

Linear Algebra Done Right

Haiweng Xu

2024.1.12



1	Vector Spaces	4
1.1	\mathbb{R}^n and \mathbb{C}^n	4
1.1.1	Complex Numbers	4
1.1.2	Lists	5
1.1.3	\mathbb{F}^n	5
1.2	Definition of Vector Space	5
1.3	Subspaces	7
1.3.1	Sum of Spaces	8
1.3.2	Direct Sums	8
2	Finite-Dimensional Vector Spaces	11
2.1	Span and Linear Independence	11
2.2	Bases	12
2.3	Dimensions	13
3	Linear Maps	14
3.1	Vector Space of Linear Maps	14
3.2	Null Spaces and Ranges	21
3.3	Matrices	24
3.3.1	Representing a Linear Map by a Matrix	24
3.3.2	Addition and Scalar Multiplication of Matrices	25
3.3.3	Matrix Multiplication	26
3.3.4	Column-Row Factorization and Rank of a Matrix	27

3.4	Invertibility and Isomorphisms	28
3.4.1	Invertible Linear Maps	28
3.4.2	Isomorphic Vector Spaces	30
3.4.3	Linear Maps Thought of as Matrix Multiplication . .	31
3.4.4	Change of Basis	32
3.5	Products and Quotients of Vector Spaces	36
3.5.1	Products of Vector Spaces	36
3.5.2	Quotient Spaces	37
3.6	Duality	45
3.6.1	Dual Space and Dual Map	45
3.6.2	Null Space and Range of Dual of Linear Map	46
3.6.3	Matrix of Dual of Linear Map	47
4	Polynomials	51
5	Eigenvalues and Eigenvectors	52
5.1	Invariant Subspaces	52
5.1.1	Eigenvalues	52
5.1.2	Polynomials Applied to Operators	53
5.2	The Minimal Polynomial	57
5.2.1	Existence of Eigenvalues on Complex Vector Spaces .	57
5.2.2	Eigenvalues and the Minimal Polynomial	58
5.2.3	Eigenvalues on Odd-Dimensional Real Vector Spaces .	59
5.3	Upper-Triangular Matrices	63
5.4	Diagonalizable Operators	65
5.4.1	Diagonal Matrices	65
5.4.2	Conditions for Diagonalizability	66
5.4.3	Gershgorin Disk Theorem	67
5.5	Commuting Operators	72
6	Inner Product Spaces	74
6.1	Inner Products and Norms	74
6.1.1	Inner Products	74
6.1.2	Norms	75

6.2	Orthonormal Bases	81
6.2.1	Orthonormal Lists and the Gram-Schmidt Procedure	81
6.2.2	Linear Functionals on Inner Product Spaces	83
6.3	Orthogonal Complements and Minimization Problems	89
6.3.1	Orthogonal Complements	89
6.3.2	Minimization Problems	91
6.3.3	Pseudoinverse	91
7	Operators on Inner Product Spaces	95
7.1	Self-Adjoint and Normal Operators	95
7.1.1	Adjoint	95
7.1.2	Self-Adjoint Operators	97
7.1.3	Normal Operators	98
7.2	Spectral Theorem	106
7.2.1	Real Spectral Theorem	106
7.2.2	Complex Spectral Theorem	107
7.3	Positive Operators	109
7.4	Isometries, Unitary Operators, and Matrix Factorization	113
7.4.1	Isometries	113
7.4.2	Unitary Operators	114
7.4.3	QR Factorization	115
7.4.4	Cholesky Factorization	116
7.5	Singular Value Decomposition	119
7.5.1	Singular Values	119
7.5.2	SVD for Linear Maps and for Matrices	120
7.6	Consequences of Singular Value Decomposition	122
7.6.1	Norms of Linear Maps	122
7.6.2	Approximation by Linear Maps with Lower-Dimensional Range	123
7.6.3	Polar Decomposition	123
7.6.4	Operators Applied to Ellipsoids and Parallelepipeds	124
7.6.5	Volume via Singular Values	125
7.6.6	Properties of an Operator as Determined by Its Eigen- values	126

8	Operators on Complex Vector Spaces	129
8.1	Generalized Eigenvectors and Nilpotent Operators	129
8.1.1	Null Spaces of Powers of an Operator	129
8.1.2	Generalized Eigenvectors	130
8.1.3	Nilpotent Operators	131
8.2	Generalized Eigenspace Decomposition	133
8.2.1	Generalized Eigenspaces	133
8.2.2	Multiplicity of an Eigenvalue	134
8.2.3	Block Diagonal Matrices	135
8.3	Consequences of Generalized Eigenspace Decomposition . . .	138
8.3.1	Square Roots of Operators	138
8.3.2	Jordan Form	139
8.4	Trace: A Connection Between Matrices and Operators	140
9	Multilinear Algebra and Determinants	142
9.1	Bilinear Forms and Quadratic Forms	142
9.1.1	Bilinear Forms	142
9.1.2	Symmetric Bilinear Forms	143
9.1.3	Quadratic Forms	145
9.2	Alternating Multilinear Forms	146
9.2.1	Multilinear Forms	146
9.2.2	Alternating Multilinear Forms and Permutations . . .	147
9.3	Determinants	149
9.4	Tensor Products	149

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 abstract objects).

Two lists are equal if and only if they have the same length and the same elements 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} :

$$\mathbb{F}^n = (x_1, \dots, x_n) : x_k \in \mathbb{F} \text{ for } k = 1, \dots, n.$$

For $(x_1, \dots, x_n) \in \mathbb{F}^n$ and $k \in 1, \dots, n$, we say that x_k is the k th coordinate of (x_1, \dots, 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}^∞ is defined to be the set of all **sequences** of elements of \mathbb{F} :

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

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

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

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

With these definitions, \mathbb{F}^∞ 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 $f, g \in \mathbb{F}^S$, 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

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

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}^∞ is the special case of \mathbb{F}^S , in the first one, $S = \{1, 2, \dots, n\}$, and in the latter one, $S = \{1, 2, \dots\}$

Some interesting facts in Exercises

Exercise

1. The empty set 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 identity. 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 $\bar{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$, and for $t \in \mathbb{R}$ 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) The set 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}^∞ .

1.3.1 Sum of Spaces

Definition 1.3.1. Suppose V_1, \dots, V_m are subspaces of V . The sum of V_1, \dots, V_m , denoted by $V_1 + \dots + V_m$, is the set of all possible sums of elements of V_1, \dots, V_m .

More precisely,

$$V_1 + \dots + V_m = \{ \mathbf{v}_1 + \dots + \mathbf{v}_m : \mathbf{v}_1 \in V_1, \dots, \mathbf{v}_m \in V_m \}$$

Theorem 1.3.1. sum of subspaces is the smallest containing subspace

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

The proof contains three parts: (1) subspace, (2) containing V_1, \dots, V_m , (3) Smallest

The third one is hard to think, but you only need to prove that “Every subspace containing V_1, \dots, V_m contains $V_1 + \dots + V_m$ ”

1.3.2 Direct Sums

Definition 1.3.2. definition: direct sum, \oplus

Suppose V_1, \dots, 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 $\mathbf{v}_1 + \cdots + \mathbf{v}_m$, where each $\mathbf{v}_k \in V_k$.

□ 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 subspaces is direct sum?

Here there are two theorems:

Theorem 1.3.2. *condition for a direct sum*

Suppose V_1, \dots, 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 $\mathbf{0}$ as a sum $\mathbf{v}_1 + \cdots + \mathbf{v}_m$, where each $\mathbf{v}_k \in V_k$, is by taking each \mathbf{v}_k equal to $\mathbf{0}$.

This theorem is the most general, 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

$$U + W \text{ is a direct sum} \Leftrightarrow U \cap W = \{\mathbf{0}\}$$

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:

$$V_1 \cap V_2 = \{\mathbf{0}\}, V_1 \cap V_3 = \{\mathbf{0}\}, V_2 \cap V_3 = \{\mathbf{0}\}$$

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

Exercise

2. A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is called 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 exercise 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!

\square

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.*

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

Definition 2.1.2. *polynomial $\mathcal{P}(\mathbb{F})$*

□ *A function $p : \mathbb{F} \rightarrow \mathbb{F}$ is called a polynomial with coefficients in \mathbb{F} if there exist $a_0, \dots, a_m \in \mathbb{F}$ such that for all $z \in \mathbb{F}$.*

$$p(z) = a_0 + a_1z + a_2z^2 + \dots + a_mz^m$$

□ *$\mathcal{P}(\mathbb{F})$ 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 $a_0, a_1, \dots, a_m \in \mathbb{F}$ with $a_m \neq 0$ such that for every $z \in \mathbb{F}$, we have

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

□ 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

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

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 $S_1 \cup S_2$ is the smallest subset of S containing S_1 and S_2	for subspaces V_1, V_2 of V , the sum $V_1 + V_2$ is the smallest subspace of V containing V_1 and V_2
$\#(S_1 \cup S_2) = \#S_1 + \#S_2 - \#(S_1 \cap S_2)$	$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)$
$\#(S_1 \cup S_2) = \#S_1 + \#S_2 \Leftrightarrow S_1 \cap S_2 = \emptyset$	$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 \Leftrightarrow V_1 \cap V_2 = \{\mathbf{0}\}$
$S_1 \cup \cdots \cup S_m$ is a disjoint union $\Leftrightarrow \#(S_1 \cup \cdots \cup S_m) = \#S_1 + \cdots + \#S_m$	$V_1 + \cdots + V_m$ 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, \dots, v_m is linearly independent in V and $w \in V$. Prove that

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

Solution

There are mainly two solutions.

(1) It's obvious that

$$v_2 - v_1, \dots, v_m - v_1 \in \text{Span}(v_1 + w, \dots, v_m + w)$$

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

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

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

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

and

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

therefore

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

□

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 $T : V \rightarrow W$ with the following properties.

additivity

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

homogeneity

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

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 to the additive identity of another (or possibly the same) vector space. To be specific, $0 \in \mathcal{L}(V, W)$ is defined by

$$0\mathbf{v} = \mathbf{0}$$

The 0 on the left side of the equation above is a function from V to W , whereas the $\mathbf{0}$ 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

$$Iv = v$$

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

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

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

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

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

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

$$\begin{aligned} T(p_1 + p_2) &= (a_{10} + a_{20}) + (a_{11} + a_{21})q(x) + \dots + (a_{1n} + a_{2n})q(x)^n \\ &= a_{10} + a_{20} + a_{11}q(x) + a_{21}q(x) + \dots + a_{1n}q(x)^n + a_{2n}q(x)^n \\ &= T(p_1) + T(p_2) \end{aligned}$$

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

$$Tv_k = w_k$$

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

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 $j = 1, \dots, m$ and $k = 1, \dots, n$ such that

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

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

Solution

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

For example,

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

and so on.

□

Exercise

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

$$\varphi(z) = \operatorname{Re} z$$

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 $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ 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>

□

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, \dots, 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

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

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

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

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 swapping two base vector or make a loop of n vectors) \square

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 solve 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, \dots, v_n is a basis of V

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

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 contradiction! \square

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

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, \dots, v_n is linearly independent in V . Prove that Tv_1, \dots, Tv_n is linearly independent in W .

A obvious facts, prove with linear independence by yourself.

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

This simple fact is significantly useful in many exercises.

any $v \in V$ can be expressed with

$$v = c_1v_1 + \cdots + c_nv_n$$

then

$$Tv = c_1Tv_1 + \cdots + c_nTv_n$$

therefore $\text{range}T \in \text{Span}(Tv_1, \dots, Tv_n)$

On the other hand $Tv_i \in \text{range}T$, so by definition, Tv_1, \dots, Tv_n spans $\text{range}T$

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

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

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

□

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

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

Meanwhile

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

Hence $\text{Span}(1, x, x^2, \dots) \subset \text{range } D$, namely $\text{range } D = \mathcal{P}(\mathbb{R})$

□

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

□

Exercise

4. Suppose V is finite-dimensional with $\dim V > 1$. Show that if $\varphi : \mathcal{L}(V) \rightarrow \mathbb{F}$ 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 $\varphi(I) = 1$ or 0

Suppose $\varphi(I) = 1$, and v_1, \dots, 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 $S_i^2 = S_i$, $\varphi(S_i) = 0$ or 1 .

$\sum_{i=1}^n S_i = (2 - n)I$, and since φ should be linear,

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

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

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

*If the bases v_1, \dots, v_n and w_1, \dots, w_m are not **clear from the context**, then the notation $\mathcal{M}(T, (v_1, \dots, v_n), (w_1, \dots, w_m))$ 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 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 addition 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 two 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 , 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: $A_{j,\dots}$, $A \cdots , k$*

Suppose A is an m -by- n matrix.

