Linear Algebra Done Right

Haiweng Xu

2024.1.12

? ⁰

1 Vector Spaces

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

1.1.1 Complex Numbers

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

1.1.2 Lists

Definition 1.1.1. *Suppose n is a nonnegative integer. A list of length n is an ordered collection of n elements (which might be numbers, other lists, or more abstractobjects).*

Twolists are equal if and only if they have the same length and the sameelements in the same order.

Notice that the length of a list is finite, while that of what we call sequence is infinite.

1.1.3 $\mathbb{F}^n$

Definition 1.1.2. *$\mathbb{F}^n$ is the **set** of all lists of length n of elements of $\mathbb{F}$:*
$\mathbb{F}^n = (x_1, \ldots, x_n) : x_k \in \mathbb{F} for k = 1, \ldots, n.$
For $(x_1, \ldots, x_n) \in \mathbb{F}^n$ and $k \in 1, \ldots, n$, we say that x_k is the kth coordinate of $(x_1, \ldots, x_n)$.

A field is a set containing at least two distinct elements called 0 and 1, along with operations of addition and multiplication satisfying all properties. Thus $\mathbb{R}$ and $\mathbb{C}$ are fields, as is the set of rational numbers along with the usual operations of addition and multiplication. Another example of a field is the set $0, 1$ with the usual operations of addition and multiplication except that $1 + 1$ is defined to equal 0.

1.2 Definition of Vector Space

Definition 1.2.1. *real vector space, complex vector space*

A vector space over $\mathbb{R}$ is called a real vector space.

A vector space over $\mathbb{C}$ is called a complex vector space.

The scalar multiplication depends on $\mathbb{F}$

Definition 1.2.2. *$\mathbb{F}^\infty$ is defined to be the set of all **sequences** of elements of $\mathbb{F}^n$:*

$$\mathbb{F}^\infty = (x_1, x_2, \cdots) : x_k \in \mathbb{F}^n \text{ for } k = 1, 2, \cdots.$$

Addition and scalar multiplication on $\mathbb{F}^\infty$ are defined as expected:

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

$$\lambda(x_1, x_2, \cdots) = (\lambda x_1, \lambda x_2, \cdots).$$

With these definitions, $\mathbb{F}^\infty$ becomes a vector space over $\mathbb{F}^n$, as you should verify. The additive identity in this vector space is the sequence of all 0's.

Definition 1.2.3. *notation: $\mathbb{F}^S$*

" *If S is a set, then $\mathbb{F}^S$ denotes the set of functions from S to $\mathbb{F}$.* " *For $fg \in \mathbb{F}^S$, the sum $f + g \in \mathbb{F}^S$ is the function defined by*

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

for all $x \in S$. " *For $\lambda \in \mathbb{F}$ and $f \in \mathbb{F}^S$, the product $\lambda f \in \mathbb{F}^S$ is the function defined by*

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

for all $x \in S$.

We can write the definition in the word of set

$$\mathbb{F}^S = \{f | f : S \to \mathbb{F}\}$$

If S is not empty, then $\mathbb{F}^S$ is a vector space over $\mathbb{F}$

Actually, the set $\mathbb{F}^n$ and $\mathbb{F}^\infty$ is the special case of $\mathbb{F}^S$, in the first one, $S = \{1, 2, \ldots, n\}$, and in the latter one, $S = \{1, 2 \ldots\}$

Some interesting facts in Exercises

Exercise

1. The emptyset is not a vector space. The empty set fails to satisfy only one of the requirements listed in the definition of a vector space. Which one?

Solution

It fails to satisfy that there exists a addition indentity. Since there is no element in it.

But why it satisfy the other properties? Because such as the scalar multiplication, if it has no element, then the requirement will never be violated.

□

Exercise

2. Define $\bar{\mathbb{R}} = \mathbb{R} \bigcup \{\infty, -\infty\}$, and for $t \in RR$ define

$$
t\infty = \begin{cases} -\infty, & \text{if} t < 0 \\ 0, & \text{if} t = 0 \\ \infty, & \text{if} t > 0 \end{cases} , \quad t(-\infty) = \begin{cases} \infty, & \text{if} t < 0 \\ 0, & \text{if} t = 0 \\ -\infty, & \text{if} t > 0 \end{cases}
$$

and

$$
t + \infty = \infty + t = \infty
$$
$$
t + (-\infty) = (-\infty) + t = -\infty
$$
$$
\infty + (-\infty) = -\infty + \infty = 0
$$

Is $\bar{\mathbb{R}}$ a vector space?

Solution

No.

$$
(2 + (-1))\infty = 1\infty = \infty, \text{but} 2\infty + (-1)\infty = \infty + (-\infty) = 0
$$

It doesn't satisfy the distributive law.

□

1.3 Subspaces

Example 1.3.1. *subspaces*

(a) *If $b \in \mathbb{F}$, then*

$$(x_1, x_2, x_3, x_4) \in \mathbb{F}^4 : x_3 = 5x_4 + b$$

is a subspace of $\mathbb{F}^4$ if and only if $b = 0$.

(b) *Theset of continuous real-valued functions on the interval $[0, 1]$ is a subspace of $\mathbb{R}^{[0,1]}$.*

(c) *The set of differentiable real-valued functions on $\mathbb{R}$ is a subspace of $\mathbb{R}$.*

(d) *The set of differentiable real-valued functions f on the interval $(0, 3)$ such that $f'(2) = b$ is a subspace of $\mathbb{R}^{(0,3)}$ if and only if $b = 0$.*

(e) *The set of all sequences of complex numbers with limit 0 is a subspace of $\mathbb{C}^\infty$.*

1.3.1 Sum of Spaces

Definition 1.3.1. *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,

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

Theorem 1.3.1. *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 $V_1, \ldots, V_m$.

The proof contains three part: (1) subspace, (2) containing $V_1, \ldots, V_m$, (3) Smallest

The third one is hard to think, but you only need to prove that "Every subspace containing $V_1, \ldots, V_m$ contains $V_1 + \cdots + V_m$"

1.3.2 Direct Sums

Definition 1.3.2. *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 $\boldsymbol{v}_1 + \cdots + \boldsymbol{v}_m$, where each $\boldsymbol{v}_k \in V_k$.*

" *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*

How can we prove that the sum of two subspace is direct sum?

Here there are two theorems:

Theorem 1.3.2. *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 $\boldsymbol{0}$ as a sum $\boldsymbol{v}_1 + \cdots + \boldsymbol{v}_m$, where each $\boldsymbol{v}_k \in V_k$, is by taking each $\boldsymbol{v}_k$ equal to 0.

This theorem is the most gereral, and for the case with only 2 subspaces, the theorem below can solve.

Theorem 1.3.3. *direct sum of two subspaces*

Suppose U and W are subspaces of V. Then

$$U + W \text{ is a direct sum} \Leftrightarrow U \bigcap W = \{\boldsymbol{0}\}$$

Notice that this theorem only work for the two subspaces case. And more importantly, for the 3 subspaces case, with V_1, V_2, V_3, it's **no enough** to just prove:

$$V_1 \bigcap V_2 = \{\boldsymbol{0}\}, V_1 \bigcap V_3 = \{\boldsymbol{0}\}, V_2 \bigcap V_3 = \{\boldsymbol{0}\}$$

Why? Imagine 3 lines in $\mathbb{R}^2$ space! (It's always useful to find a counterexample in such a simple space)

Some interesting facts in Exercises

  Exercise

1. Is $\mathbb{R}^2$ a subspace of the complex vector space $\mathbb{C}^2$?

Solution

No. When we talked about $\mathbb{C}^2$, we should notice that now we are talking about the vector space over $\mathbb{C}$, thus the scalar can be **complex number**

But of course, form the aspect of set, $\mathbb{R}^2$ is the subset of $\mathbb{C}^2$, so be aware of the difference between the two perspective. $\qquad\square$

Exercise

2. A function $f : \mathbb{R} \to \mathbb{R}$ iscalled periodic if there exists a positive number p such that $f(x) = f(x + p)$ for all $x \in \mathbb{R}$. Is the set of periodic functions from $\mathbb{R}$ to $\mathbb{R}$ a subspace of $\mathbb{R}^{\mathbb{R}}$? Explain

Solution

No. Actually this exercises is more about the property of periodic functions

$$f(x) = \sin \sqrt{2}x, g(x) = \cos x$$

thus

$$f, g \in \mathbb{R}^{\mathbb{R}}$$

However, assume that $p > 0$ is the period of $f + g$,then

$$1 = f(0) + g(0) = f(p) + g(p) = f(-p) + g(-p)$$

$$i.e. 1 = \sin \sqrt{2}p + \cos p = \sin \sqrt{2}p - \cos p$$

thus

$$\begin{cases} \sin \sqrt{2}p = 0 \\ \cos p = 1 \end{cases}$$

from which we can know that $\sqrt{2}p = l\pi, l \in \mathbb{Z}$ and $p = 2k\pi, k \in \mathbb{Z}$

$$\sqrt{2} = \frac{l\pi}{2k\pi} = \frac{l}{2k} \in \mathbb{Q}$$

which is impossible!

□

Exercise

3. Prove that the union of three subspaces of V is a subspace of V if and only if one of the subspaces contains the other two.

This exercise is not true if we replace $\mathbb{F}$ with a field containing only two elements.

Solution

See (in tex):

https://math.stackexchange.com/questions/60698/if-a-field-f-is-such-that-leftf-rightn-1-why-is-v-a-vector-space-over

for the reason

If we ignore this case, this exercises is not so hard.

□

2 Finite-Dimensional Vector Spaces

2.1 Span and Linear Independence

Some theorems and definitions have been taught in the Book "Linear Algebra and Its Application" by David C.Lay, so we ignore them for concision.

Definition 2.1.1. *A vector space is called finite-dimensional if some list of vectors in it spans the space.*

Avector space is called infinite-dimensional if it is not finite-dimensional.

Definition 2.1.2. *polynomial* $\mathcal{P}(\mathbb{F})$
    " *A function* $p : \mathbb{F} \to \mathbb{F}$ *is called a polynomial with coefficients in* $\mathbb{F}$ *if there exist* $a_0, \ldots, a_m \in \mathbb{F}$ *such that for all* $z \in \mathbb{F}$.

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

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

Definition 2.1.3. *degree of a polynomial,* $\deg p$

" *A polynomial* $p \in \mathcal{P}(\mathbb{F})$ *is said to have **degree** m if there exist scalars* $a_0, a_1, \ldots, a_m \in \mathbb{F}$ *with* $a_m \neq 0$ *such that for every* $z \in \mathbb{F}$, *we have*

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

" *The polynomial that is identically* 0 *is said to have degree* $-\infty$.
" *The degree of a polynomial* p *is denoted by* $\deg p$.

Definition 2.1.4. *notation:* $\mathcal{P}_m(\mathbb{F})$

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

Example 2.1.1. *example:* $\mathcal{P}(\mathbb{F})$ *is infinite-dimensional.*

Consider any list of elements of $\mathcal{P}(\mathbb{F})$. *Let* m *denote the highest degree of the polynomials in this list. Then every polynomial in the span of this list has degree at most* m. *Thus* z^{m+1} *is not in the span of our list. Hence no list spans* $\mathcal{P}(\mathbb{F})$. *Thus* $\mathcal{P}(\mathbb{F})$ *is infinite-dimensional*

2.2 Bases

We also skip most parts we learned on class.

Theorem 2.2.1. *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* $V = U \oplus W$

Some interesting facts in Exercises

1. Prove or give a counterexample: If p_0, p_1, p_2, p_3 is a list in $\mathcal{P}_3(\mathbb{F})$ such that none of the polynomials p_0, p_1, p_2, p_3 has degree 2, then p_0, p_1, p_2, p_3 is not a basis of $\mathcal{P}_3(\mathbb{F})$.

We should be alert to that the degree only tells us the greatest power of the polynomial, "not having polynomial degree 2 " doesn't mean that there is no term like cz^2 in the vectors!

Thus it's obvious that it can be a basis. For example

$$p_1 = 1, \quad p_2 = z, \quad p_3 = z^2 + z^3, \quad p_4 = z^3$$

2.3   Dimensions

Notice: The real vector space $\mathbb{R}^2$ has dimension two; the complex vector space $\mathbb{C}$ has dimension one. As sets, $\mathbb{R}^2$ can be identified with $\mathbb{C}$ (and addition is the same on both spaces, as is scalar multiplication by real numbers). Thus when we talk about the dimension of a vector space, the role played by the choice of $\mathbb{F}$ cannot be neglected.

Theorem 2.3.1. *dimension of a sum*

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

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

The analogy between unions of subsets (in the context of sets) and sums of subspaces (in the context of vector spaces).

set	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 $S_1 \bigcup S_2$ 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
$\#(S_1 \bigcup S_2) = \#S_1 + \#S_2 - \#(S_1 \bigcap S_2)$	$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \bigcap V_2)$
$\#(S_1 \bigcup S_2) = \#S_1 + \#S_2 \Leftrightarrow S_1 \bigcap S_2 = \varnothing$	$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 \Leftrightarrow V_1 \bigcap V_2 = \{\mathbf{0}\}$
$S_1 \bigcup \cdots \bigcup S_m$ is a disjoint union $\Leftrightarrow \#(S_1 \bigcup \cdots \bigcup S_m) = \#S_1 + \cdots + \#S_m$	$V_1 + \cdots + V_m$ is a direct sum $\Leftrightarrow \dim(V_1 + \cdots + V_m) = \dim V_1 + \cdots + \dim V_m$

Some interesting facts in Exercises

Exercise

1.Suppose $v_1, \ldots, v_m$ is linearly independent in V and $w \in V$. Prove that

$$\dim \operatorname{Span}(v_1 + w, \ldots, v_m + w) \geq m - 1$$

Solution

There are mainly two solutions.

(1) It's obvious that

$$v_2 - v_1, \cdots v_m - v_1 \in \operatorname{Span}(v_1 + w, \ldots, v_m + w)$$

and they are linearly independent, and the length of the spanning list $\geq$ linearly indenpendent list, so

$$\dim \operatorname{Span}(v_1 + w, \ldots, v_m + w) \geq m - 1$$

(2) Let V denote $\operatorname{Span}(v_1 + w, \ldots, v_m + w)$ It's also obvious that

$$\dim \operatorname{Span}(w, v_1 + w, \ldots, v_m + w) = m$$

13

and

$$\text{Span}(w, v_1 + w, \dots, v_m + w) = \text{Span}(v_1 + w, \dots, v_m + w) + \text{Span}(w)$$

therefore

$$\begin{aligned}
m &= \dim \text{Span}(w, v_1 + w, \dots, v_m + w) \\
&= \dim V + \dim \text{Span}(w) - \dim \left(V \bigcap \dim \text{Span}(w) \right) \\
&\leq \dim V + \dim \text{Span}(w) \\
&= \dim V + 1
\end{aligned}$$

$\square$

3 Linear Maps

3.1 Vector Space of Linear Maps

Definition 3.1.1. *linear map*

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

additivity

$$T(\boldsymbol{u} + \boldsymbol{v}) = T\boldsymbol{u} + T\boldsymbol{v} \, for all \, \boldsymbol{u}, \boldsymbol{v} \in V$$

homogeneity

$$T(\lambda \boldsymbol{v}) = \lambda(T\boldsymbol{v}) \forall \lambda \in \mathbb{F} \, and \forall \boldsymbol{v} \in V$$

Definition 3.1.2. *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, $\mathcal{L}(V) = \mathcal{L}(V, V)$.

Example 3.1.1. *Linear maps*

zero

In addition to its other uses, we let the symbol 0 denote the linear map that takes every element of some vector space t the additive identity of another (or possibly the same) vector space. To be specific, $0 \in \mathcal{L}(V, W)$ is defined by

$$0\boldsymbol{v} = \boldsymbol{0}$$

The 0 on the left side of the equation above is a function from V to W, whereas the $\boldsymbol{0}$ on the right side is the additive identity in W. As usual, the context should allow you to distinguish between the many uses of the symbol 0.

identity operator

The identity operator, denoted by I, is the linear map on some vector space that takes each element to itself. To be specific, $I \in \mathcal{L}(V)$ is defined by

$$I\boldsymbol{v} = \boldsymbol{v}$$

differentiation

Define $D \in \mathcal{L}(\mathcal{P}(\mathbb{R}))by$

$$Dp = p'$$

The assertion that this function is a linear map is another way of stating a basic result about differentiation: $(f+g)' = f'+g'$ and $(\lambda f)' = \lambda f'$ whenever f, g are differentiable and λ is a constant.

integration

15

Define $T \in \mathcal{L}(\mathcal{P}(\mathbb{R}), \mathbb{R})$ by

$$Tp = \int_0^1 p$$

The assertion that this function is linear is another way of stating a basic result about integration: the integral of the sum of two functions equals the sum of the integrals, and the integral of a constant times a function equals the constant times the integral of the function.

multiplication by x^2

Define a linear map $T \in \mathcal{L}(\mathcal{P}(\mathbb{R}))$ by

$$(Tp)(x) = x^2 p(x)$$

for each $x \in \mathbb{R}$.

backward shift

Recall that $\mathbb{F}^\infty$ denotes the vector space of all sequences of elements of $\mathbb{F}$. Define a linear map $T \in \mathcal{L}(\mathbb{F}^\infty)$ by

$$T(x_1, x_2, x_3, \cdots) = (x_2, x_3, \cdots)$$

composition

Fix a polynomial $q \in \mathcal{P}|(\mathbb{R})$. *Define a linear map $T \in \mathcal{L}(\mathcal{P}(\mathbb{R}))$ by*

$$(Tp)(x) = p(q(x)) = p \circ q$$

Actually, the last one is hard to believe, so we can verify it.

