Column–Row Factorization and Rank of a Matrix

3.52 definition: *column rank, row rank*

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

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

i.e. column rank A = $\dim \operatorname{Span}(A_{\cdot,1}, \dots, A_{\cdot,n})$

$\boxed{EX}$

Suppose

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

The column rank of A is the dimension of

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

The row rank of A is the dimension of

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

These seem to be the same!

Is this always so?

3.54 definition: *transpose*, A^{t}

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

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

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{t} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

3.56 *column–row factorization*

Suppose A is an m-by-n matrix with entries in $\mathbf{F}$ and column rank $c \geq 1$. Then there exist an m-by-c matrix C and a c-by-n matrix R, both with entries in $\mathbf{F}$, such that $A = CR$.

$\boxed{\text{Proof}}$ The list $A_{\cdot,1}, \dots, A_{\cdot,n}$ of columns of A can be reduced to a basis of $\operatorname{span}(A_{\cdot,1}, \dots, A_{\cdot,n})$ (i.e. a spanning list in a finite-dim. V.S. can be reduced to a basis). This basis has length $\dim(\operatorname{span}(A_{\cdot,1}, \dots, A_{\cdot,n})) = c$. The c columns that form this basis can be used to make a matrix C which has these basis elements as its columns. This matrix is m-by-c. Since any column in A can be written as a linear combo of the columns of C, let R be the matrix whose k^{th} column is the coefficients such that $A_{\cdot,k} = R_{1,k} C_{\cdot,1} + \dots + R_{c,k} C_{\cdot,c}$. Then, $A = CR$.

3.57 *column rank equals row rank*

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

$\boxed{\text{Proof}}$ Let c denote the column rank of A. Let $A = CR$ be the factorization assured by the previous theorem. Every row of A is a linear combo of R's rows, and the coefficients in this linear combo come from C's rows. Since R has c rows, this says that the row rank of A is less than or equal to c, its column rank.

Now, do the same to A^t and note that:

$\text{Column rank } A = \text{row rank } A^t \leq \text{column rank } A^t = \text{row rank } A$.

So, column rank $A \geq$ row rank A & column rank $A \leq$ row rank A. Hence, they are equal!

3.58 definition: *rank*

The *rank* of a matrix $A \in \mathbf{F}^{m,n}$ is the column rank of A. (or, equivalently, the row rank).

3D Invertibility and Isomorphisms

Invertible Linear Maps

3.59 definition: *invertible, inverse*

- A linear map $T \in \mathcal{L}(V, W)$ is called *invertible* if there exists a linear map $S \in \mathcal{L}(W, V)$ such that ST equals the identity operator on V and TS equals the identity operator on W.
- A linear map $S \in \mathcal{L}(W, V)$ satisfying $ST = I_V$ and $TS = I_W$ is called an *inverse of T* (note that the first I is the identity operator on V and the second I is the identity operator on W).

3.60 *inverse is unique*

An invertible linear map has a unique inverse.

$\boxed{\text{Proof}}$ S'pose $T \in \mathcal{L}(V, W)$ is invertible. Then, $\exists S \in \mathcal{L}(W, V)$ s.t. $TS = I_W$ & $ST = I_V$. Suppose $S' \in \mathcal{L}(W, V)$ satisfies $TS' = I_W$ & $S'T = I_V$. Then, we have

$$S = SI_W = S(TS') = (ST)S' = I_V S' = S'$$

So, $S = S'$ & inverses are unique!

3.61 notation: T^{-1}

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

3.63 *invertibility $\Longleftrightarrow$ injectivity and surjectivity*

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

$\boxed{\text{Proof}}$ First, assume $T \in \mathcal{L}(V,W)$ is invertible. Then, observe that if $w \in W$, then $T(T^{-1}w) = Iw = w$. So, T is surjective. For injectivity, suppose $u, v \in V$ are s.t. $Tu = Tv$. Then, $u = T^{-1}(Tu) = T^{-1}(Tv) = v$, so T is injective.

Now, suppose $T \in \mathcal{L}(V,W)$ is injective & surjective. For $w \in W$, define Sw to be the unique element of V s.t. $TSw = w$. This can be done because: T is surjective, so $\exists v \in V$ s.t. $Tv = w$ for whatever w we chose, and, since T is injective, if $Tv = w = Tu$, then $v = u$ (so, this choice is unique & S is well defined). By construction, $TS = I_W$. Now, observe that $T((ST)(v)) = (TS)(Tv) = Tv$. Since T is injective, $T(STv) = Tv$ implies $STv = v$. So, $ST = I_V$. All that remains is to show $S \in \mathcal{L}(W,V)$. Suppose $w_1, w_2 \in W$. Then, $T(Sw_1 + Sw_2) = TSw_1 + TSw_2 = w_1 + w_2$. Thus, $Sw_1 + Sw_2$ is the unique element that T maps to $w_1 + w_2$. By the definition of S, this says $S(w_1 + w_2) = Sw_1 + Sw_2$.

Similarly, let $\lambda \in \mathbb{F}$. Then, $T(\lambda Sw) = \lambda TSw = \lambda w$, and $S(\lambda w) = \lambda Sw$ by the definition of S. So, $S \in \mathcal{L}(W,V)$

That is, T is invertible & $S = T^{-1}$.

3.65 *injectivity is equivalent to surjectivity* (*if* $\dim V = \dim W < \infty$)

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

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

$\boxed{\text{Proof}}$ By the FToLM, we have

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

Now, if T is injective, $\operatorname{null} T = \{0\}$ & this says $\dim V = \dim \operatorname{range} T$ and, since $\dim V = \dim W$, $\dim \operatorname{range} T = \dim W$. Hence, $\operatorname{range} T = W$, so T is surjective as well. By the previous result, it is also invertible.

Likewise, if T is surjective, $\operatorname{range} T = W$ & we have $0 = \dim V - \dim \operatorname{range} T = \dim \operatorname{null} T$. So, $\operatorname{null} T = \{0\}$ & T is injective. By the previous result, it is also invertible.

$$\begin{tikzcd}
T \text{ injective} \arrow[rr, Leftrightarrow] \arrow[dr, Leftrightarrow] & & T \text{ surjective} \arrow[dl, Leftrightarrow] \\
& T \text{ invertible} &
\end{tikzcd}$$

$\boxed{\text{Ex}}$ Define $T \in \mathcal{L}(\mathcal{P}(\mathbb{R}))$ by $(Tp)(x) = ((x^2+5x+7)p(x))''$.

We will show that for any $q \in \mathcal{P}(\mathbb{R})$, there is some solution $p \in \mathcal{P}(\mathbb{R})$ to $Tp = q$. Let $m = \deg q$. Let $T' \in \mathcal{L}(\mathcal{P}_m(\mathbb{R}))$ be given by $(T'p)(x) = ((x^2+5x+7)p(x))''$.

Notice that if a polynomial has a second derivative of zero, then it is a poly of degree at most 1. Since multiplying a degree t poly by x^2+5x+7 produces a degree $t+2$ poly, $T'p = 0$ only when $p = 0$. So, T' is injective. Hence, invertible. So, $p = (T')^{-1}q$ & since T & T' agree, $T(T')^{-1}q = q$ as well.

3.68 $ST = I \iff TS = I$ *(on vector spaces of the same dimension)*

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 $TS = I$.

$\dim V = \dim W$

$\boxed{\text{Proof}}$ First, suppose $ST = I$. If $v \in V$ & $Tv = 0$, then $v = Iv = (ST)v = S(Tv) = S(0) = 0$. So, T is injective (since $\operatorname{null} T = \{0\}$). Since $\dim V = \dim W$, T is invertible. So, $T^{-1} = IT^{-1} = (ST)T^{-1} = S(TT^{-1}) = SI = S$ & $S = T^{-1}$. Thus, $TS = TT^{-1} = I$. Now, for the other direction, just reverse the roles of S & T in the previous argument.

Isomorphic Vector Spaces

3.69 definition: *isomorphism, isomorphic*

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

3.70 *dimension shows whether vector spaces are isomorphic*

Two finite-dimensional vector spaces over **F** are isomorphic if and only if they have the same dimension.

$\boxed{\text{proof}}$ $\Rightarrow$ V & W are isomorphic. Then, $\exists\, T \in \mathcal{L}(V,W)$ that is invertible and, hence, both injective & surjective. So, $\operatorname{null} T = \{0\}$ & $\operatorname{range} T = W$. The FToLM then becomes

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T = 0 + \dim W = \dim W.$$

$\Leftarrow$ $\dim V = \dim W < \infty$. Let $v_1, \ldots, v_n$ be a basis for V & let $w_1, \ldots, w_n$ be a basis for W (i.e. $n = \dim V = \dim W$).

Define $T \in \mathcal{L}(V,W)$ by $Tv = T(c_1v_1 + \cdots + c_nv_n) = c_1w_1 + \cdots + c_nw_n$ (where $c_1, \ldots, c_n \in \mathbb{F}$ are the unique coefficients in the linear combo of $v_1, \ldots, v_n$ s.t. $v = c_1v_1 + \cdots + c_nv_n$). T is surjective because $w_1, \ldots, w_n$ spans W. Furthermore, $Tv = 0 \iff c_1w_1 + \cdots + c_nw_n = 0$ which, by linear independence, implies $c_1 = \cdots = c_n = 0$ & $v = 0v_1 + \cdots + 0v_n = 0$. So, T is injective. Since T is both injective & surjective, it is an isomorphism (i.e. invertible).

Recall let $A = \mathcal{M}(T)$. Then, $Tv_k = \sum_{j=1}^{m} A_{jk} w_j$

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

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

i.e. $\dim W = m$

i.e. $\mathbb{F}^{\dim W, \dim V}$

i.e. $\dim V = n$

$\boxed{\text{Proof}}$ From a previous section, we know $\mathcal{M}$ is linear. So, to show $\mathcal{M}$ is an isomorphism, it is enough to show that $\mathcal{M}$ is injective & surjective. For injectivity, if $T \in \mathcal{L}(V, W)$ is s.t. $\mathcal{M}(T) = 0$. Then, $Tv_k = 0$ for $k = 1, \ldots, n$. Since $v_1, \ldots, v_n$ is a basis for V, this says $T = 0$. Hence, $\operatorname{null} \mathcal{M} = \{0\}$ & $\mathcal{M}$ is injective. For surjectivity, the linear map lemma says for any $A \in \mathbb{F}^{m,n}$, there exists $T \in \mathcal{L}(V, W)$ s.t. $Tv_k = \sum_{j=1}^{m} A_{jk} w_j$ for $k = 1, \ldots, n$. But this is just $\mathcal{M}(T)$, so $\mathcal{M}$ is surjective. Hence, it is an isomorphism.

3.72 $\dim \mathcal{L}(V, W) = (\dim V)(\dim W)$

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

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

$\boxed{\text{Proof}}$ By the previous result, $\mathcal{L}(V, W)$ & $\mathbb{F}^{\dim W, \dim V}$ are isomorphic. But, by the last result from last time, this means they have the same dimension. So, $\dim \mathcal{L}(V, W) = \dim \mathbb{F}^{\dim W, \dim V} = \dim W \dim V$.

Linear Maps Thought of as Matrix Multiplication

3.73 definition: *matrix of a vector, $\mathcal{M}(v)$*

Suppose $v \in V$ and $v_1, \dots, v_n$ is a basis of V. The *matrix of v* with respect to this basis is the n-by-1 matrix — i.e. $\dim V = n$

$$\mathcal{M}(v) = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

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

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

3.75 $\mathcal{M}(T)_{\cdot,k} = \mathcal{M}(Tv_k)$.

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

3.76 *linear maps act like matrix multiplication*

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

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

$\boxed{\text{Proof}}$ Suppose $v = b_1 v_1 + \dots + b_n v_n$ for $b_1, \dots, b_n \in \mathbb{F}$. Then,

$Tv = b_1 Tv_1 + \dots + b_n Tv_n$. Hence,

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

3.78 *dimension of* range T *equals 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 $\mathcal{M}(T)$.

$\boxed{\text{Proof}}$ If $v_1, \ldots, v_n$ is a basis for V & $w_1, \ldots, w_m$ is a basis for W. The map $\mathcal{M} \in \mathcal{L}(W, \mathbb{F}^{m,1})$ that takes $w \in W$ to $\mathcal{M}(w) = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$ where $w = a_1 w_1 + \cdots + a_m w_m$ is an isomorphism. Now, since range $T = \operatorname{span}(Tv_1, \ldots, Tv_n)$ for $T \in \mathcal{L}(V, W)$, take $\mathcal{M}\big|_{\text{range } T}$. This is an isomorphism from range T to $\operatorname{span}(\mathcal{M}(Tv_1), \ldots, \mathcal{M}(Tv_n))$. But, $\mathcal{M}(Tv_k) = \mathcal{M}(T)_{\cdot,k}$, so this says that range T is isomorphic to $\operatorname{span}(\mathcal{M}(T)_{\cdot,1}, \ldots, \mathcal{M}(T)_{\cdot,n})$, i.e. the range of T is the span of the columns of $\mathcal{M}(T)$. So, they have the same dimension, which is the rank of $\mathcal{M}(T)$.

Change of Basis

3.81 *matrix of product of linear maps*

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

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

$\underbrace{\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)}$

doesn't explicitly show what base is being used in the middle

3.82 *matrix of identity operator with respect to two bases*

Suppose that $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ are bases of V. Then the matrices

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

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

3.84 *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

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

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

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

$\boxed{\text{Proof}}$ Recall the following:

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

for $S \in \mathcal{L}(V, W)$, $T \in \mathcal{L}(U, V)$, $u_1, ..., u_m$ a basis for U, $v_1, ..., v_n$ a basis for V, and $w_1, ..., w_p$ a basis for W

Since $C^{-1} = \mathcal{M}(I, (v_1, ..., v_n), (u_1, ..., u_n))$, we have

$$\left.\begin{aligned} C^{-1}BC &= \mathcal{M}(I, v, u)\,\mathcal{M}(T, v, v)\,\mathcal{M}(I, u, v) \\ &= \mathcal{M}(IT, v, u)\,\mathcal{M}(I, u, v) \\ &= \mathcal{M}(ITI, u, u) \\ &= \mathcal{M}(T, u, u) \\ &= A \end{aligned}\right\}$$

here, $u = (u_1, ..., u_n)$
$v = (v_1, ..., v_n)$
for shorthand

3.86 *matrix of inverse equals inverse of matrix*

Suppose that $v_1, ..., v_n$ 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 $v_1, ..., v_n$.

$\boxed{\text{Proof}}$ Hint: $\mathcal{M}(T)^{-1}$ satisfies $\mathcal{M}(T)\mathcal{M}(T)^{-1} = I$. Could we use some bases for V in terms of the matrix of T & the matrix of its inverse?

3E Products and Quotients of Vector Spaces

Products of Vector Spaces

3.87 definition: *product of vector spaces*

Suppose $V_1, \ldots, V_m$ are vector spaces over $\mathbf{F}$. ← Same field!

- The *product* $V_1 \times \cdots \times V_m$ is defined by

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

- Addition on $V_1 \times \cdots \times V_m$ is defined by

$$(u_1, \ldots, u_m) + (v_1, \ldots, v_m) = (u_1 + v_1, \ldots, u_m + v_m).$$

- Scalar multiplication on $V_1 \times \cdots \times V_m$ is defined by

$$\lambda(v_1, \ldots, v_m) = (\lambda v_1, \ldots, \lambda v_m).$$

3.89 *product of vector spaces is a vector space*

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

3.92 *dimension of a product is the sum of dimensions*

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

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

$\boxed{\text{Proof}}$ Let's construct a basis for $V_1 \times \cdots \times V_m$.

Let $v_{1,k}, \ldots, v_{n_k,k}$ be a basis for V_k where $\dim V_k = n_k$.

We take $\underbrace{\overbrace{(v_{1,1}, 0, \ldots, 0), \ldots, (v_{n_1,1}, 0, \ldots, 0)}^{\dim V_1}, \overbrace{(0, v_{1,2}, 0, \ldots, 0), \ldots, (0, v_{n_2,2}, 0, \ldots, 0)}^{\dim V_2}, \ldots, (0, \ldots, 0, v_{n_m,m})}$ & so on

(i.e. take the list of vectors in $V_1 \times \cdots \times V_m$ where everything is 0 except for one component, which is a vector in a basis of the corresponding space).

Since $v_k \in V_k$ can be written uniquely as $v_k = a_1 v_{1,k} + \cdots + a_{n_k} v_{n_k,k}$, then $(0, \ldots, 0, v_k, 0, \ldots, 0) \in V_1 \times \cdots \times V_m$ can be written uniquely as a linear combo of

We can write every $(v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m$ uniquely as a linear combo of by breaking apart $(v_1, \ldots, v_m) = (v_1, 0, \ldots, 0) + \cdots + (0, 0, \ldots, 0, v_m)$. Hence, $\dim V_1 \times \cdots \times V_m$ is the sum of the dimensions of the V_i's.

3.93 *products and direct sums*

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

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

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

$\boxed{\text{Proof}}$ If Γ is injective, then $\Gamma(v_1, \ldots, v_m) = 0 \Rightarrow (v_1, \ldots, v_m) = 0$. So, the only way to write 0 as a sum of the form $v_1 + \cdots + v_m = 0$ with $v_k \in V_k$ is $v_1 = \cdots = v_m = 0$, which says $V_1 + \cdots + V_m$ is direct.

Now, if $V_1 \oplus \cdots \oplus V_m$, then Γ must be injective, since directness implies that $v_1 + \cdots + v_m = 0$ for $v_k \in V_k$ means $v_1 = \cdots = v_m = 0$, so $\operatorname{null} \Gamma = \{0\}$ & Γ is injective.

3.94 *a sum is a direct sum if and only if dimensions add up*

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

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

← Notice: no products!

$\boxed{\text{Proof}}$ First, note that Γ (from the previous theorem) is surjective. By the FToLM, Γ is injective if & only if $\dim(V_1 + \cdots + V_m) = \dim(V_1 \times \cdots \times V_m)$ (since $\dim V_1 \times \cdots V_m = \dim \operatorname{null} \Gamma + \dim \operatorname{range} \Gamma = \dim \operatorname{null} \Gamma + \dim(V_1 + \cdots + V_m)$).

Now, by the previous result, Γ is injective if & only if $V_1 + \cdots + V_m$ is direct. So, since $\dim(V_1 \times \cdots \times V_m) = \dim V_1 + \cdots + \dim V_m$, we have $\dim(V_1 + \cdots V_m) = \dim V_1 + \cdots + \dim V_m$ if & only if this sum is direct.

Quotient Spaces

3.95 notation: $v + U$

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

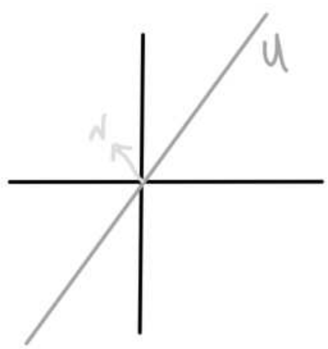
$$v + U = \{v + u : u \in U\}.$$

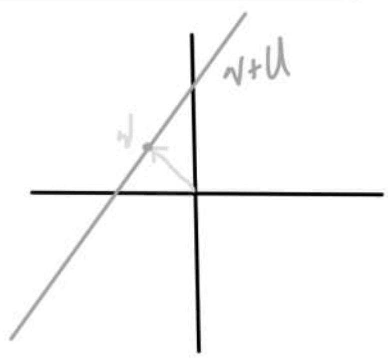
Visualize

3.97 definition: *translate*

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

Visualize



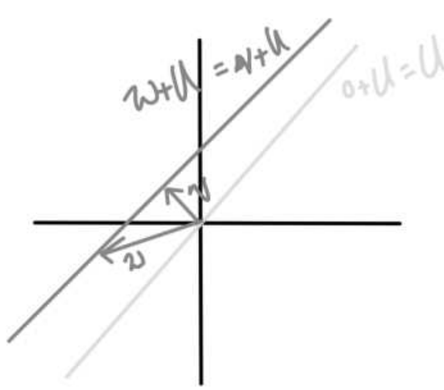


3.99 definition: *quotient space, V/U*

Suppose U is a subspace of V. Then the *quotient space* V/U is the set of all translates of U. Thus

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

Visualize



3.101 *two translates of a subspace are equal or disjoint*

Suppose U is a subspace of V and $v, w \in V$. Then

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

$\boxed{\text{Proof}}$ $\$$ $v-w \in U$. If $u \in U$, then observe that

$v+U \ni v+u = v+u+w-w = w+((v-w)+u) \in w+U$. So, $v+U \subseteq w+U$.

Likewise, repeating the argument w/ v \& w interchanged, $w+U \subseteq v+U$.

So, $v-w \in U \implies v+U = w+U$.

Now, $\$$ $v+U = w+U$. Then, $(v+U) \cap (w+U) = v+U = w+U \neq \emptyset$.

So, $v+U = w+U \Rightarrow (v+U) \cap (w+U) \neq \emptyset$.

Finally, $\$$ $(v+U) \cap (w+U) \neq \emptyset$. Then, $\exists u_1, u_2 \in U$ s.t.

$v+u_1 = w+u_2$. Rearranging, this says $v-w = u_2-u_1 \in U$.

So, $(v+U) \cap (w+U) \neq \emptyset \implies v-w \in U$.

3.102 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 $\lambda \in \mathbf{F}$.

3.103 *quotient space is a vector space*

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

$\boxed{\text{proof}}$ The problem we want to show does not happen is the following: If $v_1 + U = v_2 + U$ & $w_1 + U = w_2 + U$, then could we get $(v_1 + w_1) + U \neq (v_2 + w_2) + U$?

We show this never happens. Suppose $v_1 + U = v_2 + U$ & $w_1 + U = w_2 + U$. Then, $v_1 - v_2 \in U$ & $w_1 - w_2 \in U$. Now, this says, since U is a subspace, that $(v_1 - v_2) + (w_1 - w_2) \in U$. So, this says $(v_1 + w_1) - (v_2 + w_2) \in U$. But, we just showed that this implies $(v_1 + w_1) + U = (v_2 + w_2) + U$. So, addition makes sense!

Likewise, suppose $\lambda \in \mathbb{F}$ & $v_1 + U = v_2 + U$. Since U is a subspace, it closed under scaling, so $v_1 - v_2 \in U$ implies $\lambda(v_1 - v_2) \in U$. Hence, $\lambda v_1 - \lambda v_2 \in U$ & $\lambda v_1 + U = \lambda v_2 + U$. So, scalar mult. makes sense too!

With this, the rest is routine (exercise).

3.104 definition: *quotient map, π*

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

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

for each $v \in V$.

3.105 *dimension of quotient space*

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

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

$\boxed{\text{Proof}}$ Let π be the quotient map. Then, since $\pi(v) = v + U$ & $V/U = \{v + U : v \in V\}$, π is surjective.

Now, notice that $\pi(v) = 0 + U$ (i.e. $v \in \operatorname{null} \pi$), then $v + U = 0 + U$, which means $v = v - 0 \in U$. So, $\operatorname{null} \pi = U$.

By the FToLM, $\dim V = \dim \operatorname{range} \pi + \dim \operatorname{null} \pi$

$$\dim V = \dim V/U + \dim U$$

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

3.106 notation: $\widetilde{T}$

Suppose $T \in \mathcal{L}(V, W)$. Define $\widetilde{T}\colon V/(\operatorname{null} T) \to W$ by

$$\widetilde{T}(v + \operatorname{null} T) = Tv.$$

This makes sense: If $v + \operatorname{null} T = u + \operatorname{null} T$, then $v - u \in \operatorname{null} T$ & $T(v-u) = 0$, i.e. $\boxed{Tv = Tu.}$

So, if $v + \operatorname{null} T = u + \operatorname{null} T$, $\widetilde{T}(v + \operatorname{null} T) = Tv = Tu = \widetilde{T}(u + \operatorname{null} T)$

that is, $\widetilde{T}$ does not depend on the choice of v used to represent $v + \operatorname{null} T$.

3.107 *null space and range of* $\widetilde{T}$

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

(a) $\widetilde{T} \circ \pi = T$, where π is the quotient map of V onto $V/(\operatorname{null} T)$; $\pi : V \to V/_{\operatorname{null} T}$

(b) $\widetilde{T}$ is injective;

(c) $\operatorname{range} \widetilde{T} = \operatorname{range} T$;

(d) $V/(\operatorname{null} T)$ and $\operatorname{range} T$ are isomorphic vector spaces.

$\boxed{\text{Proof}}$

(a) If $v \in V$, then $(\widetilde{T} \circ \pi)(v) = \widetilde{T}(\pi(v)) = \widetilde{T}(v + \operatorname{null} T) = Tv$

So, $\widetilde{T} \circ \pi = T$

(b) We show injectivity by showing $\operatorname{null} \widetilde{T}$ is trivial (i.e. $\operatorname{null} \widetilde{T} = \{0 + \operatorname{null} T\}$). & $\widetilde{T}(v + \operatorname{null} T) = 0$. Then, $Tv = 0$, since $\widetilde{T}(v + \operatorname{null} T) = Tv$. So, $v \in \operatorname{null} T$ & $v + \operatorname{null} T = 0 + \operatorname{null} T$ since $v - 0 \in \operatorname{null} T$. Hence, $\widetilde{T}$ is inj.