\square *If $1 \leq j \leq m$, then $A_{j,\cdot}$ denotes the 1-by- n matrix consisting of row j of A .*

□ If $1 \leq k \leq n$, 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 $\dim \text{range } T = 1$ 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, \dots, u_n is a basis of V , and from the exercise, we know $\dim \text{null } T = n - 1$, so assume that $Tu_1 \neq 0, Tu_i = 0, i = 2, \dots, n$

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

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

Besides, Tv_1 spans $\text{span } T$, so we can extend it to a basis of W , i.e. $Tv_1, \dots, w_2, \dots, w_n$

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

□

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 **doesn'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}^∞ to \mathbb{F}^∞ is surjective (See 3A) but it is not invertible because it is not injective [the vector $(1, 0, 0, 0, \dots)$ 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, \dots, v_n is a basis of V and w_1, \dots, 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, \dots, 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, \dots, b_n are the scalars such that

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

The matrix $\mathcal{M}(v)$ of a vector $v \in V$ depends on the basis v_1, \dots, 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 **relabeling**. 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, \dots, 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. $\mathcal{M}(T)_{\cdot, k} = \mathcal{M}(T v_k)$

Suppose $T \in \mathcal{L}(V, W)$ and v_1, \dots, v_n is a basis of V and w_1, \dots, w_m is a basis of W . Let $1 \leq k \leq n$. Then the k th column of $\mathcal{M}(T)$, which is

denoted by $\mathcal{M}(T)_{.,k}$, 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, \dots, v_n is a basis of V and w_1, \dots, w_m is a basis of W . Then $\mathcal{M}(Tv) = \mathcal{M}(T)\mathcal{M}(v)$.

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 $\dim \text{range } T$ 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, \dots, v_n is a basis of V , then the notation $\mathcal{M}(T, (v_1, \dots, v_n))$ is defined by the equation

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

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, \dots, u_m is a basis of U , v_1, \dots, v_n is a basis of V , and w_1, \dots, w_p is a basis of W , then

$$\begin{aligned} \mathcal{M}(ST, (u_1, \dots, u_m), (w_1, \dots, w_p)) &= \mathcal{M}(S, (v_1, \dots, v_n), (w_1, \dots, w_p)) \\ &\quad \cdot \mathcal{M}(T, (u_1, \dots, u_m), (v_1, \dots, v_n)) \end{aligned}$$

Theorem 3.4.11. *matrix of identity operator with respect to two bases*

Suppose that u_1, \dots, u_n and v_1, \dots, v_n are bases of V . Then the matrices $\mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$ and $\mathcal{M}(I, (v_1, \dots, v_n), (u_1, \dots, u_n))$ 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, \dots, u_n and v_1, \dots, v_n are bases of V . Let $A = \mathcal{M}(T, (u_1, \dots, u_n))$ and $B = \mathcal{M}(T, (v_1, \dots, v_n))$ and $C = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$. Then

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

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

Suppose that v_1, \dots, 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, \dots, 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 is invertible $\Leftrightarrow S$ and T 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, \dots, v_n is a basis of V and $Sv_i = 0, i = 1, \dots, k$

$\text{null } A$ 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.

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

Therefore T must have

$$Tv_i = c_1v_1 + \dots + c_kv_k$$

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

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

So

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

(b)

Use the formula

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

□

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

□

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 $T : \mathcal{P}_m \rightarrow \mathcal{P}_m$ have the property that $q = Tp$

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

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

□

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 invertible, then ST, U are both invertible, and also S, T are invertible.

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

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, \dots, V_m are vector spaces over \mathbb{F} .

\square The product $V_1 \times \dots \times V_m$ is defined by

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

\square Addition on $V_1 \times \dots \times V_m$ is defined by

$$(u_1, \dots, u_m) + (v_1, \dots, v_m) = (u_1 + v_1, \dots, u_m + v_m)$$

\square Scalar multiplication on $V_1 \times \dots \times V_m$ is defined by

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

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

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

Theorem 3.5.2. *dimension of a product is the sum of dimensions*

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

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

Theorem 3.5.3. *products and direct sums*

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

$$\Gamma(v_1, \dots, v_m) = v_1 + \dots + v_m$$

Obviously, Γ is surjective.

Then $V_1 + \dots + V_m$ 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, \dots, V_m are subspaces of V . Then $V_1 + \dots + V_m$ is a direct sum if and only if $\dim(V_1 + \dots + V_m) = \dim V_1 + \dots + \dim V_m$.

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

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

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 from what we have seen, since the vectors in the space can have different expressions, so the definition make 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 $\tilde{T} : V/(\text{null}T) \rightarrow W$ 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, \dots, V_m are vector spaces. Prove that $\mathcal{L}(V_1 \times \dots \times V_m, W)$ and $\mathcal{L}(V_1, W) \times \dots \times \mathcal{L}(V_m, W)$ 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 an invertible linear map connecting the two vector spaces.

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

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

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

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

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

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

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

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

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

□

Exercise

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

Solution

Since the theorem in the section mostly focuses 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 $W \subset U$ □

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 affine 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 $A_1 \cap A_2$ is either a translate of some subspace of V or is the empty set.

Solution

Of course it can be empty

Suppose $A_1 \cap A_2 \neq \emptyset$, then $\forall x, y \in A_1 \cap A_2$, 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

$$x + (1 - \lambda)y \in A_1 \cap A_2$$

And also for the conclusion of last exercise, $A_1 \cap A_2$ is a affine set, i.e. the translate of some vector space. \square

Exercise

Suppose $v_1, \dots, v_m \in V$. Let

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

(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 $\{v_1, \dots, v_m\} \subset B$, 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 $U = \{(x_1, x_2, \dots) \in \mathbb{F}^\infty : x_k \neq 0 \text{ for only finitely many } k\}$.

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

(b) Prove that \mathbb{F}^∞/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 $m = 1, 2, \dots$ in \mathbb{F}^∞ defined by the rule:

$$e_m(p) = \begin{cases} 1, & (p-1) \equiv 0 \pmod{m} \\ 0, & \text{otherwise} \end{cases}$$

For example

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

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

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

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

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

Note that $k_1(e_1 + U) + \dots + k_m(e_m + U) = k_1e_1 + \dots + k_me_m + U$, so the equation above implies that $k_1e_1 + \dots + k_me_m = u$ 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 $e_1(p) = \dots = e_m(p) = 1$, and $u(q) = 0$ for any $q > p$

Note that $e_1(p+1) = 1, e_2(p+1) = \dots = e_m(p+1) = 0$, so $k_1e_1(p+1) + \dots + k_me_m(p+1) = k_1$, and $u(p+1) = 0$. So $k_1 = 0$. Then we can show that $k_1e_1(p+2) + \dots + k_me_m(p+2) = k_2 = 0$. Following the induction, we can show that $k_1 = \dots = k_m = 0$

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

Here is another construct.

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

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

Then $e_1 + U, \dots, e_{n+1} + U$ must be linearly-dependent. So there exists k_1, \dots, k_{n+1} which are not all zero, s.t. $k_1e_1 + \dots + k_{n+1}e_{n+1} \in U$ which means only have finite nonzero entry.

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

$$k_1r^0 + k_2r^1 + \dots + k_{n+1}r^n = 0$$

So the polynomial

$$f(x) = k_{n+1}x^n + \dots + k_2x + k_1$$

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

□

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, \dots, v_n is a basis of V , then the dual basis of v_1, \dots, v_n is the list $\varphi_1, \dots, \varphi_n$ 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, \dots, v_n is a basis of V and $\varphi_1, \dots, \varphi_n$ 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 $S \in \mathcal{L}(V, W)$;

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

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

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) $\dim \text{range}T' = \dim \text{range}T$;

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

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

Solution

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

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

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

with x as the i th component.

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

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

Then we can check φ is a isomorphism.

injectivity:

suppose $\varphi(f) = (P_1'f, \dots, P_m'f) = (0, \dots, 0)$, then $P_i'f = 0$ ($i = 1, \dots, m$),

assume $f = (x_1, \dots, x_m)$, then $x_i = 0$ ($i = 1, \dots, m$), which means $f = 0$

surjectivity:

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

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

then $\exists f$ s.t. $\varphi f = (f_1, \dots, f_m)$

□

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

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

□

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 $\varphi_1, \dots, \varphi_n$ is a basis of V' . Show that there exists a basis of V whose dual basis is $\varphi_1, \dots, \varphi_n$.

Solution

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

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

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

□

Exercise

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

$$\dim((\text{null}\varphi_1) \cap \dots \cap (\text{null}\varphi_m)) = (\dim V) - m$$

Solution

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

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

$$\text{null}\varphi_j = \{v \in V : v = c_1v_1 + \dots + c_{j-1}v_{j-1} + \dots + c_{j+1}v_{j+1} + \dots + c_nv_n\}$$

$$((\text{null}\varphi_1) \cap \dots \cap (\text{null}\varphi_m)) = \{v \in V : v = c_{m+1}v_{m+1} + \dots + c_nv_n\}$$

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

□

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 $\Lambda : V \rightarrow V''$ 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

$$\begin{aligned} (T''(\Lambda v))f &= (\Lambda v)(T'f) \\ &= (T'f)v \\ &= f(Tv) \end{aligned}$$

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 $c, \lambda_1, \dots, \lambda_m \in \mathbb{C}$

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 $Tu \in V$.*

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

$\text{range}T$ *The subspace $\text{range}T$ is invariant under T because if $u \in \text{range}T$, then $Tu \in \text{range}T$.*

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 $T^m = \underbrace{T \cdots T}_{m \text{ times}}$

□ T^0 is defined to be the identity operator I on V .

□ If T is invertible with inverse T^{-1} , then $T^{-m} \in \mathcal{L}(V)$ is defined by

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

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

$$p(z) = a_0 + a_1z + a_2z^2 + \cdots + a_mz^m$$

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

$$p(T) = a_0I + a_1T + a_2T^2 + \cdots + a_mT^m$$

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 for all $\lambda \in \mathbb{F}$ such that $0 < |\alpha - \lambda| < \delta$.

Solution

Since V is finite-dimensional, the number of eigenvalues of T is at most $\dim V$, no matter whether α is the eigenvalue or not, let $\delta = \min_{1 \leq j \leq \dim V} |\lambda_j - \alpha|$ and this δ satisfy the condition. \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 eigenvector 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. $\varphi(v) \neq 0$, $\text{range}(T - \lambda I) \neq V$, which implies $T - \lambda I$ is not surjective thus λ is a eigenvalue of T'

\Leftarrow

suppose the eigenvector 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* $\varphi \in V', \exists v \in V, s.t. \varphi v \neq 0$, so $\text{range}(T - \lambda I) \neq V$ which implies that λ is also an eigenvalue of T \square

Exercise

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

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

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, \dots, v_n .

Solution

Look up the OneNote for solution. \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 λ is an eigenvalue of T_C .

Solution

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

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

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. \square

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 . \square

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 $T - I$ or $T + I$ is not invertible and respectively indicate the eigenvalues to be 1 or -1 . Therefore, $T^2 - I$ is invertible and $T^2 + I = 0$ \square

Exercise

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

Solution

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

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

$\lambda_1, \dots, \lambda_n$ are the distinct eigenvalues of D and therefore $e^{\lambda_1 x}, \dots, e^{\lambda_n x}$ 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, \dots, v_n is a basis of V , let the eigenvalue be λ . $T - \lambda I$ is not surjective, so $\exists v_j$, s.t.

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

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

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

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

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

Therefore

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

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

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 cannot 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 vector space, 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 $\dim \text{null}(T^2 + bT + cI)$ 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

$$p(z) = a_0 + a_1z + \cdots + a_{m-1}z^{m-1} + z^m$$

we denote the eigenvalues of T with $\lambda_1, \dots, \lambda_m$. 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, \dots, 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

□

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

□

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, \dots, v_n be a basis of V and we can use induction to prove this fact

For the base case,

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

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

Suppose for $p = k$

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

then for $p = k + 1$

$$\begin{aligned} T^{k+1} v_j &= c_{1,j}(c_{1,1}v_1 + \dots + c_{n,1}v_n) + \dots + c_{n,j}(c_{1,n}v_1 + \dots + c_{n,n}v_n) \\ &= \left(\sum_{i=1}^n c_{i,j}c_{1,i} \right) v_1 + \dots + \left(\sum_{i=1}^n c_{i,j}c_{n,i} \right) v_n \quad (j = 1, 2, \dots, n) \end{aligned}$$

and have all coefficients belongs to rational numbers.

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

Exercise

Suppose $a_0, \dots, a_{n-1} \in \mathbb{F}$. 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

$$a_0 + a_1z + \cdots + a_{n-1}z^{n-1} + z^n$$

Solution

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

$$Te_1 = e_2, Te_2 = e_3, \dots, Te_{n-1} = e_n$$

Now note that

$$T^n e_1 = Te_n = -a_0 e_1 - a_1 e_2 - \cdots - a_{n-1} e_n = -a_0 e_1 - a_1 Te_1 - \cdots - a_{n-1} T^{n-1} e_1$$

which means

$$p(T)e_1 = 0$$

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

And then we get the solution we want. □

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

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

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. □

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, \dots, v_k be a basis of U and extend it to v_1, \dots, v_n , a basis of V , let λ be an eigenvalue of $T|_U$, 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.