First fix a polynomial q, which won't change in the process.

1.firstly verify $T(0) = 0(q(x)) = 0$

2.suppose $p_1 = a_{10} + a_{11}x + \cdots + a_{1n}x^n$, $p_2 = a_{20} + a_{21}x + \cdots + a_{2n}x^n$, then

$$T(p_1) = a_{10} + a_{11}q(x) + \cdots + a_{1n}q(x)^n$$

$$T(p_2) = a_{20} + a_{21}q(x) + \cdots + a_{2n}q(x)^n$$

$$T(p_1 + p_2) = (a_{10} + a_{20}) + (a_{11} + a_{21})q(x) + \cdots + (a_{1n} + a_{2n})q(x)^n$$
$$= a_{10} + a_{20} + a_{11}q(x) + a_{21}q(x) + \cdots + a_{1n}q(x)^n + a_{2n}q(x)^n$$
$$= T(p_1) + T(p_2)$$

3. Like in 2, we can verify $T(cp) = cT(p)$

Be aware of that it isn't correct if the position of p and q are changed, since not all polynomials are linear, for example, x^2 are not.

Some properties of Linear maps we have already learned on class will not be mentioned again.

Theorem 3.1.1. *Linear map lemma*

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

$$Tv_k = w_k$$

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

Theorem 3.1.2. $\mathcal{L}(V, W)$*is a vector space*

Some interesting facts in Exercises

Exercise

1. Suppose that $T \in (\mathbb{F}^n, \mathbb{F}^m)$. Show that there exist scalars $A_{j,k} \in \mathbb{F}$ for $j = 1, \ldots, m$ and $k = 1, \ldots, n$ such that

$$T(x_1, \ldots, x_n) = (A_{1,1}x_1 + \cdots + A_{1,n}x_n, \ldots, A_{m,1}x_1 + \cdots + A_{m,n}x_n)$$

for every $(x_1, \ldots, x_n) \in \mathbb{F}^n$

Solution

The exercises itself is simple, but be aware that it's always useful to describe the effect of a linear transformation with the basis.

For example,

$$T(1, 0, \ldots, 0) = (A_{1,1}, A_{2,1}, \ldots, A_{m,1})$$

and so on.

□

Exercise

2. Give an example of a function $\varphi : \mathbb{C} \to \mathbb{C}$ such that

$$\varphi(w + z) = \varphi(w) + \varphi(z)$$

for all $w, z \in \mathbb{C}$ but φ is not linear. (Here $\mathbb{C}$ is thought of as a complex vector space.)

Solution

Actually if you treat the real part and the imaginary part of the complex number differently, then you can easily find some problem.

Example 1:
$$\varphi(z) = \mathrm{Re}z$$

Example 2:
$$\varphi(z) = \bar{z}$$

if you then choose the scalar i, and you can make it.

Further Thinking:

There also exists a function $\varphi : \mathbb{R} \to \mathbb{R}$ such that φ satisfies the additivity condition above but φ is not linear. However, showing the existence of such a function involves considerably **more advanced tools**.

https://math.stackexchange.com/questions/16175/on-sort-of-linear-functions-does-fxy-fx-fy-imply-f-alpha-x-a

$\square$

Exercise

3. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T is a scalar multiple of the identity if and only if $ST = TS$ for every $S \in \mathcal{L}(V)$.

j

Solution

$\Rightarrow$ is easy, so we just ignore that part.

$\Leftarrow$:

Suppose $e_1, \ldots, e_n$ is a basis of V

Since the choice of S is random, we define $S_j \in \mathcal{L}(V)(1 \leq j \leq n)$ as

$$S_j e_i = \begin{cases} e_i, i \neq j \\ 2e_i, i = j \end{cases}$$

$$TS_j e_j = T(2e_j) = 2c_{j1}e_1 + \cdots + 2c_{ji}e_i + \cdots + 2c_{jj}e_j + \cdots + 2c_{jn}e_n$$

$$S_j T(e_j) = c_{j1}e_1 + \cdots + c_{ji}e_1 + \cdots + 2c_{jj}e_j + \cdots + c_{jn}e_n$$

So

$$c_{ji} = 0(i \neq j), \forall j, 1 \leq j \leq n$$

$$\therefore T e_j = c_{jj}e_j, \forall j, 1 \leq j \leq n$$

The exercise tells up don't be so "greedy" at a single time might help us, be focus on just one things (rather than thinking of swaping two base vector or make a loop of n vectors) □

Exercise

4. Suppose V is finite-dimensional. Show that the only two-sided ideals of $\mathcal{L}(V)$ are $\{0\}$ and $\mathcal{L}(V)$.

A subspace $\mathcal{E}$ of $\mathcal{L}(V)$ is called a **two-sided ideal** of $\mathcal{L}(V)$ if $TE \in \mathcal{E}$ and $ET \in \mathcal{E}$ $\forall E \in \mathcal{E}$ and $\forall T \in \mathcal{L}(V)$.

Here it introduce a new definition called two-sided ideal, which is a definition in the group theory.

Solution

To sole this problem, first you can verify that $\{0\}, \mathcal{L}(V)$ satisfy the definition.

Then, to prove that there is no other subspace that satisfy this property, suppose that U is a different but appropriate subspace.

Since V is a finite-dimensional space, we assume $v_1, \ldots, v_n$ is a basis of V

$$Tv_i = c_{i1}v_1 + \cdots + c_{in}v_n$$

suppose $S_i v_j = \delta_{ij} v_j$, then $S_i T \in \mathcal{E}$, then

$$I = \sum_{i=1}^{n} \frac{1}{c_{ii}} S_i T \in \mathcal{E}$$

because $U \neq \mathcal{L}(V), \exists T \in \mathcal{L}, s.t. T \notin U$, but according to the definition of the two-sided ideal, $IT = T \in U$, that is a contradition! □

3.2   Null Spaces and Ranges

Actually this section introduce nothing too new, but the exercises in the end really require a good command of the knowledge.

Theorem 3.2.1. *fundamental theorem of linear maps*

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

This theorem is also nothing new in essence, but it's from the perspective of the linear map rather than the column and rank of the matrix.

Important facts before the exercises

1. Suppose $T \in \mathcal{L}(V, W)$ is injective and $v_1, \ldots, v_n$ is linearly independent in V. Prove that $Tv_1, \ldots, Tv_n$ is linearly independent in W.

A obvious facts, prove with linear independence by yourself.

2. Suppose $v_1, \ldots, v_n$ spans V and $T \in (V, W)$. Show that $Tv_1, \ldots, Tv_n$ spans $range T$.

This simple fact is significantly useful in many exercises.

any $v \in V$ can be expressed with

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

then

$$Tv = c_1 Tv_1 + \cdots + c_n Tv_n$$

therefore $range T \in \mathrm{Span}(Tv_1, \ldots, Tv_n)$

On the other hand $Tv_i \in range T$, so by definition, $Tv_1, \ldots, Tv_n$ spans $range T$

To sum up, these two facts will play a key role in the analyze in the future.

Some interesting facts in Exercises

Exercise

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

$$\dim \text{null} ST \leq \dim \text{null} S + \dim \text{null} T$$

Solution

$$
\begin{aligned}
\dim \text{null} ST &= \dim U - \dim \text{range} ST \\
&= \dim \text{null} T + \dim \text{range} T - \dim \text{range} ST \\
&\leq \dim \text{null} T + \dim V - \dim \text{range} ST \\
&= \dim \text{null} T + \dim V - \dim \text{range} S + \dim \text{range} S - \dim \text{range} ST \\
&= \dim \text{null} T + \dim \text{null} S + \dim \text{range} S - \dim \text{range} ST
\end{aligned}
$$

Since $\text{range} ST \subset \text{range} S$

$$\dim \text{range} ST \leq \dim \text{range} S$$

$$\therefore \dim \text{null} ST \leq \dim \text{null} S + \dim \text{null} T$$

$\square$

Exercise

2. Suppose $D \in \mathcal{L}(\mathcal{P}(\mathbb{R}))$ is such that $\deg Dp = (\deg p) - 1$ for every non-constant polynomial $p \in \mathcal{P}(\mathbb{R})$. Prove that D is surjective.

Solution

For any $j \in \mathbb{N}$, there is a polynomial Dx^{j+1} with degree j, thus

$$\text{Span}(Dx, Dx^2, Dx^3, \cdots) \subset \text{range} D$$

Meanwhile

$$\text{Span}(Dx, Dx^2, Dx^3, \cdots) = \text{Span}(1, x, x^2, \cdots)$$

HenceSpan$(1, x, x^2, \cdots) \subset$ range$D$, namely range$D = \mathcal{P}(\mathbb{R})$

$\square$

Exercise

3. Suppose $p \in \mathcal{P}(\mathbb{R})$. Prove that there exists a polynomial $q \in \mathcal{P}(\mathbb{R})$ such that

$$5q'' + 3q' = p$$

Solution

Denote differentiation by D. Since $D \in \mathcal{P}(\mathbb{R}), 5D^2 + 3D \in \mathcal{P}(\mathbb{R})$

From the exercises above, we know $5D^2 + 3D$ is surjective, thus for $\forall p \in \mathcal{P}(\mathbb{R}), \exists q \in \mathcal{P}(\mathbb{R}) s.t. 5q'' + 3q' = p$

$\square$

Exercise

4.Suppose V is finite-dimensional with dim $V > 1$. Show that if $\varphi : \mathcal{L}(V) \to \mathbb{F}$ is a linear map such that $\varphi(ST) = \varphi(S)\varphi(T)$ for all $S, T \in \mathcal{L}(V)$, then $\varphi = 0$.

Hint: The description of the two-sided ideals of $\mathcal{L}(V)$

Solution

let $S = T = I$, then $\varphi(I) = \varphi(I) \cdots \varphi(I)$, namely $\varphi(I) = 1 \quad or \quad 0$

Suppose $\varphi(I) = 1$, and $v_1, \ldots, v_n$ is a basis of V, suppose $S_i \in \mathcal{L}(V)$ that satisfy

$$S_i v_j = (2\delta_{ij} - 1)v_j$$

Since $S_i^2 = S_i$, $\varphi(S_1) = 0 \quad or \quad 1$.
$\sum_{i=1}^n S_i = (2 - n)I$, and since φ should be linear,

$$\varphi(\sum_{i=1}^n S_i) = 2 - n \leq 0$$

That is not possible, then $\varphi(I) = 0, \forall T \in \mathcal{L}(V), \varphi(T) = \varphi(TI) = \varphi(T)\varphi(I) = 0$

□

3.3   Matrices

In the next two section, the content will become much harder and require a thorough understanding of the linear transformation and the matrices, and some definitions are slightly different from want we learn in class.

3.3.1   Representing a Linear Map by a Matrix

Definition 3.3.1. *definition: matrix, $A_{j,k}$*

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

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

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

Definition 3.3.2. *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 Aj, k are defined by

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

*If the bases $v_1, \ldots, v_n$ and $w_1, \ldots, w_m$ are not **clear from the context**, then the notation $\mathcal{M}(T, (v_1, \ldots, v_n), (w1, \ldots, w_m))$ is used.*

Try to understand what this mean, in the notation of matrix$\mathcal{M}$, you can find a transformation T and the two bases, the one in the left side represent the basis of the "start" vector space, and the right side represents

24

the basis of the "end" vector space, the k-th column in the matrix is the Tv_k's coefficients in each w_i

3.3.2   Addition and Scalar Multiplication of Matrices

After the definition of the adddition and scalar multiplication of the matrix, we can figure out the properties of $\mathcal{M}$

Theorem 3.3.1. *matrix of the sum of linear maps*

Suppose $S, T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$.
the matrix of a scalar times a linear map

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

And of course, with these to properties, the concept of **linear map** will come again.

notation: $\mathbb{F}^{m,n}$

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

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

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

3.3.3   Matrix Multiplication

Definition 3.3.3. *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,

25

is given by the equation

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

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.

With this definition, we can have the desirable properties.

Theorem 3.3.2. *matrix of product of linear maps*

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

Then we would like to introduce the notation of the columns and rows of the matrix

Definition 3.3.4. *notation: $A_{j,...}$, $A\cdots,k$*

Suppose A is an $m-by-n$ matrix.

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

After that, we will introduce the 4 different ways of understanding matrices multiplications, which we have already learned, so we just skip that part.

3.3.4 Column-Row Factorization and Rank of a Matrix

Now we'd like to introduce the concept of rank, but the definition is slightly different from what we have learned, so some basic proves might be different, but the actual properties are exactly the same.

Definition 3.3.5. *column rank, row rank*

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

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

" *The row rank of A is the dimension of the span of the rows of A in $\mathbb{F}^{1,n}$*

Definition 3.3.6. *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

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

Use the column-row factorization

Theorem 3.3.3. *Suppose A is an $m - by - n$ matrix with entries in $\mathbb{F}$ and column rank $c \geq 1$. Then there exist an $m - by - c$ matrix C and a $c - by - n$ matrix R, both with entries in $\mathbb{F}$, such that $A = CR$.*

we can now prove that the column rank is equal to the row rank, thus we can define the concept of rank

Definition 3.3.7. *The rank of a matrix $A \in \mathbb{F}^{m,n}$ is the column rank of A.*

Some interesting facts in Exercises
   Exercise

1.Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that $\dim \operatorname{range} T = 1$ if and only if there exist a basis of V and a basis of W such that with respect to these bases, all entries of $\mathcal{L}(T)$ equal 1.

Solution

$\Leftarrow$ is obvious.

$\Rightarrow$, suppose $u_1, \ldots, u_n$ is a basis of V, and from the exercise, we know $\dim \text{null} T = n - 1$, so assume that $Tu_1 \neq 0, Tu_i = 0, i = 2, \cdots n$

$$v_1 = u_1, v_i = v_1 + u_i, i = 2, \ldots, n$$

So $v_1, \ldots, v_n$ are linearly-independent, thus forming a base of V

Besides, Tv_1 spans $spanT$, so we can extend it to a basis of W, i.e. $Tv_1, \ldots, w_2, \ldots, w_n$

Let $w_1 = Tv_1 - w_2 - \cdots - w_n$, and also $w_1, \ldots, w_n$ are linearly-independent, and is a basis of W.

$\square$

3.4   Invertibility and Isomorphisms

3.4.1   Invertible Linear Maps

Definition 3.4.1. *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 $ST = I$ and $TS = I$ 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)*

Theorem 3.4.1. *An invertible linear map has a **unique inverse**.*

Definition 3.4.2. *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 $T^{-1}T = I$ and $TT^{-1} = I$.

Attention: The definition of inverse of linear map is different from the inverse of the matrix, since the inverse of linear map **does't require** the dimension of V and W to be the same, while the invertible matrix in $\mathbb{F}^{m,n}$ must have $m = n$

Theorem 3.4.2. *invertibility $\Leftrightarrow$ injectivity and surjectivity*

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

Theorem 3.4.3. *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 is invertible $\Leftrightarrow$ T is injective $\Leftrightarrow$ T is surjective.

Attention: the condition that $\dim V, \dim W < \infty$ is important, or we can find a counterexample:

Example 3.4.1. *The backward shift linear map from $\mathbb{F}^\infty$ to $\mathbb{F}^\infty$ is surjective (See 3A) but it is not invertible because it is not injective [the vector $(1, 0, 0, 0, \cdots)$ is in the null space].*

3.4.2 Isomorphic Vector Spaces

Definition 3.4.3. *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*

Theorem 3.4.4. *dimension shows whether vector spaces are isomorphic*

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

Theorem 3.4.5. $\mathcal{L}(V, W)$ *and* $\mathbb{F}^{m,n}$ *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 $\mathbb{F}^{m,n}$

Attention: Note that the notion $\mathcal{M}$ stands for the transformation from $\mathcal{L}(V, W)$ to $\mathbb{F}^{m,n}$, rather than a matrix!

Theorem 3.4.6. $\dim \mathcal{L}(V, W) = (\dim V)(\dim W)$

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

3.4.3 Linear Maps Thought of as Matrix Multiplication

We have already define the matrix of the linear maps, now we can define the matrix of a vector

Definition 3.4.4. *matrix of a vector,* $\mathcal{M}(v)$

*Suppose $v_1, \ldots, v_n$ is a basis of V. The matrix of v **with respect to this basis** is the $n - by - 1$ matrix*

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

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

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

The matrix $\mathcal{M}(v)$ of a vector $v \in V$ depends on the basis $v_1, \ldots, v_n$ of V, as well as on v. However, the basis should be clear from the context and thus it is not included in the notation.

This operation is like **relabling**. Occasionally,we want to think of elements of V as relabeled to be $n - by - 1$ matrices[for example we want to make a change in basis]. Once a basis $v_1, \ldots, v_n$ is chosen, the function $\mathcal{M}$ that takes $v \in V$ to $\mathcal{M}(v)$ is an isomorphism of V onto $\mathbb{F}^{n,1}$ that implements this relabeling.

Theorem 3.4.7. $\mathcal{M}(T)_{\cdot,k} = \mathcal{M}(Tv_k)$

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. Let $1 \leq k \leq n$. Then the kth column of $\mathcal{M}(T)$, which is denoted by $\mathcal{M}(T)_{\cdot,k}$, equals $\mathcal{M}(Tv_k)$.

Theorem 3.4.8. *linear maps act like matrix multiplication*

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

Theorem 3.4.9. *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 $\mathcal{M}(T)$.