(c) If $w \in \operatorname{range} T$, then $\exists v \in V$ s.t. $Tv = w$. So, $\widetilde{T}(v + \operatorname{null} T) = Tv = w$ & $w \in \operatorname{range} \widetilde{T}$ as well.

Write this argument in reverse to get $w \in \operatorname{range} T$ whenever $w \in \operatorname{range} \widetilde{T}$.

(d) From (b) & (c), the map $R : V/_{\operatorname{null} T} \to \operatorname{range} T$ defined by $R(v + \operatorname{null} T) = Tv$ is the same as $\widetilde{T}$ but with a different codomain. So, R is inj. & surject.

3F Duality

Dual Space and Dual Map

3.108 definition: *linear functional*

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

3.110 definition: *dual space, V'*

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

3.111 $\dim V' = \dim V$

Suppose V is finite-dimensional. Then V' is also finite-dimensional and

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

$\boxed{\text{Proof}}$ We know that if V & W are finite-dim vector spaces, then $\dim \mathcal{L}(V,W) = (\dim V)(\dim W)$.
So, if $W = \mathbb{F}$, $\dim W = \dim \mathbb{F} = 1$ and we get
$$\dim V' = \dim \mathcal{L}(V, \mathbb{F}) = (\dim V)(\dim \mathbb{F}) = \dim V.$$

3.112 definition: *dual basis*

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

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

Since φ_j is defined on a basis, it is defined everywhere by linearity — i.e. $\varphi_j(v) = \varphi_j(a_1v_1 + \cdots + a_nv_n) = a_1\varphi_j(v_1) + \cdots + a_n\varphi_j(v_n) = a_j\varphi_j(v_j) = a_j$.

3.114 *dual basis gives coefficients for linear combination*

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

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

for each $v \in V$.

$\boxed{\text{proof}}$ Since $v_1, \ldots, v_n$ is a basis of V, $\exists\, a_1, \ldots, a_n \in \mathbb{F}$ s.t. $v = a_1v_1 + \cdots + a_nv_n$. So, by linearity, we get

$$\varphi_j(v) = \varphi_j(a_1v_1 + \cdots + a_nv_n) = a_1\varphi_j(v_1) + \cdots + a_n\varphi_j(v_n) = a_j\varphi_j(v_j) = a_j$$

for each $j \in \{1, \ldots, n\}$. So, this says

$$v = a_1v_1 + \cdots + a_nv_n = \varphi_1(v)v_1 + \cdots + \varphi_n(v)v_n.$$

3.116 *dual basis is a basis of the dual space*

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

$\boxed{\text{proof}}$ S $v_1, \ldots, v_n$ is a basis of V. Let $\varphi_1, \ldots, \varphi_n$ denote the corresponding dual basis. For lin. indep., S $a_1, \ldots, a_n \in \mathbb{F}$ are s.t. $a_1\varphi_1 + \cdots + a_n\varphi_n = 0$ (where $0 \in \mathcal{L}(V, \mathbb{F}) = V'$ is the map $0(v) = 0\ \forall v \in V$).

But then, $0 = 0(v_j) = (a_1\varphi_1 + \cdots + a_n\varphi_n)(v_j) = a_1\varphi_1(v_j) + \cdots + a_n\varphi_n(v_j) = a_j\varphi_j(v_j) = a_j$.

So, since this holds for all $j \in \{1, \ldots, n\}$, this says $a_1 = \cdots = a_n = 0$.

Thus, $\varphi_1, \ldots, \varphi_n$ is a lin. indep. list in V' of length $\dim V'$, so it is a basis.

3.118 definition: *dual map, T'*

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

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

The claimed linearity is a consequence of the definition of T': it is just a composition of two liner maps!

Visualize

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[dr, dashed, "T' = \varphi \circ T"'] & W \arrow[d, "\varphi"] \\ & \mathbb{F} \end{tikzcd}$$

Notice: T' basically "flops" the domain & codomain of T around. Technically, it flops & dualizes, but since V & V' (also, W & W') are isomorphic if finite dim., this just relabeling.

$$T'(\varphi + \psi) = (\varphi + \psi) \circ T = \varphi \circ T + \psi \circ T = T'(\varphi) + T'(\psi)$$

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

3.120 *algebraic properties of dual maps*

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

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

(b) $(\lambda T)' = \lambda T'$ for all $\lambda \in \mathbf{F}$;

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

Proof

(a) Let $\varphi \in W'$. Then, $(S+T)'(\varphi) = \varphi \circ (S+T) = \varphi \circ S + \varphi \circ T = S'(\varphi) + T'(\varphi)$

(b) Let $\varphi \in W'$ & $\lambda \in \mathbb{F}$. Then, $(\lambda T)'(\varphi) = \varphi \circ (\lambda T) = \lambda(\varphi \circ T) = \lambda T'(\varphi)$

since φ is linear

(c) $\oint$ $S \in \mathcal{L}(W, U)$ & $\varphi \in U'$. Then, $(ST)'(\varphi) = \varphi \circ S T = (\varphi \circ S) \circ T = T'(\varphi \circ S) = T'(S'(\varphi)) = (T'S')(\varphi)$.

Null Space and Range of Dual of Linear Map

3.121 definition: *annihilator,* U^0

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

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

3.124 *the annihilator is a subspace*

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

$\boxed{\text{Proof}}$ We need only show $0 \in U^0$, U^0 is closed under addition, & U^0 is closed under scalar multiplication.

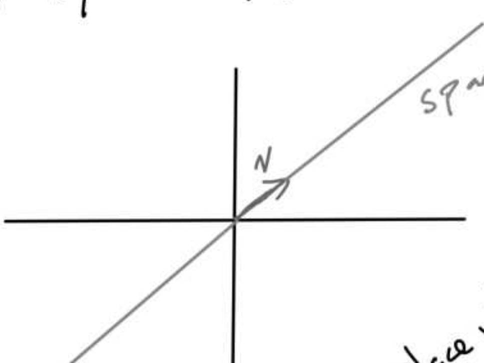
First, note that $0 \in V' = \mathcal{L}(V, \mathbb{F})$ is defined by $0(v) = 0 \ \forall v \in V$.

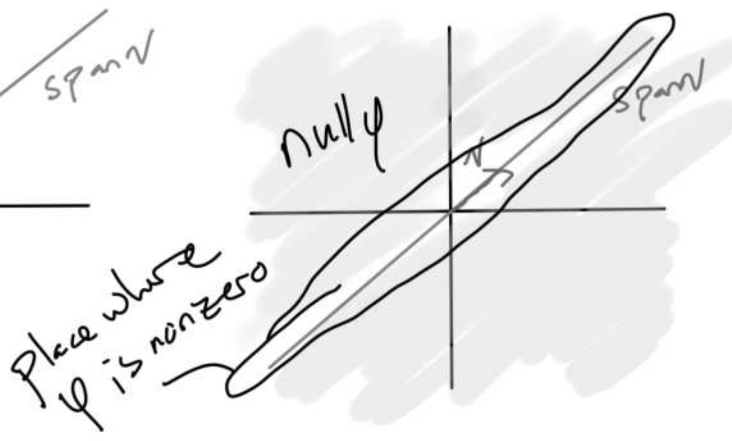
So, $0(u) = 0$ for all $u \in U$ & $0 \in U^0$.

Now, let $\varphi, \psi \in U^0$. Then, $\varphi(u) = 0$, $\psi(u) = 0 \ \forall u \in U$. Hence, we have $(\varphi + \psi)(u) = \varphi(u) + \psi(u) = 0 + 0 = 0$, so $\varphi + \psi \in U^0$ as well.

Finally, let $\varphi \in U^0$ & $\lambda \in \mathbb{F}$. Then, $(\lambda\varphi)(u) = \lambda(\varphi(u)) = \lambda(0) = 0 \ \forall u \in U$, so $\lambda\varphi \in U^0$ as well.

$\boxed{\text{Idea}}$ We can look at vectors & the subspaces they span. The idea of a dual space is to invert this:





The idea is to find something in V' that is 0 except on span v

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

$T \in \mathcal{L}(V,W)$
$T' \in \mathcal{L}(W',V')$

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[dr, dashed] & W \arrow[d, "\varphi \in W'"] \\ & \mathbb{F} \end{tikzcd}$$

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

3.125 *dimension of the annihilator*

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

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

$\boxed{\text{proof}}$ Define $i \in \mathcal{L}(U,V)$ by $i(u) = u$. So, $i' \in \mathcal{L}(V',U')$

By th FToLM, $\dim \operatorname{range} i' + \dim \operatorname{null} i' = \dim V'$. But, notice: $\operatorname{null} i' = \{\varphi \in V' : i'(\varphi) = 0\} = \{\varphi \in V' : \varphi \circ i = 0\} = \{\varphi \in V' : \varphi(u) = 0 \ \forall u \in U\}$

So, $\operatorname{null} i' = U^0$. Since we have shown $\dim V = \dim V'$ when V is finite-dimensional, we have $\dim \operatorname{range} i' + \dim U^0 = \dim V$.

Hence, if $\underline{\dim \operatorname{range} i'} = \underline{\dim U}$, we are done!

notice: $\operatorname{range} i' = \{\psi \in U' : \exists \varphi \in V' \ i'(\varphi) = \psi\} = \{\psi \in U' : \exists \varphi \in V' \ \varphi(u) = \psi(u) \ \forall u \in U\}$

Since every $\psi \in U'$ can be extended to a linear functional on V (we proved this in th homework), this says $\operatorname{range} i' = U'$ (i.e. i' is surjective).

But, $\dim U' = \dim U$, so we are finished.

3.127 *condition for the annihilator to equal* $\{0\}$ *or the whole space*

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

(a) $U^0 = \{0\} \iff U = V$;

(b) $U^0 = V' \iff U = \{0\}$.

$\boxed{\text{proof}}$ For part (a):

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

For part (b):

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

3.128 *the null space of T'*

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

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

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

$\boxed{\text{proof}}$ Suppose $\varphi \in \text{null}\, T'$. So, $\underline{0} = T'(\varphi) = \varphi \circ T$. Thus, for every $v \in V$, $\varphi(T(v)) = 0$. Hence, $\varphi \in (\text{range}\, T)^0$. So, $\text{null}\, T' \subseteq (\text{range}\, T)^0$.

Now, suppose $\psi \in (\text{range}\, T)^0$. Then, $\psi(T(v)) = 0$ for every $v \in V$, i.e. $(T'(\psi))(v) = 0$ for all $v \in V$. So, $\psi \in \text{null}\, T'$. Hence, $\text{null}\, T' \supseteq (\text{range}\, T)^0$.

Therefore, $\text{null}\, T' = (\text{range}\, T)^0$.

For the second part, observe:

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

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

3.129 *T surjective is equivalent to T' injective*

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

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

$\boxed{\text{proof}}$ If $T \in \mathcal{L}(V, W)$ is surjective, then $\text{range}\, T = W$.

But:
$$\begin{aligned}
\text{range}\, T = W &\iff (\text{range}\, T)^0 = \{0\} \\
&\iff \text{null}\, T' = \{0\} \\
&\iff T' \text{ is injective}
\end{aligned}$$

So, we are done!

3.130 *the range of T'*

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

(a) $\dim \operatorname{range} T' = \dim \operatorname{range} T$;

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

$\boxed{\text{proof}}$ For part (a), we have

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

For part (b), $\varphi \in \operatorname{range} T'$. Then, $\exists \psi \in W'$ s.t. $T'(\psi) = \varphi$. Now, if $v \in \operatorname{null} T$, then $\varphi(v) = T'(\psi)(v) = \psi(T(v)) = \psi(0) = 0$. Hence, $\varphi \in (\operatorname{null} T)^0$. So, $\operatorname{range} T' \subseteq (\operatorname{null} T)^0$.

To show $\operatorname{range} T' = (\operatorname{null} T)^0$, we show they have the same dimension.

Notice: $\dim \operatorname{range} T' = \dim \operatorname{range} T = \dim V - \dim \operatorname{null} T = \dim (\operatorname{null} T)^0$.

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

3.131 *T injective is equivalent to T' surjective*

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

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

$\boxed{\text{proof}}$

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

Matrix of Dual of Linear Map

3.132 *matrix of T' is transpose of matrix of T*

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

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

$\boxed{\text{Proof}}$ If $\psi_1, \dots, \psi_m$ is a basis for W' & $\varphi_1, \dots, \varphi_n$ is a basis for W, where $\psi_1, \dots, \psi_m$ is the dual basis of $w_1, \dots, w_m$ & $\varphi_1, \dots, \varphi_n$ is the dual basis of $v_1, \dots, v_n$, then we have

$$T'(\psi_j) = \sum_{r=1}^{n} \mathcal{M}(T')_{r,j} \varphi_r$$

Now, apply this to v_k.

$$\psi_j(Tv_k) = T'(\psi_j)(v_k) = \sum_{r=1}^{n} \mathcal{M}(T')_{r,j} \varphi_r(v_k) = \mathcal{M}(T')_{k,j}$$

We also know
$$\begin{aligned} \psi_j(T(v_k)) &= \psi_j\left(\sum_{r=1}^{m} \mathcal{M}(T)_{r,k} w_r\right) \\ &= \sum_{r=1}^{m} \mathcal{M}(T)_{r,k} \psi_j(w_r) \\ &= \mathcal{M}(T)_{j,k} \end{aligned}$$

So, $\mathcal{M}(T')_{k,j} = \mathcal{M}(T)_{j,k}$ and
$$\mathcal{M}(T') = (\mathcal{M}(T))^{T}.$$

3.133 *column rank equals row rank*

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

$x^2+1=0$
can't factor over $\mathbb{R}$

5A Invariant Subspaces

5.1 definition: *operator*

A linear map from a vector space to itself is called an *operator*.

Recall that we defined the notation $\mathcal{L}(V)$ to mean $\mathcal{L}(V,V)$.

i.e. operators are maps $T \in \mathcal{L}(V)$

5.2 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 $u \in U$.

i.e. U is invariant under T if $T|_U$ is also an operator (i.e. $T|_U \in \mathcal{L}(U)$).

5.5 definition: *eigenvalue*

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

$\boxed{\text{Idea}}$ The smallest nontrivial subspaces of a vector space V are $U = \operatorname{Span}(v)$ for $v \in V$ w/ $v \neq 0$. So, if we can "break down" V into its smallest components, i.e. $V = V_1 \oplus \cdots \oplus V_m$ w/ $V_i = \operatorname{span}(v_i)$ (w/ $v_i \neq 0$)

Then, if each subspace is invariant under T, we have a good description of what T does: Since every $v \in V$ can be written uniquely as $v = u_1 + \cdots + u_n$ w/ $u_i \in \operatorname{Span}(v_i)$, then $u_i = a_i v_i$ for some $a_i \in \mathbb{F}$. By linearity, $Tv = \underbrace{a_1 Tv_1 + \cdots + a_m Tv_m}$. So, if V_i is invariant under T, this says $T(av_i) \in V_i$, i.e. $T(av_i) = bv_i$. So, $Tv_i = \overset{\lambda_i}{\overline{\frac{b}{a}}} v_i$ if $a \neq 0$. Hence, T is just like a 1-dimensional linear map on V_i & for $v \in V$, $Tv = a_1\lambda_1 v_1 + \cdots + a_m \lambda_m v_m$

That is, we can tell exactly what T does in terms of the most simple linear map by finding 1-dimensional subspaces that are invariant under T.

5.7 *equivalent conditions to be an eigenvalue*

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

(a) λ is an eigenvalue of T.

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

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

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

Reminder: $I \in \mathcal{L}(V)$ is the identity operator. Thus $Iv = v$ for all $v \in V$.

$\boxed{\text{Proof}}$ (sketch) Notice: $Tv = \lambda v$ for some $0 \neq v \in V$ iff only if $Tv - \lambda v = 0$ which says $(T - \lambda I)v = 0$. $\iff$ $T - \lambda I$ has a nontrivial nullspace (i.e. null $T - \lambda I \neq \{0\}$)
$\iff$ $T - \lambda I$ is not injective
$\iff$ $T - \lambda I$ is not surjective
$\iff$ $T - \lambda I$ is not invertible

where the last two $\iff$'s are due to the fact that $T - \lambda I$ is a map between VS.'s of the same dimension.

5.8 definition: *eigenvector*

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

5.11 *linearly independent eigenvectors*

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

$\boxed{\text{Proof}}$ Suppose there is a linearly dependent list of eigenvectors of $T \in \mathcal{L}(V)$ corresponding to distinct eigenvalues of T. Let $v_1, \dots, v_m$ be this list and suppose further that m is minimal. Necessarily, $m \geq 2$, since a list w/ one nonzero vector is lin. indep. By our assumed linear dependence, $\exists a_1, \dots, a_m \in \mathbb{F}$ s.t. $a_1 v_1 + \dots + a_m v_m = 0$ and not all of the a_i's are zero. By minimality of m, notice that none of the a_i are zero, as if this were the case, we could remove the corresponding v_i from our list. Now, let us apply $T - \lambda_m I$ where λ_m is the eigenvalue corresponding to v_m, to both sides:

$$0 = (T - \lambda_m I)(a_1 v_1 + \dots + a_m v_m) = (T - \lambda_m I) a_1 v_1 + \dots + (T - \lambda_m I) a_m v_m = a_1 T v_1 - \lambda_m a_1 v_1 + \dots + a_{m-1} T v_{m-1} - \lambda_m a_{m-1} v_{m-1}$$
$$= a_1(\lambda_1 - \lambda_m) v_1 + \dots + a_{m-1}(\lambda_{m-1} - \lambda_m) v_{m-1}$$

But notice, since each eigenvalue is distinct, $\lambda_i - \lambda_m \neq 0$ for $i = 1, \dots, m-1$. Hence, we have a linear combo of $v_1, \dots, v_{m-1}$ that is equal to zero and has no coefficient equal to 0. This contradicts the minimality of m!

5.12 *operator cannot have more eigenvalues than dimension of vector space*

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

Polynomials Applied to Operators

5.13 notation: T^m

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

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

$$T^{-m} = (T^{-1})^m.$$

5.14 notation: $p(T)$

Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbf{F})$ is a polynomial given by

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

for all $z \in \mathbf{F}$. Then $p(T)$ is the operator on V defined by

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

5.16 definition: *product of polynomials*

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

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

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

5.17 *multiplicative properties*

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

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

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

Informal proof: When a product of polynomials is expanded using the distributive property, it does not matter whether the symbol is z or T.

$\boxed{\text{Proof}}$ The second part is easy: T commutes w/ itself \$ both P(T) \$ q(T) only involve scaled sums of powers of T.

For a longer answer: (a) $(pq)(z) = p(z)q(z)$ by definition. So, plug in T to get $(pq)(T) = p(T)q(T)$ (this is not entirely correct to say, but is basically the idea). For full clarity, $(pq)(z) = \sum_{j=0}^{m}\sum_{k=0}^{n} a_j b_k z^{j+k}$ where $p(z) = \sum_{j=0}^{m} a_j z^j$ \$ $q(z) = \sum_{k=0}^{n} b_k z^k$. Thus, $(pq)(T) = \sum_{j=0}^{m}\sum_{k=0}^{n} a_j b_k T^{j+k} = \left(\sum_{j=0}^{m} a_j T^j\right)\left(\sum_{k=0}^{n} b_k T^k\right) = p(T)\,q(T)$.

For (b) just use (a) twice: $(pq)(T) = p(T)q(T)$ \$ $(pq)(T) = (qp)(T) = q(T)p(T)$.

5.18 *null space and range of $p(T)$ are invariant under T*

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

$\boxed{\text{Proof}}$ \$ $u \in \text{null } p(T)$. Then, $p(T)u = 0$. Thus,

$p(T)(Tu) = T(p(T)u) = T0 = 0$. Hence, $Tu \in \text{null } p(T)$.

Since T is just q(T) for $q(z) = z$

\$ $v \in \text{range } P(T)$. Then, $\exists\, u \in V$ s.t. $P(T)u = v$.

$Tv = T(P(T)u) = P(T)(Tu)$. So, $Tv \in \text{range } P(T)$.

So, in either case, T does not move things in the null space (resp. range) out of the null space (resp. range) of P(T). Hence, both null P(T) \$ range P(T) are invariant under T.

5.19 *existence of eigenvalues*

Every operator on a finite-dimensional nonzero complex vector space has an eigenvalue.

$\boxed{\text{Proof}}$ $\ni$ V is a finite-dim. vector space over $\mathbb{C}$ & $T \in \mathcal{L}(V)$. Let $\dim V = n < \infty$. Choose any $0 \neq v \in V$. Then, the list $v, Tv, T^2v, \ldots, T^n v$ is linearly dependent, since this is a list of length $n+1 > n$. This says there exists coefficients $a_0, a_1, \ldots, a_n \in \mathbb{C}$ s.t. not all are zero &

$$0 = a_0 v + a_1 Tv + a_2 T^2 v + \cdots + a_n T^n v$$

Now, this gives at least one polynomial $(\text{ex. } p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n)$ s.t. $p(T)v = 0$. Let us suppose that this polynomial is chosen to be of minimal degree. So, by the fundamental theorem of algebra, there exists a poly $q \in \mathcal{P}(\mathbb{C})$ & a value $\lambda \in \mathbb{C}$ s.t.

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

So, $0 = p(T)v = (T - \lambda I) q(T) v$. Since q has degree less than p, we know $q(T)v \neq 0$ (since we supposed ⌴).

Hence, λ is an eigenvalue of T & $q(T)v$ is a corresponding eigenvector.

Eigenvalues and the Minimal Polynomial

5.21 definition: *monic polynomial*

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

$\boxed{\text{Ex}}$ $p(z) = 4z + 9z^2$ is not monic
$q(z) = 5z + 11z^2 + z^7$ is monic

5.22 *existence, uniqueness, and degree of minimal polynomial*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then there is a unique monic polynomial $p \in \mathcal{P}(\mathbf{F})$ of smallest degree such that $p(T) = 0$. Furthermore, $\deg p \leq \dim V$.

Not proven yet!

$\boxed{\text{Proof}}$ § $\dim V = n < \infty$ & $T \in \mathcal{L}(V)$.

Write $\text{poly}(T) = \{p(T) : p \in \mathcal{P}(\mathbb{F})\}$. Now, since $0 \in \text{poly}(T)$ (i.e. $0 = 0(T)$) & $q(T) + p(T) = (q+p)(T)$ & $\lambda(q(T)) = (\lambda q)(T)$, $\text{Poly}(T) \subseteq \mathcal{L}(V)$. Since $\dim V < \infty$, $\dim \mathcal{L}(V) < \infty$ & we can see that $\dim \text{Poly}(T) = m < \infty$ as well. Hence, the list

$$I = T^0, T, T^2, \dots, T^m$$

is linearly dependent since it is of length $m+1$. Therefore, there exist $\underbrace{a_0, \dots, a_m}_{\text{not all } 0} \in \mathbb{F}$ s.t. $a_0 I + a_1 T + \cdots + a_m T^m = 0$.
§ that $k \leq m$ is the largest value s.t. $a_k \neq 0$ (i.e. $a_i = 0$ for $i > k$). Then, the polynomial $\frac{1}{a_k}(a_0 + a_1 z + \cdots + a_k z^k) = p(z)$ is monic and satisfies $p(T) = \frac{1}{a_k}(a_0 I + a_1 T + \cdots + a_k T^k + \cdots + a_m T^m) = \frac{1}{a_k} 0 = 0$.
So, such a poly exists. Hence, we can choose one of lowest degree.
Now, to show uniqueness, § p & q are monic polys of minimal degree s.t. $p(T) = q(T) = 0$. Now, $(p-q)(T) = 0$ and notice that if $p - q \neq 0$, then we have a poly of lower degree (since the coefficient on the highest degree term of both p & q is 1) which can be made monic by dividing by the coefficient of the highest power of z in it. this would then contradict minimality, so $p = q$.

5.24 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 $p \in \mathcal{P}(\mathbf{F})$ of smallest degree such that $p(T) = 0$.

5.27 *eigenvalues are the zeros of the minimal polynomial*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$.

(a) The zeros of the minimal polynomial of T are the eigenvalues of T.

(b) If V is a complex vector space, then the minimal polynomial of T has the form

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

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

$\boxed{\text{Proof}}$ Let p be the minimal poly. of $T \in \mathcal{L}(V)$.
Now, $\S$ λ is a zero of p. Then, by the FToA, $p(z) = (z-\lambda)q(z)$ where q is monic $\$$ of degree less than p. Hence, since $p(T)=0$, $(T-\lambda I)q(T) = 0$ $\$$ b/c q is of lower degree than p, $q(T) \neq 0$. Hence, applying any $v \in V$ s.t. $q(T)v \neq 0$ (means this is possible), we get that $q(T)v \in \operatorname{null}(T-\lambda I)$ $\$$ λ is an e.-val. of T.
Now, $\S$ λ is an eigenvalue of T. Then, $\exists$ $0 \neq v \in V$ st $Tv = \lambda v$. Accordingly, $T^k v = T^{k-1}\lambda v = \cdots = \lambda^k v$. Hence, $p(T)v = p(\lambda)v$. Since p is the min. poly. of T, $p(T) = 0$, so $p(\lambda)v = 0$ $\$$ $p(\lambda) = 0$ since $v \neq 0$. Thus, λ is a zero of p.
For part b, just use part (a) $\$$ the FToA to factor p.