□

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 eigenvalues, but the proof of this is exceedingly trivial 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, \dots, 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 $\mathcal{M}(T, (v_1, \dots, v_n))$ 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-triangular matrix*

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, \dots, v_n is a basis of V . Then the following are equivalent.

- (a) The matrix of T with respect to v_1, \dots, v_n is upper triangular.
- (b) $\text{Span}(v_1, \dots, v_k)$ is invariant under T for each $k = 1, \dots, n$.
- (c) $Tv_k \in \text{Span}(v_1, \dots, v_k)$ for each $k = 1, \dots, n$

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 $\lambda_1, \dots, \lambda_n$. 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 $\lambda_1, \dots, \lambda_m \in \mathbb{F}$.

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

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

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 $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T . Then

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

is a direct sum. Furthermore, if V is finite-dimensional, then

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

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 $\lambda_1, \dots, \lambda_m$ 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) $V = E(\lambda_1, T) \oplus \dots \oplus E(\lambda_m, T)$.
- (d) $\dim V = \dim E(\lambda_1, T) + \dots + \dim E(\lambda_m, T)$

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 $\lambda_1, \dots, \lambda_m \in \mathbb{F}$.

5.4.3 Gershgorin Disk Theorem

Definition 5.4.4. *Gershgorin disks*

Suppose $T \in \mathcal{L}(V)$ and v_1, \dots, v_n is a basis of V . Let A denote the matrix of T with respect to this basis. A Gershgorin disk of T **with respect to the basis** v_1, \dots, v_n is a **set** of the form

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

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

Theorem 5.4.5. *Gershgorin disk theorem*

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

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

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

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 $\lambda_1, \dots, \lambda_n$ 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 $\text{null}T = \{0\}$ then the proof has done.

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

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

so

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

□

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)

□

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.

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

□

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

This exercise is really hard!

\Rightarrow is trivial.

\Leftarrow : Suppose T is not diagonalizable.

Since $\mathbb{F} = \mathbb{C}$, the minimal polynomial of T can always be expressed in the form:

$$(z - \lambda_1)^{p_1} \cdots (z - \lambda_n)^{p_n}$$

But since T isn't diagonalizable, so there exists $p_j > 1$, which indicates an exceedingly significant fact, any polynomial in the form $q = (z - \lambda_1) \cdots (z - \lambda_m)$ isn't the minimal polynomial of T and even not the polynomial multiple of the minimal polynomial of T ! This means $\exists v \in V$ s.t.

$$q(T)v \neq 0$$

If T^k is diagonalizable, then its minimal polynomial can be expressed like

$$(z - \lambda'_1) \cdots (z - \lambda'_m), (\lambda'_j \neq 0) \text{ since } T \text{ is not invertible}$$

Now we consider $(z^k - a)$ and $(z^k - b), a, b \neq 0, a \neq b$ they can be factorized into disjoint (means no common root) sets of degree one polynomial and in a single set a root only appears once.

$$(z - a_1) \cdots (z - a_k)$$

and

$$(z - b_1) \cdots (z - b_k)$$

in which $a_i \neq a_j (i \neq j), a_i \neq b_j, \forall i, j$ and $i, j = 1, \dots, k$ because the property of the complex number and a, b are not zero

Therefore, for any polynomial in the form $q = (z - \lambda_1) \cdots (z - \lambda_m)$, $q(T^k)$ can be factorized into the form like

$$(T - \lambda_{11}I) \cdots (T - \lambda_{1k}I) \cdots (T - \lambda_{m1}I) \cdots (T - \lambda_{mk}I)$$

where there is no repeated root.

But we have already proved that any polynomial $q'(T)$ in this kind won't be able to cast all the vectors into 0, which implies that the minimal polynomial of T^k can't have the form without repeated root.

□

Exercise

Suppose that $T \in \mathcal{L}(V)$ is diagonalizable. Let $\lambda_1, \dots, \lambda_m$ 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, \dots, 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 makes 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 indicates 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 $u_1, \dots, u_m \in U$, 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, \dots, u_m are linearly-independent, which also shows

$$U = U_1 \oplus \cdots \oplus U_m$$

□

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

$$T/U(v+U) = c_2 T/U(v_2+U) = c_2(Tv_2+U) = c_2(v_1+v_2+U) = c_1 v_1 + c_2 v_2 + U = v+U$$

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 \square

5.5 Commuting Operators

Definition 5.5.1. *commute*

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

\square 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, \dots, v_n is a basis of V . Then S and T commute if and only if $\mathcal{M}(S, (v_1, \dots, v_n))$ and $\mathcal{M}(T, (v_1, \dots, v_n))$ commute.

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 diagonalizability \Leftrightarrow 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. □

Exercise

Suppose V is a finite-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} \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 $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$.

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 $c = \frac{\langle u, v \rangle}{\|v\|^2}$ and $w = u - \frac{\langle u, v \rangle}{\|v\|^2} v$. 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 $\frac{\langle u, v \rangle}{\|v\|} = x + iy$, for $x, y \in \mathbb{R}$ and consider two cases $a = -\frac{x}{\|v\|}$ and $a = \frac{-iy}{\|v\|}$. 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, \dots, 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, \dots, v_n .

Solution

Let v_1, \dots, 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, \dots, v_n is a basis of V , we can write v as a linear combination of v_1, \dots, v_n , say $v = a_1 v_1 + \dots + a_n v_n$. Then, we have $\|v\|^2 = |a_1|^2 + \dots + |a_n|^2$. Now, we have

$$\begin{aligned} \|Tv\|^2 &= \|T(a_1 v_1 + \dots + a_n v_n)\|^2 \\ &= \|a_1 T(v_1) + \dots + a_n T(v_n)\|^2 \\ &= \left\| \sum_{k=1}^n a_k \sum_{h=1}^n \mathcal{M}(T)_{h,k} v_h \right\|^2 \\ &= \left\| \sum_{h=1}^n v_h \sum_{k=1}^n \mathcal{M}(T)_{h,k} a_k \right\|^2 \\ &\leq \left\| \sum_{h=1}^n v_h \sqrt{\left(\sum_{k=1}^n |\mathcal{M}(T)_{h,k}|^2 \right) \left(\sum_{k=1}^n |a_k|^2 \right)} \right\|^2 \\ &\leq \sum_{h=1}^n \left(\sum_{k=1}^n |\mathcal{M}(T)_{h,k}|^2 \right) \left(\sum_{k=1}^n |a_k|^2 \right) \|v_h\|^2 \\ &= \|v\|^2 \sum_{h=1}^n \sum_{k=1}^n |\mathcal{M}(T)_{h,k}|^2 \end{aligned}$$

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. \square

Exercise

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

Solution

Use induction.

$k = 1$ is obvious.

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

$$\begin{aligned} & \langle a_1 v_1 + \dots + a_{p+1} v_{p+1}, a_1 v_1 + \dots + a_{p+1} v_{p+1} \rangle \\ &= \langle a_1 v_1 + \dots + a_p v_p, a_1 v_1 + \dots + a_p v_p \rangle \\ &+ \langle a_{p+1} v_{p+1}, a_{p+1} v_{p+1} \rangle + 2a_{p+1} \operatorname{Re}(\langle a_1 v_1 + \dots + a_p v_p, v_{p+1} \rangle) \\ &\leq m + 1 + 2a_{p+1} \operatorname{Re}(\langle a_1 v_1 + \dots + a_p v_p, v_{p+1} \rangle) \end{aligned}$$

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

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

□

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 $2(\|(1, 0)\|^2 + \|(0, 1)\|^2) = \|(1, 1)\|^2 + \|(1, -1)\|^2 = 2 \cdot 2^{\frac{2}{p}}$ □

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

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

for all $u, v \in V$.

Solution

let $v = cu + w$

□

Exercise

A norm on a vector space U is a function

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

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 $\|u + v\| \leq \|u\| + \|v\|$ for all $u, v \in U$. Prove that a norm satisfying 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, \dots, 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, \dots, e_m of vectors in V is orthonormal if

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

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

Theorem 6.2.1. *norm of an orthonormal linear combination*

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

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

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

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

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

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, \dots, e_n is an orthonormal basis of V and $u, v \in V$. Then

- (a) $v = \langle v, e_1 \rangle e_1 + \dots + \langle v, e_n \rangle e_n$,
- (b) $\|v\|^2 = |\langle v, e_1 \rangle|^2 + \dots + |\langle v, e_n \rangle|^2$,
- (c) $\langle u, v \rangle = \langle u, e_1 \rangle \overline{\langle v, e_1 \rangle} + \dots + \langle u, e_n \rangle \overline{\langle v, e_n \rangle}$

Theorem 6.2.6. *Gram-Schmidt procedure*

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

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

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

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

$$\text{span}(v_1, \dots, v_k) = \text{span}(e_1, \dots, e_k)$$

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

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 $\lambda_1, \dots, \lambda_m \in \mathbb{F}$.*

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'*

- ☐ *A linear 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 $\varphi : \mathcal{P}_5(\mathbb{R}) \rightarrow \mathbb{R}$ 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, \dots, 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, \dots, e_n of V is chosen.

Exercise

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

$$(\cos \theta, \sin \theta), (-\sin \theta, \cos \theta) \text{ and } (\cos \theta, \sin \theta), (\sin \theta, -\cos \theta)$$

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. \square

Exercise

Suppose $f : [-\pi, \pi] \rightarrow \mathbb{R}$ 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 $f : [-\pi, \pi] \rightarrow \mathbb{R}$. However, proving that this inequality is an equality involves Fourier series techniques beyond the scope of this book

Solution

Sorry to skip this exercise. □

Exercise

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

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

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

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

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

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

for each k , but v_1, \dots, 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, \dots, a_n s.t. $a_1 v_1 + \dots + a_n v_n = 0$, therefore

$$\begin{aligned} \|a_1 e_1 + \dots + a_n e_n\| &= \|a_1 e_1 + \dots + a_n e_n - a_1 v_1 - \dots - a_n v_n + a_1 v_1 + \dots + a_n v_n\| \\ &\leq \|a_1(e_1 - v_1)\| + \dots + \|a_n(e_n - v_n)\| + \|a_1 v_1 + \dots + a_n v_n\| \\ &< \frac{1}{\sqrt{n}}|a_1| + \dots + \frac{1}{\sqrt{n}}|a_n| + |a_1 v_1 + \dots + a_n v_n| \\ &\leq \sqrt{\left(\frac{1}{n} + \dots + \frac{1}{n}\right)(a_1^2 + \dots + a_n^2)} + |a_1 v_1 + \dots + a_n v_n| \\ &= \sqrt{a_1^2 + \dots + a_n^2} + |a_1 v_1 + \dots + a_n v_n| \\ &= \|a_1 e_1 + \dots + a_n e_n\| \end{aligned}$$

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, \dots, v_n are linearly-dependent. \square

Exercise

Suppose v_1, \dots, 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, \dots, e_m in V such that $\langle v_k, e_k \rangle > 0$ and $\text{span}(v_1, \dots, v_k) = \text{span}(e_1, \dots, e_k)$ for each $k = 1, \dots, m$.

Solution

Hint: suppose there is a basis that is different, then exists $j \in \{1, \dots, m\}$ 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 $\text{span}(v_1, \dots, v_k) = \text{span}(e_1, \dots, e_k)$, these two vectors must be identical. \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

$$\begin{aligned} A^m &= (\lambda I + N)^m \\ &= \lambda^m I + m\lambda^{m-1}N + \frac{m(m-1)}{2}\lambda^{m-2}N^2 \\ &\quad + \frac{m(m-1)(m-2)}{2 \cdot 3}\lambda^{m-3}N^3 + \cdots + \frac{m(m-1) \cdots (m-n+1)}{(n-1)!}\lambda^{m-n+1}N^{n-1} \end{aligned}$$

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 $I, N, N^2, N^3, \dots, N^{n-1}$ 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 $m \rightarrow \infty$ 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. \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>

\square

Then there is an exercise about the metric space in the function analysis which is beyond our knowledge now.

Exercise

For all $u, v \in V$, 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 projection of V onto U is the operator $P_U \in \mathcal{L}(V)$ defined as follows: For each $v \in V$, write $v = u + w$, where $u \in U$ and $w \in U^\perp$. Then let $P_U v = u$.