3.4.4 Change of Basis

Definition 3.4.5. *If $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V, then the notation $\mathcal{M}(T, (v_1, \ldots, v_n))$ is defined by the equation*

$$\mathcal{M}(T, (v_1, \ldots, v_n)) = \mathcal{M}(T, (v_1, \ldots, v_n), (v_1, \ldots, v_n))$$

Definition 3.4.6. *invertible, inverse, A^{-1}*

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

Theorem 3.4.10. *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*

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

Theorem 3.4.11. *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 $\mathcal{M}(I, (u_1, \ldots, u_n), (v_1, \ldots, v_n))$ and $\mathcal{M}(I, (v_1, \ldots, v_n), (u_1, \ldots, u_n))$ are invertible, and each is the inverse of the other.*

Theorem 3.4.12. *change-of-basis formula*

Suppose $T \in \mathcal{L}(V)$. *Suppose $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ are bases of V. Let $A = \mathcal{M}(T, (u_1, \ldots, u_n))$ and $B = \mathcal{M}(T, (v_1, \ldots, v_n))$ and $C = \mathcal{M}(I, (u_1, \ldots, u_n), (v_1, \ldots, v_n))$. Then*

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

Theorem 3.4.13. *matrix of inverse equals inverse of matrix*

Suppose that $v_1, \ldots, v_n$ 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 $v_1, \ldots, v_n$

Notice the condition is in the same vector space and with the same basis.

Lemma 3.4.1. *Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$.*

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

Exercise

1. Suppose V is finite-dimensional and $S \in \mathcal{L}(V)$. Define $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ by $\mathcal{A}(T) = ST$, for $T \in \mathcal{L}(V)$.

    (a) Show that $\dim \text{null} \mathcal{A} = (\dim V)(\dim \text{null} S)$.
    (b) Show that $\dim \text{range} \mathcal{A} = (\dim V)(\dim \text{range} S)$.

Solution

(a)

Suppose $v_1, \ldots, v_n$ is a basis of V and $Sv_i = 0, i = 1, \cdots k$

null$\mathcal{A}$ is part of $\mathcal{A}$'s domain $\mathcal{L}(V, V)$, and the T in it has the property that

$$ST = 0, 0 \in \mathcal{L}(V)$$

i.e.

$$STv_i = 0, i = 1, \ldots, n$$

Therefore T must have

$$Tv_i = c_1 v_1 + \cdots + c_k v_k$$

Let $U = \text{Span}(v_1, \ldots, v_k)$, so $\dim U = \dim \text{null} S$

every $T \in \mathcal{L}(V, U)$ satisfy the property above

So

$$\dim \text{null} \mathcal{A} = \dim \mathcal{L}(V, U) = \dim V \dim U = \dim V \cdot \dim \text{null} S$$

(b)

Use the formula

$$\dim \mathcal{L}(V) = \dim V \cdot \dim V$$
$$= \dim \text{null} A + \dim \text{range} A$$

$\square$

Exercise

2.Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has the same matrix with respect to every basis of V if and only if T is a scalar multiple of the identity operator.

Solution

$\square$

Exercise

3. Suppose $q \in \mathcal{P}(\mathbb{R})$. Prove that there exists a polynomial $p \in \mathcal{P}(\mathbb{R})$ such that

$$q(x) = (x^2 + x)p''(x) + 2xp'(x) + p(3)$$

for all $x \in \mathbb{R}$

Solution

suppose $p \in \mathcal{P}_m$, then $q(x)$ also has degree m, so, let $T : \mathcal{P}_m \to \mathcal{P}_m$ have the property that $q = Tp$

$T(p) = 0 \Rightarrow p = 0$ so T is injective.

Since $\mathcal{P}_m$ is a finite-dimensional vector space, T is surjective, thus $\forall q \in \mathcal{P}_m, \exists p \in \mathcal{P}_m, s.t.q(x) = (x^2 + x)p''(x) + 2xp'(x) + p(3)$
$\square$

Exercise

4. Suppose V is finite-dimensional and $S, T, U \in \mathcal{L}(V)$ and $STU = I$. Show that T is invertible and that $T^{-1} = US$.

Show that the result in Exercise can fail without the hypothesis that V is finite-dimensional.

Solution

With the lemma we just mentioned, since $(ST)U$ is inversible, then ST, U are both invertible, and also S, T are inversible.

Therefore, $T^{-1} = US$

If V is infinite-dimensional, then there are some special linear map that do not exists with the finite-dimensional space.

Suppose U is the forward shift and add 0 at the first entry, T is the backward shift, S is the identity map.

Obviously, T, as a backward shift, is not invertible. $\square$

3.5   Products and Quotients of Vector Spaces

3.5.1   Products of Vector Spaces

Definition 3.5.1. *definition: product of vector spaces*

Suppose $V_1, \ldots, V_m$ are vector spaces over $\mathbb{F}$.

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

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

" *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*

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

Theorem 3.5.1. *product of vector spaces is **a vector space***

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

Theorem 3.5.2. *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

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

Theorem 3.5.3. *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

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

Obviously, Γ is surjective.

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

Theorem 3.5.4. *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 $\dim(V_1 + \cdots + V_m) = \dim V_1 + \cdots + \dim V_m$.

3.5.2   Quotient Spaces

Definition 3.5.2. *notation:* $v + U$

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

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

Definition 3.5.3. *translate (affine subset)*

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

Definition 3.5.4. *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 $V/U = \{v + U : v \in V\}$

Theorem 3.5.5. *two translates of a subspace are equal or disjoint*

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

$$v - w \in U \Leftrightarrow v + U = w + U \Leftrightarrow (v + U)\bigcap(w + U) \neq \varnothing$$

Definition 3.5.5. *addition and scalar multiplication on V/U*
     Suppose U is a subspace of V. Then addition and scalar multiplication are defined on V/U by

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

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

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

     This kind of definition is different form what we have seen, since the vectors in the space can have different expressions, so the definition make

37

sense only when the different expression of the same vector have the same result under operation

Proof

Let's take addition as an example.

suppose $v_1, v_2, w_1, w_2 \in V$ are such that

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

we have $v_1 - v_2 \in U$ and $w_1 - w_2 \in U$. Because U is a subspace of V and thus is closed under addition, this implies that $(v_1 - v_2) + (w_1 - w_2) \in U$. Thus $(v_1 + w_1) - (v_2 + w_2) \in U$.we see that

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

Theorem 3.5.6. *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.

Definition 3.5.6. *quotient map, π*

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

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

for each $v \in V$

Theorem 3.5.7. *dimension of quotient space*

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

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

Definition 3.5.7. *notation: $\tilde{T}$*

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

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

Theorem 3.5.8. *null space and range of $\tilde{T}$*

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

Exercise

Suppose $V_1, \ldots, V_m$ are vector spaces. Prove that $\mathcal{L}(V_1 \times \cdots \times V_m, W)$ and $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ are isomorphic vector spaces.

Solution

The exercises is easy **if we are told that the vector space is finite-dimensional**.

Since we don't know this kind of condition, so we can only return to the definition of **isomorphic**, i.e.finding a invertible linear map connecting the two vector space.

for any $f \in \mathcal{L}(V_1 \times \cdots \times V_m, W)$ and $i = 1, \ldots, m$, define $f_i : V_i \to W$ by

$$f_i(v_i) = f(0, \ldots, 0, v_i, 0, \ldots, 0)$$

Then is obvious that $f_i \in \mathcal{L}(V_i, W)$

Define $\varphi : \mathcal{L}(V_1 \times \cdots \times V_m, W) \to \mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ by

$$\varphi(f) = (f_1 \times \cdots \times f_m)$$

Now we can check that φ is a isomorphism between the two space by constructing its inverse map.

Define $\psi : \mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W) \to \mathcal{L}(V_1 \times \cdots \times V_m, W)$ by

$$\psi(f_1 \times \cdots \times f_m) = f$$

It's not hard to check that the φ, ψ are linear and the have the property that $\varphi \circ \psi = I, \psi \circ \varphi = I$

$\square$

Exercise

Suppose that v, x are vectors in V and that U, W are subspaces of V such that $v + U = x + W$. Provethat $U = W$.

Solution

Since the theorem in the section mostly focues on the case that we have already known that $U = W$, so this exercise may be tricky at first.

But we should be familiar with the way of proving two spaces are the same, i.e.to prove $U \subset W, W \subset U$

Since $v + U = x + W$, therefore $\forall p \in v + U, p \in x + W$. $0 \in U$ so $\forall v \in V, v \in v + U$ and must obey that $v \in x + W$, therefore $v = x + w_1$ for some $w_1 \in W$, therefore $v - x \in W$

For any $u \in U$, $\exists w_2 \in W$, s.t. $v + u = x + w_2$, so $u = x - v + w_2 \in W$

Thus $U \subset W$, the same is true for $W \subset U$ $\square$

Exercise

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

Solution

This exercises actually connect to two definition of translate of subspace (also affinie set as well).

$\Rightarrow$, If A is a translate of some vector space U, then $A = p + U, p \in V$, $\forall v, w \in A$ take $u_1 = v - p, u_2 = w - p \in U$. Therefore, $\lambda v + (1 - \lambda)w =$

$$p + \lambda(u_1 - u_2) \in v + U = A$$

$\Leftarrow$, Let $a \in A$, we will show that $A - a = \{x - a : x \in A\}$ is a subspace of V

For $x - a \in A - a$,

$$\lambda x + (1 - \lambda)a \in A \Rightarrow \lambda x + (1 - \lambda)a - a = \lambda(x - a) \in A - a$$

Which shows that $A - a$ is closed under scalar multiplication.

For $x - a, y - a \in A - a$, let $\lambda = \frac{1}{2}$,

$$\frac{1}{2}x + \frac{1}{2}y - a \in A - a$$

We have already proved that $A - a$ is closed under scalar multiplication, so

$$(x - a) + (y - a) \in A - a$$

Which shows that $A - a$ is closed under addition.

Obviously, $0 = a - a \in A - a$, so $A - a$ is a subspace of V

Thus $A = a + (A - a)$ is a affine set (translate from a subspace). $\square$

Exercise

Suppose $A_1 = v + U_1$ and $A_2 = w + U_2$ for some $v, w \in V$ and some subspaces U_1, U_2 of V. Prove that the intersection $A_1 \bigcap A_2$ is either a translate of some subspace of V or is the empty set.

Solution

Of course it can be empty

Suppose $A_1 \bigcap A_2 \neq \varnothing$, then $\forall x, y \in A_1 \bigcap A_2$, for the conclusion in last exercise, we know

$$x + (1 - \lambda)y \in A_1$$

$$x + (1 - \lambda)y \in A_2$$

which implies that
$$x + (1 - \lambda)y \in A_1 \bigcap A_2$$
And also for the conclusion of last exercise, $A_1 \bigcap A_2$ is a affine set, i.e.the translate of some vector space. □

Exercise
Suppose $v_1, \ldots, v_m \in V$. Let

$$A = \{\lambda_1 v_1 + \cdots + \lambda_m v_m : \lambda_1, \ldots, \lambda_m \in \mathbb{F} and \lambda_1 + \cdots + \lambda_m = 1\}$$

(a) Prove that A is a translate of some subspace of V.

(b) Prove that if B is a translate of some subspace of V and $\{v_1, \ldots, v_m\} \subset B$, then $A \subset B$.

(c) Prove that A is a translate of some subspace of V of dimension less than m

Solution
https://linearalgebras.com/3e.html
□

Exercise
Suppose $U = \{(x_1, x_2, \cdots) \in \mathbb{F}^\infty : x_k \neq 0 \text{ for only finitely many } k\}$.

(a) Show that U is a subspace of $\mathbb{F}^\infty$.

(b) Prove that $\mathbb{F}^\infty/U$ is infinite-dimensional.

Solution
The description of the set is a little confusing for my poor English.

But the exercises itself need very tricky construction and is very hard.

j

(a) Since 0 itself only have finite $x_k \neq 0$, and multiplication by a scalar won't increase the number of 0 and adding two vectors with m_1, m_2 0 respectively will get a vector with at most $m_1 + m_2$ 0

(b) Throughout this exercise, given a vector $v \in \mathbb{F}^\infty$, we'll denote by $v(p)$ the p th entry of v Consider the list of vectors e_m for the $m = 1, 2, \cdots$

in $\mathbb{F}^\infty$ defined by the rule:

$$e_m(p) = \begin{cases} 1, (p-1) \equiv 0(\mathrm{mod}\, m) \\ 0, \mathrm{otherwise} \end{cases}$$

For example

$$e_1 = (1, 1, 1, 1, 1, 1, \cdots)$$

$$e_2 = (1, 0, 1, 0, 1, 0, \cdots)$$

$$e_3 = (1, 0, 0, 1, 0, 0, \cdots)$$

We want to show this give us a linearly-independent list on $\mathbb{F}^\infty/U$. Let $k_1, \ldots, k_m \in \mathbb{F}$ such that:

$$k_1(e_1 + U) + \cdots + k_m(e_m + U) = 0 + U$$

Note that $k_1(e_1 + U) + \cdots + k_m(e_m + U) = k_1 e_1 + \cdots + k_m e_m + U$, so the equation above implies that $k_1 e_1 + \cdots k_m e_m = u$ for some $u \in U$

u only have finite nonzero entry, so it must only contain 0 from some position onward. Let L denote the last position that contain a nonzero entry, that mean $u(i) = 0$ for $i > L$.

Now let $p = tm! + 1$, where we choose $p \geq L$, and our choice tells that $e_1(p) = \cdots = e_m(p) = 1$, and $u(q) = 0$ for any $q > p$

Note that $e_1(p+1) = 1, e_2(p+1) = \cdots = e_m(p+1) = 0$, so $k_1 e_1(p+1) + \cdots + k_m e_m(p+1) = k_1$, and $u(p+1) = 0$. So $k_1 = 0$. Then we can show that $k_1 e_1(p+2) + \cdots + k_m e_m(p+2) = k_2 = 0$. Following the induction, we can show that $k_1 = \cdots = k_m = 0$

Since $\forall m$, the list we construct is linearly-independent, which means that no finite list can span the $\mathbb{F}^\infty/U$

Here is another construct.

43

Suppose $\mathbb{F}^\infty/U$ has only finite dimension n, then we construct $e_m(m = 1, 2, \cdots n+1)$ as follows

$$e_m = (1^{m-1}, 2^{m-1}, 3^{m-1}, \cdots)$$

Then $e_1 + U, \cdots e_{n+1} + U$ must be linearly-dependent. So there exists $k_1, \ldots, k_{n+1}$ which are not all zero, s.t. $k_1 e_1 + \cdots + k_{n+1} e_{n+1} \in U$ which means only have finite nonzero entry.

So $\exists \alpha \in \mathbb{Z}, \forall r > \alpha$

$$k_1 r^0 + k_2 r^1 + \cdots + k_{n+1} r^n = 0$$

So the polynomial

$$f(x) = k_{n+1} x^n + \cdots + k_2 x + k_1$$

has infinitely many zeros, which implies that $k_1 = k_2 = \cdots = k_{n+1} = 0$ and lead to a contradiction

$\square$

3.6   Duality

3.6.1   Dual Space and Dual Map

Definition 3.6.1. *linear functional*

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

Definition 3.6.2. *dual space, V'*

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

Theorem 3.6.1. $\dim V' = \dim V$

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

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

Definition 3.6.3. *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

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

Theorem 3.6.2. *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 for each $v \in V$

Theorem 3.6.3. *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'.

Definition 3.6.4. *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

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

Theorem 3.6.4. *algebraic properties of dual maps*

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

45

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

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

3.6.2 Null Space and Range of Dual of Linear Map

Definition 3.6.5. *annihilator, U^0*

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

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

Theorem 3.6.5. *the annihilator is a subspace*

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

Theorem 3.6.6. *dimension of the annihilator*

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

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

Theorem 3.6.7. *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) $U^0 = \{0\} \Leftrightarrow U = V$;
(b) $U^0 = V' \Leftrightarrow U = \{0\}$

Theorem 3.6.8. *the null space of T'*

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

Theorem 3.6.9. *T surjective is equivalent to T' injective*

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

$$T \text{ is surjective } \Leftrightarrow T' \text{is injective}$$

Theorem 3.6.10. *the range of T'*

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

Theorem 3.6.11. *T injective is equivalent to T' surjective*

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

$$T \text{ is injective } \Leftrightarrow T' \text{ is surjective}$$

3.6.3 Matrix of Dual of Linear Map

Theorem 3.6.12. *matrix of T' is transpose of matrix of T*

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

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

Exercise

Suppose that $V_1, \ldots, V_m$ are vector spaces. Prove that $(V_1 \times \cdots \times V_m)'$ and $V_1' \times \cdots \times V_m'$ are isomorphic vector spaces.

Solution

This exercise is similar to one in 3E, but we still take down the detailed solution.

Define $P_i \in \mathcal{L}(V_i, V_1 \times \cdots \times V_m)$ by

$$P_i(x) = (0, \ldots, 0, x, 0, \ldots, 0)$$

with x as the ith component.

Define $\varphi \in \mathcal{L}((V_1 \times \cdots \times V_m)', V_1' \times \cdots \times V_m')$ by

$$\varphi(f) = (P_1' f, \ldots, P_m' f)$$

Then we can check φ is a isomorphism.

injectivity:

suppose $\varphi(f) = (P_1'f, \ldots, P_m'f) = (0, \ldots, 0)$, then $P_i'f = 0 \, (i = 1, \ldots, m)$, assume $f = (x_1, \ldots, x_m)$, then $x_i = 0 \, (i = 1, \cdots m)$, which means $f = 0$

surjectivity:

for any $(f_1, \ldots, f_m) \in V_1' \times \cdots \times V_m'$, define $f \in (V_1 \times \cdots \times V_m)'$ by

$$f(x_1, \ldots, x_m) = \sum_{i=1}^{m} f_i(x_i)$$

then $\exists f \, s.t. \, \varphi f = (f_1, \ldots, f_m)$ $\qquad\square$

Exercise

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that T is invertible if and only if $T' \in \mathcal{L}(W', V')$ is invertible.