5.29 $q(T) = 0 \iff q$ *is a polynomial multiple of the minimal polynomial*

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

$\boxed{\text{proof}}$ Suppose p is the minimal polynomial of T.

If $q \in \mathcal{P}(\mathbb{F})$ is s.t. $q(T) = 0$, then doing division, there are polynomials $s, r \in \mathcal{P}(\mathbb{F})$ s.t.

$$q = ps + r$$

w/ degree of r less than the degree of p.

Then, $\underline{0 = q(T)} = \underline{p(T)}s\underline{(T)} + \underline{r(T)} = 0\,s(T) + r(T) = r(T)$.

So, $\underline{r(T) = 0}$ & $\deg r < \deg p$, so by minimality, r is the zero polynomial. That is, $q = ps$.

That is, q is a polynomial multiple of p.

Conversely, suppose $q = ps$ for some poly $s \in \mathcal{P}(\mathbb{F})$. Then,

$q(T) = p(T)s(T) = 0\,s(T) = 0$.

5.31 *minimal polynomial of a restriction operator*

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 $T|_U$.

avoid this mess

$\cancel{T : V \to V}$

$\cancel{T|_U : U \to V}$

$\cancel{\hat{T}|_U : U \to U}$

$\boxed{\text{proof}}$ Let p be the min. poly. of T.

Hence, $p(T) = 0$ & $p(T)v = 0v = 0$ for all $v \in V$.

So, $p(T)u = 0$ for all $u \in U$. Hence, $p(T|_U) = 0$ (where this is $0 \in \mathcal{L}(U)$). Thus, by the previous result, p is a poly. multiple of the min. poly. of $T|_U$.

5.32 *T not invertible $\iff$ constant term of minimal polynomial of T is 0*

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.

$\boxed{\text{Proof}}$ Let p be the mini. poly. of T.

T is not invertible $\iff$ 0 is an eigenvalue of T
$\iff$ 0 is a zero of p
$\iff$ $p(0) = 0$
$\iff$ the constant term of p is zero.

Eigenvalues on Odd-Dimensional Real Vector Spaces

5.33 *even-dimensional null space*

Suppose $\mathbf{F} = \mathbf{R}$ and V is finite-dimensional. Suppose also that $T \in \mathcal{L}(V)$ and $b, c \in \mathbf{R}$ with $b^2 < 4c$. Then $\dim \operatorname{null}(T^2 + bT + cI)$ is an even number.

$\boxed{\text{Proof}}$ Recall that the null space of $p(T)$ is invariant under T for any polynomial $p \in \mathcal{P}(\mathbb{R})$. If necessary, replace V w/ $\operatorname{null}(T^2+bT+cI)$ & T with $T|_{\operatorname{null}(T^2+bT+cI)}$ so that we are only dealing w/ $\operatorname{null}(T^2+bT+cI)$. This lets us assume $T^2+bT+cI = 0$ (since we are taking V to be $\operatorname{null}(T^2+bT+cI)$ if it is not already). We want to show $\dim(\operatorname{null}(T^2+bT+cI)) = \dim V$ is even. Suppose $\lambda \in \mathbb{R}$ & $v \in V$ is s.t. $Tv = \lambda v$. Then,

$$0 = (T^2+bT+cI)v = \lambda^2 v + b\lambda v + cv = \underbrace{\left(\left(\lambda + \frac{b}{2}\right)^2 + c - \frac{b^2}{4}\right)}_{}v$$

Since $b^2 < 4c$, $0 < 4c - b^2$ & $0 < c - \frac{b^2}{4}$. So, $\swarrow$ is a positive #. Hence, $v = 0$. So, T has no eigenvectors.

$\boxed{\text{proof}}$ (continued)

Let U be a subspace of V that is invariant under T. Moreover, $\int$ $\dim U$ is as large as possible while remaining even. If $\underline{U=V}$, we are done. If not, let $w \in V$ be some vector s.t. $w \notin U$. Let $W = \text{span}(w, Tw)$. Then, W is invariant under T, since $T(Tw) = T^2 w = -bTw - cw$ (via $T^2 + bT + cI = 0$, which means $T^2 = -bT - cI$). Moreover, since T has no eigenvectors, $Tw \neq \lambda w$ for any $\lambda \in \mathbb{F}$ $\$$ $\dim W = 2$.

Now, $\dim(U+W) = \dim U + \dim W - \dim(U \cap W) = \dim U + 2 - 0$, and $U+W$ is an even-dimensional subspace that is invariant under T, contradicting the maximality of $\dim U$. So, $U=V$ must have been the case!

5.34 *operators on odd-dimensional vector spaces have eigenvalues*

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

$\boxed{\text{proof}}$ We do induction.

Base case: & $\dim V = 1$. Then, if $T \in \mathcal{L}(V)$, every nonzero vector $v \in V$ is an eigenvector of T, since $Tv = \lambda v$ for some $\lambda \in \mathbb{R}$ (i.e. if $Tv = v$, then $Tw = \left(\frac{w}{v}\right)v$).

Induction Step: & every operator on an odd dimensional vector space w/ $\dim V \leq n$ has an eigenvalue. We show this implies any operator on an odd dimensional vector space of dimension $\dim U = n+2$. Let p be the minimal poly. of $T \in \mathcal{L}(U)$. If p is divisible by $(x-\lambda)$, then $p(\lambda) = 0$ & λ is an eigenvalue of T. So, if p is not divisible by $(x-\lambda)$ for any λ, then $p(x) = q(x)(x^2 + bx + c)$ w/ $b^2 < 4c$ (Since $x^2 + bx + c = 0$ for $x = \frac{-b \pm \sqrt{b^2 - 4c}}{2}$ & $x \notin \mathbb{R}$ when $b^2 < 4c$).

So, $0 = p(T) = q(T)(T^2 + bT + cI)$. Hence, $q(T)\big|_{\text{range}(T^2+bT+cI)} = 0$

Since $\deg q < \deg p$ & p is the min. poly. of T, $\text{range}(T^2 + bT + cI) \neq U$.

By the FToLM, $\dim U = \dim \text{null}(T^2 + bT + cI) + \dim \text{range}(T^2 + bT + cI)$.

Since $\dim V$ is odd & $\dim \text{null}(T^2 + bT + cI)$ is even, $\dim \text{range } T^2 + bT + cI$ is odd. So, our inductive hypothesis applies to $T\big|_{\text{range}(T^2+bT+cI)}$

Hence, $T\big|_{\text{range}(T^2+bT+cI)}$ has an eigenvalue, and so then does T.

5C Upper-Triangular Matrices

5.35 definition: *matrix of an operator,* $\mathcal{M}(T)$

Suppose $T \in \mathcal{L}(V)$. The *matrix of* T with respect to a basis $v_1, \ldots, v_n$ of V is the n-by-n matrix

$$\mathcal{M}(T) = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,n} \end{pmatrix}$$

whose entries $A_{j,k}$ are defined by

$$Tv_k = A_{1,k}v_1 + \cdots + A_{n,k}v_n.$$

The notation $\mathcal{M}(T, (v_1, \ldots, v_n))$ is used if the basis is not clear from the context.

$\boxed{\text{Recall}}$ If $T \in \mathcal{L}(V,W)$ & $w_1, \ldots, w_m$ is a basis of W, $v_1, \ldots, v_n$ is a basis of V, Then $\mathcal{M}(T)$ is the matrix s.t.

$$Tv_k = \mathcal{M}(T)_{1,k} w_1 + \cdots + \mathcal{M}(T)_{m,k} w_m$$

Same idea
(but w/ $W = V$ & the same basis for both)

$\boxed{\text{Ex}}$ Let $T \in \mathcal{L}(\mathbb{R}^3)$ be $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x+y \\ 5y+3z \\ 8z \end{bmatrix}$.

Then, w.r.t. the basis $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ of $\mathbb{F}^3$,

$$\mathcal{M}(T) = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{bmatrix}$$

5.37 definition: *diagonal of a matrix*

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

$\boxed{\text{Ex}}$ the diagonal of $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{bmatrix}$ is circled here

5.38 definition: *upper-triangular matrix*

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

$\boxed{EX}$ The matrix $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{bmatrix}$ is upper-triangular.

all 0's

The matrix $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ is not.

not all 0's

5.39 *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.

(a) The matrix of T with respect to $v_1, ..., v_n$ is upper triangular.

(b) $\operatorname{span}(v_1, ..., v_k)$ is invariant under T for each $k = 1, ..., n$.

(c) $Tv_k \in \operatorname{span}(v_1, ..., v_k)$ for each $k = 1, ..., n$.

$\boxed{\text{Proof}}$ (a)$\Rightarrow$(b) Suppose $\mathcal{M}(T)$ is U.T. w.r.t. the basis $v_1, \dots, v_n$. Then,

$$Tv_j = \mathcal{M}(T)_{1,j}v_1 + \dots + \mathcal{M}(T)_{n,j}v_n$$
$$= \mathcal{M}(T)_{1,j}v_1 + \dots + \mathcal{M}(T)_{j,j}v_j \quad (\text{since } \mathcal{M}(T)_{i,j} = 0 \text{ for } i > j)$$

So, $Tv_j \in \operatorname{span}(v_1, \dots, v_j) \subseteq \operatorname{span}(v_1, \dots, v_k)$ for $k \geq j$. Hence, $\operatorname{span}(v_1, \dots, v_k)$ is invariant under T for $k = 1, \dots, n$.

(b)$\Rightarrow$(c) Suppose $\operatorname{span}(v_1, \dots, v_k)$ is invariant under T for $k = 1, \dots, n$. Then, $Tv \in \operatorname{span}(v_1, \dots, v_k)$ for any $v \in \operatorname{span}(v_1, \dots, v_k)$. Since $v_k \in \operatorname{span}(v_1, \dots, v_k)$ we may conclude $Tv_k \in \operatorname{span}(v_1, \dots, v_k)$ for $k = 1, \dots, n$.

(c)$\Rightarrow$(a) Suppose $Tv_k \in \operatorname{span}(v_1, \dots, v_k)$ for $k = 1, \dots, n$.

Then, $Tv_k = a_1v_1 + \dots + a_kv_k$ for some $a_1, \dots, a_k \in \mathbb{F}$.

But this means $Tv_k = a_1v_1 + \dots + a_nv_n$ w/ $a_i = 0$ for $i > k$.

Since $v_1, \dots, v_n$ is a basis, this representation (i.e. the coefficients $a_1, \dots, a_n$) are unique. But this gives $Tv_k = \mathcal{M}(T)_{1,k}v_1 + \dots + \mathcal{M}(T)_{n,k}v_n$ has $\mathcal{M}(T)_{i,k} = 0$ for $i > k$.

That is, $\mathcal{M}(T)$ is U.T.

5.40 *equation satisfied by operator with upper-triangular matrix*

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

$$\boxed{(T - \lambda_1 I) \cdots (T - \lambda_n I) = 0.}$$

This means that $P(T) = 0$ where $P(z) = (z-\lambda_1) \cdots (z-\lambda_n)$.

This must be a poly. multiple of the min. poly. of T. So the min. poly. of T factors completely into linear terms. Moreover, some of these λ_i's are eigenvalues of T (in fact, $\lambda_1, \ldots, \lambda_n$ must contain all of T's eigenvalues)

$\boxed{\text{Proof}}$ Let $v_1, \ldots, v_n$ be a basis of T s.t. $\mathcal{M}(T)$ is upper triangular w/ diagonal entries $\lambda_1, \ldots, \lambda_n$. Since $\mathcal{M}(T)$ is U.T., the previous result tells us $Tv_k \in \operatorname{span}(v_1, \ldots, v_k)$. So, in particular, $Tv_1 = \mathcal{M}(T)_{1,1} v_1 = \lambda_1 v_1$.

Hence, $(T - \lambda_1 I) v_1 = 0$. Since $p(T) q(T) = q(T) p(T)$ for any poly's $p, q \in \mathcal{P}(\mathbb{F})$, we have $(T - \lambda_1 I) \cdots (T - \lambda_n I) v_1 = 0$.

Now, observe that $Tv_2 = \mathcal{M}(T)_{1,2} v_1 + \mathcal{M}(T)_{2,2} v_2 = \mathcal{M}(T)_{1,2} v_1 + \lambda_2 v_2$

So, $(T - \lambda_2 I) v_2 = \mathcal{M}(T)_{1,2} v_1 \in \operatorname{span}(v_1)$. Hence, $(T - \lambda_1 I)(T - \lambda_2 I) v_2 = 0$ & (using the fact that $p(T) q(T) = q(T) p(T)$ again) $(T - \lambda_1 I) \cdots (T - \lambda_n I) v_2 = 0$.

Repeat this process to show $(T - \lambda_1 I) \cdots (T - \lambda_n I) v_k = 0$ for $k = 1, \ldots, n$. This tells us that $(T - \lambda_1 I) \cdots (T - \lambda_n I) = 0$ since it is zero on a basis of V.

5.41 *determination of eigenvalues from upper-triangular matrix*

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

$\boxed{\text{proof}}$ Suppose $v_1, \ldots, v_n$ is a basis of V s.t. $T \in \mathcal{L}(V)$ has an UT matrix w.r.t. this basis.
Let $\lambda_1, \ldots, \lambda_n$ be the diagonal entries of $\mathcal{M}(T)$, i.e.

$$\mathcal{M}(T) = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

whatever

all 0's below

So, $Tv_1 = \lambda_1 v_1$. Hence, λ_1 is an eigenvalue of T (since $v_1 \neq 0$ as it is part of a basis of V). Then, for $k = 2, \ldots, n$, we have $Tv_k = \mathcal{M}(T)_{1,k} v_1 + \cdots + \mathcal{M}(T)_{k-1,k} v_{k-1} + \lambda_k v_k$ (since $\mathcal{M}(T)_{k,k} = \lambda_k$).
So, $Tv_k - \lambda_k v_k = (T - \lambda_k I) v_k \in \operatorname{span}(v_1, \ldots, v_{k-1})$. Therefore, $T - \lambda_k I$ maps $\operatorname{span}(v_1, \ldots, v_k)$ to $\operatorname{span}(v_1, \ldots, v_{k-1})$. Thus, $(T - \lambda_k I)|_{\operatorname{span}(v_1, \ldots, v_k)} \in \mathcal{L}(\operatorname{span}(v_1, \ldots, v_k))$ satisfies the following by the FToLM:

$$k = \dim \operatorname{span}(v_1, \ldots, v_k) = \dim \operatorname{null}\left(\underbrace{(T - \lambda_k I)|_{\operatorname{span}(v_1, \ldots, v_k)}}\right) + \dim \operatorname{range} \text{''}$$

$$k \leq \dim \operatorname{null}\left((T - \lambda_k I)|_{\operatorname{span}(v_1, \ldots, v_k)}\right) + (k-1)$$

$$1 \leq \dim \operatorname{null}\left((T - \lambda_k I)|_{\operatorname{span}(v_1, \ldots, v_k)}\right)$$

But this says $T - \lambda_k I$ has a nontrivial null space! That is, λ_k is an eigenvalue of T. Therefore, $\lambda_1, \ldots, \lambda_n$ are all eigenvalues of T.
Now, since the poly. $q(z) = (z - \lambda_1) \cdots (z - \lambda_n)$ satisfies $q(T) = 0$ by the previous result. So, q is a poly. mult. of the min. poly. Since the zeros of the min. poly are the eigenvalues of T, this tells us that $\lambda_1, \ldots, \lambda_n$ contains <u>all</u> the eigenvalues of T.

5.44 *necessary and sufficient condition to have an upper-triangular matrix*

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 $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$.

$\boxed{\text{proof}}$ First, $\oint$ T has an U.T. matrix wrt some basis of V. Let $\alpha_1, \ldots, \alpha_n$ denote the entries on the diagonal of $\mathcal{M}(T)$. Define $q(z) = (z-\alpha_1)\cdots(z-\alpha_n)$. Then, $q(T)=0$ by the previous results. Hence, q is a poly. mult. of the min. poly. of T. Therefore, the min poly is of the form $(z-\lambda_1)\cdots(z-\lambda_m)$ for some $\{\lambda_1, \ldots, \lambda_m\} \subseteq \{\alpha_1, \ldots, \alpha_n\}$.

For the other direction, $\oint$ the min. poly. of $T \in \mathcal{L}(V)$ is of the form $(z-\lambda_1)\cdots(z-\lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$. We do induction on m!

Base Case: Let $m=1$. Then, $T-\lambda_1 I = 0$, $\&$ $T=\lambda_1 I$ $\&$ therefore $\mathcal{M}(T)$ is upper triangular w.r.t. any basis of V. (Since $\mathcal{M}(T) = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_1 \end{bmatrix}$).

Inductive Step: $\oint$ our result is true for all values less than $m>1$. $\oint$ the min. poly. of T is $(z-\lambda_1)\cdots(z-\lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$. Then, let $\boxed{U = \text{range}(T-\lambda_m I)}$. Notice U is invariant under T since $T-\lambda_m I = p(T)$ for $p(z) = z-\lambda_m$. Hence $T|_U \in \mathcal{L}(U)$ is an operator $\&$ we can see that if $u \in U$, then $(T-\lambda_m I)v = u$ for some $v \in V$. So, $(T-\lambda_1 I)\cdots(T-\lambda_{m-1} I)u = (T-\lambda_1 I)\cdots(T-\lambda_{m-1} I)(T-\lambda_m I)v = 0v = 0$ $\&$ $(z-\lambda_1)\cdots(z-\lambda_{m-1})$ is a poly. mult. of the min. poly. of $T|_U$. So, the min poly. of $T|_U$ is of the form $(z-\beta_1)\cdots(z-\beta_t)$ for $\{\beta_1, \ldots, \beta_t\} \subseteq \{\lambda_1, \ldots, \lambda_{m-1}\}$. So, there is a basis, by our inductive hypothesis, such that $T|_U$ has an upper triangular matrix. Let $u_1, \ldots, u_M$ be this basis for U. Then, $Tu_k = T|_U u_k \in \text{Span}(u_1, \ldots, u_k)$ for $k=1, \ldots, M$. Extend $u_1, \ldots, u_M$ to a basis of V as $u_1, \ldots, u_M, v_1, \ldots, v_N$. For $k=1, \ldots, N$, we have $Tv_k = Tv_k - \lambda_m v_k + \lambda_m v_k = (T-\lambda_m I)v_k + \lambda_m v_k$. By definition, $(T-\lambda_m I)v_k \in U$ $\&$ $Tv_k \in \text{Span}(u_1, \ldots, u_M, v_1, \ldots, v_k)$. So, we are done!

5.47 *if* $\mathbf{F} = \mathbf{C}$*, then every operator on* V *has an upper-triangular matrix*

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.

5D Diagonalizable Operators

5.48 definition: *diagonal matrix*

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

$$\begin{bmatrix} * & & 0 \\ & a_{ii} & \\ 0 & & * \end{bmatrix}\begin{bmatrix}0\\ \vdots\\ 1\\ \vdots \\ 0\end{bmatrix} = a_{ii}\begin{bmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{bmatrix}$$

5.50 definition: *diagonalizable*

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

Notice! If $T \in \mathcal{L}(V)$ is diagonalizable, then $Tv_k = \mathcal{M}(T)_{kk} v_k$ for all v_k in the basis $v_1, \ldots, v_n$ diagonalizing $\mathcal{M}(T)$

5.52 definition: *eigenspace,* $E(\lambda, T)$

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

$$E(\lambda, T) = \operatorname{null}(T - \lambda I) = \{v \in V : Tv = \lambda v\}.$$

Hence $E(\lambda, T)$ is the set of all eigenvectors of T corresponding to λ, along with the 0 vector. (i.e. $E(\lambda, T) \neq \{0\} \iff \lambda$ is an eigenvalue of T)

5.54 *sum of eigenspaces is a direct sum*

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

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

is a direct sum. Furthermore, if V is finite-dimensional, then

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

Good!

$\boxed{\text{Proof}}$ To show $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ is direct, suppose $v_1 + \cdots + v_m = 0$ where $v_i \in E(\lambda_i, T)$, $i \in \{1, \ldots, m\}$. Since any nonzero v_i is an eigenvector of T corresponding to the eigenvalue λ_i, this is a sum of zero or more eigenvectors corresponding to distinct eigenvalues of T. But, eigenvectors corresponding to distinct eigenvalues are lin. indep., so this sum must have zero nonzero terms (i.e. $v_i = 0$ for $i \in \{1, \ldots, m\}$). Hence, the sum is direct. Consequently, $\dim V \geq \dim(E(\lambda_1, T) + \cdots + E(\lambda_m, T)) = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$.

5.55 *conditions equivalent to diagonalizability*

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

(a) T is diagonalizable.

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

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

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

$\boxed{\text{Proof}}$ (a)$\Rightarrow$(b) Suppose T is diagonalizable. Let $v_1, \ldots, v_n$ be a basis of V s.t. $\mathcal{M}(T) = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ for some $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$.

Then, by the definition of $\mathcal{M}(T)$, $Tv_k = \mathcal{M}(T)_{k,k} v_k = \lambda_k v_k$. So, λ_k is an eigenvalue of T since $v_k \neq 0$ (and v_k is then an eigenvector).

(b)$\Rightarrow$(c) Suppose V has a basis consisting of e-vec's of T. Then, every $w \in V$ is a linear combo of e-vec's of T. Thus, it is a sum of elements from $E(\lambda_1, T), \ldots, E(\lambda_m, T)$ where $\lambda_1, \ldots, \lambda_m$ are all the distinct eigenvalues of T. So, $V = E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ and by our previous result this sum is direct.

(c)$\Rightarrow$(d) Suppose $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$. Then, we have $\dim V = \dim\left(E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)\right) = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$.

(d)$\Rightarrow$(a) Suppose $\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$. Take a basis for each $E(\lambda_i, T)$ and put them all together as a list $v_1, \ldots, v_n$ w/ $n = \dim V$. Suppose $a_1, \ldots, a_n \in \mathbb{F}$ s.t. $a_1 v_1 + \cdots + a_n v_n = 0$. Write u_k for the portion of the sum belonging to $E(\lambda_k, T)$. So $u_1 + \cdots + u_m = 0$. Since eigenvectors corresponding to distinct e-val's are lin. indep., and since each u_k is an eigenvector if $u_k \neq 0$, this must mean $u_k = 0$ for $k = 1, \ldots, m$. So, all the a_i must be zero as well and $v_1, \ldots, v_n$ is a lin. indep. list of length $\dim V$, and hence a basis of V.

5.58 *enough eigenvalues implies diagonalizability*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$ has $\dim V$ distinct eigenvalues. Then T is diagonalizable.

$\boxed{Ex}$ Let $T \in \mathcal{L}(\mathbb{F}^3)$ defined by $T(x,y,z) = (2x+y, 5y+3z, 8z)$.

W.r.t. the basis $(1,0,0), (0,1,0), (0,0,1)$, $M(T) = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{bmatrix}$.

This is not diagonal, but it shows that the eigenvalues of T are $2, 5$, and 8. So, T has $\dim \mathbb{F}^3 = 3$ distinct eigenvalues. Hence, it can be diagonalized. To do this, just find an e-vector corresponding to each e-val here.

One choice: $\lambda_1 = 2 \rightsquigarrow v_1 = (1,0,0)$, $\lambda_2 = 5 \rightsquigarrow v_2 = (1,3,0)$, $\lambda_3 = 8 \rightsquigarrow v_3 = (1,6,6)$.

Then, if $v = a_1 v_1 + a_2 v_2 + a_3 v_3$,

$$\begin{aligned} Tv &= a_1 T v_1 + a_2 T v_2 + a_3 T v_3 \\ &= 2a_1 v_1 + 5a_2 v_2 + 8a_3 v_3 \end{aligned}$$

So, this lets us know what T does to any vector $v \in V$ since v_1, v_2, v_3 is a basis. For instance, this makes computing powers of T easy!

$$T^k v = T^{k-1}(2a_1 v_1 + 5a_2 v_2 + 8a_3 v_3) = \cdots\cdots = 2^k a_1 v_1 + 5^k a_2 v_2 + 8^k a_3 v_3$$

5.62 *necessary and sufficient condition for diagonalizability*

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 $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$.

$\boxed{\text{Proof}}$ First, $\mathcal{S}$ T is diagonalizable. Then, V has a basis of e-vec's of T. Let $v_1, \ldots, v_n$ be such a basis & let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. Then, for each v_k, there exist λ_j s.t. $(T-\lambda_j I)v_k = 0$. So, $(T-\lambda_1 I)\cdots(T-\lambda_m I) = 0$ since it is zero on a basis (as $p(T)q(T) = q(T)p(T)$ for polys p, q). Hence, the min. poly. of T is a poly. multiple of $(z-\lambda_1)\cdots(z-\lambda_m)$ & is in fact the minimal poly of T, since any divisor of it leaves some eigenspace out of its null space.