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) $\text{range } P_U = U$;
- (e) $\text{null } P_U = U^\perp$;
- (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, \dots, e_m is an orthonormal basis of U and $v \in V$, then

$$P_U v = \langle v, e_1 \rangle e_1 + \dots + \langle v, e_m \rangle e_m$$

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 $v \rightarrow \varphi v$ is a one-to-one function from V onto V' .

Caution: The function $v \rightarrow \varphi v$ 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 $T|_{(\text{null } T)^\perp}$ is a one to-one map of $(\text{null } T)^\perp$ onto $\text{range } T$.

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

$$T^\dagger w = (T|_{(\text{null } T)^\perp})^{-1} P_{\text{range } T} w$$

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) $TT^\dagger = P_{\text{range } T} = \text{the orthogonal projection of } W \text{ onto } \text{range } T$.
- (c) $T^\dagger T = P_{(\text{null } T)^\perp} = \text{the orthogonal projection of } V \text{ onto } (\text{null } T)^\perp$

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, \dots, e_n is a list of vectors in V with $\|e_k\| = 1$ for each $k = 1, \dots, n$ and

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

for all $v \in V$. Prove that e_1, \dots, e_n is an orthonormal basis of V .

Solution

let $v = e_i$ then

$$1 = 1 + |\langle e_i, e_1 \rangle|^2 + \dots + |\langle e_i, e_{i-1} \rangle|^2 + |\langle e_i, e_{i+1} \rangle|^2 + \dots + |\langle e_i, e_n \rangle|^2$$

which implies that $\langle e_i, e_j \rangle = 0$ for $i \neq j$, so e_1, \dots, e_n is an orthonormal list.

Then we show that e_1, \dots, e_n spans V .

let $U = \text{Span}(e_1, \dots, e_n)$, $V = U \oplus U^\perp$. So

$$\forall v \in V, v = c_1 e_1 + \dots + c_n e_n + w$$

such that $u = c_1 e_1 + \dots + c_n e_n \in U, w \in U^\perp$.

$$\begin{aligned} \|v\|^2 &= \langle u + w, u + w \rangle \\ &= \langle u, u \rangle + \langle u, w \rangle + \langle w, u \rangle + \langle w, w \rangle \\ &= \langle u, u \rangle + \langle w, w \rangle \\ &= |c_1|^2 + \dots + |c_n|^2 + \langle w, w \rangle \\ &= |\langle v, e_1 \rangle|^2 + \dots + |\langle v, e_n \rangle|^2 + \langle w, w \rangle \end{aligned}$$

which implies that $w = 0$, thus we finished our proof.

□

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 $u \in U$, $\exists u = Pv$

$$u = Pv = P^2v = P(Pv) = Pu$$

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.

□

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:

$$f(x) = \begin{cases} \frac{|x - x_0|}{\delta}, & x \in U(x_0, \delta) \\ 0, & x \in (-1, 0) \cup (0, 1) \setminus U(x_0, \delta) \end{cases}$$

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). □

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 $T^* : W \rightarrow V$ 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 $T \rightarrow T^*$ 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 $v \rightarrow \varphi_v$, 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

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

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

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

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

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 $j \in 1, \dots, n$ and $k \in 1, \dots, m$, 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, \dots, e_n is an orthonormal basis of V and f_1, \dots, f_m is an orthonormal basis of W . Then $\mathcal{M}(T^, (f_1, \dots, f_m), (e_1, \dots, e_n))$ is the conjugate transpose of $\mathcal{M}(T, (e_1, \dots, e_n), (f_1, \dots, f_m))$. 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 $T : V \rightarrow W$ is a linear map. Under the identification of V with V' and the identification of W with W' , the adjoint map $T^* : W \rightarrow V$ corresponds to the dual map $T' : W' \rightarrow V'$ defined in 3.118, as Exercise 32 asks you to verify.

Under this identification, the formulas for $\text{null}T^*$ and $\text{range}T^*$ [7.6(a) and (b)] then become identical to the formulas for $\text{null}T'$ and $\text{range}T'$ [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) $\text{null}T = \text{null}T^*$;*
- (b) $\text{range}T = \text{range}T^*$;*
- (c) $V = \text{null}T \oplus \text{range}T$;*
- (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 $T^*v = \bar{\lambda}v$.*

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 $\frac{n^2 + n}{2}$ entries of the matrix arbitrarily. So the dimension should be $\frac{(\dim V + 1) \dim V}{2}$ \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. \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

$$V = \text{range } P \oplus \text{null } P \quad (1)$$

(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,

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

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 orthonormal basis and the proof is not hard.

Therefore, we've established the equivalence of the three statements.

□

Exercise

Suppose $D : \mathcal{P}_8(\mathbb{R}) \rightarrow \mathcal{P}_8(\mathbb{R})$ 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.

□

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 $Tw \perp Tv$

$$\|T(v + w)\|^2 = \|Tv\|^2 + \|Tw\|^2 = 100$$

□

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

$$\begin{aligned}
 \langle w_1, T^* w_2 \rangle &= \langle T w_1, w_2 \rangle \\
 &= \langle \langle w_1, u \rangle x, w_2 \rangle \\
 &= \langle w_1, u \rangle \langle x, w_2 \rangle \\
 &= \left\langle w_1, \overline{\langle x, w_2 \rangle} u \right\rangle \\
 &= \langle w_1, \langle w_2, x \rangle u \rangle
 \end{aligned}$$

Therefore, $T^* v = \langle v, x \rangle u$

(a) Suppose T is self-adjoint. 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

$$\begin{aligned}
 \langle \langle v, u \rangle x, x \rangle u &= T^* (\langle v, u \rangle x) \\
 &= T^* T v \\
 &= T T^* v \\
 &= T (\langle v, x \rangle u) \\
 &= \langle \langle v, x \rangle u, u \rangle x
 \end{aligned}$$

The operation afterward is similar and the proof from the other side is likewise. \square

Exercise

Suppose $T \in \mathcal{L}(V)$ is normal. Prove that

$$\text{null}T^k = \text{null}T \text{ and } \text{range}T^k = \text{range}T$$

for every positive integer k .

Solution

First for S self-adjoint: suppose $S^k x = 0$. Then

$$0 = \langle S^k x, S^{k-2} x \rangle = \langle S^{k-1} x, S^{k-1} x \rangle$$

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, \dots, e_n of V such that $\|Te_k\| = \|T^*e_k\|$ for each $k = 1, \dots, n$, 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. \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

$$V = \text{span}(1, \cos x, \cos 2x, \dots, \cos nx, \sin x, \sin 2x, \dots, \sin nx)$$

(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 simpler with matrices.

From previous exercises we learned that

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \dots, \frac{\sin nx}{\sqrt{\pi}}$$

is an orthonormal 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 rigorous solution may not use the matrix.

let $e_j = \frac{\sin jx}{\sqrt{\pi}}, f_j = \frac{\cos jx}{\sqrt{\pi}}$, then for $\forall v, w \in V$

$$\begin{aligned}
\langle v, D^*w \rangle &= \langle Dv, w \rangle \\
&= \left\langle D \left(\left\langle v, \frac{1}{\sqrt{2\pi}} \right\rangle \frac{1}{\sqrt{2\pi}} + \sum_{j=1}^n (\langle v, e_j \rangle e_j + \langle v, f_j \rangle f_j) \right), w \right\rangle \\
&= \left\langle \sum_{j=1}^n (-j \langle v, e_j \rangle f_j + j \langle v, f_j \rangle e_j), \left\langle w, \frac{1}{\sqrt{2\pi}} \right\rangle \frac{1}{\sqrt{2\pi}} + \sum_{j=1}^n (\langle w, e_j \rangle e_j + \langle w, f_j \rangle f_j) \right\rangle \\
&= \sum_{j=1}^n (-j \langle v, e_j \rangle \langle w, f_j \rangle + j \langle v, f_j \rangle \langle w, e_j \rangle) \\
&= \left\langle \left\langle v, \frac{1}{\sqrt{2\pi}} \right\rangle \frac{1}{\sqrt{2\pi}} + \sum_{j=1}^n (\langle v, e_j \rangle f_j + \langle v, f_j \rangle e_j), \sum_{j=1}^n -j (\langle w, f_j \rangle e_j + j \langle w, e_j \rangle f_j) \right\rangle \\
&= \langle v, -D^*w \rangle
\end{aligned}$$

Therefore, we get $D^* = -D$

(b) Note that $T = D^2$, thus $T^* = (DD)^* = D^*D^* = (-D)(-D) = D^2 = T$ \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 $\lambda_1, \dots, \lambda_m \in \mathbb{R}$.

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 $\lambda_1, \lambda_2, \dots, \lambda_n$ be its distinct eigenvalues.

By Lagrange interpolation, we can find a polynomial

$$p(z) = a_0 + a_1z + \cdots + a_kz^k$$

such that $p(\lambda_j) = \overline{\lambda_j}$ for $j = 1, 2, \dots, n$.

So we can define

$$p(T) = a_0I + a_1T + \cdots + z_kT^k$$

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 .

□

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.

□

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.

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, \dots, e_n is an orthonormal basis of V . Prove that

T is a positive operator if and only if there exist $v_1, \dots, v_n \in V$ such that

$$\langle Te_k, e_j \rangle = \langle v_k, v_j \rangle$$

for all $j, k = 1, \dots, n$.

The numbers $\{\langle Te_k, e_j \rangle\}_{j,k=1,\dots,n}$ are the entries in the matrix of T with respect to the orthonormal basis e_1, \dots, 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

$$\langle Te_k, e_j \rangle = \langle \sqrt{T}e_k, \sqrt{T}e_j \rangle = \langle v_k, v_j \rangle$$

thus we finished our proof. \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 $\{e_1, e_2, \dots, e_n\}$, 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:

$$\begin{aligned}
\langle Tv, v \rangle &= \left\langle \sum_{j=1}^n v_j \sum_{k=1}^n \frac{1}{j+k-1} e_k, \sum_{i=1}^n v_i e_i \right\rangle \\
&= \left\langle \sum_{k=1}^n \sum_{j=1}^n \frac{1}{j+k-1} v_j e_k, \sum_{i=1}^n v_i e_i \right\rangle \\
&= \sum_{k=1}^n \left\langle \sum_{j=1}^n \frac{1}{j+k-1} v_j e_k, v_k e_k \right\rangle \\
&= \sum_{j=1}^n \sum_{k=1}^n \frac{1}{j+k-1} v_j \overline{v_k} \\
&= \int_0^1 \sum_{j=1}^n \sum_{k=1}^n v_j \overline{v_k} x^{j-1} x^{k-1} dx \\
&= \int_0^1 (v_1 + v_2 x + \cdots + v_n x^{n-1}) (\overline{v_1} + \overline{v_2} x + \cdots + \overline{v_n} x^{n-1}) dx \\
&= \int_0^1 (v_1 + v_2 x + \cdots + v_n x^{n-1}) \overline{(v_1 + v_2 x + \cdots + v_n x^{n-1})} dx \\
&= \int_0^1 |(v_1 + v_2 x + \cdots + v_n x^{n-1})|^2 dx \geq 0
\end{aligned}$$

The following part is trivial compared to the trick of transformation 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, \dots, v_n is a basis of V , thinking of a series of linear functionals $\varphi_i(u) = \langle u, v_i \rangle'$ 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. □

7.4 Isometries, Unitary Operators, and Matrix Factorization

7.4.1 Isometries

The Greek word **isos** means equal; the Greek word **metron** means measure. Thus isometry literally means equal measure.

Definition 7.4.1. *isometry*

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

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

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

Theorem 7.4.1. *characterization of isometries*

Suppose $S \in \mathcal{L}(V, W)$. Suppose e_1, \dots, e_n is an orthonormal basis of V and f_1, \dots, f_m is an orthonormal basis of W . Then the following are equivalent.

- (a) S is an isometry.
- (b) $S^*S = I$.
- (c) $\langle Su, Sv \rangle = \langle u, v \rangle$ for all $u, v \in V$.
- (d) Se_1, \dots, Se_n is an orthonormal list in W .
- (e) The columns of $\mathcal{M}(S, (e_1, \dots, e_n), (f_1, \dots, f_m))$ form an orthonormal list in \mathbb{F}^m with respect to the Euclidean inner product.

For better understanding, the matrix in (e) may look likes such long one.

$$\begin{pmatrix} & & \\ & & \\ & & \end{pmatrix}$$

7.4.2 Unitary Operators

Definition 7.4.2. *unitary operator*

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

In the case of infinite-dimensional space, the condition “invertible” can’t be left out, despite in the finite-dimensional case isometry implies injective and invertibility.

Theorem 7.4.2. *characterization of unitary operators*

Suppose $S \in \mathcal{L}(V)$. Suppose e_1, \dots, e_n is an orthonormal basis of V . Then the following are equivalent.

