1 Vector spaces

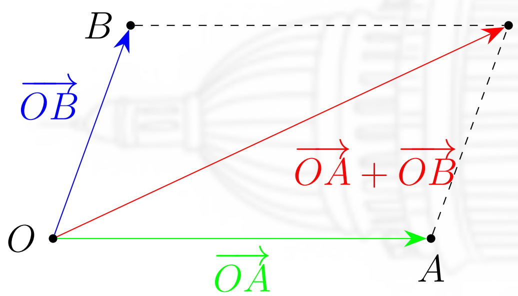
Let $F = \mathbb{R}$ or $\mathbb{C}$.

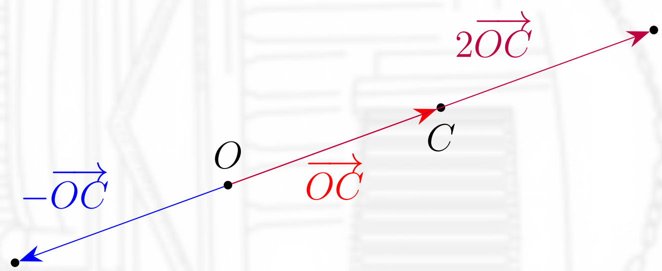
Motivating examples

Geometric vector plane

Let O be a fixed point in the plane called the **origin**. Let V be the set of directed line segments starting at O. If A is a point on the plane, the corresponding directed line segment is denoted by $\overrightarrow{OA}$.

One defines the sum $\overrightarrow{OA} + \overrightarrow{OB}$ and the scalar product of $a \in \mathbb{R}$ and $\overrightarrow{OA}$:





Properties of the sum and the scalar product:

I. $(\overrightarrow{OA} + \overrightarrow{OB}) + \overrightarrow{OC} = \overrightarrow{OA} + (\overrightarrow{OB} + \overrightarrow{OC})$ for any points A, B, C;

II. $\overrightarrow{OA} + \overrightarrow{OO} = \overrightarrow{OA}$ for any point A;

III. for any point A there is a point B such that $\overrightarrow{OA} + \overrightarrow{OB} = \overrightarrow{OO}$;

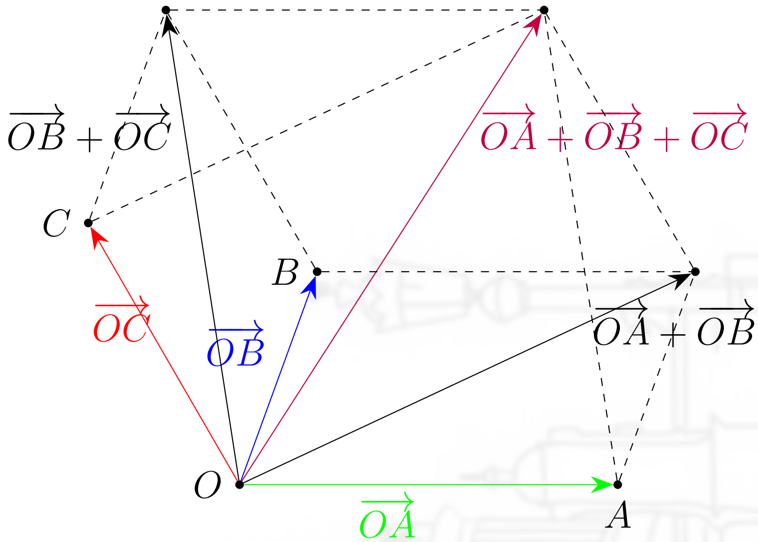
IV. $\overrightarrow{OA} + \overrightarrow{OB} = \overrightarrow{OB} + \overrightarrow{OA}$ for any points A, B;

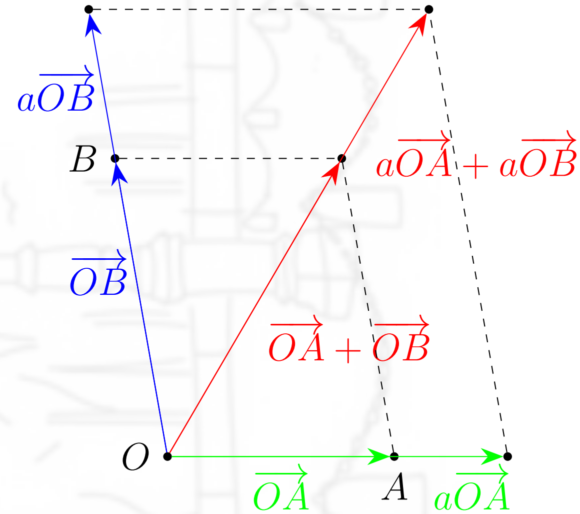
V. $a(\overrightarrow{OA} + \overrightarrow{OB}) = a\overrightarrow{OA} + a\overrightarrow{OB}$ for any points A, B and any $a \in \mathbb{R}$;

VI. $(a + b)\overrightarrow{OA} = a\overrightarrow{OA} + b\overrightarrow{OA}$ for any point A, $a, b \in \mathbb{R}$;

VII. $(ab)\overrightarrow{OA} = a(b\overrightarrow{OA})$ for any point A, $a, b \in \mathbb{R}$;

VIII. $1 \cdot \overrightarrow{OA} = \overrightarrow{OA}$ for any point A.





Polynomials over a field

Consider the polynomials over F. Here is a list of properties of their sum and product by a number:

I. $(f+g)+h = f+(g+h)$ for any $f, g, h \in F[t]$;

II. if $\mathbf{0}$ denotes the zero polynomial then $f + \mathbf{0} = f$ for any $f \in F[t]$;

III. for any $f \in F[t]$ one has $f + (-f) = \mathbf{0}$;

IV. $f + g = g + f$ for any $f, g \in F[t]$;

V. $a(f+g) = af + ag$ for any $f, g \in F[t]$, $a \in F$;

VI. $(a+b)f = af + bf$ for any $f \in F[t]$, $a, b \in F$;

VII. $(ab)f = a(bf)$ for any $f \in F[t]$, $a, b \in F$;

VIII. $1 \cdot f = f$ for any $f \in F[t]$.

Coordinate space

Let $F^n = \{(a_1, \ldots, a_n) \mid a_i \in F\}$. Define

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$$

$$a(a_1, \ldots, a_n) = (aa_1, \ldots, aa_n).$$

One can consider F^n as $1 \times n$ matrices with standard matrix addition and multiplication by a scalar. Then the following properties are satisfied:

I. $(A+B)+C = A+(B+C)$ for any $A, B, C \in F^n$;

II. $A + (0, \ldots, 0) = A$ for any $A \in F^n$;

III. for any $A = \in F^n$ $A + (-A) = (0, \dots, 0)$;

IV. $A + B = B + A$ for any $A, B \in F^n$;

V. $a(A + B) = aA + aB$ for any $A, B \in F^n$, $a \in F$;

VI. $(a + b)A = aA + bA$ for any $A \in F^n$, $a, b \in F$;

VII. $(ab)A = a(bA)$ for any $A \in F^n$, $a, b \in F$;

VIII. $1 \cdot A = A$ for any $A \in F^n$.

Definition of a vector space and further examples

Notice that the above sets of properties are similar.

Definition. A **vector space** (向量空间) over F is a set V, whose elements are called **vectors** (向量), together with two operations. The first operation, called **vector addition** (向量加法), assigns to each pair of vectors u and v a vector denoted by $u + v$, called their **sum** (加和). The second operation, called **scalar multiplication** (标量乘法), assigns to each vector v and each number $a \in F$ a vector denoted by av. The two operations are required to have the following properties:

I. $(u + v) + w = u + (v + w)$ for any $u, v, w \in V$ (*associative law of vector addition*);

II. there is $\mathbf{0} \in V$ such that $v + \mathbf{0} = v$ for any $v \in V$ (the existence of the *additive identity*);

III. for any $v \in V$ there is $-v \in V$ such that $v + (-v) = \mathbf{0}$ (the existence of the *additive inverse*);

IV. $u + v = v + u$ for any $u, v \in V$ (*commutative law of vector addition*);

V. $a(u + v) = a \cdot u + a \cdot v$ for any $u, v \in V$, $a \in F$ (*right distributivity*);

VI. $(a + b)v = a \cdot v + b \cdot v$ for any $v \in V$, $a, b \in F$ (*left distributivity*);

VII. $(a \cdot b) \cdot v = a \cdot (b \cdot v)$ for any $v \in V$, $a, b \in F$;

VIII. $1 \cdot v = v$ for any $v \in V$.

Remark. Note that the signs $+$ and $\cdot$ in the axioms have different meanings: $+$ denotes addition both in V and F, and $\cdot$ denotes scalar multiplication in V and multiplication in F.

Examples. 1. Three motivating examples

2. *Zero vector space*

   Let $V = \{O\}$ and $O + O = O, a \cdot O = O$ for $a \in F$.

3. The geometric vector space of directed line segments in the space starting at the origin.

4. The set $\mathrm{M}_{m,n}(F)$ of $m \times n$–matrices over F with matrix addition and scalar multiplication

5. The set $\mathbb{C}$ with standard addition and scalar multiplication by *real numbers*

6. *The exotic vector space*

   Consider $V = \{v \in \mathbb{R} \mid v > 0\}$ with addition and scalar multiplication given by $u+v = uv$ for $u, v \in V$ and $a \cdot v = v^a$ for $v \in V, a \in \mathbb{R}$.

   Clearly, $1 \in V$ is the additive identity ($v + 1 = v \cdot 1 = v$) and v^{-1} is the additive inverse of $v \in V$ ($v + v^{-1} = vv^{-1} = 1$). Right distributivity follows from the identities $a(u + v) = a^{uv} = a^u a^v = au + av$ for $u, v \in V, a \in \mathbb{R}$.

7. The set $C[a, b]$ of continuous functions on $[a, b]$ with

$$(f + g)(x) = f(x) + g(x), \quad (cf)(x) = cf(x), \quad f, g \in C[a, b], c \in \mathbb{R}$$

8. *Product of vector spaces*

   Let U, V be vector spaces over F. On the set $U \times V = \{(u, v) \mid u \in U, v \in V\}$ define addition

$$(u, v) + (u', v') = (u + u', v + v'), \qquad u, u' \in U, v, v' \in V$$

   and scalar multiplication

$$a(u, v) = (au, av), \qquad u \in U, v \in V, a \in F.$$

   It is easy to check that all the axioms of vector space are satisfied. For example, $(\mathbf{0}_U, \mathbf{0}_V)$ is the additive identity in $U \times V$ since

$$(u, v) + (\mathbf{0}_U, \mathbf{0}_V) = (u + \mathbf{0}_U, v + \mathbf{0}_V) = (u, v)$$

   and right distributivity follows from the identities $a((u, v) + (u', v')) = a(u + u', v + v') = (a(u + u'), a(v + v')) = (au + au', av + av') = (au, av) + (au', av') = a(u, v) + a(u', v')$.