Now, in the other direction, $\mathcal{S}$ the min. poly. of T is $(z-\lambda_1)\cdots(z-\lambda_m)$ for some list of distinct $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$.

We will do induction on m.

Base case: Let $m=1$. Then, $T - \lambda_1 I = 0$ & $\underline{T = \lambda_1 I}$ so T is a scalar multiple of the identity & $\mathcal{M}(T)$ is diagonal w.r.t. any basis of V (convince yourself of this by recalling the def'n of $\mathcal{M}(T)$).

Inductive Step: $\mathcal{S}$ our desired result is true for all values less than $m>1$, and let the min. poly. of T be $(z-\lambda_1)\cdots(z-\lambda_m)$ for some list of distinct $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$. We will show that range$(T-\lambda_m I)$ has a basis of e-vectors of T & that $\underline{V} = \underline{\text{range}(T-\lambda_m I)} \oplus \underline{\text{null}(T-\lambda_m I)}$, showing then that any extension of this basis of e-vectors for range$(T-\lambda_m I)$ adds only more eigenvectors of T, and hence that V has a basis of e-vec's of T & is therefore diagonalizable.

$\boxed{\text{Proof}}$ (continued)

First, note that $\text{range}(T-\lambda_m I)$ is an invariant subspace of T (since $\text{range}\, p(T)$ is invariant under T for any poly p).

So, $T|_{\text{range}(T-\lambda_m I)} \in \mathcal{L}(\text{range}(T-\lambda_m I))$. Now, if $u \in \text{range}(T-\lambda_m I)$, then $\exists v \in V$ s.t. $(T-\lambda_m I)v = u$. So,

$$\underline{(T-\lambda_1 I)\cdots(T-\lambda_{m-1} I)u} = \underline{(T-\lambda_1 I)\cdots(T-\lambda_{m-1} I)(T-\lambda_m I)v} = 0.$$

Hence, $(z-\lambda_1)\cdots(z-\lambda_{m-1})$ is a poly. mult. of the min. poly. of $T|_{\text{range}(T-\lambda_m I)}$. By our inductive hypothesis, this means that $T|_{\text{range}(T-\lambda_m I)}$ is diagonalizable $\therefore$ $\underline{\text{range}(T-\lambda_m I)}$ has $\underline{\text{a basis}}$ of $\underline{\text{e-vec's}}$ of $T|_{\text{range}(T-\lambda_m I)}$ (and therefore also $\underline{\text{of } T}$).

Now, to show $V = \text{range}(T-\lambda_m I) \oplus \text{null}(T-\lambda_m I)$, let $u \in \text{range}(T-\lambda_m I) \cap \text{null}(T-\lambda_m I)$. we show $u=0$ to get $\text{range}(T-\lambda_m I) \cap \text{null}(T-\lambda_m I) = \{0\}$ $\therefore$ get $V = \underline{\text{range}(T-\lambda_m I)} \oplus \underline{\text{null}(T-\lambda_m I)}$.
Since $u \in \text{null}(T-\lambda_m I)$, $Tu = \lambda_m u$. By our argument before,

$$0 = (T-\lambda_1 I)\cdots(T-\lambda_{m-1} I)u = (\lambda_m - \lambda_1)\cdots(\lambda_m - \lambda_{m-1})u$$

Since $\lambda_m - \lambda_k \neq 0$ for $k \neq m$ (as $\lambda_1, \ldots, \lambda_m$ are distinct), this means $u=0$. So, we are done!

$\underline{\text{(Wow, what a proof!)}}$

5.65 *restriction of diagonalizable operator to invariant subspace*

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.

$\boxed{\text{proof}}$ If T is diagonalizable, then the min. poly. of T is of the form $(z-\lambda_1)\cdots(z-\lambda_m)$ for some list of distinct $\lambda_1, \dots, \lambda_m \in \mathbb{F}$. So, $T|_U$ has a min. poly. of this form as well, since its min. poly. divides the min. poly. of T. Hence, it is also diagonalizable.

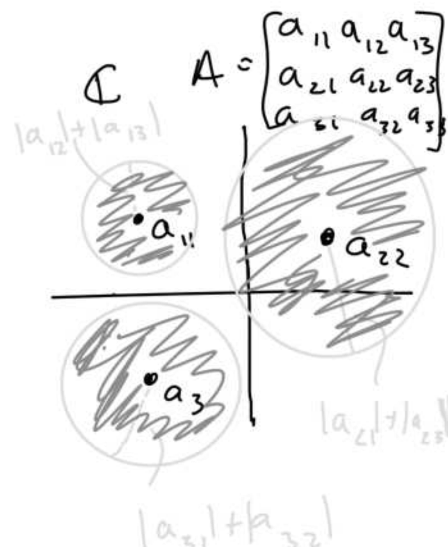
5.66 definition: *Gershgorin disks*

Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ 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 $v_1, \dots, v_n$ is a set of the form

Here, $\mathbb{F} = \mathbb{C}$ or $\mathbb{F} = \mathbb{R}$

$$\left\{ z \in \mathbf{F} : |z - A_{j,j}| \le \sum_{\substack{k=1 \\ k \neq j}}^{n} |A_{j,k}| \right\},$$

where $j \in \{1, \dots, n\}$.



5.67 *Gershgorin disk theorem*

Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V. Then each eigenvalue of T is contained in some Gershgorin disk of T with respect to the basis $v_1, \dots, v_n$.

All the e-vals live in the green part

$\boxed{\text{Proof}}$ Let $\lambda \in \mathbb{F}$ be an eigenvalue of T. Let $w \in V$ be a corresponding e-vec. Since $v_1, \dots, v_n$ is a basis of V, $\exists\, c_1, \dots, c_n \in \mathbb{F}$ s.t. $w = c_1 v_1 + \dots + c_n v_n$. Now, apply T to both sides, writing $A = \mathcal{M}(T)$ wrt. the basis $v_1, \dots, v_n$:

$$\lambda w = c_1 T v_1 + \dots + c_n T v_n = \sum_{k=1}^n c_k \sum_{j=1}^n A_{j,k} v_j = \sum_{k=1}^n \sum_{j=1}^n c_k A_{jk} v_j = \sum_{j=1}^n \left(\sum_{k=1}^n c_k A_{jk} \right) v_j.$$

so, $\sum_{k=1}^n c_k A_{jk} = \lambda c_j$. Hence, $\lambda c_j - A_{jj} c_j = \sum_{\substack{k=1 \\ k\neq j}}^n c_k A_{jk}$ & $|\lambda - A_{jj}| \le \sum_{\substack{k=1 \\ k\neq j}}^n \frac{|c_k|}{|c_j|} |A_{jk}|$.

so, if $|c_j|$ is the max of $|c_1|, \dots, |c_n|$, $\frac{|c_k|}{|c_j|} \le 1$ & we're done!

5E Commuting Operators

i.e. $S, T \in \mathcal{L}(V)$

5.71 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 $AB = BA$.

5.74 *commuting operators correspond to commuting matrices*

Suppose $S, T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V. Then S and T commute if and only if $\mathcal{M}(S, (v_1, \ldots, v_n))$ and $\mathcal{M}(T, (v_1, \ldots, v_n))$ commute.

Proof

$$
\begin{aligned}
S \;\&\; T \text{ commute} &\iff ST = TS \\
&\iff \mathcal{M}(ST) = \mathcal{M}(TS) \\
&\iff \mathcal{M}(S)\mathcal{M}(T) = \mathcal{M}(T)\mathcal{M}(S) \\
&\iff \mathcal{M}(S) \;\&\; \mathcal{M}(T) \text{ commute}
\end{aligned}
$$

5.75 *eigenspace is invariant under commuting operator*

Suppose $S, T \in \mathcal{L}(V)$ commute and $\lambda \in \mathbf{F}$. Then $E(\lambda, S)$ is invariant under T.

Proof We want to show that $v \in E(\lambda, S)$ implies $Tv \in E(\lambda, S)$.

S $v \in E(\lambda, S)$. Then,

$$S(Tv) = TSv = T(Sv) = T(\lambda v) = (\lambda Tv) \in E(\lambda, S)$$

i.e. $STv = \lambda Tv$, so $(S - \lambda I)(Tv) = 0$ \& $Tv \in E(\lambda, S)$.

Notice: $E(\lambda, T)$ is also invariant under S (just swap $S \& T$ in this proof).

5.76 *simultaneous diagonalizablity* $\iff$ *commutativity*

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.

$\boxed{\text{Proof}}$ Suppose $S, T \in \mathcal{L}(V)$ have diagonal matrices w.r.t. a basis $v_1, \ldots, v_n$ of V. Then,

$$\mathcal{M}(S)\mathcal{M}(T) = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}\begin{bmatrix} \eta_1 & & 0 \\ & \ddots & \\ 0 & & \eta_n \end{bmatrix} = \begin{bmatrix} \lambda_1\eta_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n\eta_n \end{bmatrix} = \begin{bmatrix} \eta_1\lambda_1 & & 0 \\ & \ddots & \\ 0 & & \eta_n\lambda_n \end{bmatrix} = \begin{bmatrix} \eta_1 & & 0 \\ & \ddots & \\ 0 & & \eta_n \end{bmatrix}\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} = \mathcal{M}(T)\mathcal{M}(S)$$

So, S & T commute.

In the other direction, suppose $S, T \in \mathcal{L}(V)$ are diagonalizable commuting operators. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of S. Then, we have

$$V = E(\lambda_1, S) \oplus \cdots \oplus E(\lambda_m, S) \quad (\text{since } S \text{ is diagonalizable})$$

Since $E(\lambda_k, S)$ is invariant under T for $k \in \{1, \ldots, m\}$, $T|_{E(\lambda_k, S)} \in \mathcal{L}(E(\lambda_k, S))$ & is also diagonalizable (since T is). So, $E(\lambda_k, S)$ has a basis of eigenvectors of $T|_{E(\lambda_k, S)}$ & hence of T. Combining bases for each, we get a basis of V that consists of eigenvectors of T which are also eigenvectors of S. Hence, the matrices of S & T w.r.t. this basis are both diagonal.

5.78 *common eigenvector for commuting operators*

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

$\boxed{\text{Proof}}$ Suppose V is a nonzero complex finite-dim. VS. & that $S, T \in \mathcal{L}(V)$ commute. Let λ be an eigenvalue of S (which exists since V is complex & finite-dim.). So, $E(\lambda, S) \neq \{0\}$ is invariant under T & therefore $T|_{E(\lambda, S)} \in \mathcal{L}(E(\lambda, S))$. So, $T|_{E(\lambda, S)}$ has an eigenvalue (for the same reason S does) & must therefore have an eigenvector. But, this eigenvector is in $E(\lambda, S)$, so it is also an eigenvector of S.

5.80 *commuting operators are simultaneously upper triangularizable*

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.

$\boxed{\text{Proof}}$ We do induction on $n = \dim V$. Base case: For $n=1$, every matrix of any operator is diagonal, so we are done.

Inductive Step: Suppose the result is true for all values less than $n = \dim V$. Let $S, T \in \mathcal{L}(V)$ be commuting operators. Then, by our previous result, S & T have a common eigenvector. Let v_1 be such an eigenvector of both S & T. Then, $\underline{Sv_1 \in \operatorname{Span}(v_1)}$ & $\underline{Tv_1 \in \operatorname{span}(v_1)}$.

So, let $W \subseteq V$ s.t. $V = \operatorname{span}(v_1) \oplus W$. Define $P \in \mathcal{L}(V, W)$ by $P(v) = P(av_1 + w) = w$. Then, we will show $PS|_W, PT|_W \in \mathcal{L}(W)$ (here, $S|_W \in \mathcal{L}(W, V)$ & $T|_W \in \mathcal{L}(W, V)$) commute.

$$\left((PS|_W)(PT|_W)w\right) = (PS|_W)(\underbrace{Tw - av_1}_{\substack{\text{where}\\ Tw = \tilde{w} + av_1\\ \text{for } \tilde{w} \in W}}) = P\left(STw - Sav_1\right) = P(STw - \tilde{a}v_1) = PSTw = PTSw = (PT|_W)(PS|_W)w$$

So, $PS|_W$ & $PT|_W$ are commuting operators on W & $\dim W = \dim V - 1 = n-1 < n$.

Therefore, the inductive hypothesis applies & both can be simultaneously upper-triangularized. Hence, taking $v_2, \ldots, v_n$ to be the basis of W doing this, $v_1, \ldots, v_n$ simultaneously upper-triangularizes S & T, since $Sv_k \in \operatorname{Span}(v_1, \ldots, v_k)$ & $Tv_k \in \operatorname{Span}(v_1, \ldots, v_k)$.

5.81 *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.

$\boxed{\text{Proof}}$ First, note that if A, B are upper-triangular matrices, then so is $A+B$ & AB. Moreover, the diagonals of $A+B$ and AB are the sums (resp. products) of the diagonal entries on A & B. That is,

$$A+B = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ & \ddots & \vdots \\ 0 & & a_{n,n} \end{bmatrix} + \begin{bmatrix} b_{1,1} & \cdots & b_{1,n} \\ & \ddots & \\ 0 & & b_{n,n} \end{bmatrix} = \begin{bmatrix} (a_{1,1}+b_{1,1}) & \cdots & (a_{1,n}+b_{1,n}) \\ & \ddots & \\ 0 & & (a_{n,n}+b_{n,n}) \end{bmatrix}$$

$$AB = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ & \ddots & \vdots \\ 0 & & a_{n,n} \end{bmatrix} \begin{bmatrix} b_{1,1} & \cdots & b_{1,n} \\ & \ddots & \\ 0 & & b_{n,n} \end{bmatrix} = \begin{bmatrix} (a_{1,1}b_{1,1}) & & * \\ & \ddots & \\ 0 & & (a_{n,n}b_{n,n}) \end{bmatrix}$$

So, since there exists a basis of V w.r.t. which both S & T have upper triangular matrices, and since the entries on the diagonal of an UT matrix of an operator are the eigenvalues of the operator, we are done!

6A Inner Products and Norms

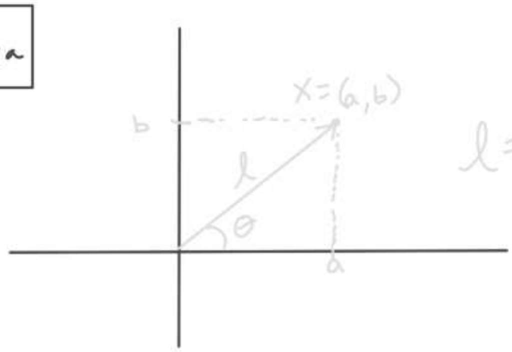
6.1 definition: *dot product*

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

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

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

$\boxed{\text{Idea}}$



$l = \sqrt{b^2 + a^2} = \sqrt{bb + aa} = \sqrt{x \cdot x}$

$\tan\theta = \frac{b}{a}$

$\theta = \tan^{-1}\left(\frac{b}{a}\right)$

6.2 definition: *inner product*

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

here, $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$

i.e. $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$

positivity

$\langle v, v\rangle \geq 0$ for all $v \in V$.

definiteness

$\langle v, v\rangle = 0$ if and only if $v = 0$.

So, $\langle v, v\rangle > 0$ for $v \neq 0$

additivity in first slot

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

homogeneity in first slot

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

So, for fixed $w \in V$, the function $v \mapsto \langle v, w\rangle$ is a linear functional

conjugate symmetry ← Symmetry if $\mathbb{F} = \mathbb{R}$, since $\overline{\langle v, u\rangle} = \langle v, u\rangle$ in $\mathbb{R}$

$\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$.

Recall: for $z = x + iy \in \mathbb{C}$, $\bar{z} = x - iy \in \mathbb{C}$ (i.e. throw a minus on the imaginary part)

6.4 definition: *inner product space*

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

6.6 *basic properties of an inner product*

(a) 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 $\mathbf{F}$.

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

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

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

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

$\boxed{\text{Proof}}$ (a) Since $\langle u+w, v\rangle = \langle u, v\rangle + \langle w, v\rangle$ & $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$, we are done.

(b) This is true for any linear map, so we get (b) from (a)

(c) Using conjugate symmetry, $\langle v, 0\rangle = \overline{\langle 0, v\rangle} = \overline{0} = 0$.

(d) $\langle u, v+w\rangle = \overline{\langle v+w, u\rangle} = \overline{\langle v, u\rangle + \langle w, u\rangle} = \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle} = \langle u, v\rangle + \langle u, w\rangle$

(e) $\langle u, \lambda v\rangle = \overline{\langle \lambda v, u\rangle} = \overline{\lambda\langle v, u\rangle} = \overline{\lambda}\ \overline{\langle v, u\rangle} = \overline{\lambda}\ \langle u, v\rangle$

6.7 definition: *norm,* $\|v\|$

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

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

6.9 *basic properties of the norm*

Suppose $v \in V$.

(a) $\|v\| = 0$ if and only if $v = 0$.

(b) $\|\lambda v\| = |\lambda|\,\|v\|$ for all $\lambda \in \mathbf{F}$.

$\boxed{\text{Proof}}$ (a) $\|v\|=0 \iff \sqrt{\langle v,v\rangle} = 0$

$\iff \langle v,v\rangle = 0$

$\iff v = 0$

(b) $\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v\rangle}$

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

$= \sqrt{\lambda\, \overline{\langle \lambda v, v\rangle}}$

$= \sqrt{\lambda \bar{\lambda}\, \overline{\langle v, v\rangle}}$

$= \sqrt{\lambda \bar{\lambda}\, \langle v, v\rangle}$

$= \sqrt{|\lambda|^2 \langle v, v\rangle}$

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

$= |\lambda|\,\|v\|$

$z = x+iy$

$\bar{z} = x-iy$

$z\bar{z} = (x+iy)(x-iy) = x^2 + \cancel{iyx} - \cancel{iyx} - i^2y^2$

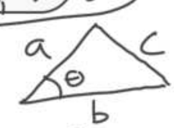
$= x^2 + y^2$

$= |z|^2$

6.10 definition: *orthogonal*

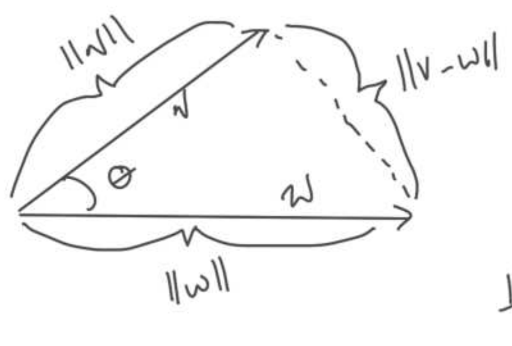
Two vectors $u, v \in V$ are called *orthogonal* if $\langle u, v\rangle = 0$. i.e. if they are at a right angle, using

$c^2 = a^2+b^2-2ab\cos\theta$



$\boxed{\text{Fact}}$ $\langle v, w\rangle = \|v\|\,\|w\| \cos\theta$

using the law of cosines,



$\|v-w\|^2 = \|w\|^2 + \|v\|^2 - 2\|v\|\,\|w\|\cos\theta$

$\langle v-w, v-w\rangle = \|w\|^2 + \|v\|^2 - 2\|v\|\,\|w\|\cos\theta$

$\cancel{\|v\|^2} + \cancel{\|w\|^2} - 2\langle v,w\rangle = \cancel{\|w\|^2} + \cancel{\|v\|^2} - 2\|v\|\,\|w\|\cos\theta$

$\langle v, w\rangle = \|v\|\,\|w\| \cos\theta$

Note: this calc assumes v, w are not scaler mult's of each other

6.11 *orthogonality and* 0

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

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

$\boxed{\text{Proof}}$ (a) $\langle 0, v\rangle = 0$ for all $v \in V$ by a previous result

(b) $\langle v, v\rangle = 0$ if & only if $v = 0$ by the definiteness of the inner product.

6.12 *Pythagorean theorem*

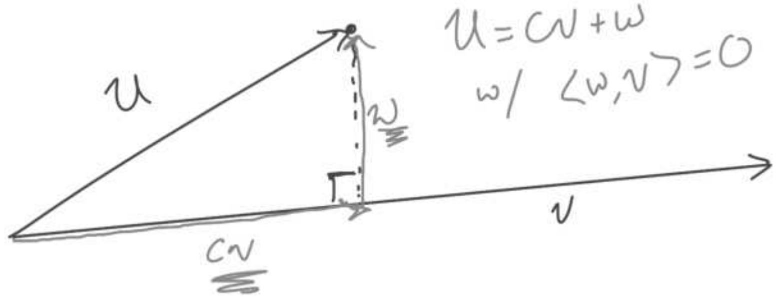
Suppose $u, v \in V$. If u and v are orthogonal, then

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

$\boxed{\text{Proof}}$ Sp $\langle u, v\rangle = 0$. Then,

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \langle u, u+v\rangle + \langle v, u+v\rangle \\
&= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\
&= \|u\|^2 + \langle u, v\rangle + \overline{\langle u, v\rangle} + \|v\|^2 \\
&= \|u\|^2 + 0 + \bar{0} + \|v\|^2 \\
&= \|u\|^2 + \|v\|^2
\end{aligned}$$

$\boxed{\text{Idea}}$



$u = cv + w$

w/ $\langle w, v\rangle = 0$

6.13 *an orthogonal decomposition*

Suppose $u, v \in V$, with $v \neq 0$. Set $c = \dfrac{\langle u, v\rangle}{\|v\|^2}$ and $w = u - \dfrac{\langle u, v\rangle}{\|v\|^2}v$. Then

$$u = cv + w \quad \text{and} \quad \langle w, v\rangle = 0.$$

$\boxed{\text{Proof}}$

$$cv + w = \underbrace{\frac{\langle u,v\rangle}{\|v\|^2} v}_{} + u - \underbrace{\frac{\langle u,v\rangle}{\|v\|^2} v}_{} = u$$

$$\begin{aligned}\langle w, v\rangle = \left\langle u - \frac{\langle u,v\rangle}{\|v\|^2} v, v\right\rangle &= \langle u, v\rangle - \left\langle \frac{\langle u,v\rangle}{\|v\|^2} v, v\right\rangle \\ &= \langle u, v\rangle - \frac{\langle u,v\rangle}{\|v\|^2}\langle v, v\rangle \\ &= \langle u, v\rangle - \langle u, v\rangle \frac{\|v\|^2}{\|v\|^2} = 0\end{aligned}$$

6.14 *Cauchy–Schwarz inequality* (Cauchy–Bunyakovsky–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.

$\boxed{\text{Proof}}$ (assume $v \neq 0$) If $u = 0$, then both sides of the inequality are zero, so it holds. If not, then $u = cv + w$ where $\langle w, v\rangle = 0$. Using the pythagorean theorem, we get

$$\begin{aligned}\|u\|^2 = \|cv + w\|^2 &= \|cv\|^2 + \|w\|^2 \quad (\text{since } \langle cv, w\rangle = c\langle v, w\rangle = 0) \\ &= \left\|\frac{\langle u,v\rangle}{\|v\|^2} v\right\|^2 + \|w\|^2 \\ &= \frac{|\langle u,v\rangle|^2}{\|v\|^4}\|v\|^2 + \|w\|^2 \\ &= \frac{|\langle u,v\rangle|^2}{\|v\|^2} + \|w\|^2 \\ &\geq \frac{|\langle u,v\rangle|^2}{\|v\|^2}\end{aligned}$$

equality if $w = 0$ & $u = cv$

So, $\|u\|^2 \geq \dfrac{|\langle u,v\rangle|^2}{\|v\|^2}$ & $\|u\|^2\|v\|^2 \geq |\langle u,v\rangle|^2$, or $|\langle u,v\rangle| \leq \|u\|\,\|v\|$.

6.17 *triangle inequality*

Suppose $u, v \in V$. Then

$$\|u + v\| \leq \|u\| + \|v\|.$$

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

$\boxed{\text{Proof}}$

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle \\
&= \|u\|^2 + \|v\|^2 + \langle u,v\rangle + \overline{\langle u,v\rangle} \\
&= \|u\|^2 + \|v\|^2 + 2\,\mathrm{Re}\left(\langle u,v\rangle\right) \\
&\leq \|u\|^2 + \|v\|^2 + 2\,|\langle v,u\rangle| \\
&\leq \|u\|^2 + \|v\|^2 + 2\,\|v\|\,\|u\| \quad (\text{via CS}) \\
&= \left(\|u\| + \|v\|\right)^2
\end{aligned}$$

$z = x + iy$
$\bar{z} = x - iy$
$z + \bar{z} = 2x$
i.e. $z + \bar{z} = 2\,\mathrm{Re}(z)$
$|z| = x^2 + y^2$
$\geq x^2$
so, $2|z| \geq 2\,\mathrm{Re}(z)^2 \geq 2\,\mathrm{Re}(z)$

So, $\|u+v\|^2 \leq \left(\|u\| + \|v\|\right)^2$ & $\|u+v\| \leq \|u\| + \|v\|$

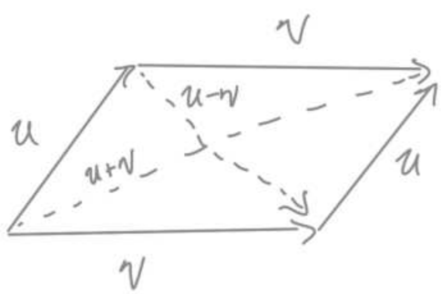
6.21 *parallelogram equality*

Suppose $u, v \in V$. Then

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

$\boxed{\text{Proof}}$

$$\begin{aligned}
\|u+v\|^2 + \|u-v\|^2 &= \langle u+v, u+v\rangle + \langle u-v, u-v\rangle \\
&= \langle u,u\rangle + \cancel{\langle u,v\rangle} + \cancel{\langle v,u\rangle} + \langle v,v\rangle \\
&\quad + \langle u,u\rangle - \cancel{\langle u,v\rangle} - \cancel{\langle v,u\rangle} + \langle v,v\rangle \\
&= 2\langle u,u\rangle + 2\langle v,v\rangle \\
&= 2\left(\|u\|^2 + \|v\|^2\right)
\end{aligned}$$



6B Orthonormal Bases

Orthonormal Lists and the Gram–Schmidt Procedure

6.22 definition: *orthonormal*

ortho – orthogonal, pairwise
normal – "normalized" i.e. norm 1

- A list of vectors is called *orthonormal* if each vector in the list has norm 1 and is orthogonal to all the other vectors in the list.
- In other words, a list $e_1, ..., e_m$ of vectors in V is orthonormal if

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

so, $\langle e_j, e_j \rangle = 1 = \|e_j\|^2$ & $\|e_j\| = 1$

for all $j, k \in \{1, ..., m\}$.