- (a) S is a unitary operator.
- (b) $S^*S = SS^* = I$.
- (c) S is invertible and $S^{-1} = S^*$.
- (d) Se_1, \dots, Se_n is an orthonormal **basis** of V .
- (e) The rows of $\mathcal{M}(S, (e_1, \dots, e_n))$ form an orthonormal basis of \mathbb{F}^n with respect to the Euclidean inner product.
- (f) S^* is a unitary operator.

Recall our analogy between \mathbb{C} and $\mathcal{L}(V)$. Under this analogy, a complex number z corresponds to an operator $S \in \mathcal{L}(V)$, and \bar{z} corresponds to S^* . The real numbers ($z = \bar{z}$) correspond to the self-adjoint operators ($S = S^*$), and the nonnegative numbers correspond to the (badly named) positive operators.

Theorem 7.4.3. *eigenvalues of unitary operators have **absolute** value 1*

Suppose λ is an eigenvalue of a unitary operator. Then $|\lambda| = 1$

Theorem 7.4.4. *description of unitary operators on complex inner product spaces*

Suppose $\mathbb{F} = \mathbb{C}$ and $S \in \mathcal{L}(V)$. Then the following are equivalent.

- (a) S is a unitary operator.
- (b) There is an orthonormal basis of V consisting of eigenvectors of S whose corresponding eigenvalues all have absolute value 1.

7.4.3 QR Factorization

Definition 7.4.3. *unitary matrix*

An n -by- n matrix is called unitary if its columns form an orthonormal list in \mathbb{F}^n .

Theorem 7.4.5. *characterizations of unitary matrices*

Suppose Q is an n -by- n matrix. Then the following are equivalent.

- (a) Q is a unitary matrix.
- (b) The rows of Q form an orthonormal list in \mathbb{F}^n .
- (c) $\|Qv\| = \|v\|$ for every $v \in \mathbb{F}^n$.
- (d) $Q^*Q = QQ^* = I$, the n -by- n matrix with 1's on the diagonal and 0's elsewhere

Theorem 7.4.6. *QR factorization*

Suppose A is a square matrix with linearly independent columns. Then there exist **unique** matrices Q and R such that Q is unitary, R is upper triangular with only positive numbers on its diagonal, and

$$A = QR$$

7.4.4 Cholesky Factorization

Theorem 7.4.7. *positive invertible operator*

A self-adjoint operator $T \in \mathcal{L}(V)$ is a positive invertible operator if and only if $\langle Tv, v \rangle > 0$ for every nonzero $v \in V$.

Definition 7.4.4. *positive definite*

A matrix $B \in \mathbb{F}^{n,n}$ is called positive definite if $B^* = B$ and

$$\langle Bx, x \rangle > 0$$

for every nonzero $x \in \mathbb{F}^n$

Theorem 7.4.8. *Cholesky factorization*

Suppose B is a positive definite matrix. Then there exists a **unique** upper triangular matrix R with only positive numbers on its diagonal such that

$$B = R^*R$$

Exercise

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Suppose every eigenvalue of T has absolute value 1 and $\|Tv\| \leq \|v\|$ for every $v \in V$. Prove that T is a unitary operator

Solution

The key is that the **unitary operator is normal**, since we know that $T^*T = TT^* = I$.

Then by complex spectral theorem, there is an orthonormal basis consists of the eigenvectors of T , we denote them by v_1, \dots, v_n . And it is easy to show that $\|Tv\| = \|v\|$ for $\forall v \in V$ \square

Exercise

Suppose $S \in \mathcal{L}(V)$. Prove that S is a unitary operator if and only if

$$\{Sv : v \in V \text{ and } \|v\| \leq 1\} = \{v \in V : \|v\| \leq 1\}$$

Solution

we denote the set on the left and right with A and B respectively.

\Rightarrow : It's easy to show that $\forall u \in A, \exists v \in V$ s.t. $\|u\| = \|Sv\| = \|v\| \leq 1$, so $A \subset B$

for $v \in B, v = S(S^{-1}v)$ for some $S^{-1}v \in V$ and $\|S^{-1}v\| = \|v\| \leq 1$, so $B \subset A$. Therefore $A = B$

\Leftarrow : If S is not unitary, then there exists $v \in V$ s.t. $\|Sv\| = |c|\|v\|$ for $|c| \neq 1$, now we suppose $|c| > 1$, then for the equation below, it means for $\forall v \in V$ s.t. $\|v\| = 1, \|Sv\| \leq 1$, but our assumption shows that $\|Sv\| > 1$, which leads to a contradiction. \square

Exercise

Prove or give a counterexample: If $S \in \mathcal{L}(V)$ is invertible and $\|S^{-1}v\| = \|Sv\|$ for every $v \in V$, then S is unitary.

Solution

The condition implies that

$$\langle S^{-1}v, S^{-1}v \rangle - \langle Sv, Sv \rangle = \langle ((S^{-1})^*S^{-1} - S^*S)v, v \rangle = 0$$

Therefore $(S^{-1})^*S^{-1} = S^*S$ i.e.

$$S^*SSS^* = SS^*S^*S = SSS^*S^* = S^*S^*SS = I$$

Thus S^2 and S^*S are unitary.

The latter one is self-adjoint, then it has a diagonal matrix with respect to some orthonormal basis consists of its eigenvectors. Since self-adjoint means that its eigenvalues are real numbers. It's also positive since $\langle S^*Sv, v \rangle = \langle Sv, Sv \rangle \geq 0 \ \forall v \in V$, so all the eigenvalues of S^*S are non-negative. So the only eigenvalue it has is 1 and it means $S^*S = I$, which implies that S is unitary.

I'm not quite confident with the proof above, so you can check carefully for any possible mistake. \square

Exercise

A square matrix A is called symmetric if it equals its transpose. Prove that if A is a symmetric matrix with real entries, then there exists a unitary matrix Q with real entries such that Q^*AQ is a diagonal matrix.

Solution

It's a trivial fact we have already learned, but we might want to use the language of this book to solve it.

Since A equals to its conjugate transpose, so we can arbitrarily choose an orthonormal basis e_1, \dots, e_n , so there exists some operator T with respect to e_1, \dots, e_n are matrix A , which is self-adjoint. T has a diagonal matrix with respect to some other orthonormal basis f_1, \dots, f_n . Since $Q = \mathcal{M}(I, (e_1, \dots, e_n), (f_1, \dots, f_n))$ satisfy the condition that $\|Qv\| = \|v\|$ for every $v \in V$, so we finished our proof. \square

Exercise

Suppose n is a positive integer. For this exercise, we adopt the notation that a typical element z of \mathbb{C}^n is denoted by $z = (z_0, z_1, \dots, z_{n-1})$. Define linear functionals $\omega_0, \omega_1, \dots, \omega_{n-1}$ on \mathbb{C}^n by

$$\omega_j(z_0, z_1, \dots, z_{n-1}) = \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} z_m e^{\frac{-2\pi i j m}{n}}$$

The discrete Fourier transform is the operator $\mathcal{F} : \mathbb{C}^n \rightarrow \mathbb{C}^n$ defined by

$$\mathcal{F}z = (\omega_0(z), \omega_1(z), \dots, \omega_{n-1}(z))$$

- (a) Show that \mathcal{F} is a unitary operator on \mathbb{C}^n .
- (b) Show that if $(z_0, \dots, z_{n-1}) \in \mathbb{C}^n$ and z_n is defined to equal z_0 , then

$$\mathcal{F}^{-1}(z_0, z_1, \dots, z_{n-1}) = \mathcal{F}(z_n, z_{n-1}, \dots, z_1)$$

- (c) Show that $\mathcal{F}^4 = I$.

The discrete Fourier transform has many important applications in data analysis. The usual Fourier transform involves expressions of the form $\int_{-\infty}^{+\infty} f(x)e^{-2\pi itx}$ for complex-valued integrable functions f defined on \mathbb{R} .

Solution

Remain to be solved.

□

7.5 Singular Value Decomposition

7.5.1 Singular Values

Theorem 7.5.1. *properties of T^*T*

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

- (a) *T^*T is a positive operator on V ;*
- (b) *$\text{null } T^*T = \text{null } T$;*
- (c) *$\text{range } T^*T = \text{range } T^*$;*
- (d) *$\dim \text{range } T = \dim \text{range } T^* = \dim \text{range } T^*T$.*

Definition 7.5.1. *singular values*

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

Theorem 7.5.2. *role of positive singular values*

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

- (a) T is injective $\Leftrightarrow 0$ is not a singular value of T ;
- (b) the number of positive singular values of T equals $\dim \text{range} T$;
- (c) T is surjective \Leftrightarrow number of positive singular values of T equals $\dim W$

Notice that (b) is important and this condition can be transformed into the rank of matrix.

Theorem 7.5.3. *isometries characterized by having all singular values equal 1*

Suppose that $S \in \mathcal{L}(V, W)$. Then S is an isometry \Leftrightarrow all singular values of S equal 1.

7.5.2 SVD for Linear Maps and for Matrices

Theorem 7.5.4. *singular value decomposition*

Suppose $T \in \mathcal{L}(V, W)$ and the positive singular values of T are s_1, \dots, s_m . Then there exist orthonormal lists e_1, \dots, e_m in V and f_1, \dots, f_m in W such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \dots + s_m \langle v, e_m \rangle f_m$$

for every $v \in V$

Attention that e_1, \dots, e_n are the eigenvectors of T^*T and

$$f_k = \frac{Te_k}{s_k} \quad k = 1, \dots, m$$

Definition 7.5.2. *diagonal matrix*

An M -by- N matrix A is called a diagonal matrix if all entries of the matrix are 0 except possibly $A_{k,k}$ for $k = 1, \dots, \min\{M, N\}$.

Theorem 7.5.5. *singular value decomposition of adjoint and pseudoinverse*

Suppose $T \in \mathcal{L}(V, W)$ and the positive singular values of T are s_1, \dots, s_m . Suppose e_1, \dots, e_m and f_1, \dots, f_m are orthonormal lists in V and W such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \dots + s_m \langle v, e_m \rangle f_m$$

for every $v \in V$. Then

$$T^*w = s_1 \langle w, f_1 \rangle e_1 + \dots + s_m \langle w, f_m \rangle e_m$$

and

$$T^\dagger w = \frac{\langle w, f_1 \rangle}{s_1} e_1 + \dots + \frac{\langle w, f_m \rangle}{s_m} e_m$$

for every $w \in W$.

Theorem 7.5.6. *matrix version of SVD*

Suppose A is an M -by- n matrix of rank $m \geq 1$. Then there exist an M -by- m matrix B with orthonormal columns, an m -by- m diagonal matrix D with positive numbers on the diagonal, and an n -by- m matrix C with orthonormal columns such that

$$A = BDC^*$$

The theorem above is a simplified version of the SVD for matrix.

Exercise

Suppose $T \in \mathcal{L}(V)$ is normal. Prove that the singular values of T^2 equal the squares of the singular values of T .

Solution

Let s be one singular value of T , then $T^*Tv = s^2v$ for some $v \in V$, therefore

$$(T^2)^*T^2v = T^*TT^*Tv = s^4v$$

so s^2 is the singular value of T^2 .

Since T and T^2 have the same number of singular values, so there exists such one-on-one corresponding relation. \square

7.6 Consequences of Singular Value Decomposition

7.6.1 Norms of Linear Maps

Theorem 7.6.1. *upper bound for $\|Tv\|$*

Suppose $T \in \mathcal{L}(V, W)$. Let s_1 be the largest singular value of T . Then

$$\|Tv\| \leq s_1 \|v\|$$

for all $v \in V$.

The lower bound can also be proved to be

$$\|Tv\| \geq s_n \|v\|$$

Definition 7.6.1. *norm of a linear map, $\|\cdot\|$*

Suppose $T \in \mathcal{L}(V, W)$. Then the norm of T , denoted by $\|T\|$, is defined by

$$\|T\| = \max\{\|Tv\| : v \in V \text{ and } \|v\| \leq 1\}$$

Theorem 7.6.2. *basic properties of norms of linear maps*

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

- (a) $\|T\| \geq 0$;
- (b) $\|T\| = 0 \Leftrightarrow T = 0$;
- (c) $\|\lambda T\| = |\lambda| \|T\|$ for all $\lambda \in \mathbb{F}$;
- (d) $\|S + T\| \leq \|S\| + \|T\|$ for all $S \in \mathcal{L}(V, W)$.

Theorem 7.6.3. *alternative formulas for $\|T\|$*

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

- (a) $\|T\| =$ the largest singular value of T ;
- (b) $\|T\| = \max\{\|Tv\| : v \in V \text{ and } \|v\| = 1\}$;
- (c) $\|T\| =$ the smallest number c such that $\|Tv\| \leq c\|v\|$ for all $v \in V$.

And a frequently used inequality:

$$\|Tv\| \leq \|T\| \|v\|$$

Theorem 7.6.4. *norm of the adjoint*

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

Since we have already proved T and T^* have the positive singular values in the exercise before.