   The resulting vector space is called the **product** (乘积空间) of U, V and is denoted by $U \times V$.

Exercise 1.1. *Prove the rest of the axioms for the exotic vector space.*

Exercise 1.2. *Prove the rest of the axioms for the product of vector spaces.*

Basic properties of vector spaces

Proposition 1.1. *Let V be a vector space over F. Then*

1. $\mathbf{0}$ is unique;

2. $-v$ is unique for any $v \in V$;

3. $0 \cdot v = \mathbf{0}$ *for any* $v \in V$;

4. $a \cdot \mathbf{0} = \mathbf{0}$ *for any* $a \in F$;

5. $(-1) \cdot v = -v$ *for any* $v \in V$;

6. *if* $av = \mathbf{0}$ *for* $v \in V, a \in F$ *then* $a = 0$ *or* $v = \mathbf{0}$.

Proof. 1. Assume there are $\mathbf{0}$ and $\mathbf{0}'$ such that $v = v + \mathbf{0}$ and $v = v + \mathbf{0}'$ for any $v \in V$. Then $\mathbf{0} = \mathbf{0} + \mathbf{0}' = \mathbf{0}'$.

2. Assume there are $-v$ and $-v'$ such that $v + (-v) = \mathbf{0}$ and $v + (-v') = \mathbf{0}$ for some $v \in V$. Then

$$-v' = \mathbf{0} + (-v') = (-v + v) + (-v') = -v + (v + (-v')) = -v + \mathbf{0} = -v.$$

3. Note that $0 \cdot v = (0+0) \cdot v = 0 \cdot v + 0 \cdot v$. There is $-0 \cdot v \in V$ such that $0 \cdot v + (-0 \cdot v) = \mathbf{0}$. Then

$$\mathbf{0} = 0 \cdot v + (-0 \cdot v) = (0 \cdot v + 0 \cdot v) + (-0 \cdot v) = 0 \cdot v + (0 \cdot v + (-0 \cdot v)) = 0 \cdot v + \mathbf{0} = 0 \cdot v.$$

4. Similarly, $a \cdot \mathbf{0} = a \cdot (\mathbf{0} + \mathbf{0}) = a \cdot \mathbf{0} + a \cdot \mathbf{0}$. Then

$$\mathbf{0} = a \cdot \mathbf{0} + (-a \cdot \mathbf{0}) = (a \cdot \mathbf{0} + a \cdot \mathbf{0}) + (-a \cdot \mathbf{0}) = a \cdot \mathbf{0} + (a \cdot \mathbf{0} + (-a \cdot \mathbf{0})) = a \cdot \mathbf{0} + \mathbf{0} = a \cdot \mathbf{0}.$$

5. By the third property

$$\mathbf{0} = 0 \cdot v = (1 + (-1)) \cdot v = 1 \cdot v + (-1) \cdot v = v + (-1) \cdot v.$$

Then

$$-v = -v + \mathbf{0} = -v + (v + (-1) \cdot v) = (-v + v) + (-1) \cdot v = \mathbf{0} + (-1) \cdot v = (-1) \cdot v.$$

6. if $a \neq 0$ then there is $a^{-1} \in F$. Then

$$v = 1 \cdot v = (a^{-1} \cdot a)v = a^{-1} \cdot (av) = a^{-1} \cdot \mathbf{0} = \mathbf{0}$$

by the forth property.

□

Remark. In above proof the following trick was used. Let $u, v, w \in V$ and $u + v = w + v$. Then

$$u = u + \mathbf{0} = u + (v + (-v)) = (u + v) + (-v) = (w + v) + (-v) = w + (v + (-v)) = w + \mathbf{0} = w$$

which is called **cancelling out**.

Exercise 1.3. *Prove that* $(v_1 + v_2) + (v_3 + v_4) = (v_2 + v_4) + (v_1 + v_3)$ *for any vectors* $v_1, v_2, v_3, v_4 \in V$.

Exercise 1.4. *Let* $v \in V, a \in F$. *Prove that if* $av = bv$ *then* $a = b$ *or* $v = \mathbf{0}$.

2 Subspaces

Definition and examples

Definition. Let V be a vector space over F. A non-empty $U \subset V$ is called a **vector subspace** (向量子空间) (or just a **subspace** (子空间)) of V if:

1. $u + v \in U$ for any $u, v \in U$;
2. $au \in U$ for any $u \in U$, $a \in F$.

Proposition 2.1. *If U is a subspace of V then U itself is a vector space under* induced operations, *i.e., under addition and scalar multiplication defined as in V.*

Proof. Notice that conditions in the definition of subspace assure us that the operations on U are correctly defined.

Let $u_0 \in U$ then $\mathbf{0} = 0 \cdot u_0 \in U$ by Proposition 1.1. The fact that $u + \mathbf{0} = u$ for any $u \in U$ is valid since it is valid in V.

If $u \in U$ then $-u = (-1) \cdot u \in U$ by Proposition 1.1, and therefore the additive inverse exists.

The other axioms are valid for vectors in U since they are valid for vectors in the bigger set V. $\square$

Examples.

1. $\{\mathbf{0}\}$ and V are trivial subspaces of V;
2. A line passing through the origin is a subspace of the geometric vector plane;
3. A ray the origin is not a subspace of the geometric vector plane;
4. $U = \{f \in F[t] \mid f(1) = 0\}$ is a subspace of $F[t]$;
5. $U = F[t]_n = \{f \in F[t] \mid \deg f \leq n\}$ is a subspace of $F[t]$;
6. $U = \{f \in F[t] \mid \deg f = n\}$ is not a subspace of $F[t]$;
7. $\{(a_1, \ldots, a_n) \in F^n \mid a_1 + \cdots + a_n = 0\}$ is a subspace of F^n;
8. $\{(a_1, \ldots, a_n) \in F^n \mid a_1 + \cdots + a_n = 1\}$ is not a subspace of F^n;
9. $\{A = (a_1, \ldots, a_n) \in F^n \mid CA^T = 0\}$ is a subspace of F^n, where $C \in \mathrm{M}_{m,n}(F)$ is a fixed matrix.

Problem 2.2. *Prove that $\{A \in \mathrm{M}_n(\mathbb{R}) \mid A^T = A\}$ is a subspace of $\mathrm{M}_n(\mathbb{R})$*

Solution. If $A, B \in \mathrm{M}_n(\mathbb{R})$ and $A^T = A, B^T = B$ then $(A + B)^T = A^T + B^T = A + B$.

If $A \in \mathrm{M}_n(\mathbb{R})$, $A^T = A$ and $a \in \mathbb{R}$ then $(aA)^T = aA^T = aA$. $\square$

Exercise 2.1. *Prove that $\{f \in C[-1, 1] \mid f(-t) = f(t)\}$ is a subspace of $C[-1, 1]$*

Exercise 2.2. *Prove that $\{(f, g) \in \mathbb{R}[t] \times \mathbb{R}[t] \mid f(0) = g(1), f'(1) = g'(0)\}$ is a subspace of $\mathbb{R}[t] \times \mathbb{R}[t]$*

Linear combinations and linear span

Definition. Let V be a vector space over F and $v_1, \dots, v_n \in V$. A **linear combination** (线性组合) of the vectors $v_1, \dots, v_n$ is a vector in V of the form $a_1v_1+\cdots+a_nv_n$, where $a_1, \dots, a_n \in F$.

The **linear span** (线性生成空间) (or just the **span** (生成空间)) of the vectors $v_1, \dots, v_n$ is

$$\mathrm{Span}(v_1, \dots, v_n) = \{a_1v_1 + \cdots + a_nv_n \mid a_1, \dots, a_n \in F\}.$$

The vectors $v_1, \dots, v_n$ **span** (生成) V or $v_1, \dots, v_n$ is a **spanning set** (生成集合) of V if $\mathrm{Span}(v_1, \dots, v_n) = V$.

Examples.

1. Let $V = F^3$. Then the vectors $v_1 = (1, 1, 0), v_2 = (1, 0, 1), v_3 = (0, 1, 1)$ span V. Indeed, let $v = (a, b, c)$ be an arbitrary element of V. Then the identity $xv_1 + yv_2 + xv_3 = v$ is equivalent to the system

$$\begin{cases} x + y = a \\ x + z = b \\ y + z = c \end{cases}$$

which is solvable.

2. Let v_1, v_2 be two vectors in the geometric vector plane then $\mathrm{Span}(v_1, v_2)$ is the line containing v_1, v_2 if v_1, v_2 are collinear and the plane containing v_1, v_2 if they are not. Thus v_1, v_2 span the plane if and only if they are not collinear.

Proposition 2.3. *Let V be a vector space and $v_1, \dots, v_n \in V$. Then*

1. $\mathrm{Span}(v_1, \dots, v_n)$ *is a subspace of V;*
2. $\mathrm{Span}(v_1, \dots, v_n)$ *is the minimum subspace of V containing $v_1, \dots, v_n$, i.e., if U is a subspace of V and $v_1, \dots, v_n \in U$ then* $\mathrm{Span}(v_1, \dots, v_n) \subset U$.

Proof. 1. If $u, u' \in \mathrm{Span}(v_1, \dots, v_n)$ then $u = a_1v_1 + \cdots + a_nv_n, u' = a'_1v_1 + \cdots + a'_nv_n$, where $a_1, a'_1 \dots, a_n, a'_n \in F$. Then

$$u+u' = a_1v_1+\cdots+a_nv_n+a'_1v_1+\cdots+a'_nv_n = (a_1+a'_1)v_1+\cdots+(a_n+a'_n)v_n \in \mathrm{Span}(v_1, \dots, v_n).$$

Similarly, for $a \in F$

$$au = a(a_1v_1 + \cdots + a_nv_n) = (aa_1)v_1 + \cdots + (aa_n)v_n \in \mathrm{Span}(v_1, \dots, v_n).$$

2. Any $u \in \mathrm{Span}(v_1, \dots, v_n)$ can be expressed as $u = a_1v_1 + \cdots + a_nv_n$ for some $a_1, \dots, a_n \in F$. If $v_1, \dots, v_n \in U$ then $u \in U$ since U is a subspace. □

Exercise 2.3. *Let $v_1 = (1, 1, 1, -1), v_2 = (0, -1, -1, 1), v_3 = (1, 1, 0, -1) \in \mathbb{R}^4$. Does v belong to* $\mathrm{Span}(v_1, v_2, v_3)$ *if*

a. $v = (0, 1, 0, 1)$

b. $v = (1, 0, 1, 0)$

Hint. The problem is equivalent to the solvability of the system

$$\begin{cases} x \phantom{{}-y} + z = a \\ x - y + z = b \\ x - y \phantom{{}+z} = c \\ -x + y - z = d \end{cases}$$

for the corresponding $a, b, c, d \in \mathbb{R}$.