6.24 *norm of an orthonormal linear combination*

Suppose $e_1, ..., e_m$ 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 $a_1, ..., a_m \in \mathbf{F}$.

$\boxed{\text{Proof}}$ Using the pythagorean theorem (which, recall, says that $\|v+w\|^2 = \|v\|^2 + \|w\|^2$ for v, w orthogonal), we get

$$\begin{aligned} \|a_1 e_1 + \cdots + a_m e_m\|^2 &= \|a_1 e_1\|^2 + \|a_2 e_2 + \cdots + a_m e_m\|^2 \\ &\vdots \\ &= \|a_1 e_1\|^2 + \cdots + \|a_m e_m\|^2 \\ &= |a_1|^2 \|e_1\|^2 + \cdots + |a_m|^2 \|e_m\|^2 \\ &= |a_1|^2 + \cdots + |a_m|^2 \end{aligned}$$

Since $\langle a_1 e_1, a_2 e_2 + \cdots + a_m e_m \rangle$
$= a_1 (a_2 \langle e_1, e_2 \rangle + \cdots + a_m \langle e_1, e_m \rangle)$
$= a_1 (a_2 (0) + \cdots + a_m (0))$
$= 0$

6.25 *orthonormal lists are linearly independent*

Every orthonormal list of vectors is linearly independent.

$\boxed{\text{Proof}}$ Let $e_1, ..., e_m$ be an ON list in V. Let $a_1, ..., a_m \in \mathbf{F}$ be a list of scalars s.t. $a_1 e_1 + \cdots + a_m e_m = 0$. Then, using the previous result, $0 = \|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$.

So, $a_1 = \cdots = a_m = 0$ and $e_1, ..., e_m$ is lin. indep.

6.26 *Bessel's inequality*

Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in V. If $v \in V$ then

$$|\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2 \le \|v\|^2.$$

$\boxed{\text{Proof}}$ Set $u = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m$. Let $w = v - u$. Then, $v = u - u + v = u + w$. Now, if $\langle w, u\rangle = 0$, then by the Pythagorean theorem, we have $\|v\|^2 = \|u\|^2 + \|w\|^2 \ge \|u\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2$ by the previous result. So, we need only show $\langle w, u\rangle = 0$. To wit,

$$\begin{aligned}
\langle w, e_k\rangle = \langle v - u, e_k\rangle &= \langle v - (\langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m), e_k\rangle \\
&= \langle v, e_k\rangle - (\langle v, e_1\rangle\langle e_1, e_k\rangle + \cdots + \langle v, e_m\rangle\langle e_m, e_k\rangle) \\
&= \langle v, e_k\rangle - \langle v, e_k\rangle\langle e_k, e_k\rangle \\
&= \langle v, e_k\rangle - \langle v, e_k\rangle\|e_k\|^2 \\
&= \langle v, e_k\rangle - \langle v, e_k\rangle \\
&= 0
\end{aligned}$$

Hence,

$$\begin{aligned}
\langle w, u\rangle &= \langle w, \langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m\rangle \\
&= \overline{\langle v, e_1\rangle}\langle w, e_1\rangle + \cdots + \overline{\langle v, e_m\rangle}\langle w, e_m\rangle \\
&= \overline{\langle v, e_1\rangle}(0) + \cdots + \overline{\langle v, e_m\rangle}(0) \\
&= 0.
\end{aligned}$$

6.27 definition: *orthonormal basis*

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

6.28 *orthonormal lists of the right length are orthonormal bases*

Suppose V is finite-dimensional. Then every orthonormal list of vectors in V of length $\dim V$ is an orthonormal basis of V.

$\boxed{\text{Proof}}$ If $e_1, \ldots, e_{\dim V}$ is orthonormal, then it is a lin. indep. list of length $\dim V$ & is therefore a basis.

6.30 *writing a vector as a linear combination of an orthonormal basis*

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

(a) $v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n$, ✓

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

(c) $\langle u, v\rangle = \langle u, e_1\rangle\overline{\langle v, e_1\rangle} + \cdots + \langle u, e_n\rangle\overline{\langle v, e_n\rangle}$.

$\boxed{\text{Proof}}$ (a) Since $e_1, \ldots, e_n$ is an ONB, the exist scalers $a_1, \ldots, a_n \in \mathbb{F}$ s.t. $v = a_1e_1 + \cdots + a_ne_n$. But, notice:

$$\begin{aligned}\langle v, e_k\rangle &= \langle a_1e_1 + \cdots + a_ne_n, e_k\rangle \\ &= a_1\langle e_1, e_k\rangle + \cdots + a_n\langle e_n, e_k\rangle \\ &= a_k\langle e_k, e_k\rangle \\ &= a_k\end{aligned}$$

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

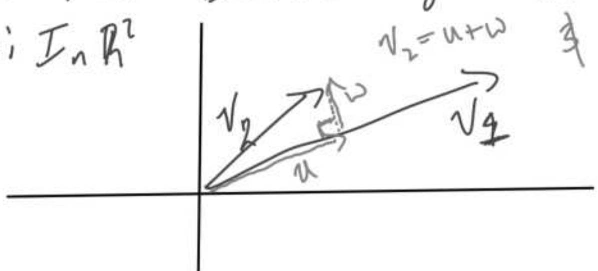
(b) Using (a), $\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2$ via the first theorem of this section.

(c) Using part (a), we get

$$\begin{aligned}\langle u, v\rangle &= \langle \langle u, e_1\rangle e_1 + \cdots + \langle u, e_n\rangle e_n, \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n\rangle \\ &= \langle \langle u, e_1\rangle e_1, \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n\rangle + \cdots + \langle \langle u, e_n\rangle e_n, \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n\rangle \\ &= \langle u, e_1\rangle\left(\langle e_1, \langle v, e_1\rangle e_1\rangle + \cdots + \langle e_1, \langle v, e_n\rangle e_n\rangle\right) + \cdots + \langle u, e_n\rangle\left(\langle e_n, \langle v, e_1\rangle e_1\rangle + \cdots + \langle e_n, \langle v, e_n\rangle e_n\rangle\right) \\ &= \langle u, e_1\rangle\left(\overline{\langle v, e_1\rangle}\langle e_1, e_1\rangle + \cdots + \overline{\langle v, e_n\rangle}\langle e_1, e_n\rangle\right) + \cdots + \langle u, e_n\rangle\left(\overline{\langle v, e_1\rangle}\langle e_n, e_1\rangle + \cdots + \overline{\langle v, e_n\rangle}\langle e_n, e_n\rangle\right) \\ &= \langle u, e_1\rangle\overline{\langle v, e_1\rangle} + \cdots + \langle u, e_n\rangle\overline{\langle v, e_n\rangle}\end{aligned}$$

So, we are done!

$\boxed{\text{Idea}}$ If we have a basis of an inner product space V, then how can we make that basis orthogonal in the easiest way? That is, visualize this: In $\mathbb{R}^2$



∋ $\langle w, v_1\rangle = 0$

So, just take v_2 and remove the part of it that isn't orthogonal to v_1.

Then, normalize! i.e. replace v_1 w/ $\frac{v_1}{\|v_1\|}$ & w w/ $\frac{w}{\|w\|}$

6.32 Gram–Schmidt procedure

Suppose $v_1, \dots, v_m$ is a linearly independent list of vectors in V. Let $f_1 = v_1$. For $k = 2, \dots, m$, 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 $k = 1, \dots, m$, let $e_k = \dfrac{f_k}{\|f_k\|}$. Then $e_1, \dots, e_m$ is an orthonormal list of vectors in V such that

$$\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k)$$

for each $k = 1, \dots, m$.

$\boxed{\text{Proof}}$ (Sketch)

Do induction on k. If $k=1$, then, $f_1 = v_1$ & $e_1 = \frac{f_1}{\|f_1\|} = \frac{v_1}{\|v_1\|}$.

So, $\|e_1\| = \left\|\frac{v_1}{\|v_1\|}\right\| = \left|\frac{1}{\|v_1\|}\right| \|v_1\| = 1$, and $0 \neq e_1 \in \operatorname{span}(v_1)$ so $\operatorname{span}(e_1) = \operatorname{span}(v_1)$.

Then, ϕ everything works up to $k-1$, i.e. that $e_1, \dots, e_{k-1}$ is an ON list s.t. $\operatorname{span}(v_1, \dots, v_{k-1}) = \operatorname{span}(e_1, \dots, e_{k-1})$. Then, build e_k & show everything is still good.

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

$\leftarrow v_k = -(\text{" "}) + f_k$ & v_k is then in $\operatorname{span}(e_1, \dots, e_k)$

$$e_k = \frac{f_k}{\|f_k\|}$$

So, $\langle e_k, e_j\rangle = \left\langle \frac{f_k}{\|f_k\|}, \frac{f_j}{\|f_j\|}\right\rangle$ (where $j < k$

$$= \frac{1}{\|f_k\|\|f_j\|}\langle f_k, f_j\rangle$$

$$= \frac{1}{\|f_k\|\|f_j\|}\left\langle v_k - \frac{\langle v_k, f_1\rangle}{\|f_1\|^2} f_1 - \cdots - \frac{\langle v_k, f_{k-1}\rangle}{\|f_{k-1}\|^2} f_{k-1}, f_j\right\rangle$$

$$= \frac{1}{\|f_k\|\|f_j\|}\left(\langle v_k, f_j\rangle - \frac{\langle v_k, f_1\rangle}{\|f_1\|^2}\langle f_1, f_j\rangle - \cdots - \frac{\langle v_k, f_{k-1}\rangle}{\|f_{k-1}\|^2}\langle f_{k-1}, f_j\rangle\right)$$

$$= \frac{1}{\|f_k\|\|f_j\|}\left(\langle v_k, f_j\rangle - \frac{\langle v_k, f_j\rangle}{\cancel{\|f_j\|^2}}\cancel{\langle f_j, f_j\rangle}\right)$$

$$= \frac{1}{\|f_k\|\|f_j\|}\left(\langle v_k, f_j\rangle - \langle v_k, f_j\rangle\right) = 0.$$

Then, show the span bit, which is not bad.

6.35 *existence of orthonormal basis*

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

$\boxed{\text{Proof}}$ Take a basis & apply GS process.

6.36 *every orthonormal list extends to an orthonormal basis*

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

$\boxed{\text{Proof}}$ Take the ON list, which is lin. indep., and extend to a basis. Then, apply GS to this basis. Only the elements added in the extension are changed, since the first part is alread ON.

6.37 *upper-triangular matrix with respect to some orthonormal basis*

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 $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$.

$\boxed{\text{Proof}}$ This is the same condition as what we had for getting an UT matrix of T w/ a regular basis, so just use the fact that we can orthonormalize this basis w/ GS & we are done (using $\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(e_1, \ldots, e_k)$).

6.38 *Schur's theorem*

Every operator on a finite-dimensional complex inner product space has an upper-triangular matrix with respect to some orthonormal basis.

Linear Functionals on Inner Product Spaces

$\boxed{\text{Idea}}$ If we take $v \in V$, we can obtain a linear functional $\varphi \in V'$ via $\varphi(u) = \langle u, v \rangle$. So, the mapping $v \mapsto \langle \cdot, v \rangle$ sends V to some portion of V'. But how much? What linear functionals are of this form?

$\boxed{\text{Ex}}$ Let $\varphi \in \mathcal{L}(\mathbb{F}^3, \mathbb{F}) = (\mathbb{F}^3)'$ be $\varphi(x_1, x_2, x_3) = 2x_1 - 5x_2 + x_3$.

Then, $\varphi(x) = \langle x, v \rangle$ for $v = (2, -5, 1)$ (if $\langle \cdot, \cdot \rangle$ is the dot product).

But what about, say, $\varphi \in (\mathcal{P}_n(\mathbb{R}))'$ w/ $\varphi(p) = \int_a^b p(t)\cos(t)\,dt$? ← ???

6.42 *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

i.e. $\varphi \in V'$

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

for every $u \in V$.

$\boxed{\text{Proof}}$ Let $e_1, \dots, e_n$ be an ONB of V. Then, if $\varphi \in V'$,

$$\begin{aligned}
\varphi(u) &= \varphi(\langle u, e_1 \rangle e_1 + \dots + \langle u, e_n \rangle e_n) \quad (\text{since } u = \langle u, e_1 \rangle e_1 + \dots + \langle u, e_n \rangle e_n) \\
&= \langle u, e_1 \rangle \varphi(e_1) + \dots + \langle u, e_n \rangle \varphi(e_n) \quad (\text{via linearity}) \\
&= \langle u, \overline{\varphi(e_1)} e_1 \rangle + \dots + \langle u, \overline{\varphi(e_n)} e_n \rangle \quad (\text{via } \langle x, \lambda y \rangle = \bar{\lambda} \langle x, y \rangle) \\
&= \langle u, \overline{\varphi(e_1)} e_1 + \dots + \overline{\varphi(e_n)} e_n \rangle \quad (\text{via additivity in } 2^{\text{nd}} \text{ slot}) \\
&= \langle u, v \rangle \quad \text{for } v = \overline{\varphi(e_1)} e_1 + \dots + \overline{\varphi(e_n)} e_n
\end{aligned}$$

So, there does exist some $v \in V$ s.t. $\varphi(u) = \langle u, v \rangle$ for all $u \in U$. For uniqueness, suppose $v_1, v_2 \in V$ are both s.t. $\varphi(u) = \langle u, v_1 \rangle$ & $\varphi(u) = \langle u, v_2 \rangle$. Then,

$$\begin{aligned}
0 &= \varphi(u) - \varphi(u) \\
&= \langle u, v_1 \rangle - \langle u, v_2 \rangle \\
&= \langle u, v_1 - v_2 \rangle
\end{aligned}$$

So, taking $u = v_1 - v_2$, we get $0 = \langle v_1 - v_2, v_1 - v_2 \rangle = \|v_1 - v_2\|^2$.

Hence, $v_1 - v_2 = 0$ & $v_1 = v_2$. So, this vector is unique!

$\boxed{\text{Ex}}$ Is there a poly. $q(t)$ s.t. $\int_{-1}^{1} p(t)\cos(\pi t)\,dt = \int_{-1}^{1} p(t)q(t)\,dt$ for all $p \in \mathcal{P}_2(\mathbb{R})$? On first glance, this seems strange to imagine, since $\cos(\pi t)$ is not a polynomial! Nonetheless, there is!

Define an inner product on $\mathcal{P}_2(\mathbb{R})$ via $\langle p,q\rangle = \int_{-1}^{1} p(t)q(t)\,dt$.
Since $\varphi(p) = \int_{-1}^{1} p(t)\cos(\pi t)\,dt$ is a linear functional on $\mathcal{P}_2(\mathbb{R})$, the Riesz rep. theorem says that there does exist $q \in \mathcal{P}_2(\mathbb{R})$ s.t. $\varphi(p) = \langle p,q\rangle$. To get it, we need an ONB of $\mathcal{P}_2(\mathbb{R})$, call it e_1, e_2, e_3, so we can construct $\boxed{q = \overline{\varphi(e_1)}e_1 + \overline{\varphi(e_2)}e_2 + \overline{\varphi(e_3)}e_3}$, since

$$\begin{aligned}
\langle p,q\rangle &= \langle p, \overline{\varphi(e_1)}e_1 + \overline{\varphi(e_2)}e_2 + \overline{\varphi(e_3)}e_3\rangle \\
&= \langle p,e_1\rangle\varphi(e_1) + \langle p,e_2\rangle\varphi(e_2) + \langle p,e_3\rangle\varphi(e_3) \\
&= \varphi(\langle p,e_1\rangle e_1 + \langle p,e_2\rangle e_2 + \langle p,e_3\rangle e_3) \\
&= \varphi(p)
\end{aligned}$$

To get this ONB, let's start w/ the basis $v_1 = 1, v_2 = x, v_3 = x^2$ then apply the GS process. First, take $f_1 = v_1 = 1$. Thus,
$\|f_1\|^2 = \int_{-1}^{1} (1)^2\,dt = 2$. So, $f_2 = v_2 - \frac{\langle v_2, f_1\rangle}{\|f_1\|^2} f_1 = x - \frac{\langle x,1\rangle}{2}1 = x - \cancel{\frac{\int_{-1}^{1}(t)(1)\,dt}{2}1}$
hence, $f_2 = x$. So, $\|f_2\|^2 = \int_{-1}^{1}(t)^2\,dt = \frac{2}{3}$. So

$$f_3 = v_3 - \frac{\langle v_3,f_1\rangle}{\|f_1\|^2} - \frac{\langle v_3, f_2\rangle}{\|f_2\|^2}f_2 = x^2 - \frac{\langle x^2,1\rangle}{2}1 - \frac{\langle x^2,x\rangle}{(2/3)}x = x^2 - \frac{1}{3}$$

Hence, $\|f_3\|^2 = \int_{-1}^{1}\left(t^2 - \frac{1}{3}\right)^2 dt = \frac{8}{45}$. So, the list e_1, e_2, e_3 w/ $e_i = \frac{f_i}{\|f_i\|}$ is an ONB & $e_1 = \sqrt{1/2}$, $e_2 = \left(\sqrt{\frac{3}{2}}\right)x$, $e_3 = \left(\sqrt{\frac{45}{8}}\right)\left(x^2 - \frac{1}{3}\right)$

Hence, the poly.
$$q(x) = \overbrace{\left(\int_{-1}^{1}\left(\sqrt{\tfrac{1}{2}}\right)\cos(\pi t)\,dt\right)}^{\varphi(e_1)}\overbrace{\sqrt{\tfrac{1}{2}}}^{e_1} + \overbrace{\left(\int_{-1}^{1}\left(\sqrt{\tfrac{3}{2}}\right)t\cos(\pi t)\,dt\right)}^{\varphi(e_2)}\overbrace{\sqrt{\tfrac{3}{2}}\,x}^{e_2}$$
$$+ \overbrace{\left(\int_{-1}^{1}\sqrt{\tfrac{45}{8}}\left(t^2 - \tfrac{1}{3}\right)\cos(\pi t)\,dt\right)}^{\varphi(e_3)}\overbrace{\sqrt{\tfrac{45}{8}}\left(x^2 - \tfrac{1}{3}\right)}^{e_3}$$
$$= \frac{15}{2\pi^2}(1 - 3x^2)$$

$\boxed{\text{Ex}}$ (continued)

What about if we want to do this for a poly of a different degree? All we have to do is repeat this process with an ONB of $\mathcal{P}_n(\mathbb{R})$ where n is our desired degree.

For example, if $n=5$, then the poly.

$$q(x) = \frac{105}{8\pi^4}\left((27-2\pi^2) + (24\pi^2-270)x^2 + (315-30\pi^2)x^4\right)$$

Satisfies $\displaystyle\int_{-1}^{1} p(t)\cos(\pi t)\,dt = \int_{-1}^{1} p(t)q(t)\,dt$

for all $p \in \mathcal{P}_5(\mathbb{R})$.

6C Orthogonal Complements and Minimization Problems

6.46 definition: *orthogonal complement,* $U^\perp$

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

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

Since $\langle \cdot, v\rangle$ is a linear functional on V, all we are doing is taking $v \in V$ s.t. $U \subseteq \operatorname{null}(\langle \cdot, v\rangle)$

6.48 *properties of orthogonal complement*

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

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

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

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

(e) If G and H are subsets of V and $G \subseteq H$, then $H^\perp \subseteq G^\perp$.

$\boxed{\text{Proof}}$ (a) If $v \in U^\perp$, then $\langle u, v\rangle = 0$ for all $u \in U$. So, $\langle u, \lambda v\rangle = \bar{\lambda}\langle u, v\rangle = \bar{\lambda}(0) = 0$ for all $u \in U$ and $\lambda v \in U^\perp$ as well for any $\lambda \in \mathbb{F}$. Similarly, if $w \in U^\perp$, then $\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle = 0 + 0 = 0$. So, $v + w \in U^\perp$ as well. Clearly, $\langle u, 0\rangle = 0$, so $0 \in U^\perp$. Hence, $U^\perp \leq V$.

(b) $\{0\}^\perp = \{v \in V : \langle 0, v\rangle = 0\} = V$ (by definition of $\langle \cdot, \cdot\rangle$)

(c) $V^\perp = \{v \in V : \langle w, v\rangle = 0 \ \forall w \in V\} \subseteq \{v \in V : \langle v, v\rangle = 0\} = \{v \in V : \|v\|^2 = 0\} = \{0\}$

(d) $U \cap U^\perp = U \cap \{v \in V : \langle u, v\rangle = 0 \ \forall u \in U\} = \{v \in U : \langle u, v\rangle = 0 \ \forall u \in U\} \subseteq \{v \in U : \langle v, v\rangle = 0\} \subseteq \{0\}$.

(e) If $G \subseteq H$, then $H^\perp = \{v \in V : \langle h, v\rangle = 0 \ \forall h \in H\} \subseteq \{v \in V : \langle g, v\rangle = 0 \ \forall g \in G\} = G^\perp$

6.49 *direct sum of a subspace and its orthogonal complement*

Suppose U is a finite-dimensional subspace of V. Then

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

$\boxed{\text{Proof}}$ Let $v \in V$ & $e_1, \dots, e_n$ is an ONB of U. Then,

$$v = v + 0 = v + (\langle v, e_1\rangle e_1 + \dots + \langle v, e_n\rangle e_n) - \underbrace{(\langle v, e_1\rangle e_1 + \dots + \langle v, e_n\rangle e_n)}_{P_U v}$$

Let $w = v - \overbrace{(\langle v, e_1\rangle e_1 + \dots + \langle v, e_n\rangle e_n)}^{u}$ & let $u = \langle v, e_1\rangle e_1 + \dots + \langle v, e_n\rangle e_n$

So that $v = u + w$. Certainly, $u \in U$, since it is a linear combo of a basis for U. Moreover,

$$\begin{aligned}
\langle w, e_k\rangle &= \langle v - u, e_k\rangle \\
&= \langle v, e_k\rangle - \langle u, e_k\rangle \\
&= \langle v, e_k\rangle - \langle \langle v, e_1\rangle e_1 + \dots + \langle v, e_n\rangle e_n, e_k\rangle \\
&= \langle v, e_k\rangle - (\langle v, e_1\rangle\langle e_1, e_k\rangle + \dots + \langle v, e_n\rangle\langle e_n, e_k\rangle) \\
&= \langle v, e_k\rangle - (\langle v, e_k\rangle\langle e_k, e_k\rangle) \\
&= \langle v, e_k\rangle - \langle v, e_k\rangle \\
&= 0
\end{aligned}$$

So, $w \in U^{\perp}$, since w is orthogonal to every element in a basis of U. Hence, $V = U + U^{\perp}$. Moreover, this sum is direct since $U \cap U^{\perp} = \{0\}$ by the previous theorem.

6.51 *dimension of orthogonal complement*

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

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

6.52 *orthogonal complement of the orthogonal complement*

Suppose U is a finite-dimensional subspace of V. Then

$$U = (U^\perp)^\perp.$$

Proof

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

So, $(U^\perp)^\perp = \{v \in V : \langle w, v\rangle = 0 \ \forall w \in U^\perp\}$

but notice that if $u \in U$, then $\langle w, u\rangle = \overline{\langle u, w\rangle} = \bar{0} = 0$ whenever $w \in U^\perp$. So, $U \subseteq (U^\perp)^\perp$.

For the reverse inclusion, suppose $v \in (U^\perp)^\perp$. By the previous result, $v = u + w$ where $u \in U$ & $w \in U^\perp$. So, $v - u = w \in U^\perp$. So, $v - u \in U^\perp$, and since $U \subseteq (U^\perp)^\perp$, $v - u \in (U^\perp)^\perp$.

Hence, $v - u \in U^\perp \cap (U^\perp)^\perp = \{0\}$, hence $v - u = 0$ & $v = u \in U$.