Solution

$$T \text{is surjective} \Leftrightarrow \text{range} T = W$$
$$\Leftrightarrow (\text{range} T)^0 = \{0\}$$
$$\Leftrightarrow \text{null} T' = \{0\}$$
$$\Leftrightarrow T' \text{is injective}$$

$$T \text{is injective} \Leftrightarrow \text{null} T = \{0\}$$
$$\Leftrightarrow (\text{null} T)^0 = V'$$
$$\Leftrightarrow \text{range} T' = V'$$
$$\Leftrightarrow T' \text{is surjective}$$

$\qquad\square$

Exercise

Suppose V is finite-dimensional and U and W are subspaces of V.

(a) Prove that $W^0 \subset U^0$ if and only if $U \subset W$.

(b) Prove that $W^0 = U^0$ if and only if $U = W$.

(c) Show that $(U + W)^0 = U^0 \cap W^0$.

(d) Show that $(U \cap W)^0 = U^0 + W^0$

Solution

This part is easy, you can look up the answer online or in OneNote   □

Exercise

Suppose V is finite-dimensional and $\varphi_1, \ldots, \varphi_n$ is a basis of V'. Show that there exists a basis of V whose dual basis is $\varphi_1, \ldots, \varphi_n$.

Solution

Recall that with a basis $v_1, \ldots, v_n$ in V, we can construct the dual basis $\varphi_1, \ldots, \varphi_n$ in V'. By linear map lemma, there exists a linear map $T : V \to V'$ defined by

$$T(c_1 v_1 + \cdots + c_n v_n) = c_1 \varphi_1 + \cdots + c_n \varphi_n$$

Since $\varphi_1, \ldots, \varphi_n$ is a basis, thus, T is surjective. And since $\dim V = \dim V'$, T is inversible. Thus given a dual basis, we can get the basis in V.

□

Exercise

Suppose V is finite-dimensional and $\varphi_1, \ldots, \varphi_m$ is a linearly independent list in V'. Prove that

$$\dim((\mathrm{null}\varphi_1) \cap \cdots \cap (\mathrm{null}\varphi_m)) = (\dim V) - m$$

Solution

Back to this exercise, first extend the $\varphi_1, \ldots, \varphi_m$ into a basis of $V', \varphi_1, \ldots, \varphi_m, \ldots, \varphi_n$, with which we can get a basis in V, namely $v_1, \ldots, v_n$ which also has the ideal property.

Therefore for $j = 1, \ldots, m$

$$\mathrm{null}\varphi_j = \{v \in V : v = c_1 v_1 + \cdots + c_{j-1} v_{j-1} + \cdots + c_{j+1} v_{j+1} + \cdots + c_n v_n\}$$

$$((\mathrm{null}\varphi_1) \cap \cdots \cap (\mathrm{null}\varphi_m)) = \{v \in V : v = c_{m+1} v_{m+1} + \cdots + c_n v_n\}$$

$$\dim((\mathrm{null}\varphi_1) \cap \cdots \cap (\mathrm{null}\varphi_m)) = n - m = \dim V - m$$

$\square$

Exercise

The double dual space of V, denoted by V'', is defined to be the dual space of V'. In other words, $V'' = (V')'$. Define $\Lambda : V \to V''$ by

$$(\Lambda v)(\varphi) = \varphi(v)$$

for each $v \in V$ and each $\varphi \in V'$.

(a) Show that Λ is a linear map from V to V''.

(b) Show that if $T \in \mathcal{L}(V)$, then $T'' \circ \Lambda = \Lambda \circ T$, where $T'' = (T')'$.

(c) Show that if V is finite-dimensional, then Λ is an isomorphism from V onto V".

Solution

(a) verify it by yourself.

(b) $\forall v \in V$, $(T'' \circ \Lambda)v = T''(\Lambda v)$ and $(\Lambda \circ T)v = \Lambda(Tv)$, the two are both the element of V'', to show they are equal, we just need to prove that $\forall f \in V'$,

$$T''(\Lambda v)f = \Lambda(Tv)f$$

We can simplify the left side

$$(T''(\Lambda v))f = (\Lambda v)(T'f)$$
$$= (T'f)v$$
$$= f(Tv)$$

And the right side

$$\Lambda(Tv)f = f(Tv)$$

(c) Since $\dim V = \dim V' = \dim V''$, thus it suffices to show that Λ is injective, which is not hard to prove $\square$

4   Polynomials

This section is short and only have limited content so we will just skip most part and keep the most important stuffs.

Theorem 4.0.1. *fundamental theorem of algebra, first version*

Every nonconstant polynomial with complex coefficients has a zero in $\mathbb{C}$.

Theorem 4.0.2. *fundamental theorem of algebra, second version*

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

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

where $c, \lambda_1, \ldots, \lambda_m \in \mathbb{C}$

5 Eigenvalues and Eigenvectors

5.1 Invariant Subspaces

5.1.1 Eigenvalues

Definition 5.1.1. *definition: operator*

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

Definition 5.1.2. *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 $u \in U$.*

Example 5.1.1. *If $T \in \mathcal{L}(V)$, then the following subspaces of V are all invariant under T.*

*{**0**} The subspace $\{0\}$ is invariant under T because if $u \in \{0\}$, then $u = 0$ and hence $Tu = 0 \in \{0\}$.*

V The subspace V is invariant under T because if $u \in V$, then $Tu \in V$.

nullT The subspace nullT is invariant under T because if $u \in$ nullT, then $Tu = 0$, and hence $Tu \in$ nullT.

rangeT The subspace rangeT is invariant under T because if $u \in$ rangeT, then $Tu \in$ rangeT.

Definition 5.1.3. *definition: eigenvalue*

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

Theorem 5.1.1. *equivalent conditions to be an eigenvalue*

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

(a) λ is an eigenvalue of T.

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

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

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

Definition 5.1.4. *definition: eigenvector*

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

Theorem 5.1.2. *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

Theorem 5.1.3. *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.

5.1.2 Polynomials Applied to Operators

Definition 5.1.5. *notation: T^m*

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

" $T^m \in \mathcal{L}(V)$ is defined by $Tm = \underbrace{T \cdots T}_{m \, times}$

" 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

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

Definition 5.1.6. *notation:* $p(T)$

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

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

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

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

Definition 5.1.7. *definition: product of polynomials*

If $p, q \in \mathcal{P}(\mathbb{F})$, *then* $pq \in \mathcal{P}(\mathbb{F})$ *is the polynomial defined by for all* $z \in \mathbb{F}$.

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

Theorem 5.1.4. *multiplicative properties*

Suppose $p, q \in \mathcal{P}(\mathbb{F})$ *and* $T \in \mathcal{L}(V)$. *Then*
(a) $(pq)(T) = p(T)q(T)$
(b) $p(T)q(T) = q(T)p(T)$.

Theorem 5.1.5. *null space and range of* $p(T)$ *are invariant under* T

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

Exercise

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

Solution

Since V is finite-dimensional, the number of eigenvalues of T is at most

53

$\dim V$, no matther whether α is the eigenvalue or not, let $\delta = \min_{1 \le j \le \dim V} |\lambda_j - \alpha|$ and this δ satisfy the condition. $\qquad \square$

Exercise

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

Solution

$\Rightarrow$

suppose the eigonvector corresponding to λ is v.

Therefore $Tv = \lambda v$, and $\forall \varphi \in V'$, $(T' - \lambda I)(\varphi)v = \varphi(T - \lambda I)v = 0$. Since $\forall v \in V$, $\exists \varphi \in V'$, s.t.$\varphi(v) \ne 0$, range$(T - \lambda I) \ne V'$, which implies $T' - \lambda I$ is not surjective thus λ is a eigenvalue of T'

$\Leftarrow$

suppose the eigonvector corresponding to λ is φ.

Therefore $T'\varphi = \lambda \varphi$, and $\forall v \in V$ we have, $T'\varphi v = \lambda \varphi v$ i.e.$\varphi(T - \lambda I)v = 0$. Since $forall\varphi \in V', \exists v \in V, s.t.\varphi v \ne 0$, so range$(T - \lambda I) \ne V$ which implies that λ is also an eigenvalue of T $\qquad \square$

Exercise

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

$$|\lambda| \le n \max |\mathcal{M}(T)_{j,k}| : 1 \le j, k \le n,$$

where $\mathcal{M}(T)_{j,k}$ denotes the entry in row j, column k of the matrix of T with respect to the basis $v_1, \ldots, v_n$.

Solution

Look up the OneNote for solution. $\qquad \square$

Exercise

Suppose $\mathbb{F} = \mathbb{R}$, $T \in \mathcal{L}(V)$,and $\lambda \in \mathbb{C}$. Prove that λ is an eigenvalue of the complexification T_C if and only if $\bar{\lambda}$ is an eigenvalue of T_C.

Solution

Assume that $\lambda = a + bi$ and therefore $\bar{\lambda} = a - bi$

$$T_c(u, v) = Tu + iTv = (a + bi)(u, v) = (a + bi)u + i(a + bi)v$$

54

from which we can get

$$\begin{cases} Tu = au - bv \\ Tv = bu + av \end{cases}$$

$$T_c(-u, v) = -Tu + iTv = -au + bv + i(bu + av) = (a - bi)(-u) + i(a - bi)v = (a - bi)(-u, v)$$

it shows that $\bar{\lambda}$ is a eigenvalue of T_c

vice versa. □

Exercise

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

Solution

One solution we are familiar with is use matrix. But we really want to figure out other way to solve the problem.

Assume λ is an eigenvalue of ST. Then there exists a non-zero vector x such that $STx = \lambda x$. And therefore $TSTx = \lambda Tx$. Let $y = Tx$. Then $TSy = \lambda y$. Therefore, λ is an eigenvalue of TS. Do not forget to also show this from the other angle!

Now, let's assume 0 is an eigenvalue of ST. This implies that ST is not invertible. Hence, S and T cannot both be invertible. Therefore, at least in the finite-dimensional case, TS is not invertible, so 0 is an eigenvalue of TS. □

Exercise

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

Solution

$T^4 - I = (T^2 + I)(T^2 - I) = 0$ which implies either $T^2 + I$ or $T^2 - I$ is not invertible (otherwise $T^4 - I$ will be invertible), and the $T^2 - I$ case can't be true since from which we can implies that $T - I$ of $T + I$ is not invertible and respectively indicate the eigenvalues to be 1 or -1. Therefore, $T^2 - I$ is inversible and $T^2 + I = 0$ □

Exercise

Suppose that $\lambda_1, \ldots, \lambda_n$ is a list of distinct real numbers. Prove that the list $e^{\lambda_1 x}, \ldots, e^{\lambda_n x}$ is linearly independent in the vector space of real-valued functions on $\mathbb{R}$.

Solution

This exercise is really interesting, define $D : \mathbb{R}^{\mathbb{R}} \to \mathbb{R}^{\mathbb{R}}$ by $Dp = p'$

Since that

$$De^{\lambda_1 x} = \lambda_1 e^{\lambda_1 x}, \ldots, De^{\lambda_n x} = \lambda_n e^{\lambda_n x}$$

$\lambda_1, \ldots, \lambda_n$ are the distinct eigenvalues of D and therefore $e^{\lambda_1 x}, \ldots, e^{\lambda_n x}$ are linearly independent.

The similar method can be applied to the sin and cos functions, despite the difference that let D represent by $Dp = p''$ $\square$

Exercise

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has an eigenvalue if and only if there exists a subspace of V of dimension $\dim V - 1$ that is invariant under T.

Solution

$\Rightarrow$, suppose $v_1, \ldots, v_n$ is a basis of V, let the eigenvalue be λ. $T - \lambda I$ is not surjective, so $\exists v_j$,s.t.

$$Tv_i = c_{i1}v_1 + \cdots c_{i,j-1}v_{j-1} + c_{i,j+1}v_{j+1} + \cdots + c_{in}v_n, i \neq j$$

which implies that T is invariant in $\mathrm{Span}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n)$, which has dimension $\dim V - 1$

$\Leftarrow$, let $\dim V = n$, denote the subspace with U, let $v_1, \ldots, v_{n-1}$ be a basis of U, and then extend it to a basis of V, $v_1, \ldots, v_n$, then we have

$$Tv_i = c_{i1}v_1 + \cdots + c_{i,n-1}v_{n-1}, i = 1, \ldots, n-1$$

$$Tv_n = c_{n1}v_1 + \cdots + c_{nn}v_n$$

Therefore

$$(T - c_{nn}I)v_i = p_{i1}v_1 + \cdots + p_{i,n-1}v_{n-1}$$

which implies that $T - c_{nn}I$ is not injective, and c_{nn} is an eigenvalue of T

□

5.2 The Minimal Polynomial

5.2.1 Existence of Eigenvalues on Complex Vector Spaces

Theorem 5.2.1. *existence of eigenvalues*

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

the finite-dimensional hypothesis in the result above also cannot be deleted.

Example 5.2.1. *Define $T \in \mathcal{L}(\mathcal{P}(\mathbb{C}))$ by $(Tp)(z) = zp(z)$. If $p \in \mathcal{P}(\mathbb{C})$ is a nonzero polynomial, then the degree of Tp is one more than the degree of p, and thus Tp canno equal a scalar multiple of p. Hence T has no eigenvalues. Because $\mathcal{P}(\mathbb{C})$ is infinite-dimensional, this example does not contradict the result above.*

5.2.2 Eigenvalues and the Minimal Polynomial

Definition 5.2.1. *monic polynomial*

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

Theorem 5.2.2. *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 $p \in \mathcal{P}(\mathbb{F})$ of **smallest degree** such that $p(T) = 0$. Furthermore, $\deg p \leq \dim V$*

Definition 5.2.2. *minimal polynomial*

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

$$p(T) = 0$$

Theorem 5.2.3. *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 vectorspace,then the minimal polynomial of T has the form $(z - \lambda_1) \cdots (z - \lambda_m)$,where $\lambda_1, \dots, \lambda_m$ is a list of all eigenvalues of T, possibly with repetitions.

Theorem 5.2.4. $q(T) = 0 \Leftrightarrow q$ *is a* **polynomial multiple** *of the minimal polynomial*

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

Theorem 5.2.5. *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 $T|_U$.

Theorem 5.2.6. T *not invertible* $\Leftrightarrow$ **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.

5.2.3 Eigenvalues on Odd-Dimensional Real Vector Spaces

Theorem 5.2.7. *even-dimensional null space*

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

Theorem 5.2.8. *operators on odd-dimensional vector spaces have eigenvalues*

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

Exercise

Suppose $\mathbb{F} = \mathbb{C}, T \in \mathcal{L}(V), p \in \mathcal{P}(\mathbb{C})$, and $\alpha \in \mathbb{C}$. Prove that α is an eigenvalue of $p(T)$ if and only if $\alpha = p(\lambda)$ for some eigenvalue λ of T.

Solution

if we assume p as

$$p(z) = a_0 + a_1 z + \cdots + a_{m-1} z^{m-1} + z^m$$

we denote the eigenvalues of T with $\lambda_1, \ldots, \lambda_m$. Then $p(\lambda_i)$ is the eigenvalues of the $p(T)$

This shows $\Leftarrow$ is correct.

$\Rightarrow$ is little tricky:

Suppose α is an eigenvalue of $p(T)$ with eigenvector v. Then indeed $(p(T) - \alpha I)v = 0$

Note that you can always factor polynomials over their complex roots. So for the polynomial $p(x) - \alpha$, there is a number $c \neq 0$ and roots $r_1, r_2, \ldots, r_n$ such that:

$$p(x) - \alpha = c(x - r_1)(x - r_2) \cdots (x - r_n)$$

In particular, $p(r_i) - \alpha = 0$ for all roots r_i. Applying this polynomial form to the equation $(p(T) - \alpha I)v = 0$ gives:

$$(p(T) - \alpha I)v = c(T - r_1 I)(T - r_2 I) \cdots (T - r_n I)v = 0$$

Now you just need to show that one of the roots r_i is an eigenvalue for T

$\square$

Exercise

Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that if at least one of S, T is invertible, then the minimal polynomial of ST equals the minimal polynomial of TS.

Hint: Show that if S is invertible and $p \in \mathcal{P}(\mathbb{F})$, then $p(TS) = S^{-1}p(ST)S$.

Solution

Follow the Hint or look up for the answer in OneNote $\square$

Exercise

Suppose $T \in \mathcal{L}(V)$ is such that with respect to some basis of V, all entries of the matrix of T are rational numbers. Explain why all coefficients of the minimal polynomial of T are rational numbers

Solution

let $v_1, \ldots, v_n$ be a basis of V and we can use induction to prove this fact

For the base case,

$$Tv_j = A_{1,j}v_1 + \cdots + A_{n,j}v_n \ (j = 1, 2, \cdots n)$$

where $A_{i,k}$ are all rational numbers.

Suppose for $p = k$

$$T^k v_j = c_{1,j}v_1 + \cdots + c_{n,j}v_n \ (j = 1, 2, \cdots n)$$

then for $p = k + 1$

$$T^{k+1}v_j = c_{1,j}(c_{1,1}v_1 + \cdots + c_{n,1}v_n) + \cdots + c_{n,j}(c_{1,n}v_1 + \cdots + c_{n,n}v_n)$$

$$= (\sum_{i=1}^{n} c_{i,j}c_{1,i})v_1 + \cdots + (\sum_{i=1}^{n} c_{i,j}c_{n,i})v_n \ (j = 1, 2, \cdots n)$$

and have all coefficients belongs to rational numbers.