Exercise 2.4. *Do the following sets of vectors span* $\mathbb{R}[t]_2 = \{f \in \mathbb{R}[t] \mid \deg f \leq 2\}$*?*

a. $f_1 = 1 + t + t^2$, $f_2 = 1 - t^2$

b. $f_1 = 1 + t + t^2$, $f_2 = 1 - t^2$, $f_3 = t + t^2$

c. $f_1 = 1 + t + t^2$, $f_2 = 1 - t^2$, $f_3 = t + 2t^2$

Hint. The problem is equivalent to the solvability of a certain system of linear equations for *any* constant terms.

Exercise 2.5. *Let V be a vector space and $v_1, v_2, v_3 \in V$ be its spanning set. Show that $v_1 + v_2, v_2 - v_3, v_1 + v_2 + v_3$ is also a spanning set of V.*

Proposition 2.4. *If U_1, U_2 are subspaces of V then $U_1 \cap U_2$ is a subspace of V.*

Proof. Suppose that $u, u' \in U_1 \cap U_2$. Then $u, u' \in U_1$. Since U_1 is a subspace, $u + u' \in U_1$. Likewise, $u + u' \in U_2$. Therefore, $u + u' \in U_1 \cap U_2$.

Similarly, if $a \in F$ and $u \in U_1 \cap U_2$ then $u \in U_1$ and $au \in U_1$ since U_1 is a subspace and $au \in U_2$ since U_2 is a subspace. Then $au \in U_1 \cap U_2$. $\square$

Sum and direct sum

Definition. Let $U_1, \ldots, U_m$ be subspaces of a vector space V. The **sum** (加和) of $U_1, \ldots, U_m$ is defined as

$$U_1 + \cdots + U_m = \{u_1 + \cdots + u_m \mid u_1 \in U_1, \ldots, u_m \in U_m\}.$$

Proposition 2.5. *Let $U_1, \ldots, U_m$ be subspaces of V. Then*

1. *$U_1 + \cdots + U_m$ is a subspace of V;*

2. *$U_1 + \cdots + U_m$ is the minimum subspace of V containing $U_1, \ldots, U_m$, i.e., if U is a subspace of V and $U_1, \ldots, U_m \subset U$ then $U_1 + \cdots + U_m \subset U$.*

Proof. 1. If $u, u' \in U_1 + \cdots + U_m$ then $u = u_1 + \cdots + u_m$, $u' = u'_1 + \cdots + u'_m$, where $u_i, u'_i \in U_i$. Then

$$u + u' = (u_1 + \cdots + u_m) + (u'_1 + \cdots + u'_m) = (u_1 + u'_1) + \cdots + (u_m + u'_m) \in U_1 + \cdots + U_m$$

since $u_i + u'_i \in U_i$.

Similarly, for $a \in F$

$$au = a(u_1 + \cdots + u_m) = au_1 + \cdots + au_m \in U_1 + \cdots + U_m$$

since $au_i \in U_i$.

2. For any $u \in U_1 + \cdots + U_m$ one has $u = u_1 + \cdots + u_m$ for some $u_i \in U_i \subset U$. Then $u \in U$ since U is a subspace. □

Exercise 2.6. *Let V be a vector space and $v_1, \ldots, v_n, v'_1, \ldots, v'_m \in V$. Prove that*

$$\mathrm{Span}(v_1, \ldots, v_n) + \mathrm{Span}(v'_1, \ldots, v'_m) = \mathrm{Span}(v_1, \ldots, v_n, v'_1, \ldots, v'_m).$$

Exercise 2.7. *Suppose that U_1, U_2 are subspaces of V. Prove that $U_1 \cup U_2$ is a subspace if and only if $U_1 \subset U_2$ or $U_2 \subset U_1$.*

Definition. Let $U_1, \ldots, U_m$ be subspaces of a vector space V. One says that V is the **direct sum** (直和) of $U_1, \ldots, U_m$ (which is denoted by $V = U_1 \oplus \cdots \oplus U_m$), if each $v \in V$ can be *uniquely* represented as $v = u_1 + \cdots + u_m$, where all $u_i \in U_i$.

Examples.

1. Let $V = F^n$. Let $U_1 = \{(*, 0, \ldots, 0)\}$, $U_2 = \{(0, *, 0, \ldots, 0)\}$,..., $U_n = \{(0, \ldots, 0, *)\}$. Then $V = U_1 \oplus \cdots \oplus U_n$.
2. The plane is the direct sum of any two unequal lines passing through the origin.
3. Let $V = F^3$, $U_1 = \{(a, a, 0) \mid a \in F\}$, $U_2 = \{(b, 0, 0) \mid b \in F\}$, $U_3 = \{(0, c, d) \mid c, d \in F\}$. Then $V = U_1 + U_2 + U_3$, but this sum is not direct. For example, $(0, 1, 0) = (1, 1, 0) + (-1, 0, 0) + (0, 0, 0)$, and, alternatively, $(0, 1, 0) = (0, 0, 0) + (0, 0, 0) + (0, 1, 0)$.

Exercise 2.8. *For V, U_1, U_2, U_3 from the above example, prove that $V = U_1 \oplus U_3$ and $V = U_2 \oplus U_3$.*

Proposition 2.6. *Let U, W be subspaces of a vector space V. Then $V = U \oplus W$ if and only if $V = U + W$ and $U \cap W = \{0\}$.*

Proof. If $V = U \oplus W$, then clearly $V = U + W$. If $v \in U \cap W$, then $0 = v + (-v)$, where $v \in U$, $(-v) \in W$. Now the uniqueness implies $v = 0$, thus $U \cap W = \{0\}$.

Assume that $V = U + W$ and $U \cap W = \{0\}$. Let $u + w = u' + w'$, where $u, u' \in U$, $w, w' \in W$. Then $u - u' = w' - w \in U \cap W$. Thus $u - u' = w' - w = 0$, whence $u = u'$ and $w = w'$. □

Exercise 2.9. *Let $V = F^n$ and $U_1 = \{(x_1, \ldots, x_n) \mid x_1 + \cdots + x_n = 0, x_i \in F\}, U_2 = \{(x, \ldots, x) \mid x \in F\}$. Prove that U_1, U_2 are subspaces of V and $V = U_1 \oplus U_2$.*

Exercise 2.10. *Let $V = \mathbb{R}[t]_n = \{f \in \mathbb{R}[x] \mid \deg f \le n\}$ and $U_1 = \{f \in \mathbb{R}[t]_n \mid f(-t) = f(t)\}, U_2 = \{f \in \mathbb{R}[t]_n \mid f(-t) = -f(t)\}$. Prove that U_1, U_2 are subspaces of V and $V = U_1 \oplus U_2$.*

Exercise 2.11. *Let $V = \mathrm{M}_2(\mathbb{R})$ and*

$$U_1 = \left\{\begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix} \mid x \in F\right\}, U_2 = \left\{\begin{pmatrix} a & b \\ c & c \end{pmatrix} \mid a+b+c=0, a,b,c \in F\right\}, U_3 = \left\{\begin{pmatrix} y & y \\ y & 0 \end{pmatrix} \mid y \in F\right\}$$

Prove that U_1, U_2, U_3 are subspaces of V and $V = U_1 \oplus U_2 \oplus U_3$.

3 Basis

Linear independence and dependence

Definition. Let V be a vector space over F. Vectors $v_1, \dots, v_n$ are **linearly dependent** (线性相关) if there exist $a_1, \dots, a_n \in F$ *not all zero* such that $a_1v_1 + \cdots + a_nv_n = 0$ (**non-trivial zero linear combination** (非零的线性组合)). Otherwise $v_1, \dots, v_n$ are **linearly independent** (线性无关).

Examples. 1. Let $V = F^3$. Then the vectors $v_1 = (1,1,0), v_2 = (1,0,1), v_3 = (0,1,1)$ are linearly independent. Indeed, one can see that the identity $xv_1 + yv_2 + zv_3 = 0$ is equivalent to the system

$$\begin{cases} x+y=0 \\ y+z=0 \\ x+z=0 \end{cases}$$

whose unique solution is $x = y = z = 0$.

2. The polynomials $f_1 = t^3 + t^2 + 1, f_2 = 2t^3 + t^2 - t, f_3 = t^2 + t + 2 \in \mathbb{R}[t]$ are linearly dependent because $2f_1 - f_2 - f_3 = 0$.

3. Non-zero vectors $v, u \in V$ are linearly dependent if and only if $v = au$ for some $a \in F$.

4. A set of vectors that contains the zero vector is linearly dependent.

Exercise 3.1. *Let $V = \mathbb{R}^3$. Are the following vectors linearly independent?*

a. $v_1 = (1,0,-1), v_2 = (0,1,1), v_3 = (1,-1,0)$

b. $v_1 = (1,1,0), v_2 = (-1,0,1), v_3 = (0,-1,-1)$

Exercise 3.2. *Let $V = M_2(\mathbb{R})$. Prove that the following matrices are linearly independent:*

$$\begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

Exercise 3.3. *Let V be a vector space and $v_1, v_2, v_3 \in V$ be linearly independent. Show that $v_1 + v_2, v_2 - v_3, v_1 + v_2 + v_3$ are also linearly independent.*

Exercise 3.4. *Show that a set of vectors containing two equal vectors is linearly dependent.*

Lemma 3.1. *Let V be a vector space, $X \subset Y$ be finite sets of vectors in V. If Y is linearly independent then X is linearly independent; if X is linearly dependent, then Y is linearly dependent.*

Proof. A non-trivial zero linear combination of elements of X is a non-trivial zero linear combination of elements of Y. □

Lemma 3.2 (Linear Dependence Lemma)**.** *Let V be a vector space over F. Let $v_1, \dots, v_n \in V$ be linearly dependent and $v_1 \neq 0$. Then there exists $2 \leq j \leq n$ such that $v_j \in \operatorname{Span}(v_1, \dots, v_{j-1})$.*

Moreover, $\operatorname{Span}(v_1, \dots, v_n) = \operatorname{Span}(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n)$.

Proof. There are $a_1, \dots, a_n \in F$ not all zero such that $a_1 v_1 + \dots + a_n v_n = 0$. Let a_j be the nonzero coefficient with the maximum index. Then $v_j = -\frac{a_1}{a_j} v_1 - \dots - \frac{a_{j-1}}{a_j} v_{j-1}$. If $j = 1$ then $a_2 = \dots = a_n = 0$ whence $v_1 = 0$, a contradiction. Thus $j \geq 2$.