So $(U^\perp)^\perp \subseteq U$. Therefore, $U = (U^\perp)^\perp$.

6.54 $U^\perp = \{0\} \iff U = V$ *(for U a finite-dimensional subspace of V)*

Suppose U is a finite-dimensional subspace of V. Then

$$U^\perp = \{0\} \iff U = V.$$

Proof

$$U^\perp = \{0\} \iff (U^\perp)^\perp = \{0\}^\perp$$

$$\iff U = V \quad (\text{since } (U^\perp)^\perp = U \ \& \ \{0\}^\perp = V$$

6.55 definition: *orthogonal projection,* P_U

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 $P_U v = u$.

For an ONB $e_1, \dots, e_n$ of U, $P_u v$ is just $P_u v = \langle v, e_1\rangle e_1 + \dots + \langle v, e_n\rangle e_n$

6.57 *properties of orthogonal projection* P_U

Suppose U is a finite-dimensional subspace of V. Then

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

(b) $P_U u = u$ for every $u \in U$; $\leftarrow$ directly from this

(c) $P_U w = 0$ for every $w \in U^\perp$; $\leftarrow$ easy, from definition of P_u

(d) range $P_U = U$; Use (b) & (c) w/ $V = U \oplus U^\perp$

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

(f) $v - P_U v \in U^\perp$ for every $v \in V$; $\leftarrow$ $v = u + w$, $u \in U$, $w \in U^\perp$; $v = P_u v + w \Rightarrow v - P_u v = w \in U^\perp$

(g) $P_U^2 = P_U$; If $v = u + w$, $u \in U$, $w \in U^\perp$, $P_u v = u$ & $P_u^2 v = P_u u = u$

(h) $\|P_U v\| \le \|v\|$ for every $v \in V$; $\|u\| \le \|u + w\| = \|v\|$ since u & w are orthogonal

(i) if $e_1, \dots, e_m$ is an orthonormal basis of U and $v \in V$, then

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

$\boxed{\text{Proof}}$ (h) Since $V = U \oplus U^\perp$, write $v \in V$ as $v = u + w$ where $u \in U$ & $w \in U^\perp$. Since $P_u v = u$, and since u & w are orthogonal, we get $\|P_u v\|^2 = \|u\|^2 \le \|u\|^2 + \|w\|^2 = \|u + w\|^2$ via the Pythagorean theorem. Taking a square root, $\|P_u v\| \le \|u + w\| = \|v\|$.

(i) If $u \in U$, then $u = \langle u, e_1\rangle e_1 + \dots + \langle u, e_m\rangle e_m$ for any ONB $e_1, \dots, e_m$ of U by a previous result. So, since $P_u v = P_u(u + w) = u$, we get that $\underline{P_u v = \langle u, e_1\rangle e_1 + \dots + \langle u, e_m\rangle e_m}$

But notice: since $w \in U^\perp$, $\langle w, e_k\rangle = 0$, so

$$\begin{aligned} P_u v &= \langle u, e_1\rangle e_1 + \dots + \langle u, e_m\rangle e_m \\ &= (\langle u, e_1\rangle + 0) e_1 + \dots + (\langle u, e_m\rangle + 0) e_m \\ &= (\langle u, e_1\rangle + \langle w, e_1\rangle) e_1 + \dots + (\langle u, e_m\rangle + \langle w, e_m\rangle) e_m \\ &= \langle v, e_1\rangle e_1 + \dots + \langle v, e_m\rangle e_m. \end{aligned}$$

6.58 *Riesz representation theorem, revisited*

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 $v \mapsto \varphi_v$ is a one-to-one function from V onto V'.

Note If $\mathbb{F} = \mathbb{R}$, then $\varphi_{\lambda v}(u) = \langle u, \lambda v \rangle = \bar{\lambda}\langle u, v\rangle = \lambda \langle u, v \rangle = \lambda \varphi_v(u)$

So, in the real case, the map $T(v) = \varphi_v$ is linear! i.e. $T \in \mathcal{L}(V, V')$ if V is a real inner product space. But, if $\mathbb{F} = \mathbb{C}$, then T is not linear.

Proof (if $\mathbb{F} = \mathbb{R}$) Let $T \in \mathcal{L}(V, V')$ be $Tv = \varphi_v$. Then,

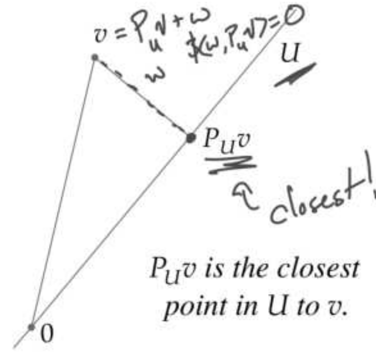
Notice that $\operatorname{null} T = \{v \in V : Tv = 0 \in V'\} = \{v \in V : \langle u, v \rangle = 0 \ \forall u \in V\} \subseteq \{v \in V : \langle v, v \rangle = 0\} = \{0\}$

Hence, $\operatorname{null} T = \{0\}$ & T is injective. Since $\dim V = \dim V'$, T is then also surjective & an isomorphism.

(If $\mathbb{F} = \mathbb{C}$) In this case, the above works to show that the set of vectors $v \in V$ s.t. $\varphi_v = 0$ is just $\{0\}$.

For surjectivity, all we need to do is note that $\varphi_0 = 0$ & for $v \in V$, $v \neq 0$, $\operatorname{null} \varphi_v$ has dimension $\dim V - 1$. So, $(\operatorname{null} \varphi_v)^{\perp}$ is one dimensional. The same is true of any $\varphi \in V'$, so all we need to do is "match up" the null spaces' orthogonal complements & scale appropriately. For details, see page 216 & 217 in the book.

Minimization Problems



$P_U v$ *is the closest point in* U *to* v.

6.61 *minimizing distance to a subspace*

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

$$\|v - P_U v\| \le \|v - u\|.$$

i.e. $\operatorname{dist}(v, P_U v) \le \operatorname{dist}(v, u)$ for all $u \in U$

Furthermore, the inequality above is an equality if and only if $u = P_U v$.

$\boxed{\text{Proof}}$ $0 \le \|P_U v - u\|^2$, so we have

$$\|v - P_U v\|^2 - \|v - P_U v\|^2 \le \|P_U v - u\|^2$$

$$\|v - P_U v\|^2 \le \|\underbrace{P_U v - u}_{\text{this is in } U}\|^2 + \|\underbrace{v - P_U v}_{\text{this is in } U^\perp}\|^2$$

$$= \|(\cancel{P_U v} - u) + (v - \cancel{P_U v})\|^2 \quad \text{(via pythagorean)}$$

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

So, we are done!

$\boxed{\text{Idea}}$

We can use this to solve minimation problems! That is, if we want to minimize some quantity described by a distance (as induced by an inner product), then we just use a projection to get the minimizing value.

$\boxed{\text{Ex}}$ Suppose we want to find a poly. that approximates a non-poly function on a closed interval as well as possible given a constraint on its degree. For instance, suppose we want a poly of degree at most 5 to approx sin as well as possible on $[-\pi, \pi]$, in the following sense: $p(x)$, the poly, should minimize

$$\int_{-\pi}^{\pi} |\sin(x) - p(x)|^2 dx$$

Let $C([-\pi,\pi])$ denote the vector space of continuous real-valued functions on $[-\pi,\pi]$ & define the inner product $\langle f,g\rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$ (this is an inner product, which you can check). Let U be the subspace of $C[-\pi,\pi]$ that is $\mathcal{P}_5(\mathbb{R})$. Then, our task becomes finding $p(x) \in U$ s.t. $\|p(x) - \sin(x)\|$ is as small as possible, since $\|p(x)-\sin(x)\| = \langle p(x)-\sin(x), p(x)-\sin(x)\rangle$

$$= \int_{-\pi}^{\pi} |p(x) - \sin(x)|^2 dx.$$

So, we just use the previous result & find our answer is $P_U(\sin(x)) = p(x)$. To actually do this computation, we would GS the basis $1, x, x^2, x^3, x^4, x^5$ of U, then compute $P_U \sin(x)$ as $P_U \sin(x) = \langle \sin(x), e_1\rangle e_1 + \cdots + \langle \sin(x), e_6\rangle e_6$ where $e_1, \ldots, e_6$ is the ONB we obtained via GS. This gives $p(x) \approx 0.987862x - 0.155271x^3 + 0.00564312x^5$, which is a far better approx. than the Taylor poly $x - \frac{x^3}{3!} + \frac{x^5}{5!}$

Pseudoinverse

$\boxed{\text{Idea}}$ $\oint$ $T \in \mathcal{L}(V,W)$, where V is an inner product space, and $\oint$ we want to solve $Tx = b$ for some $b \in W$. If T is invertible, $x = T^{-1}b$ is the unique solution. If not, then there are either many solutions or none. If there are many, then we could ask for the "best one" in the sense of finding $\|x\|$ as small as possible or, if W is also an inner product space, finding x s.t. $\|Tx - b\|$ is as small as possible if there are no solutions.

6.67 *restriction of a linear map to obtain a one-to-one and onto map*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Then $T|_{(\operatorname{null} T)^\perp}$ is a one-to-one map of $(\operatorname{null} T)^\perp$ onto range T. i.e. $T|_{(\operatorname{null} T)^\perp} \in \mathcal{L}((\operatorname{null} T)^\perp, \operatorname{range} T)$ is invertible

$\boxed{\text{Proof}}$ (Sketch) If $v \in (\operatorname{null} T)^\perp$ & $T|_{(\operatorname{null} T)^\perp} v = 0$, then $v \in \operatorname{null} T \cap (\operatorname{null} T)^\perp = \{0\}$ & $v = 0$. So, $T|_{(\operatorname{null} T)^\perp}$ has trivial null space & is injective.
For surjectivity, if $\hat{w} = Tv \in \operatorname{range} T$, then write $v = u + w$ where $u \in \operatorname{null} T$ & $w \in (\operatorname{null} T)^\perp$. So, $T|_{(\operatorname{null} T)^\perp} w = Tw = Tw + 0 = Tw + Tu = Tv = \hat{w}$ & $\hat{w} \in \operatorname{range}\left(T|_{(\operatorname{null} T)^\perp}\right)$. Hence, $\operatorname{range} T \subseteq \operatorname{range} T|_{(\operatorname{null} T)^\perp}$ & the other direction is immediate, so they are equal.

6.68 definition: *pseudoinverse, $T^\dagger$*

Suppose that 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

$$\underbrace{T^\dagger}_{\text{this is the pseudoinverse of } T} w = (T|_{(\operatorname{null} T)^\perp})^{-1} P_{\operatorname{range} T}\, w$$

for each $w \in W$.

6.69 *algebraic properties of the pseudoinverse*

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

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

(b) $TT^\dagger = P_{\operatorname{range} T}$ = the orthogonal projection of W onto range T.

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

Projections are identities when restricted (i.e. if $P_U \in \mathcal{L}(V)$ w/ $U \subseteq V$, then $P_U|_U = I$)

$\boxed{\text{Proof}}$ (a) If T is invertible, then $\operatorname{null} T = \{0\}$ & $(\operatorname{null} T)^\perp = V$, as well as range $T = W$.

So, $T^\dagger = \left(T|_{(\operatorname{null} T)^\perp}\right)^{-1} P_{\operatorname{range} T} = \left(T|_V\right)^{-1} P_W = (T)^{-1} I = T^{-1}$

(b) $TT^\dagger = T\left(T|_{(\operatorname{null} T)^\perp}\right)^{-1} P_{\operatorname{range} T}$, so if $w \in \operatorname{range} T$, we get $TT^\dagger w = T\left(T|_{(\operatorname{null} T)^\perp}\right)^{-1} P_{\operatorname{range} T}\, w = T\left(T|_{(\operatorname{null} T)^\perp}\right)^{-1} w = w = P_{\operatorname{range} T}\, w$

Likewise, if $w \in (\operatorname{range} T)^\perp$, then $TT^\dagger w = T\left(T|_{(\operatorname{null} T)^\perp}\right)^{-1} P_{\operatorname{range} T}\, w = 0 = P_{\operatorname{range} T}\, w$

So, $TT^\dagger = P_{\operatorname{range} T}$ since they agree on both range T & $(\operatorname{range} T)^\perp$.

(c) (basically the same process as part (b))

6.70 *pseudoinverse provides best approximate solution or best solution*

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

(a) If $x \in V$, then

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

with equality if and only if $x \in T^\dagger b + \operatorname{null} T$.

i.e. minimize $\|Tx - b\|$ when there is no solution

(b) If $x \in T^\dagger b + \operatorname{null} T$, then

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

with equality if and only if $x = T^\dagger b$.

i.e. minimize $\|x\|$ when there are many solutions

7A Self-Adjoint and Normal Operators

Adjoints

7.1 definition: *adjoint*, T^*

Suppose $T \in \mathcal{L}(V, W)$. The *adjoint* of T is the function $T^*: W \to V$ such that
$$\langle Tv, w\rangle_W = \langle v, T^*w\rangle_V$$
for every $v \in V$ and every $w \in W$.

Idea: How do we know such a T^* exists? Consider the linear functionals $\varphi_w \in W'$ given by $\varphi_w(u) = \langle u, w\rangle$. Then, $\varphi_w \circ T \in V'$, since the composition of linear maps is linear. Hence, by the Riesz rep theorem, there is some $v_w \in V$ s.t. $(\varphi_w \circ T)(v) = \langle v, v_w\rangle$, which says $\langle Tv, w\rangle = \langle v, v_w\rangle$. So, $T^*w = v_w$ is how we define T^*. Now, is it true that $T^* \in \mathcal{L}(W, V)$? Yes!

7.4 *adjoint of a linear map is a linear map*

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

proof: S. $T \in \mathcal{L}(V, W)$. Now, if $v \in V$ & $w_1, w_2 \in W$, then
$$\begin{aligned}\langle Tv, w_1 + w_2\rangle &= \langle Tv, w_1\rangle + \langle Tv, w_2\rangle \\ &= \langle v, T^*w_1\rangle + \langle v, T^*w_2\rangle \\ &= \langle v, T^*w_1 + T^*w_2\rangle\end{aligned}$$
So, $\langle Tv, w_1 + w_2\rangle = \langle v, T^*(w_1 + w_2)\rangle = \langle v, T^*w_1 + T^*w_2\rangle$ and we may conclude (via the uniqueness ensured by Riesz rep) that $T^*(w_1 + w_2) = T^*w_1 + T^*w_2$.

We do the same for homogeneity, taking $\lambda \in \mathbb{F}$ & noting that
$$\langle v, T^*(\lambda w)\rangle = \langle Tv, \lambda w\rangle = \bar{\lambda}\langle Tv, w\rangle = \bar{\lambda}\langle v, T^*w\rangle = \langle v, \lambda(T^*w)\rangle$$
So, $T^*(\lambda w) = \lambda T^*w$ & T^* is linear!

7.5 *properties of the adjoint*

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

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

(b) $(\lambda T)^* = \bar{\lambda} T^*$ for all $\lambda \in \mathbf{F}$;

(c) $(T^*)^* = T$;

(d) $(ST)^* = T^* S^*$ for all $S \in \mathcal{L}(W, U)$ (here U is a finite-dimensional inner product space over $\mathbf{F}$);

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

(f) if T is invertible, then T^* is invertible and $(T^*)^{-1} = (T^{-1})^*$.

$\boxed{\text{Proof}}$ (a)

$$\langle (S+T)v, w\rangle = \langle Sv + Tv, w\rangle$$

$$\langle v, (S+T)^* w\rangle$$

$$= \langle Sv, w\rangle + \langle Tv, w\rangle$$

$$= \langle v, S^* w\rangle + \langle v, T^* w\rangle$$

$$= \langle v, S^* w + T^* w\rangle$$

$$= \langle v, (S^* + T^*) w\rangle$$

So, $(S+T)^* = S^* + T^*$.

(b) $$\langle v, (\lambda T)^* w\rangle = \langle (\lambda T) v, w\rangle = \lambda \langle Tv, w\rangle$$

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

$$= \langle v, \bar{\lambda} T^* w\rangle \quad \text{So, } (\lambda T)^* = \bar{\lambda} T^*$$

(c) $$\langle Tv, w\rangle = \langle v, T^* w\rangle = \overline{\langle T^* w, v\rangle} = \overline{\langle w, (T^*)^* v\rangle}, \langle (T^*)^* v, w\rangle$$

Hence, $Tv = (T^*)^* v$ for all $v \in U$, i.e. $T = (T^*)^*$

(d) $$\langle v, (ST)^* w\rangle = \langle STv, w\rangle = \langle Tv, S^* w\rangle = \langle v, T^* S^* w\rangle$$

So, $(ST)^* = T^* S^*$

(e) $$\langle u, I^* v\rangle = \langle Iu, v\rangle = \langle u, v\rangle = \langle u, Iv\rangle$$

So, $I^* = I$

(f) If T is invertible. Then, $T^{-1} T = I$ & $(T^{-1} T)^* = T^* \underbrace{(T^{-1})^*}_{} = I^* = I$

Hence, T^* is invertible & $(T^*)^{-1} = (T^{-1})^*$ same!

(Technically, we should also note that $TT^{-1} = I$ & $(TT^{-1})^* = \overbrace{(T^{-1})^*}^{} T^* = I^* = I$ as well)

7.6 *null space and range of T^**

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

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

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

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

(d) $\operatorname{range} T = (\operatorname{null} T^*)^{\perp}$.

$\boxed{\text{Proof}}$ (a) $w \in \operatorname{null} T^* \iff T^* w = 0$

$\iff \langle v, T^* w \rangle = 0$ for all $v \in V$

$\iff \langle Tv, w \rangle = 0$ for all $v \in V$

$\iff w \in (\operatorname{range} T)^{\perp}$ (since it is orthogonal to every Tv)

(b) This is part (d) w/ T replaced by T^*, since $(T^*)^* = T$

(c) This is part (a) w/ T replaced by T^*, since $(T^*)^* = T$

(d) Using (a), $\operatorname{null} T^* = (\operatorname{range} T)^{\perp}$, so $(\operatorname{null} T^*)^{\perp} = \left((\operatorname{range} T)^{\perp}\right)^{\perp} = \operatorname{range} T$.

7.7 definition: *conjugate transpose, A^**

The *conjugate transpose* of an m-by-n matrix A is the n-by-m matrix A^* obtained by interchanging the rows and columns and then taking the complex conjugate of each entry. In other words, if $j \in \{1, \ldots, n\}$ and $k \in \{1, \ldots, m\}$, then

$$(A^*)_{j,k} = \overline{A_{k,j}}.$$

7.9 *matrix of T^* equals conjugate transpose of matrix of T*

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

$$\mathcal{M}(T^*) = (\mathcal{M}(T))^*.$$

$\boxed{\text{Proof}}$ (sketch)

$$Te_k = \langle Te_k, f_1 \rangle f_1 + \cdots + \langle Te_k, f_m \rangle f_m$$

So, $\mathcal{M}(T)_{j,k} = \langle Te_k, f_j \rangle = \langle e_k, T^* f_j \rangle = \overline{\langle T^* f_j, e_k \rangle} = \overline{\mathcal{M}(T^*)_{k,j}}$

for the same reason, i.e.,

$$T^* f_j = \langle T^* f_j, e_1 \rangle e_1 + \cdots + \langle T^* f_j, e_n \rangle e_n$$

Self-Adjoint Operators

7.10 definition: *self-adjoint*

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

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

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

for all $v, w \in V$.

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

So $T = T^*$ requires $V = W$

i.e. $\langle Tv, w\rangle - \langle v, Tw\rangle = 0$

$\langle v, T^*w\rangle - \langle v, Tw\rangle = 0$

$\langle v, (T^*-T)w\rangle = 0$

$T^*-T = 0$

$T = T^*$

$\langle (T^*-T)v, w\rangle = 0$

$\langle (T-T^*)v, w\rangle = 0$

So, $\langle (T-T^*)v, v\rangle = 0$

$\therefore T - T^* = 0$

or $T = T^*$

7.12 *eigenvalues of self-adjoint operators*

Every eigenvalue of a self-adjoint operator is real.

$\boxed{\text{Proof}}$ Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Let λ be an eigenvalue of T, and let $v \in V$ be a corresponding eigenvector (which means $v \neq 0$). Then,

$$\lambda\|v\|^2 = \lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Tv, v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\langle v, v\rangle = \bar{\lambda}\|v\|^2$$

So, $\lambda\|v\|^2 = \bar{\lambda}\|v\|^2$, and dividing by $\|v\|^2 \neq 0$, we get $\lambda = \bar{\lambda}$.

So, λ is real!

7.13 *Tv is orthogonal to v for all $v \iff T = 0$ (assuming $\mathbf{F} = \mathbf{C}$)*

Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Then

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

$\boxed{\text{proof}}$ If $u, w \in V$, then

$$\boxed{\langle Tu, w\rangle = \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4} + \frac{\langle T(u+iw), u+iw\rangle - \langle T(u-iw), u-iw\rangle}{4} i}$$

The RHS is of the form $\langle Tv, v\rangle$ in each term, so $\langle Tu, w\rangle = 0 \ \forall u, w \in V$. Hence, $Tu = 0$ $\therefore T = 0$.

$$= \frac{\langle Tu + Tw, u+w\rangle - \langle Tu - Tw, u-w\rangle}{4} + \frac{\langle Tu + iTw, u+iw\rangle - \langle Tu - iTw, u-iw\rangle}{4} i$$

$$= \frac{\cancel{\langle Tu, u\rangle} + \langle Tu, w\rangle + \cancel{\langle Tw, u\rangle} + \cancel{\langle Tw, w\rangle} - \cancel{\langle Tu, u\rangle} + \langle Tu, w\rangle + \cancel{\langle Tw, u\rangle} - \cancel{\langle Tw, w\rangle}}{4}$$

$$+ \frac{\cancel{\langle Tu, u\rangle} + \langle Tu, w\rangle - \cancel{\langle Tw, u\rangle} + \cancel{\langle Tw, w\rangle} - \cancel{\langle Tu, u\rangle} + \langle Tu, w\rangle - \cancel{\langle Tw, u\rangle} - \cancel{\langle Tw, w\rangle}}{4}$$

$$= \frac{4\langle Tu, w\rangle}{4} = \langle Tu, w\rangle$$

7.14 *$\langle Tv, v\rangle$ is real for all $v \iff T$ is self-adjoint (assuming $\mathbf{F} = \mathbf{C}$)*

Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Then

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

$\boxed{\text{proof}}$ If $v \in V$, then $\langle T^*v, v\rangle = \langle v, Tv\rangle = \overline{\langle Tv, v\rangle}$. As such, we get

$$\begin{aligned} T \text{ is self-adjoint} &\iff T = T^* \\ &\iff T - T^* = 0 \\ &\iff \langle (T-T^*)v, v\rangle = 0 \text{ for all } v \in V \\ &\iff \langle Tv, v\rangle - \overline{\langle Tv, v\rangle} = 0 \text{ for all } v \in V \\ &\iff \langle Tv, v\rangle \in \mathbb{R} \text{ for all } v \in V \end{aligned}$$

If $z = x+iy$, $\bar{z} = x - iy$ & $z - \bar{z} = x+iy-x+iy = 2iy$

7.16 *T self-adjoint and $\langle Tv, v\rangle = 0$ for all $v \iff T = 0$*

Suppose T is a self-adjoint operator on V. Then

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

$\boxed{\text{proof}}$ We have already proved this in the case that $\mathbb{F} = \mathbb{C}$. So, assume V is a real inner product space. Then, if $u, w \in V$, we have

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

So $\langle Tu, w\rangle$ is of the form of a sum of inner products of the form $\langle Tv, v\rangle$ for several v's

(which follows from $\langle Tw, u\rangle \underset{\text{self adjoint}}{=} \langle w, Tu\rangle \underset{\text{since } V \text{ is a real inner product space}}{=} \langle Tu, w\rangle$)

So, suppose $\langle Tv, v\rangle = 0$ for all $v \in V$. Then, $\langle Tu, w\rangle = 0$ for all $u, w \in V$ and this implies that $T = 0$.

The other direction is easy: If $T = 0$, then $\langle Tv, v\rangle = \langle 0, v\rangle = 0$, $\forall v \in V$.

Normal Operators

7.18 definition: *normal*

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

$\boxed{\text{Idea}}$ Every self-adjoint operator satisfies $T=T^*$, so certainly $T^*T = T^2 = TT^*$ (i.e. every self-adjoint operator is normal). So, in some sense, normality is just a way of weakening self-adjointness.

7.20 *T is normal if and only if Tv and T^*v have the same norm*

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

$$T \text{ is normal} \iff \|Tv\| = \|T^*v\| \text{ for every } v \in V.$$

$\boxed{\text{Proof}}$

$$\begin{aligned}
T \text{ normal} &\iff TT^* = T^*T \\
&\iff T^*T - TT^* = 0 \\
&\iff \langle (T^*T - TT^*)v, v\rangle = 0 \text{ for all } v \in V \\
&\iff \langle T^*Tv, v\rangle = \langle TT^*v, v\rangle \text{ for all } v \in V \\
&\iff \langle Tv, Tv\rangle = \langle T^*v, T^*v\rangle \text{ for all } v \in V \\
&\iff \|Tv\|^2 = \|T^*v\|^2 \text{ for all } v \in V \\
&\iff \|Tv\| = \|T^*v\| \text{ for all } v \in V
\end{aligned}$$

Notice that
$$\underline{(T^*T - TT^*)^*} = T^*(T^*)^* - (T^*)^*T^* = \underline{T^*T - TT^*}$$
Same! So, $T^*T - TT^*$ is self-adjoint!