The steps afterwards are obvious and we stop now. $\square$

Exercise

Suppose $a_0, \ldots, a_{n-1} \in \mathbb{F}$. Let T be the operator on $\mathbb{F}^n$ whose matrix (with respect to the standard basis) is

$$\begin{pmatrix} 0 & & & & & -a_0 \\ 1 & 0 & & & & -a_1 \\ & 1 & \ddots & & & -a_2 \\ & & \ddots & & & \vdots \\ & & & & 0 & -a_{n-2} \\ & & & & 1 & -a_{n-1} \end{pmatrix}$$

Here all entries of the matrix are 0 except for all 1's on the line under the diagonal and the entries in the last column (some of which might also be 0).

Show that the minimal polynomial of T is the polynomial

$$a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n$$

Solution

Let $e_1, e_2, \ldots, e_n$ be the standard basis of $\mathbb{C}^n$. Note that

$$T e_1 = e_2, T e_2 = e_3, \ldots, T e_{n-1} = e_n$$

Now note that

$$T^n e_1 = T e_n = -a_0 e_1 - a_1 e_2 - \cdots - a_{n-1} e_n = -a_0 e_1 - a_1 T e_1 - \cdots - a_{n-1} T^{n-1} e_1$$

which means

$$p(T) e_1 = 0$$

$$p(T) e_{i+1} = p(T) T^i e_1 = 0 (i = 1, 2, \ldots, n-1)$$

And then we get the solution we want. $\qquad\square$

Exercise

25 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V that is invariant under T.

(a) Prove that the minimal polynomial of T is a polynomial multiple of the minimal polynomial of the quotient operator T/U.

(b) Prove that

(minimal polynomial of $T|_U$) $\times$ (minimal polynomial of T/U)

is a polynomial multiple of the minimal polynomial of T.

Solution

(a) is easy since in $5A$ we proved that every eigenvalue of T/U is one of that of T

(b) $\forall \lambda$ of T, if the eigenvector v corresponding to it $\in U$, then it is also the eigenvalue of $T|_U$ and if it $\notin U$, [look up OneNote of 5D 19 if find it hard to figure out] then it is the eigenvalue of T/U, which means that the we finish the proof.

$\square$

Exercise

26 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V that is invariant under T. Prove that the set of eigenvalues of T equals the union of the set of eigenvalues of $T|U$ and the set of eigenvalues of T/U.

Solution

let $v_1, \ldots, v_k$ be a basis of U and extend it to $v_1, \ldots, v_n$, a basis of V, let λ be an eigenvalue $T|_U$, then it must be an eigenvalue of T, let λ be an eigenvalue of T/U, then we have for some $v + U \neq 0 + U$ (which implies that $v \notin U$), $(T/U)(v + U) = Tv + U = \lambda v + U$, which means that $Tv - \lambda v = (T - \lambda I)v \in U$, thus $T - \lambda I$ is not surjective, and λ is an eigenvalue of T.

Also for the conclusion of last exercise, we can get to our conclusion.

$\square$

Exercise

27 Suppose $\mathbb{F} = \mathbb{R}$, V is finite-dimensional, and $T \in \mathcal{L}(V)$. Prove that the minimal polynomial of T_C equals the minimal polynomial of T.

Solution

Actually T and T_C have the same eigonvalues, but the proof of this is exceedingly trival and you can refer to the OneNote of $5A$ ☐

Exercise

28 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the minimal polynomial of $T' \in \mathcal{L}(V')$ equals the minimal polynomial of T.

Solution

Just look at the solution in $5A$ and you will find the answer easily. ☐

5.3   Upper-Triangular Matrices

Definition 5.3.1. *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

Notice the coefficients is read in **column**!

Definition 5.3.2. *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.

Definition 5.3.3. *upper-triangularmatrix*

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

Theorem 5.3.1. *conditions for upper-triangular matrix*

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

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

(b) $Span(v_1, \ldots, v_k)$ is invariant under T for each $k = 1, \ldots, n$.

(c) $Tv_k \in Span(v_1, \ldots, v_k)$ for each $k = 1, \ldots, n$

Theorem 5.3.2. *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

$$(T - \lambda_1 I) \cdots (T - \lambda_n I) = 0$$

Theorem 5.3.3. *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.

Theorem 5.3.4. *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 $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$.

Theorem 5.3.5. *if $\mathbb{F} = \mathbb{C}$, 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.

The exercises in the section mostly have a strong connection with the former sections, so you can just review them and find the connection with exercises here.

5.4 Diagonalizable Operators

5.4.1 Diagonal Matrices

Definition 5.4.1. *diagonal matrix*

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

Definition 5.4.2. *diagonalizable*

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

Definition 5.4.3. *eigenspace, $E(\lambda, T)$*

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

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

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

Theorem 5.4.1. *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

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

5.4.2 Conditions for Diagonalizability

Theorem 5.4.2. *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) $\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$

Theorem 5.4.3. *enough eigenvalues implies diagonalizability*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$ has $\dim V$ **distinct** eigenvalues. Then T is diagonalizable.

Theorem 5.4.4. *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 $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$.

5.4.3 Gershgorin Disk Theorem

Definition 5.4.4. *Gershgorin disks*

Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ 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** $v_1, \ldots, v_n$ is a **set** of the form

$$\{z \in \mathbb{F} : |z - A_{j,j}| \leq \sum_{k=1, k \neq j}^{n} |A_{j,k}|\}$$

where $j \in \{1, \ldots, n\}$

Theorem 5.4.5. *Gershgorin disk theorem*

Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V. Then each eigenvalue of T is contained in some Gershgorin disk of T with respect to the basis $v_1, \ldots, v_n$

This theorem tells us that if the nondiagonal entries of A are small, then each eigenvalue of T is **near** a diagonal entry of A.

Exercise

Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if

$$V = null(T - \lambda I) \oplus range(T - \lambda I) \text{ for every } \lambda \in \mathbb{C}$$

Solution

First you can show that T is diagonalizable $\Leftrightarrow T - \lambda I$ is diagonalizable.

Assume that $(z - \lambda_1) \cdots (z - \lambda_n)$ is the minimal polynomial of T and $\lambda_1, \ldots, \lambda_n$ are distinct eigenvalues.

Therefore $(T - \lambda_1 I) \cdots (T - \lambda_n I) = 0$, $(T - \lambda I - (\lambda_1 - \lambda)I) \cdots (T - \lambda I - (\lambda_n - \lambda)I) = 0$ which implies that $T - \lambda I$ is diagonalizable $\forall \lambda \in \mathbb{C}$

If $null T = \{0\}$ then the proof has done.

So you can assume $v_1, \ldots, v_k$ is corresponding to eigenvalue 0, and the $v_{k+1}, \ldots, v_n$ corresponding to the non-zero eigenvalues $\lambda_1, \ldots, \lambda_m$, notice that

$$\begin{aligned} \text{range} T = \text{Span}(Tv_{k+1}, \ldots, Tv_n) &= \text{Span}(\lambda_{i_{k+1}} v_{k+1}, \ldots, \lambda_{i_n} v_n) \\ &= \text{Span}(v_{k+1}, \ldots, v_n) \end{aligned}$$

so

$$\begin{aligned} V = \text{Span}(v_1, \ldots, v_k) \oplus \text{Span}(v_{k+1}, \ldots, v_n) &= \text{null} T \oplus \text{range} T \\ &= \text{null}(T - \lambda I) \oplus \text{range}(T - \lambda I) \end{aligned}$$

$\square$

Exercise

Find $R, T \in \mathcal{L}(\mathbb{F}^4)$ such that R and T each have $2, 6, 7$ as eigenvalues, R and T have no other eigenvalues, and there does not exist an invertible operator $S \in \mathcal{L}(\mathbb{F}^4)$ such that $R = S^{-1}TS$.

Solution

Refer to https://linearalgebras.com/5c.html

(13)

$\square$

Exercise

Suppose $T \in \mathcal{L}(\mathbb{C}^3)$ is such that 6 and 7 are eigenvalues of T. Furthermore,

suppose T does not have a diagonal matrix with respect to any basis of $\mathbb{C}^3$. Prove that there exists $(z_1, z_2, z_3) \in \mathbb{C}^3$ such that

$$T(z_1, z_2, z_3) = (6 + 8z_1, 7 + 8z_2, 13 + 8z_3)$$

Solution

This exercise is quite hard to solve, since T does not have a diagonal matrix with respect to any basis of $\mathbb{C}^3$, so 8 is not an eigenvalue of T, which means $T - 8I$ is surjective. So there exists $v \in \mathbb{C}^3, s.t.$

$$(T - 8I)(z_1 z_2, z_3) = (6, 7, 13)$$

i.e.
$$T(z_1, z_2, z_3) = (6 + 8z_1, 7 + 8z_2, 13 + 8z_3)$$

$\square$

Exercise

Suppose $\mathbb{F} = \mathbb{C}$, k is a positive integer, and $T \in \mathcal{L}(V)$ is invertible. Prove that T is diagonalizable if and only if T^k is diagonalizable

Solution

Remained unsolved. $\square$

Exercise

Suppose that $T \in \mathcal{L}(V)$ is diagonalizable. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T. Prove that a subspace U of V is invariant under T if and only if there exist subspaces $U_1, \ldots, U_m$ of V such that $U_k \subset E(\lambda_k, T)$ for each k and $U = U_1 \oplus \cdots \oplus U_m$.

Solution

$\Leftarrow$ is quite obvious.

$\Rightarrow$: let $U_k = U \cap E(\lambda_k, T)$, this make sense because the intersection of two subspaces is still a subspace.

Since T is diagonalizable, then

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

So $\forall u \in U, \exists E(\lambda_k, T), s.t. u \in E(\lambda_k, T)$ therefore $u \in U_k$, which indicate that $U \subset (U_1 + \cdots + U_m)$

On the other hand, for all $u \in (U_1 + \cdots + U_m)$, $u = u_1 + \cdots + u_m$, and $u_1, \ldots, u_m \in U$, which means that $u \in U$ and thus $(U_1 + \cdots + U_m) \subset U$
Therefore
$$U = U_1 + \cdots + U_m$$

Since the eigenvectors corresponding to different eigenvalues are linearly-independent, so $u_1, \ldots, u_m$ are linearly-independent, which also shows
$$U = U_1 \oplus \cdots \oplus U_m$$

$\square$

Exercise
Prove or give a counter example: If $T \in \mathcal{L}(V)$ and there exists a subspace U of V that is invariant under T such that $T|_U$ and T/U are both diagonalizable, then T is diagonalizable.

Solution
Counter example:

Suppose $\dim V = 2$, and let v_1, v_2 be a basis of V, $U = \text{Span}(v_1)$. Define T by
$$Tv_i = \begin{cases} v_1, i = 1 \\ v_1 + v_2, i = 2 \end{cases}$$
Therefore, by definition,
$$\forall v \in U, T|_U v = v$$
and $\forall v \in V, v = c_1 v_1 + c_2 v_2$, therefore
$$T/U(v+U) = c_2 T/U(v_2+U) = c_2(Tv_2+U) = c_2(v_1+v_2+U) = c_1v_1+c_2v_2+U = v+U$$

Which implies that $T|_U$ and T/U are both diagonalizable.

However, let $v = c_1 v_1 + c_2 v_2$,
$$v = c_1 v_1 + c_2 v_2$$

$$Tv = (c_1 + c_2)v_1 + c_2 v_2$$

$$T^2 v = (c_1 + 2c_2)v_1 + c_2 v_2$$

Therefore

$$T^2 - 2T + I = 0$$

So the minimal polynomial of T is $p = (z - 1)^2$, which means T is not diagonalizable.

Why that? Think about the property of eigenvalues once again, we know that

$$(\text{minimal polynomial of } T|_U) \times (\text{minimal polynomial of } T/U)$$

is a polynomial multiple of the minimal polynomial of T, but what if $T|_U$ and T/U have common eigenvalues? That might lead to that the minimal polynomial of T have some $(z - \lambda_i)$ with power greater than 1 $\qquad\square$

5.5 Commuting Operators

Definition 5.5.1. *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 $AB = BA$.*

Theorem 5.5.1. *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.

Theorem 5.5.2. *eigenspace is invariant under commuting operator*

Suppose $S, T \in \mathcal{L}(V)$ commute and $\lambda \in \mathbb{F}$. Then $E(\lambda, S)$ is invariant under T

Theorem 5.5.3. *simultaneous diagonalizablity ⇔ 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.*

Theorem 5.5.4. *common eigenvector for commuting operators*

Every pair** of commuting operators on a finite-dimensional nonzero complex vector space has **a common eigenvector

Theorem 5.5.5. *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.*

Theorem 5.5.6. *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.*

notice that if V is over real numbers, then S, T can have no eigenvalue but $ST, S + T$ have eigenvalues.

Exercise

Suppose D_x, D_y are the commuting partial differentiation operators on $\mathcal{P}_3(\mathbb{R}^2)$ from that example. Find a basis of $\mathcal{P}_3(\mathbb{R}^2)$ with respect to which D_x and D_y each have an upper-triangular matrix.

Solution

Since D_x, D_y are commute, so we just need to find one basis with respect to which has upper-triangle matrix.

Let $1, x, y, xy, x^2, y^2, x^2y, xy^2, x^3, y^3$ be a basis, not hard to verify that this basis is what we want. $\qquad\square$

Exercise

Suppose V isafinite-dimensional nonzero complex vector space. Suppose that $\mathcal{E} \subset \mathcal{L}(V)$ is such that S and T commute for all $S, T \in \mathcal{E}$.

(a) Prove that there is a vector in V that is an eigenvector for every element of $\mathcal{E}$.

(b) Prove that there is a basis of V with respect to which every element of $\mathcal{E}$ has an upper-triangular matrix.

This exercise extends 5.78 and 5.80, which consider the case in which $\mathcal{E}$ contains only two elements. For this exercise, $\mathcal{E}$ may contain any number of elements, and $\mathcal{E}$ may even be an infinite set.

Solution

(a) let λ be an eigenvalue of S, and for any $T \in \mathcal{E}$, $E(\lambda, S)$ is invariant under T, let v be an eigenvector in $E(\lambda, S)$, this is also an eigenvector for all the $T \in \mathcal{E}$

(b)Suppose with respect to $\mathcal{B}$, the matrix of S is upper-triangle, than all $T \in \mathcal{E}$, the matrix of T is upper-triangle with respect to $\mathcal{B}$ $\qquad\square$

6 Inner Product Spaces

6.1 Inner Products and Norms

6.1.1 Inner Products

Definition 6.1.1. *dot product*

For $x, y \in \mathbb{R}^n$, 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 $y = (y_1, \ldots, y_n)$.

Definition 6.1.2. *inner product*

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

positivity

$\langle v, v \rangle \geq 0$ *for all $v \in V$.*

definiteness

$\langle v, v \rangle = 0$ *if and only if $v = 0$.*

additivity in first slot*(Notice it's the 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 $\lambda \in \mathbb{F}$ and all $u, v \in V$.*

conjugate symmetry

$\langle u, v \rangle = \overline{\langle v, u \rangle}$ *for all $u, v \in V$*

Definition 6.1.3. *inner product space*

An inner product space is a vector space V along with an inner product on V

Definition 6.1.4. *notation: V, W*

For the rest of this chapter and the next chapter, V and W denote inner product spaces over $\mathbb{F}$.

Theorem 6.1.1. *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 $\mathbb{F}$.

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

6.1.2 Norms

Definition 6.1.5. *norm, $\|v\|$*

For $v \in V$, the norm of v, denoted by $\|v\|$, is defined by

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

Theorem 6.1.2. *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 $\lambda \in \mathbb{F}$.

Definition 6.1.6. *orthogonal*

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

Theorem 6.1.3. *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*

Theorem 6.1.4. *Pythagorean theorem*

Suppose $u, v \in V$. If u and v are orthogonal, then $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.

Theorem 6.1.5. *an orthogonal decomposition*

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

Theorem 6.1.6. *Cauchy-Schwarz inequality*

Suppose $u, v \in V$. Then

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

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

With this inequality and the vector space of functions, we can prove some integral inequalities easily.

If f, g are continuous real-valued functions on $[-1, 1]$, then

$$\left| \int_{-1}^{1} fg \right| \leq \int_{-1}^{1} f^2 \int_{-1}^{1} g^2$$

as follows from applying the Cauchy-Schwarz inequality to Example 6.3(c).

Theorem 6.1.7. *triangle inequality*

Suppose $u, v \in V$. Then in this triangle, the length of $u + v$ is less than the length of u plus the length of v.

$$\|u + v\| \leq \|u\| + \|v\|$$

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

Theorem 6.1.8. *parallelogram equality*

Suppose $u, v \in V$. Then the diagonals of this parallelogram are $u + v$ and $u - v$.

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

.

Exercise

Suppose $u, v \in V$. Prove that $\langle u, v \rangle = 0 \Leftrightarrow \|u\| \leq \|u + av\|$ for all $a \in \mathbb{F}$.

Solution

As you have calculated, the inequality $\|u\|^2 \leq \|u + av\|^2$ constrains the real part of $Re(\bar{a} \langle u, v \rangle)$, because

$$\|u\|^2 \leq \|u + av\|^2 = \langle u + av, u + av \rangle = \|u\|^2 + 2Re(\bar{a} \langle u, v \rangle) + |a|^2 \|v\|^2$$

gives us

$$0 \leq 2Re(\bar{a} \langle u, v \rangle) + |a|^2 \|v\|^2$$

We can use the freedom in choosing a to compute both the real and imaginary parts of $\langle u, v \rangle$. We accomplish this by choosing real a in the first case and imaginary a in the second.