7.6.2 Approximation by Linear Maps with Lower-Dimensional Range

Theorem 7.6.5. *best approximation by linear map whose range has dimension $\leq k$*

Suppose $T \in \mathcal{L}(V, W)$ and $s_1 \geq \cdots \geq s_m$ are the positive singular values of T . Suppose $1 \leq k < m$. Then

$$\min\{\|T - S\| : S \in \mathcal{L}(V, W) \text{ and } \dim \text{range} S \leq k\} = s_{k+1}.$$

Furthermore, if

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

is a singular value decomposition of T and $T_k \in \mathcal{L}(V, W)$ is defined by

$$T_k v = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_k \langle v, e_k \rangle f_k$$

for each $v \in V$, then $\dim \text{range} T_k = k$ and $\|T - T_k\| = s_{k+1}$.

7.6.3 Polar Decomposition

Theorem 7.6.6. *polar decomposition*

Suppose $T \in \mathcal{L}(V)$. Then there exists a unitary operator $S \in \mathcal{L}(V)$ such that

$$T = S\sqrt{T^*T}.$$

7.6.4 Operators Applied to Ellipsoids and Parallelepipeds

Definition 7.6.2. *ball, B*

The ball in V of radius 1 centered at 0, denoted by B , is defined by

$$B = \{v \in V : \|v\| < 1\}.$$

Definition 7.6.3. *ellipsoid, $E(s_1 f_1, \dots, s_n f_n)$, principal axes*

Suppose that f_1, \dots, f_n is an orthonormal basis of V and s_1, \dots, s_n are positive numbers. The ellipsoid $E(s_1 f_1, \dots, s_n f_n)$ with principal axes $s_1 f_1, \dots, s_n f_n$ is defined by

$$E(s_1 f_1, \dots, s_n f_n) = \left\{ v \in V : \frac{|\langle v, f_1 \rangle|^2}{s_1^2} + \dots + \frac{|\langle v, f_n \rangle|^2}{s_n^2} < 1 \right\}$$

Definition 7.6.4. *notation: $T(\Omega)$*

For T a function defined on V and $\Omega \subset V$, define $T(\Omega)$ by

$$T(\Omega) = \{Tv : v \in \Omega\}.$$

Theorem 7.6.7. *invertible operator takes ball to ellipsoid*

Suppose $T \in \mathcal{L}(V)$ is invertible. Then T maps the ball B in V onto an ellipsoid in V

Theorem 7.6.8. *invertible operator takes ellipsoids to ellipsoids*

Suppose $T \in \mathcal{L}(V)$ is invertible and E is an ellipsoid in V . Then $T(E)$ is an ellipsoid in V .

Definition 7.6.5. *$P(v_1, \dots, v_n)$, parallelepiped*

Suppose v_1, \dots, v_n is a basis of V . Let

$$P(v_1, \dots, v_n) = \{\alpha_1 v_1 + \dots + \alpha_n v_n : \alpha_1, \dots, \alpha_n \in (0, 1)\}.$$

*A **parallelepiped** is a set of the form $u + P(v_1, \dots, v_n)$ for some $u \in V$. The vectors v_1, \dots, v_n are called the **edges** of this parallelepiped.*

Theorem 7.6.9. *invertible operator takes parallelepipeds to parallelepipeds*

Suppose $u \in V$ and v_1, \dots, v_n is a **basis** of V . Suppose $T \in \mathcal{L}(V)$ is invertible. Then

$$T(u + P(v_1, \dots, v_n)) = Tu + P(Tv_1, \dots, Tv_n).$$

Definition 7.6.6. *box*

A box in V is a set of the form

$$u + P(r_1 e_1, \dots, r_n e_n)$$

where $u \in V$ and r_1, \dots, r_n are positive numbers and e_1, \dots, e_n is an **orthonormal basis** of V .

Theorem 7.6.10. *every invertible operator takes some boxes to boxes*

Suppose $T \in \mathcal{L}(V)$ is invertible. Suppose T has singular value decomposition

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \dots + s_n \langle v, e_n \rangle f_n$$

where s_1, \dots, s_n are the singular values of T and e_1, \dots, e_n and f_1, \dots, f_n are orthonormal bases of V and the equation above holds for all $v \in V$. Then T maps the box $u + P(r_1 e_1, \dots, r_n e_n)$ onto the box $Tu + P(r_1 s_1 f_1, \dots, r_n s_n f_n)$ for all positive numbers r_1, \dots, r_n and all $u \in V$.

7.6.5 Volume via Singular Values

Definition 7.6.7. *volume of a box*

Suppose $\mathbb{F} = \mathbb{R}$. If $u \in V$ and r_1, \dots, r_n are positive numbers and e_1, \dots, e_n is an orthonormal basis of V , then

$$\text{volume}(u + P(r_1 e_1, \dots, r_n e_n)) = r_1 \times \dots \times r_n$$

To define the volume of a subset of V , approximate the subset by a finite collection of disjoint boxes, and then add up the volumes of the

approximating collection of boxes. As we approximate a subset of V more accurately by disjoint unions of more boxes, we get a better approximation to the volume.

These ideas should remind you of how the Riemann integral is defined by approximating the area under a curve by a disjoint collection of rectangles. This discussion leads to the following nonrigorous but intuitive definition.

Definition 7.6.8. *volume*

Suppose $\mathbb{F} = \mathbb{R}$ and $\Omega \subset V$. Then the volume of Ω , denoted by $\text{volume}\Omega$, is approximately the sum of the volumes of a collection of disjoint boxes that approximate Ω .

Theorem 7.6.11. *volume changes by a factor of the product of the singular values*

Suppose $\mathbb{F} = \mathbb{R}$, $T \in \mathcal{L}(V)$ is invertible, and $\Omega \subset V$. Then

$$\text{volume}T(\Omega) = (\text{product of singular values of } T)(\text{volume}\Omega).$$

Suppose $T \in \mathcal{L}(V)$. As we will see when we get to determinants, the product of the singular values of T equals $|\det T|$;

7.6.6 Properties of an Operator as Determined by Its Eigenvalues

properties of a normal operator	eigenvalues are contained in
invertible	$\mathbb{C} \setminus \{0\}$
self-adjoint	\mathbb{R}
skew	$\{\lambda \in \mathbb{C} : \text{Re}\lambda = 0\}$
orthogonal projection	$\{0, 1\}$
positive	$[0, \infty)$
unitary	$\{\lambda \in \mathbb{C} : \lambda = 1\}$
norm is less than 1	$\{\lambda \in \mathbb{C} : \lambda < 1\}$

Exercise

Show that defining $d(S, T) = \|S - T\|$ for $S, T \in \mathcal{L}(V, W)$ makes d a metric on $\mathcal{L}(V, W)$.

This exercise is intended for readers who are familiar with metric spaces.

Solution

This will remain until I learned the function analysis. □

Exercise

(a) Prove that if $T \in \mathcal{L}(V)$ and $\|I - T\| < 1$, then T is invertible.

(b) Suppose that $S \in \mathcal{L}(V)$ is invertible. Prove that if $T \in \mathcal{L}(V)$ and $\|S - T\| < 1/\|S^{-1}\|$, then T is invertible.

This exercise shows that the set of invertible operators in $\mathcal{L}(V)$ is an open subset of $\mathcal{L}(V)$, using the metric defined in Exercise above

Solution

□

Exercise

Suppose $T \in \mathcal{L}(V)$. Prove that for every $\varepsilon > 0$, there exists an invertible operator $S \in \mathcal{L}(V)$ such that $0 < \|T - S\| < \varepsilon$.

Solution

□

Exercise

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Prove that for every $\varepsilon > 0$ there exists a diagonalizable operator $S \in \mathcal{L}(V)$ such that $0 < \|T - S\| < \varepsilon$.

Solution

□

Exercise

Suppose e_1, \dots, e_n is an orthonormal basis of V and $T \in \mathcal{L}(V, W)$.

(a) Prove that $\max\{\|Te_1\|, \dots, \|Te_n\|\} \leq \|T\| \leq (\|Te_1\|^2 + \dots + \|Te_n\|^2)^{\frac{1}{2}}$.

(b) Prove that $\|T\| = (\|Te_1\|^2 + \cdots + \|Te_n\|^2)^{\frac{1}{2}}$ if and only if $\dim \text{range } T \leq 1$.

Here e_1, \dots, e_n is an arbitrary orthonormal basis of V , not necessarily connected with a singular value decomposition of T . If s_1, \dots, s_n is the list of singular values of T , then the right side of the inequality above equals $(s_1^2 + \cdots + s_n^2)^{\frac{1}{2}}$, as was shown in Exercise 11(a) in Section 7E.

Solution

□

Exercise

Suppose $T \in \mathcal{L}(V)$ is normal. Prove that $\|T^k\| = \|T\|^k$ for every positive integer k .

Solution

□

Exercise

Suppose $\dim V > 1$ and $\dim W > 1$. Prove that the norm on $\mathcal{L}(V, W)$ does not come from an inner product. In other words, prove that there does not exist an inner product on $\mathcal{L}(V, W)$ such that

$$\max\{\|Tv\| : v \in V \text{ and } \|v\| \leq 1\} = \sqrt{\langle T, T \rangle}$$

for all $T \in \mathcal{L}(V, W)$.

Solution

□

Exercise

Suppose $T \in \mathcal{L}(V, W)$. Show that T is uniformly continuous with respect to the metrics on V and W that arise from the norms on those spaces (see Exercise 23 in Section 6B).

Solution

□

Exercise

Fix $u, x \in V$ with $u \neq 0$. Define $T \in \mathcal{L}(V)$ by $Tv = \langle v, u \rangle x$ for every $v \in V$. Prove that

$$\sqrt{T^*T}v = \frac{\|x\|}{\|u\|} \langle v, u \rangle u$$

for every $v \in V$.

Solution

□

8 Operators on Complex Vector Spaces

8.1 Generalized Eigenvectors and Nilpotent Operators

8.1.1 Null Spaces of Powers of an Operator

Theorem 8.1.1. *sequence of increasing null spaces*

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

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

Theorem 8.1.2. *equality in the sequence of null spaces*

Suppose $T \in \mathcal{L}(V)$ and m is a nonnegative integer such that

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

Then

$$\text{null}T^m = \text{null}T^{m+1} = \text{null}T^{m+2} = \text{null}T^{m+3} = \cdots$$

Theorem 8.1.3. *null spaces stop growing*

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

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

Theorem 8.1.4. *V is the direct sum of $\text{null}T^{\dim V}$ and $\text{range}T^{\dim V}$*

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

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

8.1.2 Generalized Eigenvectors

Definition 8.1.1. *generalized eigenvector*

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

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

for some positive integer k

Generalized eigenvalues are not defined because doing so would not lead to anything new. Reason: if $(T - \lambda I)^k$ is not injective for some positive integer k , then $T - \lambda I$ is not injective, and hence λ is an eigenvalue of T .

Theorem 8.1.5. *a basis of generalized eigenvectors*

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then there is a basis of V consisting of generalized eigenvectors of T

Theorem 8.1.6. *generalized eigenvector corresponds to a unique eigenvalue*

Suppose $T \in \mathcal{L}(V)$. Then each generalized eigenvector of T corresponds to only one eigenvalue of T .

Theorem 8.1.7. *linearly independent generalized eigenvectors*

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

8.1.3 Nilpotent Operators

Definition 8.1.2. *nilpotent*

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

Theorem 8.1.8. *nilpotent operator raised to dimension of domain is 0*

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

Theorem 8.1.9. *eigenvalues of nilpotent operator*

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

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

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

Theorem 8.1.10. *minimal polynomial and upper-triangular matrix of nilpotent operator*

Suppose $T \in \mathcal{L}(V)$. Then the following are equivalent.

(a) T is nilpotent.

(b) The minimal polynomial of T is z^m for some positive integer m .

(c) There is a basis of V with respect to which the matrix of T has the form

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

*where all entries **on and below** the diagonal equal 0*

Exercise

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

$$\dim \text{null} T^m \leq m \dim \text{null} T$$

Hint: Exercise 21 in Section 3B may be useful.

Solution

The exercise below is what it mentioned in the exercise.

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

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

The lemma is not hard to prove:

We skip the part that prove $S = \{v \in V : Tv \in U\}$ is a subspace.

We consider $T|_S \in \mathcal{L}(S, U)$, we have

$$\dim S = \dim \text{null}T|_S + \dim \text{range}T|_S$$

Since $0 \in U$, if $v \notin S$, $Tv \neq 0$, therefore

$$\text{null}T = \text{null}T|_S$$

Likely, if $v \in S \Leftrightarrow Tv \in \text{range}T$ and $Tv \in U$, so

$$\text{range}T|_S = \text{range}T \cap U$$

Then with these we can finish the proof of lemma.