7.21 *range, null space, and eigenvectors of a normal operator*

Suppose $T \in \mathcal{L}(V)$ is normal. Then

(a) $\operatorname{null} T = \operatorname{null} T^*$;

(b) $\operatorname{range} T = \operatorname{range} T^*$;

(c) $V = \operatorname{null} T \oplus \operatorname{range} T$;

(d) $T - \lambda I$ is normal for every $\lambda \in \mathbf{F}$;

(e) if $v \in V$ and $\lambda \in \mathbf{F}$, then $Tv = \lambda v$ if and only if $T^*v = \bar{\lambda} v$.

$\boxed{\text{Proof}}$ (a) Using the last result, T is normal iff $\|Tv\| = \|T^*v\|$ for all $v \in V$. So, $v \in \operatorname{null} T \iff Tv = 0 \iff \|Tv\| = 0 \iff \|T^*v\| = 0 \iff T^*v = 0 \iff v \in \operatorname{null} T^*$

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

(c) $V = \operatorname{null} T \oplus (\operatorname{null} T)^\perp = \operatorname{null} T \oplus \operatorname{range} T^* = \operatorname{null} T \oplus \operatorname{range} T$

(d) Let $\lambda \in \mathbb{F}$. Then,

$$\begin{aligned}(T-\lambda I)(T-\lambda I)^* &= (T-\lambda I)(T^* - \bar{\lambda} I) \\ &= TT^* - \bar{\lambda} I T - \lambda I T^* + \lambda \bar{\lambda} I^2 \\ &= T^*T - \bar{\lambda} T - \lambda T^* + \lambda \bar{\lambda} I \\ &= (T^* - \bar{\lambda} I)(T - \lambda I) \\ &= (T-\lambda I)^*(T-\lambda I)\end{aligned}$$

(e) Let $v \in V$ & $\lambda \in \mathbb{F}$. Then, $\|(T-\lambda I)v\| = \|(T-\lambda I)^* v\| = \|(T^* - \bar{\lambda} I)v\|$.

So, $(T-\lambda I)v = 0$ if & only if $(T^* - \bar{\lambda} I)v = 0$. That is, $Tv = \lambda v$ if & only if $T^* v = \bar{\lambda} v$.

7.22 *orthogonal eigenvectors for normal operators*

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

$\boxed{\text{proof}}$ Suppose $\alpha, \beta \in \mathbb{F}$ are distinct e-vals of T & $u, w \in V$ are corresponding e-vecs (so that $Tu = \alpha u$, $Tw = \beta w$). Then, $T^* w = \bar{\beta} w$ (since $Tv = \lambda v \Rightarrow T^* v = \bar{\lambda} v$). Thus, we have

$$\begin{aligned}(\alpha - \beta)\langle u, w\rangle &= \alpha\langle u, w\rangle - \beta\langle u, w\rangle \\ &= \langle \alpha u, w\rangle - \langle \beta u, w\rangle \\ &= \langle \alpha u, w\rangle - \langle u, \bar{\beta} w\rangle \\ &= \langle Tu, w\rangle - \langle u, T^* w\rangle \\ &= \langle Tu, w\rangle - \langle Tu, w\rangle \\ &= 0\end{aligned}$$

Since $\alpha \neq \beta$, $\langle u, w\rangle = 0$ must hold.

7.23 *T is normal $\iff$ the real and imaginary parts of T commute*

Suppose $\mathbf{F} = \mathbf{C}$ 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 $T = A + iB$.

$\boxed{\text{Proof}}$ First, suppose T is normal. Then, let

$$A = \frac{T+T^*}{2}, \quad B = \frac{T-T^*}{2i}$$ & notice $A + iB = \frac{T+T^*}{2} + \frac{T-T^*}{2} = \frac{2T}{2} = T$.

Since $A^* = \frac{T^*+T}{2} = \frac{T+T^*}{2} = A$ & $B^* = \overline{\left(\frac{1}{2i}\right)}(T^* - T) = \left(\frac{-1}{2i}\right)(T^* - T) = \frac{T-T^*}{2i} = B$, both A & B are self-adjoint. Moreover,

$$AB = \left(\frac{T^*+T}{2}\right)\left(\frac{T-T^*}{2i}\right) = \left(\frac{1}{4i}\right)\left(T^*T - T^*T^* + TT - TT^*\right) = \frac{TT - T^*T^*}{4i}$$

$$BA = \left(\frac{T-T^*}{2i}\right)\left(\frac{T+T^*}{2}\right) = \left(\frac{1}{4i}\right)\left(TT + TT^* - T^*T - T^*T^*\right) = \frac{TT - T^*T^*}{4i}$$

So, we are done!

Now, in the other direction, if $T = A + iB$ for A, B self adjoint commuting operators, then $T^* = A^* - iB^* = A - iB$. So,

$$T^*T = (A - iB)(A + iB) = A^2 + iAB - iBA + B^2 = A^2 + B^2$$
$$TT^* = (A + iB)(A - iB) = A^2 - iAB + iBA + B^2 = A^2 + B^2$$ & T is normal.

Real Spectral Theorem

7.26 *invertible quadratic expressions*

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

$$T^2 + bT + cI$$

is an invertible operator.

$\boxed{\text{proof}}$ Suppose $v \neq 0$. Then,

$$\begin{aligned}
\langle (T^2 + bT + cI)v, v \rangle &= \langle T^2 v, v\rangle + b\langle Tv, v\rangle + c\langle v, v\rangle \\
&= \langle Tv, T^* v\rangle + b\langle Tv, v\rangle + c\langle v, v\rangle \\
&= \langle Tv, Tv\rangle + b\langle Tv, v\rangle + c\langle v, v\rangle \\
&= \|Tv\|^2 + b\langle Tv, v\rangle + c\|v\|^2 \\
&\geq \|Tv\|^2 - |b||\langle Tv, v\rangle| + c\|v\|^2 \\
&\geq \|Tv\|^2 - |b|\,\|Tv\|\,\|v\| + c\|v\|^2 \\
&= \left(\|Tv\| - \frac{|b|\,\|v\|}{2}\right)^2 + \left(c - \frac{b^2}{4}\right)\|v\|^2 \\
&\geq \left(c - \frac{b^2}{4}\right)\|v\|^2 \\
&> 0
\end{aligned}$$

via C.S.
$-|\langle u, v\rangle| \geq -\|u\|\,\|v\|$

Since $b^2 < 4c$, i.e.
$0 < 4c - b^2$
$0 < c - b^2/4$

So, $(T^2 + bT + cI)v \neq 0$ if $v \neq 0$.
Hence, it is injective & thus invertible.

7.27 *minimal polynomial of self-adjoint operator*

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

$\boxed{\text{proof}}$ First, if $\mathbb{F} = \mathbb{C}$, then the zeros of the min. poly. of T are the eigenvalues of T, all of which are real (as we proved last time). So, since polynomials factor completely over $\mathbb{C}$, the min. poly. of T is of the desired form.

Now, if $\mathbb{F} = \mathbb{R}$, the min. poly. of T is of the form $(z-\lambda_1)\cdots(z-\lambda_m)(z^2+b_1z+c_1)\cdots(z^2+b_nz+c_n)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ & $b_1, c_1, \ldots, b_n, c_n \in \mathbb{R}$ s.t. $b_k^2 < 4c_k$ (since the roots of $z^2+b_kz+c_k$ are $z = \frac{-b_k \pm \sqrt{b_k^2-4c_k}}{2}$, which is real unless $b_k^2 - 4c_k < 0$, i.e. $b_k^2 < 4c_k$). But this says
$$(T-\lambda_1 I)\cdots(T-\lambda_m I)(T^2+b_1T+c_1I)\cdots(T^2+b_nT+c_nI) = 0$$
and since $T^2+b_nT+c_nI$ is invertible, we get (if there are any such quadratic terms) that $(T-\lambda_1 I)\cdots(T-\lambda_m I)(T^2+b_1T+c_1I)\cdots(T^2+b_{n-1}T+c_{n-1}I)=0$. This would contradict the minimality of the degree of the minimal polynomial. Hence, there must be no quadratic terms & the min. poly. must be of the form $(z-\lambda_1)\cdots(z-\lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$.

7.29 *real spectral theorem*

Suppose $\mathbf{F} = \mathbf{R}$ 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.

So, self-adjoint is the same as diagonalizable by an orthonormal basis of eigenvectors

$\boxed{\text{Proof}}$ Suppose (a) is true. Then, T has an UT matrix w.r.t. an orthonormal basis of V since its min. poly. is of the form $(z-\lambda_1)\cdots(z-\lambda_n)$ for some $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ (this is an old result we proved in Ch 6). W.r.t. this basis, $\mathcal{M}(T)$ is U.T. and, since T is self adjoint, we get $\mathcal{M}(T) = \mathcal{M}(T^*) = \mathcal{M}(T)^t$. However, $\mathcal{M}(T)^t$ is lower triangular, while $\mathcal{M}(T)$ is UT, so $\mathcal{M}(T)$ must be diagonal. Therefore, (a) $\Rightarrow$ (b).

Now, suppose (b). Then, $\mathcal{M}(T)^t = \mathcal{M}(T^*) = \mathcal{M}(T)$ (since $\mathcal{M}(T) = \mathcal{M}(T)^t$)
Hence, $T = T^*$ and (b) $\Rightarrow$ (a).

That (c) $\Leftrightarrow$ (b) was shown in Chapter 5
(i.e. diagonalizability of T $\Leftrightarrow$ $\exists$ of a basis for V of e-vectors of T).

Complex Spectral Theorem

7.31 *complex spectral theorem*

Suppose $\mathbf{F} = \mathbf{C}$ 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.

$\boxed{\text{Proof}}$ $\oint$ (a). By Schur's theorem (from Ch 6), there is an orthonormal basis of V, $e_1, \dots, e_n$, s.t. $\mathcal{M}(T)$ is upper triangular w.r.t. this basis. Let $\mathcal{M}(T) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{bmatrix}$.

Then, $Te_1 = a_{11}e_1$, so $\|Te_1\|^2 = |a_{11}|^2$ & since $\mathcal{M}(T^*) = \mathcal{M}(T)^*$ (where the second $*$ denotes the conjugate transpose) $T^*e_1 = \overline{a_{11}}e_1 + \dots + \overline{a_{1n}}e_n$.

So, since $\|Tv\| = \|T^*v\|$ when T is normal, we get

$$|a_{11}|^2 = \|Te_1\|^2 = \|T^*e_1\|^2 = |a_{11}|^2 + \dots + |a_{1n}|^2$$

Hence, $a_{12}, \dots, a_{1n}$ are all 0. Repeating this process for $e_2, \dots, e_n$ yields that the off-diagonal entries of $\mathcal{M}(T)$ are all 0. Hence, it is diagonal & (a) $\Rightarrow$ (b).

Now, $\oint$ (b). Then, T and T^* both have diagonal matrices wrt some orthonormal basis of V (in fact, w.r.t. the same orthonormal basis of V). Since any two diagonal matrices commute, $\mathcal{M}(TT^*) = \mathcal{M}(T)\mathcal{M}(T^*) = \mathcal{M}(T^*)\mathcal{M}(T) = \mathcal{M}(T^*T)$. Therefore, $TT^* = T^*T$ & T commutes w/ its adjoint (that is, T is normal). So, (b) $\Rightarrow$ (a).

(c) $\Leftrightarrow$ (b) is the same as in the real case.

7C Positive Operators

7.34 definition: *positive operator*

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

$$\boxed{\underbrace{\langle Tv, v\rangle}_{\mathbb{R}}} \geq 0$$

for all $v \in V$.

$\boxed{\text{Note}}$ If V is an inner product space over $\mathbb{C}$, then we can drop the requirement that T is self-adjoint since $\langle Tv, v\rangle \in \mathbb{R}$ for all $v \in V$ implies this in that case. Be careful to not make a mistake and forget this! $\mathbb{C}$ doesn't have "$\leq$" like $\mathbb{R}$ does!

7.36 definition: *square root*

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

$\boxed{\text{Ex}}$ Let $T \in \mathcal{L}(\mathbb{F}^3)$ be $T(x_1, x_2, x_3) = (x_3, 0, 0)$.

Let $R \in \mathcal{L}(\mathbb{F}^3)$ be $R(x_1, x_2, x_3) = (x_2, x_3, 0)$. Then,

$R^2(x_1, x_2, x_3) = R(x_2, x_3, 0) = (x_3, 0, 0) = T(x_1, x_2, x_3)$.

So, $R^2 = T$ & R is a square root of T.

Notice: If $R'(x_1, x_2, x_3) = (-x_2, -x_3, 0)$, then

$(R')^2(x_1, x_2, x_3) = R'(-x_2, -x_3, 0) = (x_3, 0, 0) = T(x_1, x_2, x_3)$,

so R' is a square root of T as well.

7.38 *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 $R \in \mathcal{L}(V)$.

$\boxed{\text{Proof}}$ ((a) ⇒ (b)) That a positive operator is self adjoint is immediate from the definition. Suppose T is positive and λ is an e-val of T w/ e-vec v. Then $0 \le \langle Tv, v\rangle = \langle \lambda v, v\rangle = \lambda \|v\|^2$

So, $\lambda \ge 0$.

((b) ⇒ (c)) Spectral Theorem

((c) ⇒ (d)) By the linear map lemma, there is a $R \in \mathcal{L}(V)$ s.t. $Re_k = \sqrt{\lambda_k}\, e_k$ where $e_1, \dots, e_n$ is the assumed ONB of V s.t. $\mathcal{M}(T)$ is diagonal w/ λ_k its kth diagonal entry. Certainly, R is positive: $\langle Re_k, e_k\rangle = \sqrt{\lambda_k}\langle e_k, e_k\rangle = \sqrt{\lambda_k}\,\|e_k\|^2 = \sqrt{\lambda_k} \ge 0$.

Moreover, $R^2 e_k = \sqrt{\lambda_k}^{\,2} e_k = \lambda_k e_k = Te_k$, so $T = R^2$.

((d) ⇒ (e)) Positive operators are self adjoint.

((e) ⇒ (f)) Suppose $T = R^2$ and R is self-adjoint. Then, $R = R^*$, so $R^2 = R^*R = T$.

((f) ⇒ (a)) Suppose $T = R^*R$. Then, $T^* = R^*(R^*)^* = R^*R = T$. So, T is self-adjoint. Moreover, $\langle Tv, v\rangle = \langle R^*Rv, v\rangle$

$$= \langle Rv, Rv\rangle = \|Rv\|^2 \ge 0.$$

So, T is positive.

7.39 *each positive operator has only one positive square root*

Every positive operator on V has a unique positive square root.

$\boxed{\text{Proof}}$ Suppose $T \in \mathcal{L}(V)$ is a positive operator. Let R be a positive square root of T, which we know exists by the last result. We show uniqueness. Let λ be an eigenvalue of T w/ eigenvector v. We will show that $Rv = \sqrt{\lambda}\, v$ which shows R is unique, since V has a basis of eigenvectors of T (by the spectral theorem).

That $Rv = \sqrt{\lambda} v$ is true follows from applying the spectral theorem to R to get an ONB of e-vec's of R for V. Let $e_1, \dots, e_n$ be this basis. Since R is positive, $\exists$ nonnegative $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ s.t. $Re_k = \sqrt{\lambda_k}\, e_k$. So, write $v = a_1 e_1 + \dots + a_n e_n$, and notice $R^2 v = a_1 \lambda_1 e_1 + \dots + a_n \lambda_n e_n$.

& $Tv = \lambda v$ implies $\lambda v = a_1 \lambda e_1 + \dots + a_n \lambda e_n = Tv = R^2 v = a_1 \lambda_1 e_1 + \dots + a_n \lambda_n e_n$

So, $a_1(\lambda - \lambda_1)e_1 + \dots + a_n(\lambda - \lambda_n)e_n = 0$ & since $e_1, \dots, e_n$ is a basis, if $a_i \neq 0$, then $\lambda = \lambda_i$. So, $Rv = a_1\sqrt{\lambda} e_1 + \dots + a_n \sqrt{\lambda} e_n = \sqrt{\lambda} v$

(since, if $a_i = 0$, we can still replace $\sqrt{\lambda_k}$ w/ $\sqrt{\lambda}$ here w/o changing anything).

7.40 notation: $\sqrt{T}$

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

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

Suppose T is a positive operator on V and $v \in V$ is such that $\langle Tv, v\rangle = 0$. Then $Tv = 0$.

$\boxed{\text{Proof}}$ $0 = \langle Tv, v\rangle = \langle \sqrt{T}\sqrt{T}v, v\rangle = \langle \sqrt{T}v, \sqrt{T}v\rangle = \|\sqrt{T}v\|$

So, $\sqrt{T}v = 0$. Thus, $Tv = \sqrt{T}\sqrt{T}v = \sqrt{T}0 = 0$.

$\boxed{\text{Ex}}$ Let $T \in \mathcal{L}(\mathbb{R}^2)$ be $T(x,y) = (x+y, x+y)$.

$\langle T(x,y), (x,y)\rangle = \langle (x+y, x+y), (x,y)\rangle = x(x+y) + y(x+y) = (x+y)^2 \geq 0$

So, T is positive, since w.r.t. the standard basis $(1,0), (0,1)$ of $\mathbb{R}^2$, $\mathcal{M}(T) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \mathcal{M}(T)^t = \mathcal{M}(T^*)$. Now, $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$ is a basis of $\mathbb{R}^2$ that consists only of e-vectors of T w/ corresponding e-vals $2, 0$. Wrt. this basis, T has a matrix $\mathcal{M}(T) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$, so $\mathcal{M}(\sqrt{T}) = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix}$ wrt. this basis.

Wrt. the standard basis of $\mathbb{R}^2$, $\mathcal{M}(\sqrt{T}) = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$.

7D Isometries, Unitary Operators, and Matrix Factorization

Isometries

7.44 definition: *isometry*

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

$$\|Sv\|_W = \|v\|_V$$

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

$\boxed{\text{Note}}$ Suppose $Sv = 0$. Then, $\|v\| = \|Sv\| = 0$, so $v = 0$.
That is, null $S = \{0\}$ & S is injective!

7.49 *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 $\mathbf{F}^m$ with respect to the Euclidean inner product.

$\boxed{\text{Proof}}$ $\boxed{(a) \Rightarrow (b)}$ If $v \in V$, then
$$\langle (I - S^*S)v, v \rangle = \langle v, v \rangle - \langle S^*Sv, v \rangle = \langle v, v \rangle - \langle Sv, Sv \rangle = \|v\|^2 - \|Sv\|^2 = 0$$

So, $I - S^*S$ is the zero operator, since $I - S^*S$ is self-adjoint & satisfies $\langle (I - S^*S)v, v \rangle = 0$ for $\forall v \in V$. That is $S^*S = I$.

$\boxed{(b) \Rightarrow (c)}$ Suppose $S^*S = I$. $\langle Su, Sv \rangle = \langle S^*Su, v \rangle = \langle u, v \rangle$

$\boxed{(c) \Rightarrow (d)}$ If $e_1, \ldots, e_n$ is an ONB of V & $\langle Su, Sv \rangle = \langle u, v \rangle$ $\forall u, v \in V$, then $\langle Se_i, Se_j \rangle = \langle e_i, e_j \rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$. So, $Se_1, \ldots, Se_n$ is an orthonormal list in W.

$\boxed{(d) \Rightarrow (e)}$ Let $A = \mathcal{M}(S)$. Then, for $k, r \in \{1, \ldots, n\}$, we have

$$\underline{A_{\cdot,k} \bullet A_{\cdot,r} = \sum_{j=1}^{n} A_{j,k} \overline{A_{j,r}}} = \left\langle \sum_{j=1}^{m} A_{jk} f_j, \sum_{i=1}^{m} A_{ir} f_i \right\rangle = \langle Se_k, Se_r \rangle = \begin{cases} 1 & k = r \\ 0 & k \neq r \end{cases}$$

$$\left\langle \sum_{j=1}^{m} A_{jk} f_j, \sum_{i=1}^{m} A_{ir} f_i \right\rangle = \sum_{j=1}^{m} A_{jk} \left\langle f_j, \sum_{i=1}^{m} A_{ir} f_i \right\rangle = \sum_{j=1}^{m} A_{jk} \sum_{i=1}^{m} \overline{A_{ir}} \langle f_j, f_i \rangle$$

$\boxed{\text{Proof}}$ (continued)

$\boxed{(c) \Rightarrow (a)}$ If the columns of $\mathcal{M}(S)$ are an orthonormal list in $\mathbb{F}^m$, then $Se_1, \dots, Se_n$ is an orthonormal list in W.

So, for $v \in V$, $v = \langle v, e_1 \rangle e_1 + \dots + \langle v, e_n \rangle e_n$ \$ $Sv = \langle v, e_1 \rangle Se_1 + \dots + \langle v, e_n \rangle Se_n$. So, $\|v\|^2 = |\langle v, e_1 \rangle|^2 + \dots + |\langle v, e_n \rangle|^2$ \$ $\|Sv\|^2 = |\langle v, e_1 \rangle|^2 + \dots + |\langle v, e_n \rangle|^2$ \$ S is an isometry.

Unitary Operators

7.51 definition: *unitary operator*

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

7.53 *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 of V.

(e) The rows of $\mathcal{M}(S, (e_1, \ldots, e_n))$ form an orthonormal basis of $\mathbf{F}^n$ with respect to the Euclidean inner product.

(f) S^* is a unitary operator.

Proof (a)⇒(b) If S is unitary, it is an isometry. So, $S^*S = I$. Since unitary operators are invertible, this gives $S^* = S^*SS^{-1} = S^{-1}$. So, $S^*S = SS^* = I$.

(b)⇒(c) $\$$ $S^*S = SS^* = I$. So, by definition, $S^{-1} = S^*$

(c)⇒(d) Since $S^{-1} = S^*$ implies $S^*S = I$, the previous result delivers this via (b)⇔(d) in the last result.

(d)⇒(e) This is a direct consequence of the last result (particularly, (d)⇒(e) there).

(e)⇒(f) $M(S)$ has ~~cols~~ rows an ONB implies $M(S^*) = \overline{M(S)}^t$ cols an ONB. So, S^* is an isometry & given both the rows & columns of $M(S)$ are ONBs, S^* is invertible. (doing this again w/ S^* replacing S)

(f)⇒(a)

just do (a)⇒(f) to S^* to get $(S^*)^* = S$ is unitary

7.54 *eigenvalues of unitary operators have absolute value* 1

Suppose λ is an eigenvalue of a unitary operator. Then $|\lambda| = 1$.

$\boxed{\text{Proof}}$ Suppose S is unitary and $Sv = \lambda v$ for $v \neq 0$ and $\lambda \in \mathbb{F}$. Then,

$$1 = \frac{\|Sv\|}{\|v\|} = \frac{\|\lambda v\|}{\|v\|} = \frac{|\lambda|\,\|v\|}{\|v\|} = |\lambda| \quad , \text{since } \|Sv\| = \|v\|.$$

7.55 *description of unitary operators on complex inner product spaces*

Suppose $\mathbf{F} = \mathbf{C}$ 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.

$\boxed{\text{Proof}}$ Suppose S is unitary. Then, $S^*S = I = SS^*$, so S is normal. Thus, by the complex spectral theorem $\exists$ an ONB of e-vec of S. Since all e-values of S have modulus 1, (a)$\Rightarrow$(b).

Now, suppose $\exists$ an ONB of e-vectors of S w/ corresponding e-vals all modulus 1. Then, $\lambda_1 e_1, \dots, \lambda_n e_n$ is also an ONB, since $\langle \lambda_i e_i, \lambda_k e_k \rangle = \langle e_i, e_k \rangle \lambda_i \overline{\lambda_k} = \begin{cases} 0 & i \neq k \\ \underset{\substack{\| \\ 1}}{|\lambda_i|^2} & i = k \end{cases}$

So, since $\lambda_1 e_1, \dots, \lambda_n e_n$ is just $Se_1, \dots, Se_n$, S is unitary by the previous result.

QR Factorization

7.56 definition: *unitary matrix*

An n-by-n matrix is called *unitary* if its columns form an orthonormal list in $\mathbf{F}^n$.

$\leftarrow$ i.e. if it is the matrix of a unitary operator w.r.t. an ONB.

7.57 *characterizations 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 $\mathbf{F}^n$.

(c) $\|Qv\| = \|v\|$ for every $v \in \mathbf{F}^n$.

(d) $Q^*Q = QQ^* = I$, the n-by-n matrix with 1's on the diagonal and 0's elsewhere.