Specifically, let $\dfrac{\langle u, v \rangle}{\|v\|} = x + iy$, for $x, y \in \mathbb{R}$ and consider two cases $a = -\dfrac{x}{\|v\|}$ and $a = \dfrac{-iy}{\|v\|}$. Substituting, we get

$$0 \leq -2Re(x(x + iy)) + x^2 = -2x^2 + x^2 = -x^2$$

$$0 \leq 2Re(iy(x + iy)) + y^2 = -2y^2 + y^2 = -y^2$$

which means that $x^2 = y^2 = 0$ and hence $\langle u, v \rangle = 0$. □

Exercise

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

$$|\lambda|^2 \leq \sum_{h=1}^{n} \sum_{k=1}^{n} |\mathcal{M}(T)_{h,k}|^2$$

where $\mathcal{M}(T)_{h,k}$ denotes the entry in row j, column k of the matrix of T with respect to the basis $v_1, \ldots, v_n$.

Solution

Let $v_1, \ldots, v_n$ be a basis of V, which have $\|v_i\| = 1$ and $T \in \mathcal{L}(V)$. Suppose λ is an eigenvalue of T with eigenvector v. Then, we have $T(v) = \lambda v$. Taking the norm of both sides, we get $|\lambda|^2 \|v\|^2 = \|\lambda v\|^2 = \|T(v)\|^2$. Since $v_1, \ldots, v_n$ is a basis of V, we can write v as a linear combination of $v_1, \ldots, v_n$, say $v = a_1 v_1 + \cdots + a_n v_n$. Then, we have $\|v\|^2 = |a_1|^2 + \cdots + |a_n|^2$. Now,

we have

$$\|Tv\|^2 = \|T(a_1v_1 + \cdots + a_nv_n)\|^2$$

$$= \|a_1T(v_1) + \cdots + a_nT(v_n)\|^2$$

$$= \left\|\sum_{k=1}^{n} a_k \sum_{h=1}^{n} \mathcal{M}(T)_{h,k} v_h\right\|^2$$

$$= \left\|\sum_{h=1}^{n} v_h \sum_{k=1}^{n} \mathcal{M}(T)_{h,k} a_k\right\|^2$$

$$\leq \left\|\sum_{h=1}^{n} v_h \sqrt{\left(\sum_{k=1}^{n} |\mathcal{M}(T)_{h,k}|^2\right)\left(\sum_{k=1}^{n} |a_k|^2\right)}\right\|^2$$

$$\leq \sum_{h=1}^{n} \left(\sum_{k=1}^{n} |\mathcal{M}(T)_{h,k}|^2\right)\left(\sum_{k=1}^{n} |a_k|^2\right)\|v_h\|^2$$

$$= \|v\|^2 \sum_{h=1}^{n}\sum_{k=1}^{n} |\mathcal{M}(T)_{h,k}|^2$$

Since

$$\|\lambda v\|^2 = |\lambda|^2\|v\|^2 = \|Tv\|^2 \leq \|v\|^2 \sum_{h=1}^{n}\sum_{k=1}^{n} |\mathcal{M}(T)_{h,k}|^2$$

we finished our proof. $\qquad\square$

Exercise

Suppose $v_1, \ldots, v_m \in V$ are such that $\|v_k\| \leq 1$ for each $k = 1, \ldots, m$. Show that there exist $a_1, \ldots, a_m \in \{1, -1\}$ such that $\|a_1v_1 + \cdots + a_mv_m\| \leq \sqrt{m}$.

Solution

Use induction.

$k = 1$ is obvious.

suppose $k = p$ satisfy this inequality, then when $k = p + 1$,

$$\langle a_1v_1 + \cdots + a_{p+1}v_{p+1}, a_1v_1 + \cdots + a_{p+1}v_{p+1}\rangle$$

$$= \langle a_1v_1 + \cdots + a_pv_p, a_1v_1 + \cdots + a_pv_p\rangle$$

$$+ \langle a_{p+1}v_{p+1}, a_{p+1}v_{p+1}\rangle + 2a_{p+1}Re(\langle a_1v_1 + \cdots + a_pv_p, v_{p+1}\rangle)$$

$$\leq m + 1 + 2a_{p+1}Re(\langle a_1v_1 + \cdots + a_pv_p, v_{p+1}\rangle)$$

if $Re(\langle a_1 v_1 + \cdots + a_p v_p, v_{p+1} \rangle) \geq 0$, then let $a_{p+1} = -1$,

if $Re(\langle a_1 v_1 + \cdots + a_p v_p, v_{p+1} \rangle) < 0$, then let $a_{p+1} = 1$ $\qquad\square$

Exercise

Suppose $p > 0$. Prove that there is an inner product on $\mathbb{R}^2$ such that the associated norm is given by

$$\|(x, y)\| = (|x|^p + |y|^p)^{\frac{1}{p}}$$

for all $(x, y) \in \mathbb{R}^2$ if and only if $p = 2$.

Solution

$\Leftarrow$ is obvious

$\Rightarrow$, due to the positivity, p must be a even number, and norm associated with the inner product must obey the parallelogram equality, so $2(\|(1,0)\|^2 + \|(0,1)\|^2) = \|(1,1)\|^2 + \|(1,-1)\|^2 = 2 \cdot 2^{\frac{2}{p}}$ $\qquad\square$

Exercise

Suppose V is a real inner product space. Prove that

$$\langle u, v \rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}$$

for all $u, v \in V$.

Suppose V is a complex inner product space. Prove that

$$\langle u, v \rangle = \frac{\|u+v\|^2 - \|u-v\|^2 + \|u+iv\|^2 i - \|u-iv\|^2 i}{4}$$

for all $u, v \in V$.

Solution

let $v = cu + w$ $\qquad\square$

Exercise

A norm on a vector space U is a function

$$\| \cdot \| : U \to [0, \infty)$$

such that $\|u\| = 0$ if and only if $u = 0$, $\|\alpha u\| = |\alpha| \|u\|$ for all $\alpha \in \mathbb{F}$ and all $u \in U$, and $\|u + v\| \leq \|u\| + \|v\|$ forall $u, v \in U$. Prove that a norm satisfying

78

the parallelogram equality comes from an inner product (in other words, show that if $\| \cdot \|$ is a norm on U satisfying the parallelogram equality, then there is an inner product $\langle \cdot, \cdot \rangle$ on U such that $\|u\| = \langle u, u \rangle^{\frac{1}{2}}$ for all $u \in U$).

Solution

This part is hard, go to

https://zhuanlan.zhihu.com/p/597194291   □

Exercise

Fix a positive integer n. The Laplacian Δp of a twice differentiable real valued function p on $\mathbb{R}^n$ is the function on $\mathbb{R}^n$ defined by

$$\Delta p = \frac{\partial^2 p}{\partial x_1^2} + \cdots + \frac{\partial^2 p}{\partial x_n^2}$$

The function p is called harmonic if $\Delta p = 0$.

A polynomial on $\mathbb{R}^n$ is a linear combination (with coefficients in $\mathbb{R}$) of functions of the form $x_1^{m_1} \cdots x_n^{m_n}$, where $m_1, \ldots, m_n$ are nonnegative integers.

Suppose q is a polynomial on $\mathbb{R}^n$. Prove that there exists a harmonic polynomial p on $\mathbb{R}^n$ such that $p(x) = q(x)$ for every $x \in \mathbb{R}^n$ with $\|x\| = 1$.

The only fact about harmonic functions that you need for this exercise is that if p is a harmonic function on $\mathbb{R}^n$ and $p(x) = 0$ for all $x \in \mathbb{R}^n$ with $\|x\| = 1$, then $p = 0$.

Hint: A reasonable guess is that the desired harmonic polynomial p is of the form $q + (1 - \|x\|^2)r$ for some polynomial r. Prove that there is a polynomial r on $\mathbb{R}^n$ such that $q + (1 - \|x\|^2)r$ is harmonic by defining an operator T on a suitable vector space by $Tr = \Delta((1 - \|x\|^2)r)$ and then showing that T is injective and hence surjective.

Solution

You can just follow the hint.   □

6.2   Orthonormal Bases

6.2.1   Orthonormal Lists and the Gram-Schmidt Procedure

Definition 6.2.1. *orthonormal*

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

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

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

for all $j, k \in \{1, \ldots, m\}$

Theorem 6.2.1. *norm of an orthonormal linear combination*

Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in V. Then

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

for all $a_1, \ldots, a_m \in \mathbb{F}$

Theorem 6.2.2. *orthonormal lists are linearly independent*

Every orthonormal list of vectors is linearly independent.

Theorem 6.2.3. *Bessel's inequality*

Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in V. If $v \in V$ then

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

Definition 6.2.2. *orthonormal basis*

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

Theorem 6.2.4. *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

Theorem 6.2.5. *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) $\langle u, v \rangle = \langle u, e_1 \rangle \overline{\langle v, e_1 \rangle} + \cdots + \langle u, e_n \rangle \overline{\langle v, e_n \rangle}$

Theorem 6.2.6. *Gram-Schmidt procedure*

Suppose $v_1, \ldots, v_m$ is a linearly independent list of vectors in V. Let $f_1 = v_1$.

For $k = 2, \ldots, m$, define f_k inductively by

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

For each $k = 1, \ldots, m$, let $e_k = \dfrac{f_k}{\|f_k\|}$. Then $e_1, \ldots, e_m$ is an orthonormal list of vectors in V such that

$$span(v1, \ldots, vk) = span(e_1, \ldots, e_k)$$

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

Theorem 6.2.7. *existence of orthonormal basis*

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

Theorem 6.2.8. *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.

Theorem 6.2.9. *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 $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$.*

Theorem 6.2.10. *Schur's theorem*

*Every operator on a finite-dimensional **complex** inner product space has an upper-triangular matrix with respect to some orthonormal basis*

6.2.2 Linear Functionals on Inner Product Spaces

Definition 6.2.3. *linear functional, dual space, V'*

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

Example 6.2.1. *example: linear functional on $\mathcal{P}_5(\mathbb{R})$*

The function $\varphi : \mathcal{P}_5(\mathbb{R}) \to \mathbb{R}$ defined by

$$\varphi(p) = \int_{-1}^{1} p(t) \cos(\pi t)\, dt$$

is a linear functional on $\mathcal{P}_5(\mathbb{R})$.

Theorem 6.2.11. *Riesz representation theorem*

*Suppose V is finite-dimensional and φ is a linear functional on V. Then there is a **unique** vector $v \in V$ such that*

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

for every $u \in V$.
In it

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

The right side of the equation above **seems to** depend on the orthonormal basis $e_1, \ldots, e_n$ as well as on φ. However, v is uniquely determined by φ. Thus the right side of the equation above is the same regardless of which orthonormal basis $e_1, \ldots, e_n$ of V is chosen.

Exercise

(a) Suppose $\theta \in \mathbb{R}$. Show that both

$$(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta) \, and \, (\cos\theta, \sin\theta), (\sin\theta, -\cos\theta)$$

are orthonormal bases of $\mathbb{R}^2$.

(b) Show that each orthonormal basis of $\mathbb{R}^2$ is of the form given by one of the two possibilities in (a)

Solution

(a) is trivial, we skip this.

(b) suppose $u = (\cos\theta, \sin\theta)$, then $v = (\cos(\theta+\frac{\pi}{2}), \sin(\theta+\frac{\pi}{2})))$ or $(\cos(\theta-\frac{\pi}{2}), \sin(\theta-\frac{\pi}{2})))$, then just simplify.

So this exercise is easy but quite hard to think. $\qquad \square$

Exercise

Suppose $f : [-\pi, \pi] \to \mathbb{R}$ is continuous. For each nonnegative integer k, define

$$a_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x)\cos(kx)\,dx \text{ and } b_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x)\sin(kx)\,dx$$

Prove that

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty}(a_k^2 + b_k^2) \leq \int_{-\pi}^{\pi} f^2$$

The inequality above is actually an equality for all continuous functions $f : [-\pi, \pi] \to \mathbb{R}$. However, proving that this inequality is an equality involves Fourier series techniques beyond the scope of this book

Solution

Sorry to skip this exercise. $\qquad \square$

Exercise

Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V.

(a) Prove that if $v_1, \ldots, v_n$ are vectors in V such that

$$\|e_k - v_k\| < \frac{1}{\sqrt{n}}$$

for each k, then $v_1, \ldots, v_n$ is a basis of V.

(b) Show that there exist $v_1, \ldots, v_n \in V$ such that

$$\|e_k - v_k\| \leq \frac{1}{\sqrt{n}}$$

for each k, but $v_1, \ldots, v_n$ is not linearly independent.

This exercise states in (a) that an appropriately small perturbation of an orthonormal basis is a basis. Then (b) shows that the number $\frac{1}{\sqrt{n}}$ on the right side of the inequality in (a) cannot be improved upon.

Solution

(a)Suppose there are non-zero $a_1, \ldots, a_n$ s.t. $a_1 v_1 + \cdots + a_n v_n = 0$, therefore

$$
\begin{aligned}
\|a_1 e_1 + \cdots + a_n e_n\| &= \|a_1 e_1 + \cdots + a_n e_n - a_1 v_1 - \cdots - a_n v_n + a_1 v_1 + \cdots + a_n v_n\| \\
&\leq \|a_1(e_1 - v_1)\| + \cdots + \|a_n(e_n - v_n)\| + \|a_1 v_1 + \cdots + a_n v_n\| \\
&< \frac{1}{\sqrt{n}}|a_1| + \cdots + \frac{1}{\sqrt{n}}|a_n| + |a_1 v_1 + \cdots + a_n v_n| \\
&\leq \sqrt{\left(\frac{1}{n} + \cdots + \frac{1}{n}\right)(a_1^2 + \cdots + a_n^2)} + |a_1 v_1 + \cdots + a_n v_n| \\
&= \sqrt{a_1^2 + \cdots + a_n^2} + |a_1 v_1 + \cdots + a_n v_n| \\
&= \|a_1 e_1 + \cdots + a_n e_n\|
\end{aligned}
$$

which resulted in contradiction.

(b)let $v_k = e_k - \frac{1}{n}(e_1 + \cdots + e_n)$, we can verify that $\|v_k - e_k\| = \frac{1}{\sqrt{n}}$, but

$$v_1 + \cdots + v_n = (e_1 + \cdots + e_n) - n \cdot \frac{1}{n}(e_1 + \cdots + e_n) = 0$$

which shows that $v_1, \ldots, v_n$ are linearly-dependent. $\qquad\square$

Exercise

Suppose $v_1, \ldots, v_m$ is a linearly independent list in V. Explain why the orthonormal list produced by the formulas of the Gram-Schmidt procedure is the only orthonormal list $e_1, \ldots, e_m$ in V such that $\langle v_k, e_k \rangle > 0$ and $span(v_1, \ldots, v_k) = span(e_1, \ldots, e_k)$ for each $k = 1, \ldots, m$.

Solution

Hint: suppose there is a basis that is diffierent, then exists $j \in \{1, \ldots, m\}$

that is the first vector that is different. Then prove that in order to suffice the condition $\langle v_k, e_k \rangle > 0$ and $span(v_1, \ldots, v_k) = span(e_1, \ldots, e_k)$, these two vectors must be identical. $\qquad \square$

Exercise

Suppose $\mathbb{F} = \mathbb{C}$, V is finite-dimensional, $T \in \mathcal{L}(V)$, and all eigenvalues of T have absolute value less than 1. Let $\varepsilon > 0$. Prove that there exists a positive integer m such that $\|T^m v\| \leq \varepsilon \|v\|$ for every $v \in V$.

Solution

[The solution written by the author!]

Here is what I had in mind when I added that exercise to the third edition of Linear Algebra Done Right:

By Schur's Theorem (which is in the same section as this exercise), there is an orthonormal basis of V such that T has an upper-triangular matrix A with respect to that basis. Orthonormality is important so that we can compute norms. We will work with A instead of T and $z \in \mathbb{C}^n$ instead of v.

The entries on the diagonal of A are the eigenvalues of T. Thus all the diagonal entries of A have absolute value less than 1.

If we replace each entry in A **with its absolute value** and each coordinate of z with its absolute value, then $\|A^m z\|$ gets larger or remains the same for each positive integer m, and $\|z\|$ remains the same. Thus we can assume without loss of generality that each entry of A is nonnegative.

If we now replace each entry on the diagonal of A with the **largest element** on the diagonal of A, then $\|A^m z\|$ gets larger or remains the same for each positive integer m. Thus we can assume without loss of generality that there exists a nonnegative number $\lambda < 1$ such that every entry of the diagonal A equals λ.

We can write

$$A = \lambda I + N$$

where N is an upper-triangular matrix whose diagonal entries all equal 0. Thus $N^n = 0$.

Let m be a positive integer with $m > n$. Then

$$A^m = (\lambda I + N)^m$$
$$= \lambda^m I + m\lambda^{m-1} N + \frac{m(m-1)}{2}\lambda^{m-2} N^2$$
$$+ \frac{m(m-1)(m-2)}{2 \cdot 3}\lambda^{m-3} N^3 + \cdots + \frac{m(m-1)\cdots(m-n+1)}{(n-1)!}\lambda^{m-n+1} N^{n-1}$$

where the other terms in the binomial expansion do not appear because $N^k = 0$ for $k \geq n$. In the expression above, the coefficients of the n matrices $I, N, N^2, N^3, \ldots, N^{n-1}$ are all bounded by

$$\frac{m^n \lambda^m}{\lambda^{n-1}}$$

Think of n and λ as fixed. Then the limit of the expression above as $m \to \infty$ equals 0 (because $0 \leq \lambda < 1$).

Because the sum above has only a fixed number n of terms, we thus see that by taking m sufficiently large, we can make all the entries of the matrix A^m have absolute value as small as we wish. In particular, there exists a positive integer m such every entry of A^m has absolute value less than ε/n.