Let U be $\text{null}T^k$, then

$$\dim \text{null}T^{k+1} = \dim \text{null}T + \dim(\text{null}T^k \cap \text{range}T)$$

Since it's trivial that

$$\dim(\text{null}T^k \cap \text{range}T) \leq \dim \text{null}T^k$$

Therefore,

$$\dim \text{null}T^{k+1} \leq \dim \text{null}T + \dim \text{null}T^k \leq \cdots \leq (k+1) \dim \text{null}T$$

from which we can finish out proof. \square

Exercise

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

$$\text{null}T^m = \text{null}T^{m+1} \Leftrightarrow \text{range}T^m = \text{range}T^{m+1}$$

Solution

This exercise might be easier than you think, since

$$\dim V = \dim \text{null}T^m + \dim \text{range}T^m = \dim \text{null}T^{m+1} + \dim \text{range}T^{m+1}$$

The things afterwards is trivial because

$$\text{null}T^k \subset \text{null}T^{k+1}, \text{range}T^{k+1} \subset \text{range}T^k$$

□

Exercise

Suppose that $T \in \mathcal{L}(V)$. Prove that there is a basis of V consisting of generalized eigenvectors of T if and only if the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \dots, \lambda_m \in \mathbb{F}$.

Assume $\mathbb{F} = \mathbb{C}$ because the case $\mathbb{F} = \mathbb{C}$ follows from 5.27(b) and 8.9.

This exercise states that the condition for there to be a basis of V consisting of generalized eigenvectors of T is the same as the condition for there to be a basis with respect to which T has an upper-triangular matrix (see 5.44).

Caution: If T has an upper-triangular matrix with respect to a basis v_1, \dots, v_n of V , then v_1 is an eigenvector of T but it is not necessarily true that v_2, \dots, v_n are generalized eigenvectors of T .

Solution

You can try to use the proof in 8.9 (In 8A)

□

8.2 Generalized Eigenspace Decomposition

8.2.1 Generalized Eigenspaces

Definition 8.2.1. *generalized eigenspace, $G(\lambda, T)$*

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

$$G(\lambda, T) = \{v \in V : (T - \lambda I)^k v = 0 \text{ for some positive integer } k\}$$

Thus $G(\lambda, T)$ is the set of generalized eigenvectors of T corresponding to λ , along with the 0 vector.

Theorem 8.2.1. *description of generalized eigenspaces*

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

Theorem 8.2.2. *generalized eigenspace decomposition*

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T . Then

- (a) $G(\lambda_k, T)$ is invariant under T for each $k = 1, \dots, m$;
- (b) $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent for each $k = 1, \dots, m$;
- (c) $V = G(\lambda_1, T) \oplus \dots \oplus G(\lambda_m, T)$.

8.2.2 Multiplicity of an Eigenvalue

Definition 8.2.2. *multiplicity*

□ Suppose $T \in \mathcal{L}(V)$. The multiplicity of an eigenvalue λ of T is defined to be the dimension of the corresponding generalized eigenspace $G(\lambda, T)$.

□ In other words, the multiplicity of an eigenvalue λ of T equals $\dim \text{null}(T - \lambda I)^{\dim V}$.

Theorem 8.2.3. *sum of the multiplicities equals $\dim V$*

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then the sum of the multiplicities of all eigenvalues of T equals $\dim V$.

Notice:

algebraic multiplicity of $\lambda = \dim \text{null}(T - \lambda I)^{\dim V} = \dim G(\lambda, T)$,

geometric multiplicity of $\lambda = \dim \text{null}(T - \lambda I) = \dim E(\lambda, T)$

Definition 8.2.3. *characteristic polynomial*

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

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

is called the characteristic polynomial of T .

Theorem 8.2.4. *degree and zeros of characteristic polynomial*

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then

- (a) the characteristic polynomial of T has degree $\dim V$;
- (b) the zeros of the characteristic polynomial of T are the eigenvalues of T .

Theorem 8.2.5. *Cayley-Hamilton theorem*

Suppose $\mathbb{F} = \mathbb{C}$, $T \in \mathcal{L}(V)$, and q is the characteristic polynomial of T . Then $q(T) = 0$.

Theorem 8.2.6. *characteristic polynomial is a multiple of minimal polynomial*

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

Theorem 8.2.7. *multiplicity of an eigenvalue equals number of times on diagonal*

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Suppose v_1, \dots, v_n is a basis of V such that $\mathcal{M}(T, (v_1, \dots, v_n))$ is upper triangular. Then the number of times that each eigenvalue λ of T appears on the diagonal of $\mathcal{M}(T, (v_1, \dots, v_n))$ equals the multiplicity of λ as an eigenvalue of T .

8.2.3 Block Diagonal Matrices

Definition 8.2.4. *block diagonal matrix*

A block diagonal matrix is a **square** matrix of the form

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

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

Theorem 8.2.8. *block diagonal matrix with upper-triangular blocks*

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T , with multiplicities d_1, \dots, d_m . Then there is a basis of V with respect to which T has a block diagonal matrix of the form

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

where each A_k is a d_k -by- d_k upper-triangular matrix of the form

$$A_k = \begin{pmatrix} \lambda_k & & * \\ & \ddots & \\ 0 & & \lambda_k \end{pmatrix}$$

Exercise

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Prove that there exist $D, N \in \mathcal{L}(V)$ such that $T = D + N$, the operator D is diagonalizable, N is nilpotent, and $DN = ND$.

Solution

By the conclusion we have on general eigenvalues decomposition, we know that

$$V = G(\lambda_1, T) \oplus \dots \oplus G(\lambda_m, T)$$

And in each general eigenspace $G(\lambda_k, T) = G(\lambda_k, T|_{G(\lambda_k, T)})$. Since $\mathbb{F} = \mathbb{C}$, there exists some basis with respect to which the matrix of T is upper triangle. We have that

$$T|_{G(\lambda_k, T)} = \lambda_k I + (T|_{G(\lambda_k, T)} - \lambda_k I)$$

The first term on the right is a diagonalizable operator, and the second term is a nilpotent as you can verify.

Since the general eigenvalues decomposition holds, we can find the ideal D and N . □

Exercise

Suppose V is a complex inner product space, e_1, \dots, e_n is an orthonormal basis of T , and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of T , each included as many times as its multiplicity. Prove that

$$|\lambda_1|^2 + \dots + |\lambda_n|^2 \leq \|Te_1\|^2 + \dots + \|Te_n\|^2$$

See the comment after Exercise 5 in Section 7A

Solution

From the exercise 5 in 7A we know:

Suppose $T \in \mathcal{L}(V, W)$. Suppose e_1, \dots, e_n is an orthonormal basis of V and f_1, \dots, f_m is an orthonormal basis of W .

$$\|Te_1\|^2 + \dots + \|Te_n\|^2 = \|T^*f_1\|^2 + \dots + \|T^*f_m\|^2.$$

Remained to be solved. □

Exercise

Suppose that $\mathbb{F} = \mathbb{C}$ and V_1, \dots, V_m are nonzero subspaces of V such that

$$V = V_1 \oplus \dots \oplus V_m$$

Suppose $T \in \mathcal{L}(V)$ and each V_k is invariant under T . For each k , let p_k denote the characteristic polynomial of $T|_{V_k}$. Prove that the characteristic polynomial of T equals $p_1 \cdots p_m$.

Solution

It's obvious that $\forall v \in V, \exists v_1 \in V_1, \dots, v_m \in V_m$

$$v = v_1 + \dots + v_m$$

You can verify that

$$Tv = T|_{V_1}v_1 + \dots + T|_{V_m}v_m$$

Therefore $\forall \lambda \in \mathbb{F}$

$$\dim G(\lambda, T) = \dim G(\lambda, T|_{V_1}) + \cdots + \dim G(\lambda, T|_{V_m})$$

With this equality, you can cope with every eigenvalue with its multiplicity.

□

Exercise

Suppose $\mathbb{F} = \mathbb{R}$, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{C}$.

(a) Show that $u + iv \in G(\lambda, T_{\mathbb{C}})$ if and only if $u - iv \in G(\lambda, T_{\mathbb{C}})$.

(b) Show that the multiplicity of λ as an eigenvalue of $T|_{\mathbb{C}}$ equals the multiplicity of λ as an eigenvalue of $T|_{\mathbb{C}}$.

(c) Use (b) and the result about the sum of the multiplicities (8.25) to show that if $\dim V$ is an odd number, then $T|_{\mathbb{C}}$ has a real eigenvalue.

(d) Use (c) and the result about real eigenvalues of $T|_{\mathbb{C}}$ (Exercise 17 in Section 5A) to show that if $\dim V$ is an odd number, then T has an eigenvalue (thus giving an alternative proof of 5.34).

See Exercise 33 in Section 3B for the definition of the complexification $T|_{\mathbb{C}}$.

Solution

□

8.3 Consequences of Generalized Eigenspace Decomposition

8.3.1 Square Roots of Operators

Theorem 8.3.1. *identity plus nilpotent has a square root*

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

Theorem 8.3.2. *over \mathbb{C} , invertible operators have square roots*

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

8.3.2 Jordan Form

Specifically, there is a finite collection of vectors $v_1, \dots, v_n \in V$ such that there is a basis of V consisting of the vectors of the form $T^j v_k$, as k varies from 1 to n and j varies (in reverse order) from 0 to the largest nonnegative integer m_k such that $T^{m_k} v_k \neq 0$. With respect to this basis, the matrix of T looks like the matrix in the previous example. More specifically, T has a block diagonal matrix with respect to this basis, with each block a square matrix that is 0 everywhere except on the line above the diagonal.

In the next definition, the diagonal of each A_k is filled with some eigenvalue λ_k of T , the line directly above the diagonal of A_k is filled with 1's, and all other entries in A_k are 0 (to understand why each λ_k is an eigenvalue of T , see 5.41). The λ_k 's need not be distinct. Also, A_k may be a 1-by-1 matrix (λ_k) containing just an eigenvalue of T . If each λ_k is 0, then the next definition captures the behavior described in the paragraph above (recall that if T is nilpotent, then 0 is the only eigenvalue of T).

Definition 8.3.1. *Jordan basis*

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

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

in which each A_k is an upper-triangular matrix of the form

$$A_k = \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{pmatrix}$$

Theorem 8.3.3. *every nilpotent operator has a Jordan basis*

Suppose $T \in \mathcal{L}(V)$ is nilpotent. Then there is a basis of V that is a Jordan basis for T .

Theorem 8.3.4. *Jordan form*

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

Exercise

Suppose $T \in \mathcal{L}(\mathbb{C}^3)$ is the operator defined by $T(z_1, z_2, z_3) = (z_2, z_3, 0)$. Prove that T does not have a square root

Solution

$\dim \text{null} T = 1$, suppose there exists a square root R , then $\dim \text{null} R^2 = 1$, which is against what we learned in Chapter 8, since $\dim \text{null} R^4 = 2$, which require that the null space of R^k grows at least when $k \leq 4$. So $\dim \text{null} R^2 \geq 2$ \square

8.4 Trace: A Connection Between Matrices and Operators

Definition 8.4.1. *trace of a matrix*

Suppose A is a square matrix with entries in \mathbb{F} . The trace of A , denoted by $\text{tr} A$, is defined to be the sum of the diagonal entries of A .

Theorem 8.4.1. *trace of AB equals trace of BA*

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

$$\text{tr}(AB) = \text{tr}(BA)$$

Theorem 8.4.2. *trace of matrix of operator does not depend on basis*

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

$$\text{tr} \mathcal{M}(T, (u_1, \dots, u_n)) = \text{tr} \mathcal{M}(T, (v_1, \dots, v_n))$$

Definition 8.4.2. *trace of an operator*

Suppose $T \in \mathcal{L}(V)$. The trace of T , denote $\text{tr}T$, is defined by

$$\text{tr}T = \text{tr}\mathcal{M}(T, (v_1, \dots, v_n)),$$

where v_1, \dots, v_n is any basis of V .

Theorem 8.4.3. on complex vector spaces, trace equals sum of eigenvalues

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then $\text{tr}T$ equals the sum of the eigenvalues of T , with each eigenvalue included as many times as its multiplicity.

Theorem 8.4.4. trace and characteristic polynomial

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then $\text{tr}T$ equals the negative of the coefficient of z^{n-1} in the characteristic polynomial of T .

Theorem 8.4.5. trace on an inner product space

Suppose V is an inner product space, $T \in \mathcal{L}(V)$, and e_1, \dots, e_n is an orthonormal basis of V . Then

$$\text{tr}T = \langle Te_1, e_1 \rangle + \dots + \langle Te_n, e_n \rangle.$$

Theorem 8.4.6. trace is linear

The function $\text{tr} : \mathcal{L}(V) \rightarrow \mathbb{F}$ is a linear functional on $\mathcal{L}(V)$ such that

$$\text{tr}(ST) = \text{tr}(TS)$$

for all $S, T \in \mathcal{L}(V)$.