It is obvious that $\operatorname{Span}(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n) \subset \operatorname{Span}(v_1, \dots, v_n)$. To show the inverse inclusion, pick up $u \in \operatorname{Span}(v_1, \dots, v_n)$. Then for some $c_1, \dots, c_n \in F$

$$u = c_1 v_1 + \dots + c_n v_n = c_1 v_1 + \dots + c_{j-1} v_{j-1} + c_j\left(-\frac{a_1}{a_j} v_1 - \dots - \frac{a_{j-1}}{a_j} v_{j-1}\right) + c_{j+1} v_{j+1} + \dots + c_n v_n,$$

whence u is a linear combination of $v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n$. □

Corollary 3.3. *Let $v_1, \dots, v_n \in V$ be linearly independent and $v \in V$. The vectors $v_1, \dots, v_n, v$ are linearly dependent if and only if $v \in \operatorname{Span}(v_1, \dots, v_n)$.*

Proof. If $v_1, \dots, v_n, v$ are linearly dependent then some of these vectors is expressed as a linear combination of the preceding ones by Lemma 3.2. It cannot be one of $v_1, \dots, v_n$ due to their linear independence. The inverse statement is obvious. □

Theorem 3.4. *Let V be a vector space. If $u_1, \dots, u_m \in V$ are linearly independent and $v_1, \dots, v_n \in V$ are a spanning set, then $m \leq n$.*

Proof. By Corollary 3.3 the vectors $u_1, v_1, \dots, v_n$ are linearly dependent. By Lemma 3.2, there is a vector v_j, $1 \leq j \leq n$, such that $\operatorname{Span}(u_1, v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n) = V$. Now the vectors $u_1, u_2, v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n$ are linearly dependent, therefore one of these vectors can be expressed via the preceding ones. Clearly, it is not u_2: that would have implied that u_1, u_2 are linearly dependent, so it is one of $v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n$. Lemma 3.2 states that it can be dropped and the remaining vectors still span V.

One can continue this process: at the ith step one has a spanning set $u_1, \dots, u_{i-1}, v_{j_1}, \dots$ of n vectors. Appending u_i results in a linearly dependent set $u_1, \dots, u_i, v_{j_1}, \dots$. By Lemma 3.2 a certain vector from this set can be expressed via the preceding ones. It cannot be one of $u_1, \dots, u_i$ due to linear independence of $u_1, \dots, u_m$. Therefore it is one of v_i; it can be excluded and the remaining vectors still span V.

Note that at each step a vector from $v_1, \ldots, v_n$ is replaced by a vector from $u_1, \ldots, u_m$. If $m > n$, after the nth step one has a spanning set $u_1, \ldots, u_n$. Appending u_{n+1} will result a linearly dependent set, which is a subset of the linearly independent set $u_1, \ldots, u_m$, a contradiction. □

Exercise 3.5. *Show that any four vectors in $\mathbb{R}^3$ are linearly dependent.*

Proposition 3.5. *Let V be a vector space. If $v_1, \ldots v_n \in V$ span V, then there are linearly independent $v_{i_1}, \ldots v_{i_m}$ that span V.*

Proof. If $v_1, \ldots v_n$ are linearly independent, there is nothing to prove. Suppose that $v_1, \ldots v_n$ are linearly dependent. By Lemma 3.2 a certain vector from this set can be excluded so that the remaining vectors still span V. If they are linearly independent, we are done, otherwise repeat our argument. Since the initial set of vectors is finite, the process terminates in a finite number of steps. □

Exercise 3.6. *Following the proof of Proposition 3.5 find a linearly independent spanning subset of $\{(1,0,0),(1,1,0),(0,1,0),(0,0,1),(1,1,1)\}$ in $\mathbb{R}^3$.*

Finite-dimensional vector spaces

Definition. A vector space V is **finite-dimensional** (有限维的) if it has a finite spanning set.

Examples.

1. F^n is finite-dimensional as $F^n = \mathrm{Span}(e_1, \ldots, e_n)$, where $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ has 1 at ith position

2. $F[t]$ is not finite-dimensional. Indeed, assume that $f_1, \ldots, f_n$ span $F[t]$. Denote $m = \max_{1 \le i \le n} \deg f_i$. Then $t^{m+1} \notin \mathrm{Span}(f_1, \ldots, f_n)$, a contradiction

3. $F[t]_n$ is finite-dimensional as $F[t]_n = \mathrm{Span}(1, t, \ldots, t^n)$.

Exercise 3.7. *Prove that the vector space $V = \{A \in \mathrm{M}_2(\mathbb{R}) \mid A^T = A\}$ is finite-dimensional.*

Proposition 3.6. *A subspace of a finite-dimensional vector space is finite-dimensional.*

Proof. Let U be a subspace of a finite-dimensional vector space V. If $U = \{0\}$, there is nothing to prove. If $U \neq \{0\}$, choose a nonzero $v_1 \in U$. Obviously, $\mathrm{Span}(v_1) \subset U$. If $\mathrm{Span}(v_1) = U$, we are done. Otherwise one can choose $v_2 \in U$ such that $v_2 \notin \mathrm{Span}(v_1)$. Then $\mathrm{Span}(v_1, v_2) \subset U$. Continue the process: at the ith step one has a set $v_1, \ldots, v_{i-1}$ such that $\mathrm{Span}(v_1, \ldots, v_{i-1}) \subset U$. If these sets coincide, U is finite-dimensional, as required. Otherwise choose $v_i \in U$ such that $v_i \notin \mathrm{Span}(v_1, \ldots, v_{i-1})$. Note that $v_1, \ldots, v_i$ are linearly independent by Corollary 3.3.

Theorem 3.4 implies that the number of linearly independent vectors in V can not exceed the number of elements in its finite spanning set. Therefore, the process terminates in a finite number of steps. □

Exercise 3.8. *Following the proof of Proposition 3.6 find a finite spanning set in*

$$U = \{(a_1, a_2, a_3, a_4) \in F^4 \mid a_1 + a_2 + a_3 + a_4 = 0\}.$$

Bases

Definition. Let V be a vector space. A **basis** (基) of V is a finite linearly independent spanning set of vectors in V.

Remark. By definition, if a vector space has a basis, it is finite-dimensional. Conversely, any finite-dimensional vector space has a basis by Proposition 3.5.

From now on, all the vector spaces are assumed to be finite-dimensional!

Theorem 3.7. *Vectors $u_1, \dots, u_n$ in a vector space V over F form a basis if and only if any vector in V can be uniquely represented as a linear combination of $u_1, \dots, u_n$.*

基 ⇔ 任意向量唯一线性表出

Proof. If $u_1, \dots, u_n$ is a basis, then any vector in V can be represented as its linear combination. If $v \in V$ has two such representations $a_1u_1 + \cdots + a_nu_n = v = b_1u_1 + \cdots + b_nu_n$ for some $a_1 \dots, a_n, b_1, \dots, b_n \in F$, then $(a_1 - b_1)u_1 + \cdots + (a_n - b_n)u_n = 0$, and linear independence of $u_1, \dots, u_n$ implies that all the coefficients are 0, whence $a_i = b_i$ for every $1 \le i \le n$.

Conversely, if any vector V is represented as a linear combination of $u_1, \dots, u_n$, then these vectors span V. If they are linearly dependent, there is a nontrivial zero linear combination:

$$a_1v_1 + \cdots + a_nv_n = 0 = 0 \cdot v_1 + \cdots + 0 \cdot v_n.$$

It gives two different representations of the same vector $0 \in V$, a contradiction. □

Examples.

1. $e_1, \dots, e_n$ is a basis of F^n, where $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ has 1 at ith position.
2. $1, t, \dots, t^n$ is a basis of $F[t]_n$.
3. $\{E_{i,j}\}_{1\le i\le m, 1\le j\le n}$ is a basis of $\mathrm{M}_{m,n}(F)$, where $E_{i,j}$ has 1 at (i, j)th position and zeros at the other positions.
4. $1, i$ is a basis of $\mathbb{C}$ as a vector space over $\mathbb{R}$. Any non-zero element is a basis of F over F.

Problem 3.8. *Show that $f_1 = t + 1, f_2 = t^2 - 1, f_3 = -t^2 + t$ form a basis of $\mathbb{R}[t]_2$*

Solution. First, the equality $xf_1 + yf_2 + zf_3 = 0$ is equivalent to the system

$$\begin{cases} x - y = 0 \\ x + z = 0 \\ y - z = 0 \end{cases},$$

which implies $x = y = z = 0$.

Further, for any $a, b, c \in \mathbb{R}$ the equality $xf_1 + yf_2 + zf_3 = a + bt + ct^2$ is equivalent to the system

$$\begin{cases} x - y = a \\ x + z = b \\ y - z = c \end{cases},$$

whose solution is given by $x = \frac{a+b+c}{2}, y = \frac{-a+b+c}{2}, z = \frac{-a+b-c}{2}$. □

Exercise 3.9. *Show that $f_1 = t + 1, f_2 = t^3 - t^2 - 1, f_3 = t^3 + t^2, f_4 = t^3 - t$ form a basis of $\mathbb{R}[t]_3$*

Exercise 3.10. *Show that $g_1, \dots, g_n$ form a basis of $F^n, n \geq 2$, where $g_i = (1, \dots, 1, 0, 1, \dots, 1)$ has 0 at ith position.*

exotic space. $v = a * v_0 = v_0^a$ $a = \log_{v_0} v$. (in exotic space $v \in \mathbb{R}$)

Exercise 3.11. *Show that any $v > 0, v \neq 1$ forms a basis of the exotic vector space*

Theorem 3.9. *Any linearly independent set of vectors can be extended to a basis. In other words, if $v_1, \dots, v_n$ are linearly independent in a vector space V then there are $v_{n+1}, \dots, v_m \in V$ such that $v_1, \dots, v_m$ form a basis of V.*

Proof. If $v_1, \dots, v_n$ span V, then a basis is found. Otherwise, there are $v_{n+1} \in V \backslash \mathrm{Span}(v_1, \dots, v_n)$. By Corollary 3.3 the vectors $v_1, \dots, v_{n+1}$ are also linearly independent, therefore this procedure can be repeated. Since V is finite-dimensional, Theorem 3.4 implies that the process terminates in a finite number of steps. □

Remark. Proposition 3.5 can be restated as the following: a basis can be extracted from any spanning set of vectors.

Problem 3.10. *Find a basis for $V = \{f \in \mathbb{R}[t]_4 \mid f(-1) = 0, f'(1) = 0\}$.*

(Provide a formula to find a basis)

Solution. If $f \in \mathbb{R}[t]_4$, then $f = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4$ and $f \in V$ if and only if

$$\begin{cases} a_0 - a_1 + a_2 - a_3 + a_4 = 0 \\ a_1 + 2a_2 + 3a_3 + 4a_4 = 0 \end{cases}$$

⇒ two restriction. five variable

· dimension is 3.