The definition of matrix multiplication and the Cauchy-Schwarz Inequality imply that for each j, the entry in row j, column 1 of $A^m z$ has absolute value at most $\varepsilon \|z\|/\sqrt{n}$. Thus $\|A^m z\| \leq \varepsilon \|z\|$, as desired.

This exercise is, in my opinion, one of the harder exercises in the third edition of Linear Algebra Done Right if one uses only the tools available at that stage of the book. The proof becomes much clearer and cleaner if one uses the tools associated with the spectral radius in a Banach algebra, which shows the power of those tools. $\quad\square$

Exercise

Suppose $C[-1, 1]$ is the vector space of continuous real-valued functions on the interval $[-1, 1]$ with inner product given by

$$\langle f, g \rangle = \int_{-1}^{1} fg$$

for all $f, g \in C[-1, 1]$. Let φ be the linear functional on $C[-1, 1]$ defined by $\varphi(f) = f(0)$. Show that there does not exist $g \in C[-1, 1]$ such that

$$\varphi(f) = \langle f, g \rangle$$

for every $f \in C[-1, 1]$.

This exercise shows that the Riesz representation theorem (6.42) does not hold on infinite-dimensional vector spaces without additional hypotheses on V and φ.

Solution

I think it's more about the knowledge of the constructing functions, so I didn't copy the answer.

https://linearalgebras.com/6b.html $\qquad\square$

Then there is an exercise about the metric space in the function analysis which is beyond our knowledge now.

Exercise

Forall $u, v \in V$, define $d(u, v) = \|u - v\|$.

(a) Show that d is a metric on V.

(b) Show that if V is finite-dimensional, then d is a complete metric on V (meaning that every Cauchy sequence converges).

(c) Show that every finite-dimensional subspace of V is a closed subset of V (with respect to the metric d).

6.3 Orthogonal Complements and Minimization Problems

6.3.1 Orthogonal Complements

Definition 6.3.1. *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:

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

Theorem 6.3.1. *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 $U \cap U^\perp \subset 0$.*
(e) *If G and H are subsets of V and $G \subset H$, then $H^\perp \subset G^\perp$*

Theorem 6.3.2. *direct sum of a subspace and its orthogonal complement*

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

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

Theorem 6.3.3. *dimension of orthogonal complement*

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

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

Theorem 6.3.4. *orthogonal complement of the orthogonal complement*
Suppose U is a finite-dimensional subspace of V. Then

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

the result below can fail without the hypothesis that the subspace U is finite dimensional.

Theorem 6.3.5. *$U^\perp = \{0\} \Leftrightarrow U = V$ (for U a finite-dimensional subspace of V)*

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

$$U^\perp = \{0\} \Leftrightarrow U = V$$

Definition 6.3.2. *orthogonal projection, P_U*

Suppose U is a finite-dimensional subspace of V. The orthogonal pro-
jection 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 $P_U v = u$.

Theorem 6.3.6. *properties of orthogonal projection P_U*

Suppose U is a finite-dimensional subspace of V. Then
(a) $P_U \in \mathcal{L}(V)$;
(b) $P_U u = u$ for every $u \in U$;
(c) $P_U w = 0$ for every $w \in U^\perp$;
(d) $range P_U = U$;
(e) $null P_U = U^\perp$;
(f) $v - P_U v \in U^\perp$ for every $v \in V$;
(g) $P_U^2 = P_U$;
(h) $\|P_U v\| \leq \|v\|$ for every $v \in V$;
(i) if $e_1, \ldots, e_m$ is an orthonormal basis of U and $v \in V$, then

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

Theorem 6.3.7. *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 $v \to \varphi v$ is a one-to-one function from V onto V'.

Caution: The function $v \to \varphi_v$ is a **linear mapping** from V to V'
if $\mathbb{F} = \mathbb{R}$. However, this function is **not linear** if $\mathbb{F} = \mathbb{C}$ because $\varphi_{\lambda v} = \bar{\lambda} \varphi_v$ if $\lambda \in \mathbb{C}$.

6.3.2 Minimization Problems

Theorem 6.3.8. *minimizing distance to a subspace*

Suppose U is a finite-dimensional subspace of V, $v \in V$, and $u \in U$. Then

$$\|v - P_U v\| \leq \|v - u\|$$

Furthermore, the inequality above is an equality if and only if $u = P_U v$

We can use this to find the polynomial approximation of the function $\sin x$ on the interval $[-\pi, \pi]$, with better approximation than the Taylor series!

6.3.3   Pseudoinverse

Theorem 6.3.9. *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 $T|_{(nullT)^\perp}$ is a one to-one map of $(nullT)^\perp$ onto $rangeT$.

Definition 6.3.3. *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

$$T^\dagger w = (T|_{(nullT)^\perp})^{-1} P_{rangeT} w$$

for each $w \in W$.

Theorem 6.3.10. *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) $TT^\dagger = P_{rangeT} =$ the orthogonal projection of W onto $rangeT$.
(c) $T^\dagger T = P_{(nullT)^\perp} =$ the orthogonal projection of V onto $(nullT)^\perp$

Theorem 6.3.11. *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

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

with equality if and only if $x \in T^\dagger b + \text{null} T$.
(b) If $x \in T^\dagger b + \text{null} T$, then

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

with equality if and only if $x = T^\dagger b$.

Exercise

Suppose $e_1, \ldots, e_n$ is a list of vectors in V with $\|e_k\| = 1$ for each $k = 1, \ldots, n$
and

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

for all $v \in V$. Prove that $e_1, \ldots, e_n$ is an orthonormal basis of V.

Solution

let $v = e_i$ then

$$1 = 1 + |\langle e_i, e_1 \rangle|^2 + \cdots + |\langle e_i, e_{i-1} \rangle|^2 + |\langle e_i, e_{i+1} \rangle|^2 + \cdots + |\langle e_i, e_n \rangle|^2$$

which implies that $\langle e_i, e_j \rangle = 0$ for $i \neq j$, so $e_1, \ldots, e_n$ is an orthonormal list.

Then we show that $e_1, \ldots, e_n$ spans V.
let $U = \text{Span}(e_1, \ldots, e_n)$, $V = U \oplus U^\perp$. So

$$\forall v \in V, v = c_1 e_1 + \cdots + c_n e_n + w$$

such that $u = c_1 e_1 + \cdots + c_n e_n \in U, w \in U^\perp$.

$$\begin{aligned}
\|v\|^2 &= \langle u + w, u + w \rangle \\
&= \langle u, u \rangle + \langle u, w \rangle + \langle w, u \rangle + \langle w, w \rangle \\
&= \langle u, u \rangle + \langle w, w \rangle \\
&= |c_1|^2 + \cdots + |c_n|^2 + \langle w, w \rangle \\
&= |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2 + \langle w, w \rangle
\end{aligned}$$

which implies that $w = 0$, thus we finished our proof.

$\square$

Exercise

Suppose V is finite-dimensional and $P \in (V)$ is such that $P^2 = P$ and

$$\|Pv\| \leq \|v\|$$

for every $v \in V$. Prove that there exists a subspace U of V such that $P = P_U$.

Solution

Let $U = \text{range} P$, so for $u \in U$, $\exists u = Pv$

$$u = Pv = P^2 v = P(Pv) = Pu$$

We have $V = \text{range} P \oplus \text{null} P$, for any $u \in \text{range} P$ and $w \in \text{null} P$

$$\|w\| = \|Pw\| = \|P(\lambda u + w)\| \leq \|\lambda u + w\|$$

which implies that $\langle u, w \rangle = 0$. Since u, w are arbitrarily chosen, $\text{range} P \perp \text{null} P$, the following is easy.

But actually this method still implicitly use the property of the eigenvalues. $\square$

Exercise

Suppose $C[-1, 1]$ is the vector space of continuous real-valued functions on the interval $[-1, 1]$ with inner product given by

$$\langle f, g \rangle = \int_{-1}^{1} fg$$

for all $f, g \in C[-1, 1]$. Let U be the subspace of $C[-1, 1]$ defined by

$$U = \{f \in C[-1, 1] : f(0) = 0\}$$

(a) Show that $U^\perp = \{0\}$.

(b) Show that 6.49 and 6.52 do not hold without the finite-dimensional hypothesis

[6.49 and 6.52 is "direct sum of a subspace and its orthogonal complement and" and " orthogonal complement of the orthogonal complement"]

Solution

This exercise need to construct function and use the method in analysis.

(a)suppose $g \in U^\perp$ and $\exists x_0 \in (-1,0) \cup (0,1)$ s.t. $g(x_0) = c \neq 0$. Then $\exists U(x_0, \delta) \subset (-1,0) \cup (0,1)$ s.t. $x \in U(x_0, \delta), g(x) > \frac{c}{2}$ (we can assume that $c > 0$), we can construct $f \in U$ as follows:

$$f(x) = \begin{cases} \dfrac{|x - x_0|}{\delta}, x \in U(x_0, \delta) \\ 0, x \in (-1,0) \cup (0,1) \setminus U(x_0, \delta) \end{cases}$$

Therefore we get

$$\langle f, g \rangle > \frac{c\delta}{2} > 0$$

Which means $\forall x \in (-1,0) \cup (0,1), g(x) = 0$, and then for the continuity of g, $g = 0$

(b) you can use the conclusion of (a). $\square$

7 Operators on Inner Product Spaces

7.1 Self-Adjoint and Normal Operators

7.1.1 Adjoints

Definition 7.1.1. *adjoint, T^**

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

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

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

a common technique for computing T^*: start with a formula for $\langle Tv, w \rangle$ then manipulate it to get just v in the first slot; the entry in the second slot will then be T^*w.

Theorem 7.1.1. *adjoint of a linear map is a linear map*

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

Theorem 7.1.2. *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) $(\lambda T)^* = \bar{\lambda} T^* \, for all \lambda \in \mathbb{F}$;

(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 $\mathbb{F}$);

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

(f) if T is invertible, then T^* is invertible and $(T^*)^{-1} = (T^{-1})^*$.

Notice: If $\mathbb{F} = \mathbb{R}$, then the map $T \to T^*$ is a linear map from $\mathcal{L}(V, W)$ to $\mathcal{L}(W, V)$, as follows from (a) and (b) of the result above. However, if $\mathbb{F} = \mathbb{C}$, then this map is not linear because of the complex conjugate that appears in (b).

[It's like the map $v \to \varphi_v$, which let $\varphi_v(u) = \langle u, v \rangle$ is not a linear map]

Theorem 7.1.3. *nullspace and range of T^**

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

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

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

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

(d) $range T = (null T^*)^\perp$.

Definition 7.1.2. *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

$$(A^*)_{j,k} = \overline{A_{k,j}}$$

The adjoint of a linear map **does not depend on a choice of basis**. Thus we frequently emphasize adjoints of linear maps instead of transposes

or conjugate transposes of matrices.

Caution: With respect to nonorthonormal bases, the matrix of T^* does not necessarily equal the conjugate transpose of the matrix of T.

Theorem 7.1.4. *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,*

$$\mathcal{M}(T^*) = (\mathcal{M}(T))^*$$

The Riesz representation theorem as stated in 6.58 provides an identification of V with its **dual space** V' defined in 3.110. Under this identification, the orthogonal complement $U^\perp$ of a subset $U \subset V$ corresponds to the annihilator U^0 of U. If U is a subspace of V, then the formulas for the dimensions of $U^\perp$ and U^0 become identical under this identification ━see 3.125 and 6.51

Suppose $T : V \to W$ is a linear map. Under the identification of V with V' and the identification of W with W', the adjoint map $T^* : W \to V$ corresponds to the dual map $T' : W' \to V'$ defined in 3.118, as Exercise 32 asks you to verify.

Under this identification, the formulas for nullT^* and rangeT^* [7.6(a) and (b)] then become identical to the formulas for nullT' and rangeT' [3.128(a) and 3.130(b)]. Furthermore, the theorem about the matrix of T^* (7.9) is **analogous** to the theorem about the matrix of T' (3.132).

Because orthogonal complements and adjoints are easier to deal with than annihilators and dual maps, there is no need to work with annihilators and dual maps in the context of **inner product spaces**.

7.1.2 Self-Adjoint Operators

Definition 7.1.3. *self-adjoint*

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

A good analogy to keep in mind is that the adjoint on $\mathcal{L}(V)$ plays a role similar to that of the **complex conjugate** on $\mathbb{C}$

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

Theorem 7.1.5. *eigenvalues of self-adjoint operators*

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

This theorem has no value when $\mathbb{F} = \mathbb{R}$

Theorem 7.1.6. Tv *is orthogonal to* v *for all* $v \Leftrightarrow T = 0$ **(assuming** $\mathbb{F} = \mathbb{C}$**)**

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

The next result provides another good example of how self-adjoint operators behave like real numbers

Theorem 7.1.7. $\langle Tv, v \rangle$ *is real for all* $v \Leftrightarrow T$ *is self-adjoint* **(assuming** $\mathbb{F} = \mathbb{C}$**)**

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

Theorem 7.1.8. T **self-adjoint** *and* $\langle Tv, v \rangle = 0$ *for all* $v \Leftrightarrow T = 0$

Suppose T is a self-adjoint operator on V. Then

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

7.1.3 Normal Operators

Definition 7.1.4. *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 $TT^* = T^*T$*

Theorem 7.1.9. *T is normal **if and only if** Tv and T^*v have the same norm*

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

$$T \text{ is normal} \Leftrightarrow \|Tv\| = \|T^*v\| \text{ for every } v \in V$$

Theorem 7.1.10. *range, null space, and eigenvectors of a normal operator*

    Suppose $T \in \mathcal{L}(V)$ is normal. Then
(a) $null\,T = null\,T^*$;
(b) $range\,T = range\,T^*$;
(c) $V = null\,T \oplus range\,T$;
(d) $T - \lambda I$ is normal for every $\lambda \in \mathbb{F}$;
(e) if $v \in V$ and $\lambda \in \mathbb{F}$, then $Tv = \lambda v$ if and only if $T^*v = \overline{\lambda}v$.

Theorem 7.1.11. *orthogonal eigenvectors for normal operators*

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

Theorem 7.1.12. *T is normal $\Leftrightarrow$ the real and imaginary parts of T commute*

    Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then T is normal if and only if there exist commuting self-adjoint operators A and B such that $T = A + iB$

$$A = \frac{T + T^*}{2}, B = \frac{T - T^*}{2i}$$

Exercise

Suppose $\mathbb{F} = \mathbb{R}$.

(a) Show that the set of self-adjoint operators on V is a subspace of $\mathcal{L}(V)$.

(b) What is the dimension of the subspace of $\mathcal{L}(V)$ in(a) [in terms of $\dim V$]?

Solution

(a) just verify it.

(b) We can use the perspective of the matrix.

suppose $\dim V = n$, then $\mathcal{M}(T)$ would be a $n - by - n$ matrix. Since the self-adjoint operator satisfies that

$$\mathcal{M}(T)_{j,k} = \overline{\mathcal{M}(T^*)_{k,j}}$$

In this context, it just need

$$\mathcal{M}(T)_{j,k} = \mathcal{M}(T^*)_{k,j}$$

which means you can choose $\dfrac{n^2 + n}{2}$ entries of the matrix arbitrarily. So the dimension should be $\dfrac{(\dim V + 1)\dim V}{2}$ $\qquad \square$

Exercise

Suppose $T \in \mathcal{L}(V)$ and $\|T^*v\| \leq \|Tv\|$ for every $v \in V$. Prove that T is normal.

This exercise fails on infinite-dimensional inner product spaces, leading to what are called **hyponormal operators**, which have a well-developed theory.

Solution

Remained to be solved. $\qquad \square$

Exercise

Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that the following are equivalent.

(a) P is self-adjoint.

(b) P is normal.

(c) There is a subspace U of V such that $P = P_U$.

This exercise deserves your attention since it connects many interesting things.

Solution

Since $P^2 = P$, we can learn that

$$V = \text{range}P \oplus \text{null}P \tag{1}$$

$(1 \Rightarrow 2)$: If P is self-adjoint, then P is normal because $PP^* = P^*P$.

$(2 \Rightarrow 1)$:

From the property of normal operator, we have $\text{range}P = \text{range}P^*$, $\text{null}P = \text{null}P^*$, so let $u \in \text{range}P$, $\exists v, w \in V$ s.t. $u = Pv, u = P^*w$,

$$Pu = P^2v = Pv = u, P^*u = (P^*)^2w = P^*w = u$$

for $\forall v_0 \in V$, $v_0 = u_0 + w_0, u_0 \in \text{range}P, w_0 \in \text{null}P$

$$Pv_0 = P(u_0 + w_0) = Pu_0 = P^*u_0 = P^*(u_0 + w_0) = P^*v_0$$

Therefore $P = P^*$ and P is self-adjoint

$(2 \Rightarrow 3)$: By the property of adjoint operator and that of normal operator,

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

so there exists $U = \text{range}T$, s.t. $P = P_U$ (by equation (1) and (2))

$(3 \Rightarrow 1/2)$: You can use the language of the orthogonormal basis and the proof is not hard.

Therefore, we've established the equivalence of the three statements.

$\square$

Exercise

Suppose $D : \mathcal{P}_8(\mathbb{R}) \to \mathcal{P}_8(\mathbb{R})$ is the differentiation operator defined by $Dp = p'$. Prove that there does not exist an inner product on $\mathcal{P}_8(\mathbb{R})$ that makes D a normal operator.