7.58 *QR factorization*

Suppose A is a square matrix with linearly independent columns. Then there exist $\boxed{\text{unique}}$ matrices Q and R such that Q is unitary, R is upper triangular with only positive numbers on its diagonal, and

$$A = QR.$$

True, but we won't prove it

$\boxed{\text{proof}}$ Let $A \in \mathbb{F}^{n \times n}$ have linearly indep. cols, $v_1, \ldots, v_n$. Use GS to get $e_1, \ldots, e_n$ an ONB of $\mathbb{F}^n$ s.t. $\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(e_1, \ldots, e_k)$. Let R be the matrix w/ $R_{jk} = \langle v_k, e_j \rangle$. Then, R is UT, since for $k < j$, e_j is orthogonal to everything in $\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(e_1, \ldots, e_k)$. Then, if Q is the matrix w/ columns $e_1, \ldots, e_n$, it is unitary (since its cols are an ONB) & the kth col of QR is just $\langle v_k, e_1 \rangle e_1 + \cdots + \langle v_k, e_k \rangle e_k = v_k$. So, $QR = A$.

$\boxed{\text{Ex}}$ Let $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -4 \\ 0 & 3 & 2 \end{bmatrix}$. Applying GS to

$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $v_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$, $v_3 = \begin{bmatrix} 1 \\ -4 \\ 2 \end{bmatrix}$, we get

$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 1 \\ 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$, $e_3 = \begin{bmatrix} 0 \\ 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$. Then,

$$Q = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1/\sqrt{10} & 3/\sqrt{10} \\ 0 & -3/\sqrt{10} & 1/\sqrt{10} \end{bmatrix} \text{ \& } R = \begin{bmatrix} \langle v_1, e_1 \rangle & \langle v_2, e_1 \rangle & \langle v_3, e_1 \rangle \\ \langle v_1, e_2 \rangle & \langle v_2, e_2 \rangle & \langle v_3, e_2 \rangle \\ \langle v_1, e_3 \rangle & \langle v_2, e_3 \rangle & \langle v_3, e_3 \rangle \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & \sqrt{10} & \frac{\sqrt{10}}{5} \\ 0 & 0 & \frac{7\sqrt{10}}{5} \end{bmatrix}$$

You can check that $A = QR$.

This is nice: If we want to solve $Ax = b$, then $QRx = b$ & $Rx = Q^* b$ since Q is unitary.

As $Rx = y$ is easy to solve (R is UT), we can do this quickly!

$$Rx = \begin{bmatrix} 1 & 2 & 1 \\ 0 & \sqrt{10} & \frac{\sqrt{10}}{5} \\ 0 & 0 & \frac{7\sqrt{10}}{5} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a + 2b + c \\ \sqrt{10}\, b + \frac{\sqrt{10}}{5} c \\ \frac{7\sqrt{10}}{5} c \end{bmatrix}$$

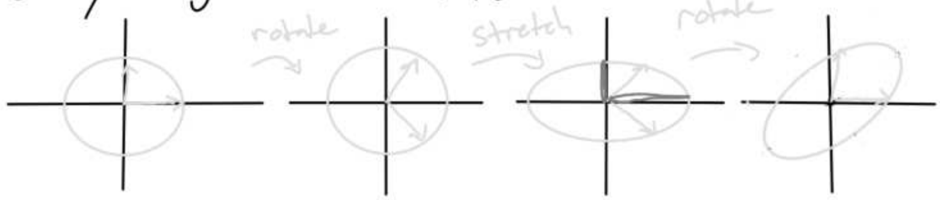
So, we can solve $Rx = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$ via $c = \frac{5}{7\sqrt{10}} z_3$

$$b = \frac{\left(z_2 - \left(\frac{\sqrt{10}}{5}\right)\left(\frac{5}{7\sqrt{10}}\right) z_3 \right)}{\sqrt{10}}$$

$a =$ (same process)

7E Singular Value Decomposition

Idea For linear maps between $\mathbb{R}^n$ & $\mathbb{R}^m$, everything works like this:



Today, we will show this is true & that an analog of it holds for maps in $\mathcal{L}(V,W)$ for V,W inner product spaces of finite dim.

7.64 *properties of T^*T*

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

(a) T^*T is a positive operator on V;

(b) $\operatorname{null} T^*T = \operatorname{null} T$;

(c) $\operatorname{range} T^*T = \operatorname{range} T^*$;

(d) $\dim \operatorname{range} T = \dim \operatorname{range} T^* = \dim \operatorname{range} T^*T$.

Proof **(a)** Notice $(T^*T)^* = T^*(T^*)^* = T^*T$, so T^*T is self-adjoint. If $v \in V$, $\langle T^*Tv, v\rangle = \langle Tv, Tv\rangle = \|Tv\|^2 \geq 0$. Hence, T^*T is positive.

(b) If $v \in \operatorname{null} T$, $Tv = 0$ so $T^*Tv = T^*0 = 0$ & $v \in \operatorname{null} T^*T$. Hence, $\underline{\operatorname{null} T \subseteq \operatorname{null} T^*T}$. Now, $\oint$ $v \in \operatorname{null} T^*T$. Then, we have $0 = \langle 0, v\rangle = \langle T^*Tv, v\rangle = \langle Tv, Tv\rangle = \|Tv\|^2$. So, $Tv = 0$ & $v \in \operatorname{null} T$. Hence $\underline{\operatorname{null} T^*T \subseteq \operatorname{null} T}$. So, $\operatorname{null} T = \operatorname{null} T^*T$.

(c) Since T^*T is self-adjoint (by part (a)), we have $\operatorname{range} T^*T = (\operatorname{null} T^*T)^\perp = (\operatorname{null} T)^\perp = \operatorname{range} T^*$.

(d) By part (c), $\dim \operatorname{range} T^* = \dim \operatorname{range} T^*T$. Now, observe $\dim \operatorname{range} T = \dim (\operatorname{null} T^*)^\perp = \dim W - \dim \operatorname{null} T^* = \dim \operatorname{range} T^*$.

7.65 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 included as many times as the dimension of the corresponding eigenspace of T^*T.

i.e. if $\lambda_1, \dots, \lambda_K$ are the eigenvalues of T^*T, and $\dim E(\lambda_i, T^*T) = n_i$, then $\underbrace{\sqrt{\lambda_1}, \dots, \sqrt{\lambda_1}}_{n_1}, \dots, \underbrace{\sqrt{\lambda_K}, \dots, \sqrt{\lambda_K}}_{n_K}$ are the singular values of T, assuming that $\lambda_1 \geq \dots \geq \lambda_K$.

7.68 *role of positive singular values*

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

(a) T is injective $\iff$ 0 is not a singular value of T;

(b) the number of positive singular values of T equals dim range T;

(c) T is surjective $\iff$ number of positive singular values of T equals dim W.

$\boxed{\text{Proof}}$ $\boxed{\text{(a)}}$ A liner map $T \in \mathcal{L}(V, W)$ is injective iff null $T = \{0\}$.

So, T is injective $\iff$ null $T = \{0\}$ $\iff$ null $T^*T = \{0\}$ $\iff$ 0 not an eigenvalue of T^*T $\iff$ 0 not a singular value of T

$\boxed{\text{(b)}}$ Using the spectral theorem on T^*T, there is a basis of e-vectors of T^*T s.t. $M(T)$ is diagonal w/ its eigenvalues on the diagonal. Hence, dim range T^*T = dim span(columns of $M(T)$)

= # of columns − # of times 0 appears on the diagonal of $M(T)$

= # of positive e-vals on diagonal of $M(T)$

= # of positive singular values

Now, just use dim range T = dim range T^*T

$\boxed{\text{(c)}}$ Obvious from (b).

list of eigenvalues	list of singular values
context: vector spaces	context: inner product spaces
defined only for linear maps from a vector space to itself (i.e. operators)	defined for linear maps from an inner product space to a possibly different inner product space
can be arbitrary real numbers (if **F** = **R**) or complex numbers (if **F** = **C**)	are nonnegative numbers
can be the empty list if **F** = **R**	length of list equals dimension of domain
includes 0 $\iff$ operator is not invertible	includes 0 $\iff$ linear map is not injective
no standard order, especially if **F** = **C**	always listed in decreasing order

bookkeeping, not really a "fact" in any substantial sense.

7.69 *isometries characterized by having all singular values equal* 1

Suppose that $S \in \mathcal{L}(V, W)$. Then

$$S \text{ is an isometry} \iff \text{all singular values of } S \text{ equal } 1.$$

$\boxed{\text{proof}}$ S is an isometry $\iff S^*S = I$
$\iff S^*S$ has all eigenvalues equal to 1
$\iff S$ has all singular values equal to 1

(the first equivalence is something we proved last time).

7.70 *singular value decomposition*

Suppose $T \in \mathcal{L}(V, W)$ and the positive singular values of T are $s_1, \ldots, s_m$. Then there exist orthonormal lists $e_1, \ldots, e_m$ in V and $f_1, \ldots, f_m$ in W such that

7.71 $$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

for every $v \in V$.

$\boxed{\text{proof}}$ Let $s_1, \ldots, s_n$ be the singular values of T (hence, $n = \dim V$). By the spectral theorem applied to the positive operator T^*T, there is an ONB $e_1, \ldots, e_n$ of V s.t. $T^*Te_k = s_k^2 e_k$ (since s_k is the square root of an eigenvalue of T^*T). Now, for $k \in \{1, \ldots, m\}$ (where $s_1, \ldots, s_m$ are positive & $s_{m+1}, \ldots, s_n$ are all 0), set $f_k = \frac{Te_k}{s_k}$.

Then, if $j, k \in \{1, \ldots, m\}$, we have

$$\langle f_j, f_k\rangle = \left\langle \tfrac{1}{s_j}Te_j, \tfrac{1}{s_k}Te_k\right\rangle = \frac{1}{s_j s_k}\langle e_j, T^*Te_k\rangle = \frac{s_k^2}{s_j s_k}\langle e_j, e_k\rangle = \begin{cases} 1 & j = k \\ 0 & j \neq k. \end{cases}$$

So, $f_1, \ldots, f_m$ is an orthonormal list in W.

Let $v \in V$. Since $e_1, \ldots, e_n$ is an ONB, we get

$$v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n$$
$$Tv = \langle v, e_1\rangle Te_1 + \cdots + \langle v, e_n\rangle Te_n$$
$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

since $s_k f_k = Te_k$ & this works if $s_k = 0$ as well since $Te_k = 0$ b/c $\operatorname{null} T = \operatorname{null} T^*T$

spectral theorem	singular value decomposition
describes only self-adjoint operators (when $\mathbf{F} = \mathbf{R}$) or normal operators (when $\mathbf{F} = \mathbf{C}$)	describes arbitrary linear maps from an inner product space to a possibly different inner product space
produces a single orthonormal basis	produces two orthonormal lists, one for domain space and one for range space, that are not necessarily the same even when range space equals domain space
different proofs depending on whether $\mathbf{F} = \mathbf{R}$ or $\mathbf{F} = \mathbf{C}$	same proof works regardless of whether $\mathbf{F} = \mathbf{R}$ or $\mathbf{F} = \mathbf{C}$

7.75 *singular value decomposition of adjoint and pseudoinverse*

Suppose $T \in \mathcal{L}(V, W)$ and the positive singular values of T are $s_1, \ldots, s_m$. Suppose $e_1, \ldots, e_m$ and $f_1, \ldots, f_m$ are orthonormal lists in V and W such that

7.76
$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

for every $v \in V$. Then

7.77
$$T^*w = s_1\langle w, f_1\rangle e_1 + \cdots + s_m\langle w, f_m\rangle e_m$$

and

7.78
$$T^\dagger w = \frac{\langle w, f_1\rangle}{s_1}e_1 + \cdots + \frac{\langle w, f_m\rangle}{s_m}e_m$$

for every $w \in W$.

7.80 *matrix version of 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-m diagonal matrix D with positive numbers on the diagonal, and an n-by-m matrix C with orthonormal columns such that

$$A = BDC^*.$$

This is the picture from the start of class!

8D Trace: A Connection Between Matrices and Operators

8.47 definition: *trace of a matrix*

Suppose A is a square matrix with entries in $\mathbf{F}$. The *trace* of A, denoted by $\operatorname{tr} A$, is defined to be the sum of the diagonal entries of A.

$\boxed{\text{Ex}}$ Let $A = \begin{bmatrix} 3 & -1 & 2 \\ 3 & 2 & -3 \\ 1 & 2 & 0 \end{bmatrix}$. Then, $\operatorname{tr} A = 3+2+0=5$.

Easy stuff!

8.49 *trace of AB equals trace of BA*

Suppose A is an m-by-n matrix and B is an n-by-m matrix. Then

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

$\boxed{\text{Proof}}$ Idea: If $A = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix}$ & $B = \begin{bmatrix} B_{11} & \cdots & B_{1n} \\ \vdots & & \\ B_{n1} & \cdots & B_{nn} \end{bmatrix}$

Then $\operatorname{tr} A = \operatorname{tr} \hat{A}$ & $\operatorname{tr} B = \operatorname{tr} \hat{B}$ w/ $\hat{A} = \begin{bmatrix} A_{11} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & A_{nn} \end{bmatrix}$ $\hat{B} = \begin{bmatrix} B_{11} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & B_{nn} \end{bmatrix}$

Since $\hat{A}$ & $\hat{B}$ commute, we are done!

More precisely: $(AB)_{jj} = \sum_{k=1}^{n} A_{jk} B_{kj}$, so

$$\operatorname{tr} AB = \sum_{j=1}^{m} \sum_{k=1}^{n} A_{jk} B_{kj} = \sum_{k=1}^{n} \sum_{j=1}^{m} B_{kj} A_{jk} = \operatorname{tr} BA.$$

8.50 *trace of matrix of operator does not depend on basis*

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

$$\operatorname{tr} \mathcal{M}(T, (u_1, \ldots, u_n)) = \operatorname{tr} \mathcal{M}(T, (v_1, \ldots, v_n)).$$ ✓

$\boxed{\text{Proof}}$ Let $A = \mathcal{M}(T, (u_1, \ldots, u_n))$ & $B = \mathcal{M}(T, (v_1, \ldots, v_n))$.

By the change of Basis formula (back in ch3), there is an invertible matrix C such that $A = C^{-1}BC$. Thus, using the previous result,

$$\operatorname{tr}(A) = \operatorname{tr}(C^{-1}BC) = \operatorname{tr}(C^{-1}CB) = \operatorname{tr}(B).$$

8.51 definition: *trace of an operator*

Suppose $T \in \mathcal{L}(V)$. The *trace* of T, denote $\operatorname{tr} T$, is defined by

$$\operatorname{tr} T = \operatorname{tr} \mathcal{M}(T, (v_1, ..., v_n)),$$

where $v_1, ..., v_n$ is any basis of V.

8.52 *on complex vector spaces, trace equals sum of eigenvalues*

Suppose $\mathbf{F} = \mathbf{C}$ 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.

← This says that if $\lambda_1, \ldots, \lambda_m$ are the distinct e-vals of T, then $\operatorname{tr} T = \sum_{i=1}^{m} \lambda_i \dim\left((T-\lambda_i I)^{\dim V}\right) = \sum_{i=1}^{m} \lambda_i \dim G(\lambda_i, T)$

$\boxed{\text{Proof}}$ Using the Jordan form, there is a basis of V for which the matrix of T is upper triangular & has on its diagonal its eigenvalues repeated as many times as their multiplicities. So, since the trace does not depend on the choice of basis, using this basis delivers the result!

Jordan Form | $\mathcal{M}(T) = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{bmatrix}$ w/ $A_i = \underbrace{\begin{bmatrix} \lambda_i & & * \\ & \ddots & \\ & & \lambda_i \end{bmatrix}}_{\dim G(\lambda_i, T)}$

8.54 *trace and characteristic polynomial*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then $\operatorname{tr} T$ equals the negative of the coefficient of z^{n-1} in the characteristic polynomial of T.

$\boxed{\text{Proof}}$ S $\mathbb{F} = \mathbb{C}$, $T \in \mathcal{L}(V)$, $n = \dim V$, & $\lambda_1, \ldots, \lambda_m$ are the eigenvalues of T. Then, the characteristic poly. of T is $(z-\lambda_1)^{\dim G(\lambda_1, T)} \cdots (z-\lambda_m)^{\dim G(\lambda_m, T)}$.

Expanding, we get a poly of the form

$$z^n - \underbrace{(\underbrace{\lambda_1 + \cdots + \lambda_1}_{\dim G(\lambda_1, T)} + \cdots + \underbrace{\lambda_m + \cdots + \lambda_m}_{\dim G(\lambda_m, T)})}_{\operatorname{tr} T} z^{n-1} + \cdots + ((\underbrace{\lambda_1 \cdots \lambda_1}_{\dim G(\lambda_1, T)}) \cdots (\underbrace{\lambda_m \cdots \lambda_m}_{\dim G(\lambda_m, T)}))(-1)^n$$

not really showing this bit

Or, if we let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of T w/ repetition according to multiplicity, we get the slightly cleaner form:

$$z^n - (\lambda_1 + \cdots + \lambda_n) z^{n-1} + \cdots + (-1)^n (\lambda_1 \cdots \lambda_n)$$

9C Determinants

Defining the Determinant

Def'n Suppose $T \in \mathcal{L}(V)$. If $\mathbb{F} = \mathbb{C}$, then the determinant of T, denoted $\det T$, is the product of the eigenvalues of T, each repeated according to their multiplicity. If $\mathbb{F} = \mathbb{R}$, then $\det(T) = \det(T_{\mathbb{C}})$.

Theorem The determinant of $T \in \mathcal{L}(V)$ is, up to a sign, the constant term of the characteristic poly of T.

Proof (Look at the last one)

Theorem An operator is invertible if & only if its determinant is nonzero.

proof By definition, $\det T = 0$ for $T \in \mathcal{L}(V)$ if & only if 0 is an eigenvalue of T (since a product of complex #'s is zero if & only if one is).
Hence, $\det T \neq 0 \Leftrightarrow 0$ not an e-val of T
$\Leftrightarrow T$ is invertible.

Theorem Suppose $T \in \mathcal{L}(V)$. Then, the characteristic poly. of T is $\det(zI - T)$.

Proof Suppose V is a $\mathbb{C}$-V.S. & $T \in \mathcal{L}(V)$, $\longleftarrow$ If $\mathbb{F} = \mathbb{R}$, then the same argument holds w/ T replaced by $T_{\mathbb{C}}$

λ is an e-val of $T \Leftrightarrow z - \lambda$ is an e-val of $zI - T$ (since $-(T - \lambda I) = (zI - T) - (z - \lambda)I$). So,
$\det(zI - T) = (z - \lambda_1) \cdots (z - \lambda_n)$ where the λ_i's are e-vals of T repeated according to their multiplicity. But this is just the characteristic poly of T!

Theorem Let $T \in \mathcal{L}(V)$ & define $g: \mathbb{F} \to \mathbb{F}$ by
$$g(x) = \det(I + xT).$$
Then, $g'(0) = \operatorname{tr} T$.

proof Notice that $I + xT$ has eigenvalues $1 + x\lambda_1, \ldots, 1 + x\lambda_n$ (each repeated according to their multiplicity)

So, $g(x) = (1 + x\lambda_1) \cdots (1 + x\lambda_n)$ and

$$g'(0) = \lim_{x \to 0} \frac{g(x) - g(0)}{x - 0}$$
$$= \lim_{x \to 0} \frac{(1 + x\lambda_1) \cdots (1 + x\lambda_n) - 1}{x}$$
$$= \lim_{x \to 0} \frac{\cancel{1} + (x\lambda_1 + \cdots + x\lambda_n) + \overbrace{\cdots + (x\lambda_1) \cdots (x\lambda_n)}^{\text{has } x \text{ terms of deg 2 or more}} - \cancel{1}}{x}$$
$$= \lim_{x \to 0} (\lambda_1 + \cdots + \lambda_n) + (\text{stuff w/ } x\text{'s})$$
$$= \lambda_1 + \cdots \lambda_n$$
$$= \operatorname{tr} T.$$

So, the trace is the rate of change of the determinant as it starts moving from the identity in the direction of your operator

---

Fact: There is a fancy version of this known as Jacobi's formula, which says that if $A : \mathbb{R} \to \mathcal{L}(V)$, then $\frac{d}{dt} \det A(t) = \operatorname{tr}\left(\operatorname{adj}(A(t)) \frac{dA}{dt}(t)\right)$ where $\operatorname{adj} A$ is the adjugate of A.

$\boxed{\text{Def'n}}$ A permutation of $(1,\dots,n)$ is a list $(m_1,\dots,m_n)$ where each of $1,\dots,n$ appears exactly once. The set of all permutations of $(1,\dots,n)$ is denoted S_n.

$\boxed{\text{Def'n}}$ The sign of a permutation $(m_1,\dots,m_n)\in S_n$ is the quantity $(-1)^t$ where t is the # of pairs (j,k) w/ $1\le j\le k\le n$ s.t. j appears after k in $(m_1,\dots,m_n)$.

$\boxed{\text{Ex}}$

- $(2,1,3,4)\in S_4$ has sign -1, since only $(2,1)$ is out of order
- $(2,3,\dots,n,1)\in S_n$ has sign $(-1)^{n-1}$

$\boxed{\text{Def'n}}$ Suppose A is an $n\times n$ matrix. Then,

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

$\boxed{\text{Ex}}$

$$\text{"det"} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \sum_{(m_1,m_2)\in S_2} \text{sign}(m_1,m_2)\, A_{m_1,1} A_{m_2,2}$$

$$= \text{sign}(\underline{1},\underline{2})\, A_{\underline{1},1} A_{\underline{2},2} + \text{sign}(\underline{2},\underline{1})\, A_{\underline{2},1} A_{\underline{1},2}$$

$$= A_{1,1} A_{2,2} - A_{2,1} A_{1,2}$$

Lemma Suppose A is upper triangular. Then, "det" $A = A_{11} \cdots A_{nn}$.

proof The permutation $(1, \dots, n) \in S_n$ contributes the term $\text{sign}(1, \dots, n) A_{1,1} \cdots A_{n,n} = A_{1,1} \cdots A_{n,n}$. For any other permutation $(m_1, \dots, m_n) \in S_n$, there is some entry m_k s.t. $k < m_k$. But, $\underline{A_{m_k,k} = 0}$ in this case since A is upper triangular. So, $\text{sign}(m_1, \dots, m_n) A_{m_1,1} \cdots A_{m_n,n} = 0$. Therefore,

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

Lemma "det"(AB) = "det"(BA) for A, B $n \times n$ matrices.

proof (sketch) Just write it out & use permutations to reorder everything.

Theorem "det"$(\mathcal{M}(T, (u_1, \dots, u_n)))$ = "det"$(\mathcal{M}(T, (v_1, \dots, v_n)))$ for $T \in \mathcal{L}(V)$ & $(v_1, \dots, v_n)$, $(u_1, \dots, u_n)$ two bases of V.

Proof By the change of basis formula, there is an invertible matrix C s.t. $A = C^{-1} B C$ where $A = \mathcal{M}(T, (u_1, \dots, u_n))$ & $B = \mathcal{M}(T, (v_1, \dots, v_n))$.
So, using the last lemma, "det"(A) = "det"$(C^{-1}BC)$ = "det"$(BC^{-1}C)$ = "det"(B).

Theorem "det" = det (i.e. $\det(T)$ = "det"$(\mathcal{M}(T))$)

proof Use Jordan normal form along w/ our lemma re: upper triangular matrices!

Theorem If V is an inner product space and $S \in \mathcal{L}(V)$ is unitary. Then, $|\det S| = 1$.

Proof $|\det S| = |\lambda_1 \cdots \lambda_n| = |\lambda_1| \cdots |\lambda_n| = 1 \cdots 1 = 1$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of S repeated according to their multiplicity.

Theorem If V is an inner product space and $T \in \mathcal{L}(V)$. Then, $\det T = s_1 \cdots s_n$ where $s_1, \dots, s_n$ are the singular values of T.

Proof By SVD, $\exists$ ONBs $e_1, \dots, e_n$ and $f_1, \dots, f_n$ of V s.t. $Tv = s_1 \langle v, e_1 \rangle f_1 + \dots + s_n \langle v, e_n \rangle f_n$. Take $S \in \mathcal{L}(V)$ so that $Se_k = f_k$ (this defines S since $e_1, \dots, e_n$ is a basis via the linear map lemma). Then,

$$
\begin{aligned}
\|Sv\|^2 &= \| S(\langle v, e_1 \rangle e_1 + \dots + \langle v, e_n \rangle e_n) \|^2 \\
&= \| \langle v, e_1 \rangle f_1 + \dots + \langle v, e_n \rangle f_n \|^2 \\
&= |\langle v, e_1 \rangle|^2 + \dots + |\langle v, e_n \rangle|^2 \\
&= \|v\|^2
\end{aligned}
$$

So, S is unitary! Moreover, $T = S\sqrt{T^*T}$.

Hence, $\det T = \det(S\sqrt{T^*T}) = (\det S)(\det \sqrt{T^*T}) = \det(\sqrt{T^*T}) = s_1 \cdots s_n$.

Theorem If $T \in \mathcal{L}(\mathbb{R}^n)$ and $\Omega \subseteq \mathbb{R}^n$. Then,

$$\text{volume } T\Omega = |\det T| \text{ volume } \Omega.$$

Picture

lengths are singular values

So, $|\det T|$ measures change in volume!