The solution of this system is

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \alpha \begin{pmatrix} -3 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -2 \\ -3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} -5 \\ -4 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

whence $f_1 = -3 - 2t + t^2, f_2 = -2 - 3t + t^3, f_3 = -5 - 4t + t^4$ form a basis for V. □

also a variable. vector $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ \boxed{0} \end{pmatrix} \to (t^3+1, \underline{\underline{0}})$

Exercise 3.12. *Find a basis for $V = \{(f, a) \in \mathbb{R}[t]_3 \times \mathbb{R} \mid f(1) = \boxed{a}, f(-1) = 0\}$.*

Exercise 3.13. *Find a basis for*

$$V = \left\{ (A, B) \in \mathrm{M}_2(\mathbb{R}) \times \mathrm{M}_2(\mathbb{R}) \mid A \begin{pmatrix} 1 \\ 1 \end{pmatrix} = B \begin{pmatrix} 0 \\ 1 \end{pmatrix}, A^T \begin{pmatrix} -1 \\ 0 \end{pmatrix} = B \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}.$$

Exercise 3.14. *Find a basis for* $\mathrm{Span}(v_1, v_2, v_3, v_4, v_5)$ *if* $v_1 = (1, 1, 1, 1, 0)$, $v_2 = (1, 1, -1, -1, -1)$, $v_3 = (2, 2, 0, 0, -1)$, $v_4 = (1, 1, 5, 5, 2)$, $v_5 = (1, -1, -1, 0, 0)$.

Hint. Consider a matrix with rows v_1, v_2, v_3, v_4, v_5 and reduce it to a row echelon form using elementary row operations. Notice that any elementary row operation does not change the span of the rows. Finding a basis for the span of the rows of a row echelon matrix is easy.

Exercise 3.15. *Extend the vectors* $v_1 = (1,1,0,1,0)$, $v_2 = (1,-1,1,-1,-1)$, $v_3 = (1,2,0,1,-1)$ *to a basis of* $\mathbb{R}^5$.

Hint. Consider a matrix with rows v_1, v_2, v_3 and reduce it to a row echelon form using elementary row operations. Notice that any elementary row operation does not change the span of the rows. Extending the rows of a row echelon matrix to a basis is easy.

Proposition 3.11. *Let* U, V *be vector spaces over* F. *If* $u_1, \dots, u_n$ *is a basis of* U *and* $v_1, \dots, v_m$ *is a basis of* V *then* $(u_1, 0), \dots, (u_n, 0), (0, v_1), \dots, (0, v_m)$ *is a basis of* $U \times V$.

Proof. If $u \in U, v \in V$, then $u = a_1u_1 + \cdots + a_nu_n$, $v = b_1v_1 + \cdots + b_mv_m$ for some $a_1, \dots, a_n, b_1, \dots, b_m \in F$. It gives

$$(u, v) = (a_1u_1+\cdots+a_nu_n, b_1v_1+\cdots+b_mv_m) = a_1(u_1, 0)+\cdots+a_n(u_n, 0)+b_1(0, v_1)+\cdots+b_m(0, v_m).$$

If

$$(0,0) = a_1(u_1, 0)+\cdots+a_n(u_n, 0)+b_1(0, v_1)+\cdots+b_m(0, v_m) = (a_1u_1+\cdots+a_nu_n, b_1v_1+\cdots+b_mv_m),$$

then $a_1u_1 + \cdots + a_nu_n = 0$ and $b_1v_1 + \cdots + b_mv_m = 0$, whence $a_1 = \cdots = a_n = 0$ and $b_1 = \cdots = b_m = 0$. □

Exercise 3.16. *Let* V *be a vector space and* $\mathcal{B}$ *be its finite subset. Prove that the following statements are equivalent*

1. $\mathcal{B}$ *is a basis of* V
2. $\mathcal{B}$ *is a maximal linearly independent set, i.e.,* $\mathcal{B}$ *is linearly independent and including any vector in* $\mathcal{B}$ *would make it linearly dependent*
3. $\mathcal{B}$ *is minimal spanning set, i.e.,* $\mathcal{B}$ *spans* V *and excluding any vector from* $\mathcal{B}$ *would turn it into a set that does not span* V

Hint. Prove $1 \Leftrightarrow 2$ and $1 \Leftrightarrow 3$.

Exercise 3.17. *Let* $\mathcal{B}$ *be a basis of* V *and* $v_1, \dots, v_n \in V$ *be linearly independent. Prove that there are* $v_{n+1}, \dots, v_m \in \mathcal{B}$ *such that* $v_1, \dots, v_m$ *form a basis of* V.

Hint. Follow the proof of Theorem 3.9.

Coordinates

Definition. Let V be a vector space with basis $v_1, \cdots, v_n$. Then any vector $v \in V$ may be written uniquely as a linear combination $v = a_1v_1 + \cdots + a_nv_n$. Then the numbers $(a_1, \ldots, a_n)$ are called the **coordinates** (坐标) of v relative to the basis $v_1, \cdots, v_n$.

Remark. The coordinates depend on the ordering of the vectors in the basis.

Problem 3.12. *Find the coordinates of $f = (t+1)^2 \in \mathbb{R}[t]_2$ relative to the basis $t-1, t^2+1, t^2-t$.*

Solution. Let (a, b, c) be the coordinates of f. Then $a(t-1) + b(t^2+1) + c(t^2-t) = (t+1)^2$ which is equivalent to the system

$$\begin{cases} -a + b = 1 \\ a - c = 2 \\ b + c = 1 \end{cases}$$

whose solution is $a = 1, b = 2, c = -1$. ∎

Exercise 3.18. *Find the coordinates of $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \in \mathrm{M}_2(\mathbb{R})$ relative to the basis*

$$\begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Exercise 3.19. *Let (a_1, a_2, a_3) be the coordinates of a vector v relative to a basis v_1, v_2, v_3. Find the coordinates of v relative to the basis $v_1 - v_3, -v_2, v_1 + v_2 + v_3$.*

4 Dimension

Definition of dimension and basic properties

Theorem 4.1. *Any two bases of a vector space V contain the equal number of elements.*

Proof. Let $v_1, \ldots, v_n$ and $u_1, \ldots, u_m$ be two bases of V. Then $v_1, \ldots, v_n$ is a linearly independent set, and $u_1, \ldots, u_m$ is a spanning set; Theorem 3.4 implies that $n \leq m$. On the other hand, $u_1, \ldots, u_m$ is a linearly independent set, and $v_1, \ldots, v_n$ is a spanning set, so $n \geq m$. Therefore $m = n$. ∎

Definition. Let V be a vector space over F. The number of elements in any of its bases is called its **dimension** (维度) and is denoted by $\dim_F V$ or simply by $\dim V$.

1维度：任意一组基的向量数目.

Examples.

1. $\dim F^n = n$
2. $\dim \mathrm{M}_{m,n}(F) = mn$
3. The geometric vector plane is of dimension 2; the geometric vector space is of dimension 3

4. $\dim_{\mathbb{R}} \mathbb{C} = 2, \dim_F F = 1$

5. $\dim F[t]_n = n + 1$

6. $\dim V \times U = \dim V + \dim U$ by Proposition 3.11

Exercise 4.1. *Find the dimension of the vector space* $V = \{f \in \mathbb{R}[x]_6 \mid f(-x) = f(x)\}$.

Proposition 4.2. *Let V be a vector space and U be a subspace of V. Then* $\dim U \leq \dim V$. *Moreover, if* $\dim U = \dim V$*, then $U = V$.*

Proof. Let $u_1, \ldots, u_n$ be a basis of U and $\dim V = m$. Note that $u_1, \ldots, u_n$ is a linearly independent set in V. Theorem 3.9 implies it can be extended to a basis of V which contains m elements.

If $\dim U = \dim V$, then this extended basis must be of the same size as the set $u_1, \ldots, u_n$. Therefore $u_1, \ldots, u_n$ is a basis of V, and $U = \operatorname{Span}(u_1, \ldots, u_n) = V$. □

Proposition 4.3. *Let V be a vector space,* $\dim V = n$ *and $v_1, \ldots, v_n \in V$ Then $v_1, \ldots, v_n$ is a basis of V if they are linearly independent OR span V.*

Proof. If $v_1, \ldots, v_n$ are linearly independent, they can be extended to a basis by Theorem 3.9. Since this basis has n elements, like the original system, $v_1, \ldots, v_n$ itself form a basis.

Similarly, if $v_1, \ldots, v_n$ span V then a basis an be extracted from this set. Since this basis will have n elements, $v_1, \ldots, v_n$ itself form a basis. □

Exercise 4.2. *Let $v_1, \ldots, v_n$ be vectors in a vector space V. Show that $v_1, \ldots, v_n$ are linearly independent if and only if* $\dim \operatorname{Span}(v_1, \ldots, v_n) = n$.

Proposition 4.4. *Let $v_1, \ldots, v_n$ be a basis of a vector space V over F and $A = (a_{ij}) \in \mathrm{M}_n(F)$. Let $u_i = \sum_j a_{ij} v_j, 1 \leq i \leq n$. Then $u_1, \ldots, u_n$ form a basis for V if and only if A is invertible.*

Proof. Since $\dim V = n$, the vectors $u_1, \ldots, u_n$ form a basis of V if and only if they are linearly independent, i.e. any zero linear combination $\sum_i c_i u_i = 0$ is trivial. But

$$\sum_i c_i u_i = \sum_{i,j} c_i a_{ij} v_j = \sum_j \left(\sum_i c_i a_{ij} \right) v_j = 0$$

if and only if $\sum_i c_i a_{ij} = 0$ for all $1 \leq j \leq n$, which is in turn equivalent to linear independence of the rows of A. This is one of conditions equivalent to the invertibility of A. □

Exercise 4.3. *Using Proposition 4.4 show that the vectors from Exercise 3.9 form a basis.*

Theorem 4.5. *Let V be a vector space and U_1, U_2 be its subspaces. Then*

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim U_1 \cap U_2.$$

Proof. Let $u_1, \dots, u_m$ be an basis of $U_1 \cap U_2$. Then $u_1, \dots, u_m$ are linearly independent vectors in U_1, and thus there exist $v_1, \dots, v_l \in U_1$ such that $u_1, \dots, u_m, v_1, \dots, v_l$ is a basis of U_1. Similarly, $u_1, \dots, u_m$ are linearly independent vectors in U_2, and thus there exist $w_1, \dots, w_n \in U_2$ such that $u_1, \dots, u_m, w_1, \dots, w_n$ is a basis of U_2.

We are done if we show that $u_1, \dots, u_m, v_1, \dots, v_l, w_1, \dots, w_n$ form a basis of $U_1 + U_2$ since then $\dim(U_1 + U_2) = m + l + n$, $\dim U_1 \cap U_2 = m$, $\dim U_1 = m + l$, $\dim U_2 = m + n$. Take any vector in $U_1 + U_2$, it is the sum of a vector from U_1 and a vector from U_2, and each of these two vectors is a linear combination of $u_1, \dots, u_m, v_1, \dots, v_l, w_1, \dots, w_n$, so this set spans $U_1 + U_2$.