Theorem 8.4.7. identity operator is not the difference of ST and TS

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

Exercise

Prove that the trace is the **only** linear functional $\tau : \mathcal{L}(V) \rightarrow \mathbb{F}$ such that

$$\tau(ST) = \tau(TS)$$

for all $S, T \in \mathcal{L}(V)$ and $\tau(I) = \dim V$.

Hint: Suppose that v_1, \dots, v_n is a basis of V . For $j, k \in \{1, \dots, n\}$, define $P_{j,k} \in \mathcal{L}(V)$ by

$$P_{j,k}(a_1v_1 + \dots + a_nv_n) = a_kv_j$$

Prove that

$$\tau(P_{j,k}) = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases}$$

Then for $T \in \mathcal{L}(V)$, use the equation

$$T = \sum_{k=1}^n \sum_{j=1}^n \mathcal{M}(T)_{j,k} P_{j,k}$$

to show that $\tau(T) = \text{tr} T$.

9 Multilinear Algebra and Determinants

9.1 Bilinear Forms and Quadratic Forms

9.1.1 Bilinear Forms

Definition 9.1.1. *bilinearform*

A bilinear form on V is a function $\beta : V \times V \rightarrow \mathbb{F}$ such that

$$v \rightarrow \beta(v, u) \text{ and } v \rightarrow \beta(u, v)$$

are both linear functionals on V for every $u \in V$.

Definition 9.1.2. $V^{(2)}$

The set of bilinear forms on V is denoted by $V^{(2)}$

Definition 9.1.3. *matrix of a bilinear form, $\mathcal{M}(\beta)$*

Suppose β is a bilinear form on V and e_1, \dots, e_n is a basis of V . The matrix of β with respect to this basis is the n -by- n matrix $\mathcal{M}(\beta)$ whose entry

$\mathcal{M}(\beta)_{j,k}$ in row j , column k is given by

$$\mathcal{M}(\beta)_{j,k} = \beta(e_j, e_k)$$

If the basis e_1, \dots, e_n is not clear from the context, then the notation $\mathcal{M}(\beta, (e_1, \dots, e_n))$ is used.

Theorem 9.1.1. $\dim V^{(2)} = (\dim V)^2$

Suppose e_1, \dots, e_n is a basis of V . Then the map $\beta \rightarrow \mathcal{M}(\beta)$ is an **isomorphism** of $V^{(2)}$ onto $\mathbb{F}^{n,n}$. Furthermore, $\dim V^{(2)} = (\dim V)^2$.

Theorem 9.1.2. composition of a bilinear form and an operator

Suppose β is a bilinear form on V and $T \in \mathcal{L}(V)$. Define bilinear forms α and ρ on V by

$$\alpha(u, v) = \beta(u, Tv) \text{ and } \rho(u, v) = \beta(Tu, v).$$

Let e_1, \dots, e_n be a basis of V . Then

$$\mathcal{M}(\alpha) = \mathcal{M}(\beta)\mathcal{M}(T) \text{ and } \mathcal{M}(\rho) = \mathcal{M}(T)\mathcal{M}(\beta)$$

Theorem 9.1.3. change-of-basis formula

Suppose $\beta \in V^{(2)}$. Suppose e_1, \dots, e_n and f_1, \dots, f_n are bases of V . Let

$$A = \mathcal{M}(\beta, (e_1, \dots, e_n)) \text{ and } B = \mathcal{M}(\beta, (f_1, \dots, f_n))$$

and $C = \mathcal{M}(I, (e_1, \dots, e_n), (f_1, \dots, f_n))$. Then

$$A = C^t B C$$

9.1.2 Symmetric Bilinear Forms

Definition 9.1.4. symmetric bilinear form, $V_{sym}^{(2)}$

A bilinear form $\rho \in V^{(2)}$ is called symmetric if

$$\rho(u, w) = \rho(w, u)$$

for all $u, w \in V$. The set of symmetric bilinear forms on V is denoted by $V_{sym}^{(2)}$.

Definition 9.1.5. *symmetric matrix*

A square matrix A is called symmetric if it equals its transpose.

Theorem 9.1.4. *symmetric bilinear forms are diagonalizable*

Suppose $\rho \in V^{(2)}$. Then the following are equivalent.

- (a) ρ is a symmetric bilinear form on V .
- (b) $\mathcal{M}(\rho, (e_1, \dots, e_n))$ is a symmetric matrix for every basis e_1, \dots, e_n of V .
- (c) $\mathcal{M}(\rho, (e_1, \dots, e_n))$ is a symmetric matrix for some basis e_1, \dots, e_n of V .
- (d) $\mathcal{M}(\rho, (e_1, \dots, e_n))$ is a diagonal matrix for some basis e_1, \dots, e_n of V .

Theorem 9.1.5. *diagonalization of a symmetric bilinear form by an orthonormal basis*

Suppose V is a real inner product space and ρ is a symmetric bilinear form on V . Then ρ has a diagonal matrix with respect to some orthonormal basis of V .

Definition 9.1.6. *alternating bilinear form, $V_{alt}^{(2)}$*

A bilinear form $\alpha \in V^{(2)}$ is called alternating if

$$\alpha(v, v) = 0$$

for all $v \in V$. The set of alternating bilinear forms on V is denoted by $V_{alt}^{(2)}$.

Theorem 9.1.6. *characterization of alternating bilinear forms*

A bilinear form α on V is alternating if and only if

$$\alpha(u, w) = -\alpha(w, u)$$

for all $u, w \in V$.

Theorem 9.1.7. $V^{(2)} = V_{sym}^{(2)} \oplus V_{alt}^{(2)}$

The sets $V_{sym}^{(2)}$ and $V_{alt}^{(2)}$ are subspaces of $V^{(2)}$. Furthermore,

$$V^{(2)} = V_{sym}^{(2)} \oplus V_{alt}^{(2)}$$

9.1.3 Quadratic Forms

Definition 9.1.7. quadratic form associated with a bilinear form, q_β

For β a bilinear form on V , define a function $q_\beta : V \rightarrow \mathbb{F}$ by

$$q_\beta(v) = \beta(v, v).$$

A function $q : V \rightarrow \mathbb{F}$ is called a quadratic form on V if there exists a bilinear form β on V such that $q = q_\beta$

Theorem 9.1.8. quadratic forms on \mathbb{F}^n

Suppose n is a positive integer and q is a function from \mathbb{F}^n to \mathbb{F} . Then q is a quadratic form on \mathbb{F}^n if and only if there exist numbers $A_{j,k} \in \mathbb{F}$ for $j, k \in \{1, \dots, n\}$ such that

$$q(x_1, \dots, x_n) = \sum_{k=1}^n \sum_{j=1}^n A_{j,k} x_j x_k$$

for all $(x_1, \dots, x_n) \in \mathbb{F}^n$

Theorem 9.1.9. characterization of quadratic forms

Suppose $q : V \rightarrow \mathbb{F}$ is a function. The following are equivalent.

- (a) q is a quadratic form.
- (b) There exists a unique symmetric bilinear form ρ on V such that $q = q_\rho$.
- (c) $q(\lambda v) = \lambda^2 q(v)$ for all $\lambda \in \mathbb{F}$ and all $v \in V$, and the function

$$(u, w) \rightarrow q(u + w) - q(u) - q(w)$$

is a symmetric bilinear form on V .

(d) $q(2v) = 4q(v)$ for all $v \in V$, and the function

$$(u, w) \rightarrow q(u + w) - q(u) - q(w)$$

is a symmetric bilinear form on V .

Theorem 9.1.10. *diagonalization of quadratic form*

Suppose q is a quadratic form on V .

(a) There exist a basis e_1, \dots, e_n of V and $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ such that

$$q(x_1e_1 + \dots + x_ne_n) = \lambda_1x_1^2 + \dots + \lambda_nx_n^2$$

for all $x_1, \dots, x_n \in \mathbb{F}$.

(b) If $\mathbb{F} = \mathbb{R}$ and V is an inner product space, then the basis in (a) can be chosen to be an orthonormal basis of V .

The exercise in this Section is not hard with the definition and theorem here.

9.2 Alternating Multilinear Forms

9.2.1 Multilinear Forms

Definition 9.2.1. V^m

For m a positive integer, define V^m by

$$V^m = \underbrace{V \times \dots \times V}_{m \text{ times}}$$

Definition 9.2.2. m -linear form, $V^{(m)}$, multilinear form

□ For m a positive integer, an m -linear form on V is a function $\beta : V^m \rightarrow \mathbb{F}$ that is linear in each slot when the other slots are held fixed. This means that for each $k \in \{1, \dots, m\}$ and all $u_1, \dots, u_m \in V$, the function

$$v \rightarrow \beta(u_1, \dots, u_{k-1}, v, u_{k+1}, \dots, u_m)$$

is a linear map from V to \mathbb{F} .

- The set of m -linear forms on V is denoted by $V^{(m)}$.
- A function β is called a multilinear form on V if it is an m -linear form on V for some positive integer m .

Definition 9.2.3. *alternating forms, $V_{alt}^{(m)}$*

Suppose m is a positive integer.

- An m -linear form α on V is called alternating if $\alpha(v_1, \dots, v_m) = 0$ whenever v_1, \dots, v_m is a list of vectors in V with $v_j = v_k$ for some two distinct values of j and k in $\{1, \dots, m\}$.

$$V_{alt}^{(m)} = \{\alpha \in V^{(m)} : \alpha \text{ is an alternating } m\text{-linear form on } V\}.$$

Theorem 9.2.1. *alternating multilinear forms and linear dependence*

Suppose m is a positive integer and α is an alternating m -linear form on V . If v_1, \dots, v_m is a linearly dependent list in V , then

$$\alpha(v_1, \dots, v_m) = 0.$$

Theorem 9.2.2. *no nonzero alternating m -linear forms for $m > \dim V$*

Suppose $m > \dim V$. Then 0 is the only alternating m -linear form on V .

9.2.2 Alternating Multilinear Forms and Permutations

Theorem 9.2.3. *swapping input vectors in an alternating multilinear form*

Suppose m is a positive integer, α is an alternating m -linear form on V , and v_1, \dots, v_m is a list of vectors in V . Then swapping the vectors in any two slots of $\alpha(v_1, \dots, v_m)$ changes the value of α by a factor of -1 .

Definition 9.2.4. *permutation, perm*

Suppose m is a positive integer.

- A permutation of $(1, \dots, m)$ is a list (j_1, \dots, j_m) that contains each of the numbers $1, \dots, m$ exactly once.

□ The set of all permutations of $(1, \dots, m)$ is denoted by permm .

Definition 9.2.5. *sign of a permutation*

The sign of a permutation (j_1, \dots, j_m) is defined by

$$\text{sign}(j_1, \dots, j_m) = (-1)^N,$$

where N is the number of pairs of integers (k, l) with $1 \leq k < l \leq m$ such that k appears after l in the list (j_1, \dots, j_m)

Theorem 9.2.4. *swapping two entries in a permutation*

Swapping two entries in a permutation multiplies the sign of the permutation by -1 .

Theorem 9.2.5. *permutations and alternating multilinear forms*

Suppose m is a positive integer and $\alpha \in V_{\text{alt}}^{(m)}$. Then

$$\alpha(v_{j_1}, \dots, v_{j_m}) = (\text{sign}(j_1, \dots, j_m))\alpha(v_1, \dots, v_m)$$

for every list v_1, \dots, v_m of vectors in V and all $(j_1, \dots, j_m) \in \text{permm}$

Theorem 9.2.6. *formula for $(\dim V)$ -linear alternating forms on V*

Let $n = \dim V$. Suppose e_1, \dots, e_n is a basis of V and $v_1, \dots, v_n \in V$. For each $k \in \{1, \dots, n\}$, let $b_{1,k}, \dots, b_{n,k} \in \mathbb{F}$ be such that

$$v_k = \sum_{j=1}^n b_{j,k} e_j$$

Then

$$\alpha(v_1, \dots, v_n) = \sum_{(j_1, \dots, j_n) \in \text{permm}} (\text{sign}(j_1, \dots, j_n)) b_{j_1,1} \cdots b_{j_n,n}$$

for every alternating n -linear form α on V

Theorem 9.2.7. $\dim V^{(\dim V)}_{\text{alt}} = 1$

The vector space $V^{(\dim V)}_{\text{alt}}$ has dimension one.

Theorem 9.2.8. *alternating $(\dim V)$ -linear forms and linear independence*

Let $n = \dim V$. Suppose α is a nonzero alternating n -linear form on V and e_1, \dots, e_n is a list of vectors in V . Then

$$\alpha(e_1, \dots, e_n) \neq 0$$

if and only if e_1, \dots, e_n is linearly independent

9.3 Determinants

9.4 Tensor Products