Solution

the condition $\mathcal{P}_8(\mathbb{R}) = \text{null}D \oplus \text{range}D$ can't be satisfied. $\qquad\square$

Exercise

Suppose T is a normal operator on V. Suppose also that $v, w \in V$ satisfy the equations

$$\|v\| = \|w\| = 2, Tv = 3v, Tw = 4w$$

Show that $\|T(v + w)\| = 10$.

Solution

Since T is normal, and from the conditions above we learned that w, v are the eigenvectors of T corresponding to different eigenvalues, thus $w \perp v$, therefore $Tw \perp Tv$

$$\|T(v + w)\|^2 = \|Tv\|^2 + \|Tw\|^2 = 100$$

$\qquad\square$

Exercise

Fix $u, x \in V$. Define $T \in \mathcal{L}(V)$ by $Tv = \langle v, u \rangle x$ for every $v \in V$.

(a) Prove that if V is a real vector space, then T is self-adjoint if and only if the list u, x is linearly dependent.

(b) Prove that T is normal if and only if the list u, x is linearly dependent.

Solution

Let $w_1, w_2 \in V$, then we have

$$\langle w_1, T^*w_2 \rangle = \langle Tw_1, w_2 \rangle$$
$$= \langle \langle w_1, u \rangle x, w_2 \rangle$$
$$= \langle w_1, u \rangle \langle x, w_2 \rangle$$
$$= \left\langle w_1, \overline{\langle x, w_2 \rangle}u \right\rangle$$
$$= \langle w_1, \langle w_2, x \rangle u \rangle$$

Therefore, $T^*v = \langle v, x \rangle u$

(a) Suppose T is self-ajoint. Then

$$\langle v, u \rangle x - \langle v, x \rangle u = Tv - T^* v$$

for $\forall v \in V$.

We can assume that $u \neq 0$ (or the proof is trivial), and let $v = u$ thus $\langle v, u \rangle \neq 0$, which implies that u, x are linearly dependent.

Conversely, we can also assume that x, u are non-zero, and $x = cu$, the proof afterward is easy.

(b) Again we can assume that u, x are non-zero.

We have

$$\langle \langle v, u \rangle x, x \rangle u = T^* \left(\langle v, u \rangle x \right)$$
$$= T^* T v$$
$$= T T^* v$$
$$= T \left(\langle v, x \rangle u \right)$$
$$= \langle \langle v, x \rangle u, u \rangle x$$

The operation afterward is similar and the proof from the other side is likewise. □

Exercise

Suppose $T \in \mathcal{L}(V)$ is normal. Prove that

$$\text{null} T^k = \text{null} T \text{ and range} T^k = \text{range} T$$

for every positive integer k.

Solution

First for S self-adjoint: suppose $S^k x = 0$. Then

$$0 = \left\langle S^k x, S^{k-2} x \right\rangle = \left\langle S^{k-1} x, S^{k-1} x \right\rangle$$

so by positive definiteness of inner product, $S^{k-1} x = 0$, and we can continue down to $Sx = 0$.

If T is normal, suppose $T^k x = 0$. Then

$$(T^*T)^k x = (T^*)^k (T^k x) = 0$$

(The key is $(T^*T)^k = (T^*)^k T^k$ since T is normal). So by the first part (since T^*T is self adjoint)

$$0 = \langle T^*Tx, x \rangle = \langle Tx, Tx \rangle$$

so $Tx = 0$.

It's trivial to show that if $x \in V, Tx = 0, T^k x = 0$

so

$$\text{null}T^k = \text{null}T$$

It's also simple to show that

$$\text{range}T^k = \text{range}T$$

$\square$

Exercise

Suppose $T \in \mathcal{L}(V)$ is normal. Prove that if $\lambda \in \mathbb{F}$, then the minimal polynomial of T is not a polynomial multiple of $(x - \lambda)^2$.

Solution

From the conclusion above and the property of normal operator, we have

$$\text{null}(T - \lambda I)^2 = \text{null}(T - \lambda I)$$

however if T is a polynomial multiple of $(x - \lambda)^2$, then it means $\exists v \in V$ s.t. $(T - \lambda I)^2 v = 0$ but $(T - \lambda I)v \neq 0$, which lead to a contradiction.

The conclusion we get from the exercise is amazing since it tells us **a normal operator is diagonalizable!** $\square$

Exercise

Prove or give a counterexample: If $T \in \mathcal{L}(V)$ and there is an orthonormal basis $e_1, \ldots, e_n$ of V such that $\|Te_k\| = \|T^*e_k\|$ for each $k = 1, \ldots, n$, then T is normal

Solution

let

$$Te_k = \begin{cases} e_{k+1}, k \neq n \\ 0, k = n \end{cases}$$

then

$$T^*e_k = \begin{cases} 0, k = 0 \\ e_{k-1}, k \neq 0 \end{cases}$$

Then we find that

$$(T^*T - TT^*)e_1 \neq 0$$

Therefore T is not normal. $\qquad\qquad\square$

Exercise

Fix a positive integer n. In the inner product space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

$$\langle f, g \rangle = \int_{-\pi}^{\pi} fg$$

let

$$V = span(1, \cos x, \cos 2x, \ldots, \cos nx, \sin x, \sin 2x, \ldots, \sin nx)$$

(a) Define $D \in \mathcal{L}(V)$ by $Df = f'$. Show that $D^* = -D$. Conclude that D is normal but not self-adjoint.

(b) Define $T \in \mathcal{L}(V)$ by $Tf = f''$. Show that T is self-adjoint.

Solution

This exercise is simplier with matrices.

From previous exercises we learned that

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \ldots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \ldots, \frac{\sin nx}{\sqrt{\pi}}$$

is an orthogonormal basis.

(a) With respect to this basis, the matrix of D is

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & n \\ 0 & -1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -n & 0 & \cdots & 0 \end{pmatrix}$$

It's not hard to verify that $D^* = -D$ and therefore D is normal but not self-adjoint.

(b) It's similar to (a), except the matrix now become

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & n^2 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & n^2 & 0 & \cdots & 0 \end{pmatrix}$$

But the rigorious solution may not use the matrix.

let $e_j = \dfrac{\sin jx}{\sqrt{\pi}}, f_j = \dfrac{\cos jx}{\sqrt{\pi}}$, then for $\forall v, w \in V$

$$\langle v, D^* w \rangle = \langle Dv, w \rangle$$

$$= \left\langle D \left(\left\langle v, \frac{1}{\sqrt{2\pi}} \right\rangle \frac{1}{\sqrt{2\pi}} + \sum_{j=1}^{n} \left(\langle v, e_j \rangle e_j + \langle v, f_j \rangle f_j \right) \right), w \right\rangle$$

$$= \left\langle \sum_{j=1}^{n} \left(-j \langle v, e_j \rangle f_j + j \langle v, f_j \rangle e_j \right), \left\langle w, \frac{1}{\sqrt{2\pi}} \right\rangle \frac{1}{\sqrt{2\pi}} + \sum_{j=1}^{n} \left(\langle w, e_j \rangle e_j + \langle w, f_j \rangle f_j \right) \right\rangle$$

$$= \sum_{j=1}^{n} \left(-j \langle v, e_j \rangle \langle w, f_j \rangle + j \langle v, f_j \rangle \langle w, e_j \rangle \right)$$

$$= \left\langle \left\langle v, \frac{1}{\sqrt{2\pi}} \right\rangle \frac{1}{\sqrt{2\pi}} + \sum_{j=1}^{n} \left(\langle v, e_j \rangle f_j + \langle v, f_j \rangle e_j \right), \sum_{j=1}^{n} -j \left(\langle w, f_j \rangle e_j + j \langle w, e_j \rangle f_j \right) \right\rangle$$

$$= \langle v, -D^* w \rangle$$

Therefore, we get $D^* = -D$

(b) Note that $T = D^2$, thus $T^* = (DD)^* = D^*D^* = (-D)(-D) = D^2 = T$ $\qquad\qquad\square$

7.2 Spectral Theorem

7.2.1 Real Spectral Theorem

Theorem 7.2.1. *invertible quadratic expressions*

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

$$T^2 + bT + cI$$

is an invertible operator

Theorem 7.2.2. *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 $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$.

Theorem 7.2.3. *real spectral theorem*

Suppose $\mathbb{F} = \mathbb{R}$ and $T \in \mathcal{L}(V)$. Then the following are equivalent.

(a) T is self-adjoint.

(b) T has a diagonal matrix with respect to some orthonormal basis of V.

(c) V has an orthonormal basis consisting of eigenvectors of T.

7.2.2 Complex Spectral Theorem

Theorem 7.2.4. *complex spectral theorem*

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then the following are equivalent.

(a) T is normal.

(b) T has a diagonal matrix with respect to some orthonormal basis of V.

(c) V has an orthonormal basis consisting of eigenvectors of T.

Notice that the proof use

$$\|Te_1\|^2 = |a_{1,1}|^2,$$

$$\|T^*e_1\|^2 = |a_{1,1}|^2 + |a_{1,2}|^2 + \cdots + |a_{1,n}|^2$$

and

$$\|Te_1\|^2 = \|T^*e_1\|^2$$

Exercise

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Prove that T is normal if and only if there exists a polynomial $p \in \mathcal{P}(\mathbb{C})$ such that $T^* = p(T)$.

Solution

$\Rightarrow$:

Assume that T is normal, i.e., $TT^* = T^*T$.Since T is normal, it is diagonalizable. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be its distinct eigenvalues.

By Lagrange interpolation, we can find a polynomial

$$p(z) = a_0 + a_1 z + \cdots + a_k z^k$$

such that $p(\lambda_j) = \overline{\lambda_j}$ for $j = 1, 2, \ldots, n$.

So we can define

$$p(T) = a_0 I + a_1 T + \cdots + z_k T^k$$

then

$$p(T)v = p(\lambda_i)v = \overline{\lambda_i}v$$

and v is any eigenvector of T.

So the matrix of $p(T)$ is the conjugate transpose of T, which can be deduced from the fact that they are diagonal with respect to the same basis and the eigenvalues are the conjugate of the others.

$\Leftarrow$:

Trivial, since $p(T)$ commute with T.

$\square$

Exercise

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ is normal. Prove that if S is an operator on V that commutes with T, then S commutes with T^*.

Solution

This can be proved using the conclusion of last exercise.

Since T^* can be expressed like $p(T)$. So if S commute with T, then it commute with $T^* = p(T)$ □

Exercise

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

(a) Prove that $U^\perp$ is invariant under T.

(b) Prove that $T|_U \in \mathcal{L}(U)$ is self-adjoint.

(c) Prove that $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ is self-adjoint.

Solution

(a) for every $u \in U, w \in U^\perp$

$$\langle Tu, w \rangle = \langle u, Tw \rangle = 0$$

Therefore $Tw \in U^\perp$, so $U^\perp$ is invariant under T

(b) We know that T is self-adjoint. Now let's consider the restriction of T to U, denoted as $T|_U$. We want to show that $T|_U$ is also self-adjoint.

For any vectors $u, v \in U$, we have:

$$\langle T|_U(u), v \rangle = \langle T(u), v \rangle = \langle u, T(v) \rangle = \langle u, T|_U(v) \rangle$$

(c) similar.

□

Exercise

Suppose $T \in \mathcal{L}(V)$ is normal and U is a subspace of V that is invariant under T. [Maybe it need the condition that $\mathbb{F} = \mathbb{C}$?]

(a) Prove that $U^\perp$ is invariant under T.

(b) Prove that U is invariant under T^*.

(c) Prove that $(T|_U)^* = (T^*)|_U$.

(d) Prove that $T|_U \in \mathcal{L}(U)$ and $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ are normal operators.

This exercise can be used to give yet another proof of the complex spectral theorem (use induction on $\dim V$ and the result that T has an eigenvector)

Solution

There's no solution yet.

$\square$

7.3   Positive Operators

Definition 7.3.1. *positive operator*

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

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

for all $v \in V$.

If V is a **complex vector space**, then the requirement that T be self-adjoint can be dropped from the definition above (by 7.14).

Definition 7.3.2. *square root*

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

Theorem 7.3.1. *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 $R \in \mathcal{L}(V)$.

Theorem 7.3.2. *each positive operator has only one **positive** square root*

Every positive operator on V has a unique positive square root.

108

A positive operator can have infinitely many square roots (although only one of them can be positive). For example, the identity operator on V has infinitely many square roots if $\dim V > 1$. (Just imagine rotating around an arbitrarily chosen axis)

Definition 7.3.3. *notation:* $\sqrt{T}$

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

Theorem 7.3.3. *T positive and* $\langle Tv, v \rangle = 0 \Rightarrow Tv = 0$

Suppose T is a positive operator on V and $v \in V$ is such that $\langle Tv, v \rangle = 0$. Then $Tv = 0$

Exercise

Suppose $T \in \mathcal{L}(V)$ and $e_1, \ldots, e_n$ is an orthonormal basis of V. Prove that T is a positive operator if and only if there exist $v_1, \ldots, v_n \in V$ such that

$$\langle Te_k, e_j \rangle = \langle v_k, v_j \rangle$$

forall $j, k = 1, \ldots, n$.

The numbers $\{\langle Te_k, e_j \rangle\} j, k = 1, \ldots, n$ are the entries in the matrix of T with respect to the orthonormal basis $e_1, \ldots, e_n$.

Solution

Since T is a positive operator, by theorem we know there exists a unique positive square root $\sqrt{T}$. Let $v_j = \sqrt{T}e_j$, which satisfy that

$$\langle Te_k, e_j \rangle = \left\langle \sqrt{T}e_k, \sqrt{T}e_j \right\rangle = \langle v_k, v_j \rangle$$

thus we finished our proof. $\qquad\qquad\square$

Exercise

Suppose n is a positive integer. The $n-by-n$ Hilbert matrix is the $n-by-n$ matrix whose entry in row j,column k is $\frac{1}{j+k-1}$. Suppose $T \in \mathcal{L}(V)$ is an operator whose matrix with respect to some orthonormal basis of V is the $n-by-n$ Hilbert matrix. Prove that T is a positive invertible operator.

Example: The $4 - by - 4$ Hilbert matrix

$$
\begin{pmatrix}
1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\
\frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\
\frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\
\frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7}
\end{pmatrix}
$$

Solution

Given an orthonormal basis $\{e_1, e_2, \ldots, e_n\}$, we can express any vector $v \in V$ as $v = \sum_{i=1}^{n} v_i e_i$. Then, the inner product $\langle Tv, v \rangle$ becomes:

$$
\begin{aligned}
\langle Tv, v \rangle &= \left\langle \sum_{j=1}^{n} v_j \sum_{k=1}^{n} \frac{1}{j+k-1} e_k, \sum_{i=1}^{n} v_i e_i \right\rangle \\
&= \left\langle \sum_{k=1}^{n} \sum_{j=1}^{n} \frac{1}{j+k-1} v_j e_k, \sum_{i=1}^{n} v_i e_i \right\rangle \\
&= \sum_{k=1}^{n} \left\langle \sum_{j=1}^{n} \frac{1}{j+k-1} v_j e_k, v_k e_k \right\rangle \\
&= \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{1}{j+k-1} v_j \overline{v_k} \\
&= \int_0^1 \sum_{j=1}^{n} \sum_{k=1}^{n} v_j \overline{v_k} x^{j-1} x^{k-1} \, dx \\
&= \int_0^1 (v_1 + v_2 x + \cdots v_n x^{n-1})(\overline{v_1} + \overline{v_2} x + \cdots + \overline{v_n} x^{n-1}) \, dx \\
&= \int_0^1 (v_1 + v_2 x + \cdots v_n x^{n-1})\overline{(v_1 + v_2 x + \cdots + v_n x^{n-1})} \, dx \\
&= \int_0^1 \left| (v_1 + v_2 x + \cdots v_n x^{n-1}) \right|^2 \, dx \geq 0
\end{aligned}
$$

The following part is trivial compared to the trick of transformtion above.

https://math.stackexchange.com/questions/4309056/how-to-show-a-hilbert-matrix-is-invertible

https://math.stackexchange.com/questions/1754729/prove-the-positive-definiteness-of-hilbert-matrix

The link above is some solution but is a little irrelevant to the content of this book. I prefer the first one, which also tells how to prove positiveness □

Exercise

For $T \in \mathcal{L}(V)$ and $u, v \in V$, define $\langle u, v \rangle_T$ by $\langle u, v \rangle_T = \langle Tu, v \rangle$.

(a) Suppose $T \in \mathcal{L}(V)$. Prove that $\langle \cdot, \cdot \rangle_T$ is an inner product on V if and only if T is an invertible positive operator (with respect to the original inner product $\langle \cdot, \cdot \rangle$).

(b) Prove that every inner product on V is of the form $\langle \cdot, \cdot \rangle_T$ for some positive invertible operator $T \in \mathcal{L}(V)$.

Solution

(a) is simple, you can verify it yourself.

(b) Now we just prove the finite-dimensional one.

Suppose $v_1, \ldots, v_n$ is a basis of V, thinking of a series of linear functionals $\varphi_i(u) = \langle u, v_1 \rangle'$ from V to $\mathbb{F}$, for any $u \in V$.

By Riesz represents theorem, there exists w_i s.t.

$$\langle u, v_i \rangle' = \langle u, w_i \rangle$$

And by linear map lemma, there exists linear map s.t.

$$Tv_i = w_i$$

So

$$\langle u, v_i \rangle' = \langle u, Tv_i \rangle = \langle u, v_i \rangle_{T^*}$$

from which we can imply that

$$\langle u, v \rangle' = \langle u, v \rangle_{T^*}, \forall u, v \in V$$

And because $\langle \cdot, \cdot \rangle'$ is an inner product, then $\langle \cdot, \cdot \rangle_{T^*}$ is an inner product, thus T is an invertible positive operator.

I'm not sure with my solution, please check it carefully for any possible mistake. □