It remains to check that $u_1, \dots, u_m, v_1, \dots, v_l, w_1, \dots, w_n$ are linearly independent. Suppose that $a_1u_1 + \dots + a_mu_m + b_1v_1 + \dots + b_lv_l + c_1w_1 + \dots + c_nw_n = 0$ which imples

$$c_1w_1 + \dots + c_nw_n = -a_1u_1 - \dots - a_mu_m - b_1v_1 - \dots - b_lv_l.$$

Note that the left-hand part lies in U_2 and the right-hand part lies in U_1. Therefore $c_1w_1 + \dots + c_nw_n \in U_1 \cap U_2$. Since $u_1, \dots, u_m$ is a basis in $U_1 \cap U_2$, one has

$$c_1w_1 + \dots + c_nw_n = d_1u_1 + \dots + d_mu_m.$$

The vectors $u_1, \dots, u_m, w_1, \dots, w_n$ are linearly independent; therefore all the coefficients here are 0. In particular, $c_1 = \dots = c_n = 0$, whence

$$a_1u_1 + \dots + a_mu_m + b_1v_1 + \dots + b_lv_l = 0.$$

The vectors $u_1, \dots, u_m, v_1, \dots, v_l$ are also linearly independent, and therefore $a_1 = \dots = a_m = b_1 = \dots = b_l = 0$; so the initial linear combination is trivial. □

Exercise 4.4. *Prove that two three-dimensional subspaces in a five-dimensional vector space have a non-zero intersection.*

Direct sum revisited

Proposition 4.6. *Let $U_1, \dots, U_m$ be subspaces of a vector space V. The following conditions are equivalent:*

1. $V = U_1 \oplus \dots \oplus U_m$;
2. $V = U_1 + \dots + U_m$ *and* $U_j \cap (U_1 + \dots + U_{i-1} + U_{i+1} + \dots + U_m) = \{0\}$ *for any* $1 \le j \le m$.
3. $V = U_1 + \dots + U_m$ *and if* $u_1 + \dots + u_m = 0$ *for* $u_i \in U_i$, *then* $u_1 = \dots = u_m = 0$;
4. *the union of bases for* U_i *is a basis for* V;
5. $V = U_1 + \dots + U_m$ *and* $\dim V = \dim U_1 + \dots + \dim U_m$;

Proof.

1 ⇒ 2 Let $u_j \in U_j \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_m)$. Then $u_j = u_1 + \cdots + u_{j-1} + u_{j+1} + \cdots + u_m$ for $u_i \in U_i, 1 \leq i \leq m$. It can be rewritten as

$$0 + \cdots + 0 + u_j + 0 + \cdots + 0 = u_1 + \cdots + u_{j-1} + 0 + u_{j+1} + \cdots + u_m,$$

whence $u_j = 0$.

2 ⇒ 3 Assume $u_1 + \cdots + u_m = 0$ for $u_i \in U_i$. Then for any $1 \leq j \leq m$ one has $u_j = -u_1 - \cdots - u_{j-1} - u_{j+1} - \cdots - u_m$, whence $u_j \in U_j \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_m)$ and $u_j = 0$.

3 ⇒ 4 Let $\mathcal{B}_i$ be a basis of $U_i, 1 \leq i \leq m$. For any $v \in V$ there exist $u_i \in U_i$ such that $v = u_1 + \cdots + u_m$. Each u_i is a linear combination of the vectors from $\mathcal{B}_i$, thus $\cup_i \mathcal{B}_i$ spans V. In a zero linear combination of the vectors from $\cup_i \mathcal{B}_i$ group together multiples of the vectors from $\mathcal{B}_1, \ldots, \mathcal{B}_m$ and obtain that the sum is each group equals 0. It gives zero linear combinations of the vectors in each of $\mathcal{B}_1, \ldots, \mathcal{B}_m$, thus all the coefficients equals 0.

always possible since we can let $v_i = cv_j$ they correspond but different

4 ⇒ 5 Choose a basis $\mathcal{B}_i$ in each U_i such that $\mathcal{B}_i \cap \mathcal{B}_j = \varnothing$, and denote $\mathcal{B} = \cup_i \mathcal{B}_i$. Since $\mathcal{B}$ is a basis for V, then $\dim V = |\mathcal{B}| = |\mathcal{B}_1| + \cdots + |\mathcal{B}_m| = \dim U_1 + \cdots + \dim U_m$.

Any $v \in V$ is expressed as a linear combination of the vectors from $\cup_i \mathcal{B}_i$. Group together multiples of the vectors from $\mathcal{B}_1, \ldots, \mathcal{B}_m$ and obtain the expression of v as the sum of vectors from $U_1, \ldots, U_m$.

5 ⇒ 1 Let $u_1 + \cdots + u_m = u_1' + \cdots + u_m'$ for some $u_i, u_i' \in U_i$. If $u_1 \neq u_1'$, then $u_1 - u_1' = (u_2' - u_2) + \cdots + (u_m' - u_m) \neq 0$, whence $U_1 \cap (U_2 + \cdots + U_m) \neq 0$. By Theorem 4.5

$$\dim(U_1 + U_2 + \cdots + U_m) + \dim U_1 \cap (U_2 + \cdots + U_m) = \dim U_1 + \dim(U_2 + \cdots + U_m),$$

whence $\dim U_2 + \cdots + \dim U_m < \dim(U_2 + \cdots + U_m)$, which is impossible. Indeed, choose a basis for each subspace $U_2, \ldots, U_m$, their union spans $U_2 + \cdots + U_m$ but has less elements than its dimension. □

Exercise 4.5. *Let $V = \mathbb{R}^5$ and*

$$U_1 = \operatorname{Span}((1,1,-1,1,0),(1,1,0,-1,1)),$$
$$U_2 = \operatorname{Span}((1,-1,1,0,1),(1,1,-1,0,1)),$$
$$U_3 = \operatorname{Span}((-1,0,1,1,1),(0,1,1,1,-1),(1,1,0,0,-2)).$$

Use each equivalent condition from Proposition 4.6 to show that the sum $U_1 + U_2 + U_3$ is not direct.

5 Linear transformations

Definition and examples

Definition. Let V, W be vector spaces over F. A map $L\colon V \to W$ is called a **linear transformation** (线性变换) if

1. $L(u+v) = L(u) + L(v)$;

2. $L(av) = aL(v)$ for all $a \in k$, $v \in V$.

The set of all linear transformations from V to W is denoted by $\mathcal{L}(V, W)$, and in the case of $V = W$ just $\mathcal{L}(V)$.

Remark. (2) implies $L(0) = 0$.

Examples.

1. Define $0(v) = 0$ for all $v \in V$, then $0 \in \mathcal{L}(V, W)$.

2. $\mathrm{id}_V \in \mathcal{L}(V)$.

3. If $L\colon F[t] \to F[t], L(f) = f'$ then $L \in \mathcal{L}(F[t])$.

4. For $c \in F$, define $L\colon F[t] \to F, L(f) = f(c)$, then $L \in \mathcal{L}(F[t], F)$.

5. For $C \in \mathrm{M}_{l,n}(F)$, define
$$L\colon \mathrm{M}_{m,l}(F) \to \mathrm{M}_{m,n}(F), L(A) = AC,$$
then $L \in \mathcal{L}(\mathrm{M}_{m,l}(F), \mathrm{M}_{m,n}(F))$.

6. If $L\colon \mathbb{R}[x] \to \mathbb{R}$,
$$L(f) = \int_0^1 f(x)\,dx,$$
then $L \in \mathcal{L}(\mathbb{R}[t], \mathbb{R})$.

7. Rotation around the origin is a linear transformation from the geometric plane to itself.

Exercise 5.1. *Which of the following maps are linear transformations:*

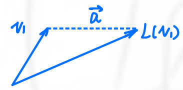
a. $L\colon \mathrm{M}_{m,n}(F) \to \mathrm{M}_{n,m}(F), L(A) = A^T$

b. $L\colon \mathrm{M}_n(F) \to F, L(A) = \det A$ ⇒ depends. n=1 is / n=2 not.

c. $L\colon F[t]_n \to F^{n+1}, L(a_0 + a_1 t + \cdots a_n t^n) = (a_0, a_1, \ldots, a_n)$

d. *A reflection across an axis passing through the origin of the geometric plane*

e. *A translation on the geometric plane*



def: translation $L(v) = \vec{v} + \vec{a}$

Theorem 5.1. *Let V, W be vector spaces over F, $v_1, \ldots, v_n$ be a basis of V, and $w_1, \ldots, w_n \in W$. There exists a unique $L \in \mathcal{L}(V, W)$ such that $L(v_i) = w_i$ for all $i = 1, \ldots, n$.*

给定基，像的向量组 ⇒ 线性变换唯一

Proof. Assume there is $L \in \mathcal{L}(V, W)$ such that $L(v_i) = w_i$ for all $i = 1, \ldots, n$. For any $v \in V$ one has $v = a_1 v_1 + \cdots + a_n v_n$ for some $a_1, \ldots, a_n \in F$. If $L(v_i) = w_i$ for $i = 1, \ldots, n$, then

$$L(v) = L(a_1 v_1 + \cdots + a_n v_n) = a_1 L(v_1) + \cdots + a_n L(v_n) = a_1 w_1 + \cdots + a_n w_n.$$

Thus, L is unique.

To show the existence of L, define $L(v) = a_1 w_1 + \cdots + a_n w_n$ for $v \in V, v = a_1 v_1 + \cdots + a_n v_n$. It remains to prove that it is a linear transformation. Let $u, v \in V$, where $v = a_1 v_1 + \cdots + a_n v_n$ and $u = b_1 v_1 + \cdots + b_n v_n$. Then by definition $L(v) = a_1 w_1 + \cdots + a_n w_n$, $L(u) = b_1 w_1 + \cdots + b_n w_n$ and $L(u + v) = (a_1 + b_1) w_1 + \cdots + (a_n + b_n) w_n$ since $u + v = (a_1 + b_1) v_1 + \cdots + (a_n + b_n) v_n$. This gives $L(u + v) = L(u) + L(v)$. If $a \in F$, then $av = (aa_1) v_1 + \cdots + (aa_n) v_n$, and $L(av) = (aa_1) w_1 + \cdots + (aa_n) w_n$ whence $L(av) = aL(v)$. □

Exercise 5.2. *Let W be vector spaces over F and $L \in \mathcal{L}(F^2, W)$. Let $v_1 = (1, -1), v_2 = (2, 1), v_3 = (1, 1)$ and $L(v_1) = w_1, L(v_2) = w_2$. Find $L(v_3)$.*

Exercise 5.3. *Let V, W be vector spaces, $v_1, \ldots, v_n$ be a basis of V. Show that if $L, L' \in \mathcal{L}(V, W)$ are such that $L(v_i) = L'(v_i)$ for all $i = 1, \ldots, n$ then $L = L'$.*

To show linear transformation are equal. their domain, codomain are equal + $\forall v \in$ domain $L(v) = L'(v)$.

Operations on linear transformations

Let V, W be vector spaces over F. If $L', L \in \mathcal{L}(V, W)$, define $L' + L \colon V \to W$ by formula $(L' + L)(v) = L'(v) + L(v)$ for all $v \in V$. It is easy to check that $L' + L$ is a linear transformation. Indeed, for any $u, v \in V$ and $a \in F$

$$\begin{aligned}(L' + L)(u + v) &= L'(u + v) + L(u + v) = L'(u) + L'(v) + L(u) + L(v) \\ &= L'(u) + L(u) + L'(v) + L(v) = (L' + L)(u) + (L' + L)(v),\end{aligned}$$

$$(L' + L)(av) = L'(av) + L(av) = aL'(v) + aL(v) = a(L'(v) + L(v)) = a(L' + L)(v).$$

If $L \in \mathcal{L}(V, W)$ and $a \in F$, define $aL \colon V \to W$ by formula $(aL)(v) = a(L(v))$. Similarly, $aL \in \mathcal{L}(V, W)$.

Let U, V, W be vector spaces over F. For $L \in \mathcal{L}(U, V)$ and $L' \in \mathcal{L}(V, W)$, consider their composition $L' \circ L \colon U \to W$. Then $L' \circ L \in \mathcal{L}(U, W$. Indeed, by definition $(L' \circ L)(u) = L'(L(u))$ for all $u \in U$. Therefore,

$$\begin{aligned}(L' \circ L)(u_1 + u_2) &= L'(L(u_1 + u_2)) = L'(L(u_1) + L(u_2)) \\ &= L'(L(u_1)) + L'(L(u_2)) = (L' \circ L)(u_1) + (L' \circ L)(u_2)\end{aligned}$$

for all $u_1, u_2 \in U$. If $u \in U$, $a \in F$, then

$$(L' \circ L)(au) = L'(L(au)) = L'(a(L(u))) = aL'(L(u)) = a(L' \circ L)(u)$$

whence $L' \circ L \in \mathcal{L}(U, W)$.

Remark. The composition $L' \circ L$ is often written as $L'L$ and is called informally the *product* of L' and L.

Proposition 5.2. *$\mathcal{L}(V,W)$ with respect to the above defined addition and scalar multiplication is a vector space.*

In addition, the product of linear transformations satisfies the following properties:

1. *$L''(L'L) = (L''L)L$ for $L \in \mathcal{L}(U,V), L' \in \mathcal{L}(V,W), L'' \in \mathcal{L}(W,X)$*
2. *$L\,\mathrm{id}_V = \mathrm{id}_W\,L = L$ for $L \in \mathcal{L}(V,W)$*
3. *$(L'_1 + L'_2)L = L'_1L + L'_2L$ and $L'(L_1 + L_2) = L'L_1 + L'L_2$ for $L, L_1, L_2 \in \mathcal{L}(U,V)$ and $L', L'_1, L'_2 \in \mathcal{L}(V,W)$.*

Proof. The zero linear transformation in $\mathcal{L}(V,W)$ is its additive identity. The inverse of $L \in \mathcal{L}(V,W)$ is $-L$.

Let $L', L \in \mathcal{L}(V,W)$, $a \in F$. Then for any $v \in V$

$$\begin{aligned}(a(L'+L))(v) &= a((L'+L)(v)) = a(L'(v)+L(v)) = a(L'(v)) + a(L(v)) \\ &= (aL')(v) + (aL)(v) = (aL'+aL)(v),\end{aligned}$$

thus the maps $a(L'+L)$, $aL'+aL$ from V to W coincide.

The remaining properties are checked in a similar way. □

Remark. The composition of linear transformations is noncommutative: for example, if $L \in \mathcal{L}(F[t], F[t])$, $L(f) = f'$ and $L' \in \mathcal{L}(F[t], F[t]), L'(f) = tf$, then $(L'L)(f) = tf'(t)$ and $((LL')(f)) = (tf(t))' = tf'(t) + f(t)$.

Exercise 5.4. *Prove the remaining properties in Proposition 5.2.*

Kernel and image

Definition. The **kernel** (核) of $L \in \mathcal{L}(V,W)$ is defined by

$$\mathrm{Ker}(L) = \{v \in V \mid L(v) = 0\}.$$

Examples.

1. The kernel of $0 \in \mathcal{L}(V,W)$ is V
2. $\mathrm{Ker}(\mathrm{id}_V) = \{0\}$
3. Let $A \in \mathrm{M}_{m,n}(F)$ and $L \in \mathcal{L}(\mathrm{M}_{n,1}(F), \mathrm{M}_{m,1}(F)), L(X) = AX$. The kernel of L equals the set of the solutions of the system of linear equations $AX = 0$.
4. If $L \in \mathcal{L}(F[t]), L(f) = f'$ then $\mathrm{Ker}(L) = \{f \in F[t] \mid f' = 0\} = F$.

Proposition 5.3. *If $L \in \mathcal{L}(V,W)$, then* $\mathrm{Ker}(L)$ *is a subspace in V.*

Proof. If $u, v \in \mathrm{Ker}(L)$, then by definition $L(u) = L(v) = 0$. Then $L(u+v) = L(u) + L(v) = 0 + 0 = 0$, that is, $u + v \in \mathrm{Ker}(L)$. If $u \in \mathrm{Ker}(L)$ and $a \in F$, then $L(u) = 0$ and $L(au) = a(L(u)) = a \cdot 0 = 0$, whence $au \in \mathrm{Ker}(L)$. □

Proposition 5.4. *$L \in \mathcal{L}(V, W)$ is injective if and only if* $\operatorname{Ker}(L) = \{0\}$.

Proof. Suppose L is injective. Since $L(0) = 0$, the condition $L(v) = 0$ that $v = 0$; therefore $\operatorname{Ker}(L) = \{0\}$.

Conversely, suppose $\operatorname{Ker}(L) = \{0\}$. If $v_1, v_2 \in V$ are such that $L(v_1) = L(v_2)$, then $L(v_1 - v_2) = L(v_1) - L(v_2) = 0$. Thus $v_1 - v_2 \in \operatorname{Ker}(L) = \{0\}$ which implies $v_1 - v_2 = 0$. □

Definition. The **image** (像) of $L \in \mathcal{L}(V, W)$ is defined by

$$\operatorname{Im}(L) = \{L(v) \mid v \in V\}.$$

Proposition 5.5. *If $L \in \mathcal{L}(V, W)$, then $\operatorname{Im}(L)$ is a subspace of W.*

Proof. If $w_1, w_2 \in \operatorname{Im}(L)$, there are $v_1, v_2 \in V$ such that $L(v_1) = w_1$ and $L(v_2) = w_2$. Then $L(v_1 + v_2) = L(v_1) + L(v_2) = w_1 + w_2$ and thus $w_1 + w_2 \in \operatorname{Im}(L)$.

If $w \in \operatorname{Im}(L)$ then $L(v) = w$ for some $v \in V$. For $a \in k$ one has $L(av) = aL(v) = aw$ and $aw \in \operatorname{Im}(L)$. □

Theorem 5.6. *Let V be a vector space over F, $L \in \mathcal{L}(V, W)$. Then*

$$\dim V = \dim \operatorname{Ker}(L) + \dim \operatorname{Im}(L).$$

Proof. Let $u_1, \ldots, u_m$ be a basis of $\operatorname{Ker}(L)$. It can be extended to a basis $u_1, \ldots, u_m, v_1, \ldots, v_n$ of V. Thus, $\dim \operatorname{Ker}(L) = m$ and $\dim V = m + n$ and it remains to prove that $\dim \operatorname{Im}(L) = n$.

Consider the vectors $L(v_1), \ldots, L(v_n) \in \operatorname{Im}(L)$ and show that they form a basis of $\operatorname{Im}(L)$. If $w \in \operatorname{Im}(L)$, then $w = L(v)$ for some $v \in V$ and

$$v = a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_n v_n$$

for some $a_1, \ldots, a_m, b_1, \ldots, b_n \in F$. Then

$$w = L(v) = L(a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_n v_n) = b_1 L(v_1) + \cdots + b_n L(v_n),$$

whence $w \in \operatorname{Span}(L(v_1), \ldots, L(v_n))$.

Finally, show that $L(v_1), \ldots, L(v_n)$ are linearly independent. If $c_1 L(v_1) + \cdots + c_n L(v_n) = 0$ then $0 = L(c_1 v_1 + \cdots + c_n v_n)$ whence $c_1 v_1 + \cdots + c_n v_n \in \operatorname{Ker}(L)$ and $c_1 v_1 + \cdots + c_n v_n = d_1 u_1 + \cdots + d_m u_m$ for some $d_1, \ldots d_m \in F$. Since $u_1, \ldots, u_m, v_1, \ldots, v_n$ are linearly independent, one has $c_1 = \cdots = c_n = d_1 = \cdots = d_m = 0$. □

Corollary 5.7. *A homogeneous system of linear equations with more variables than equations has nonzero solutions.*

Proof. Consider a system of linear equations $AX = 0$, where $A \in \mathrm{M}_{n,m}(F)$ and $m > n$. The set of its solutions is equal to $\operatorname{Ker}(L)$, where $L \in \mathcal{L}(\mathrm{M}_{m,1}(F), \mathrm{M}_{n,1}(F)), L(X) = AX$. Then $\dim \operatorname{Ker}(L) = m - \dim \operatorname{Im}(L) \geq m - n > 0$ by Theorem 5.6. □

Exercise 5.5. *Find the kernel and the image of $L \in \mathcal{L}(\mathrm{M}_n(F)), L(X) = X - X^T$. Check the identity from Theorem 5.6.*

Exercise 5.6. *Find a basis for the subspace* $U = \{XA - BX \mid X \in \mathrm{M}_2(\mathbb{R})\}$ *and*

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}.$$

Hint. Define $L \in \mathcal{L}(\mathrm{M}_2(\mathbb{R}))$ by $L(X) = XA - BX$ and find a basis of $\operatorname{Ker} L$, then follow the proof of Theorem 5.6.

Exercise 5.7. *Let V and W be vector spaces. Prove that there exists an injective linear transformation from V to W if and only if* $\dim V \leq \dim W$.

Hint. If $\dim V \leq \dim W$, construct an injective linear transformation using Theorem 5.1. If there is an injective linear transformation, apply Theorem 5.6.

6 Matrices

Matrix of linear transformation

Definition. Let V, W be two vector spaces over F with bases $\mathcal{B} = \{v_1, \ldots, v_n\}$ and $\mathcal{B}' = \{w_1, \ldots, w_m\}$, respectively. Let $L \in \mathcal{L}(V, W)$. If

$$L(v_j) = a_{1j}w_1 + a_{2j}w_2 + \cdots + a_{mj}w_m,$$

then the coefficients $(a_{ij})_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$ forms a $m \times n$ matrix which is called the **matrix of L relative to the bases $\mathcal{B}$, $\mathcal{B}'$** (L的相对于基$\mathcal{B}$,$\mathcal{B}'$的矩阵) and is denoted by $[L]_{\mathcal{B},\mathcal{B}'}$ or by $[L]_{\mathcal{B}}$ if $V = W$ and $\mathcal{B} = \mathcal{B}'$.

Example. Let $V = \mathbb{R}[t]_2, W = \mathbb{R}[t]_1$ and $\mathcal{B} = \{1, t, t^2\}, \mathcal{B}' = \{t, 1\}$. Then for $L \in \mathcal{L}(V, W), L(f) = f'(t)$ one has

$$\begin{aligned} L(1) &= 0 = 0 \cdot t + 0 \cdot 1 \\ L(t) &= 1 = 0 \cdot t + 1 \cdot 1 \\ L(t^2) &= 2t = 2 \cdot t + 0 \cdot 1 \end{aligned}$$

Thus

$$[L]_{\mathcal{B},\mathcal{B}'} = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 1 & 0 \end{pmatrix}.$$

Exercise 6.1. *Find the matrix of L from the above example if $\mathcal{B} = \{(t+1)^2, t+1, 1\}, \mathcal{B}' = \{1, t-1\}$.*

Exercise 6.2. *Let V, W be two vector spaces with bases $\mathcal{B} = \{v_1, v_2, v_3\}$ and $\mathcal{B}' = \{w_1, w_2, w_3\}$, respectively. For $L \in \mathcal{L}(V, W)$ find $L(v_1 - 2v_3)$ if*

$$[L]_{\mathcal{B},\mathcal{B}'} = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Theorem 6.1. *Let V, W be vector spaces over F with bases $\mathcal{B} = \{u_1, \dots, u_l\}$,$\mathcal{B}' = \{v_1, \dots, v_m\}$, respectively. Then $[L+L']_{\mathcal{B},\mathcal{B}'} = [L]_{\mathcal{B},\mathcal{B}'} + [L']_{\mathcal{B},\mathcal{B}'}$ for any $L, L' \in \mathcal{L}(V, W)$ and $[aL]_{\mathcal{B},\mathcal{B}'} = a[L]_{\mathcal{B},\mathcal{B}'}$ for any $L \in \mathcal{L}(V, W)$ and $a \in F$.*

In other words, the map $\varphi\colon \mathcal{L}(V, W) \to \mathrm{M}_{m,n}(F)$ defined by $\varphi(L) = [L]_{\mathcal{B},\mathcal{B}'}$, is a linear transformation.

Proof. Let $[L]_{\mathcal{B},\mathcal{B}'} = (a_{ij})$, $[L']_{\mathcal{B},\mathcal{B}'} = (b_{ij})$. By definition, $L(v_j) = \sum_{i=1}^m a_{ij}w_i$, $L'(v_j) = \sum_{i=1}^m b_{ij}w_i$. Then

$$(L + L')(v_j) = L(v_j) + L'(v_j) = \sum_{i=1}^m a_{ij}w_i + \sum_{i=1}^m b_{ij}w_i = \sum_{i=1}^m (a_{ij} + b_{ij})w_i,$$

whence $[L + L']_{\mathcal{B},\mathcal{B}'} = (a_{ij} + b_{ij}) = [L]_{\mathcal{B},\mathcal{B}'} + [L']_{\mathcal{B},\mathcal{B}'}$.

Similarly, let $[L]_{\mathcal{B},\mathcal{B}'} = (a_{ij})$, then $L(v_j) = \sum_{i=1}^m a_{ij}w_i$ and

$$(aL)(v_j) = aL(v_j) = a\sum_{i=1}^m a_{ij}w_i = \sum_{i=1}^m (aa_{ij})w_i,$$

whence $[aL]_{\mathcal{B},\mathcal{B}'} = (aa_{ij}) = a[L]_{\mathcal{B},\mathcal{B}'}$. □

Theorem 6.2. *Let U, V, W be three vector spaces over F with bases $\mathcal{B} = \{u_1, \dots, u_l\}$,$\mathcal{B}' = \{v_1, \dots, v_m\}$, $\mathcal{B}'' = \{w_1, \dots, w_n\}$, respectively, and $L' \in \mathcal{L}(U, V)$, $L \in \mathcal{L}(V, W)$. Then*

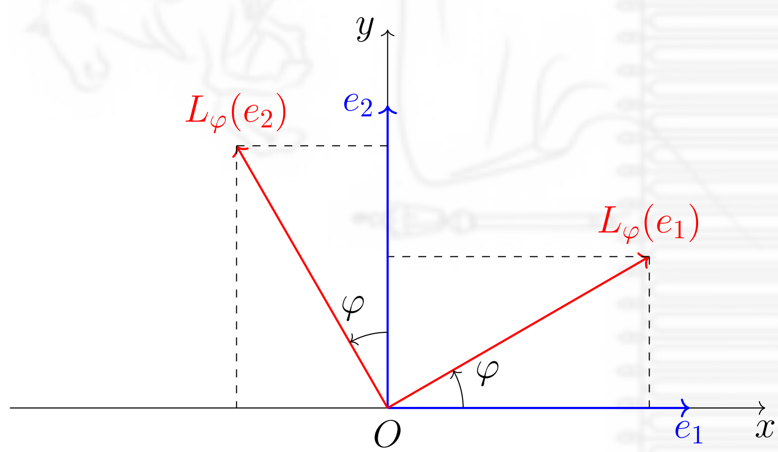
$$[LL']_{\mathcal{B},\mathcal{B}''} = [L]_{\mathcal{B}',\mathcal{B}''}[L']_{\mathcal{B},\mathcal{B}'}.$$

Proof. Let $[L]_{\mathcal{B}',\mathcal{B}''} = (a_{ij}) \in \mathrm{M}_{n,m}(F)$, $[L']_{\mathcal{B},\mathcal{B}'} = (b_{ij}) \in \mathrm{M}_{m,l}(F)$, whence $L(v_j) = \sum_{i=1}^n a_{ij}w_i$ and $L'(u_p) = \sum_{j=1}^m b_{jp}v_j$. Then

$$(LL')(u_p) = L(L'(u_p)) = L\left(\sum_{j=1}^m b_{jp}v_j\right) = \sum_{j=1}^m b_{jp}L(v_j) = \sum_{j=1}^m\sum_{i=1}^n b_{jp}a_{ij}w_i = \sum_{i=1}^n\left(\sum_{j=1}^m a_{ij}b_{jp}\right)w_i.$$

Thus the entry of $[LL']_{\mathcal{B},\mathcal{B}''}$ in the (i, p) position is $\sum_{j=1}^m a_{ij}b_{jp}$, the same for the product of $[L]_{\mathcal{B}',\mathcal{B}''}$ and $[L']_{\mathcal{B},\mathcal{B}'}$. □

Example. Let L_φ be the counterclockwise rotation about the origin by φ.



Let $e_1 = (1,0), e_2 = (0,1)$. Then $L_\varphi(e_1) = (\cos\varphi, \sin\varphi), L_\varphi(e_2) = (-\sin\varphi, \cos\varphi)$ and for $\mathcal{B} = \{e_1, e_2\}$,

$$[L_\varphi]_{\mathcal{B}} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

Then

$$\begin{aligned}
[L_\varphi]_{\mathcal{B}}[L_\psi]_{\mathcal{B}} &= \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix} \\
&= \begin{pmatrix} \cos\varphi\cos\psi - \sin\varphi\sin\psi & -\cos\varphi\sin\psi - \sin\varphi\cos\psi \\ \sin\varphi\cos\psi + \cos\varphi\sin\psi & -\sin\varphi\sin\psi + \cos\varphi\cos\psi \end{pmatrix} \\
&= \begin{pmatrix} \cos(\varphi+\psi) & -\sin(\varphi+\psi) \\ \sin(\varphi+\psi) & \cos(\varphi+\psi) \end{pmatrix} = [L_{\varphi+\psi}]_{\mathcal{B}} = [L_\varphi L_\psi]_{\mathcal{B}}
\end{aligned}$$

Exercise 6.3. *Let $U = W = \mathbb{R}[t]_2, V = \mathbb{R}[t]_1$ and $\mathcal{B} = \{1, t, t^2\}, \mathcal{B}' = \{t, t+1\}, \mathcal{B}'' = \{t^2, t, t+1\}$. For $L' \in \mathcal{L}(U, V), L'(f) = f'(t)$ and $L \in \mathcal{L}(V, W), L(f) = tf(t)$, find $[L']_{\mathcal{B},\mathcal{B}'}$, $[L]_{\mathcal{B}',\mathcal{B}''}$, $[LL']_{\mathcal{B},\mathcal{B}''}$ and check the identity from Theorem 6.2.*

Exercise 6.4. *Find the composition of two reflections across the lines passing through the origin if the angle between them is α.*

Let $\mathcal{B} = (v_1, \ldots, v_n)$ be a basis of a vector space V over F. Let $v \in V$ and $v = a_1v_1 + \cdots + a_nv_n, a_i \in F$. Denote

$$[v]_{\mathcal{B}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = (a_1 \ldots a_n)^T \in \mathrm{M}_{n,1}(F).$$

Theorem 6.3. *Let V be a vector space over F with basis $\mathcal{B} = \{v_1, \ldots, v_n\}$. Then $[v + v']_{\mathcal{B}} = [v]_{\mathcal{B}} + [v']_{\mathcal{B}}$ for any $v, v' \in V$ and $[av]_{\mathcal{B}} = a[v]_{\mathcal{B}}$ for any $a \in F, v \in V$.*

In other words, the map

$$\begin{gathered} V \to \mathrm{M}_{n,1}(F), \\ v \mapsto [v]_{\mathcal{B}} \end{gathered}$$

is a linear transformation.

Proof. Let $[v]_{\mathcal{B}} = (a_1 \ldots a_n)^T, [v']_{\mathcal{B}} = (b_1 \ldots b_n)^T$. By definition, $v = a_1v_1 + \cdots + a_nv_n, v' = b_1v_1 + \cdots + b_nv_n$. Then $v + v' = (a_1 + b_1)v_1 + \cdots + (a_n + b_n)v_n$ and

$$[v + v']_{\mathcal{B}} = (a_1 + b_1 \ldots a_n + b_n)^T = (a_1 \ldots a_n)^T + (b_1 \ldots b_n)^T = [v]_{\mathcal{B}} + [v']_{\mathcal{B}}.$$

Similarly, let $[v]_{\mathcal{B}} = (a_1 \ldots a_n)^T$, whence $v = a_1v_1 + \cdots + a_nv_n$ and $av = a(a_1v_1 + \cdots + a_nv_n) = (aa_1)v_1 + \cdots + (aa_n)v_n$. Then

$$[av]_{\mathcal{B}} = (aa_1 \ldots aa_n)^T = a(a_1 \ldots a_n)^T = a[v]_{\mathcal{B}}.$$

$\square$

Theorem 6.4. *Let V, W be vector spaces over F with bases $\mathcal{B}, \mathcal{B}'$, respectively, and $L \in \mathcal{L}(V, W)$. Then*

$$[L(v)]_{\mathcal{B}'} = [L]_{\mathcal{B},\mathcal{B}'}[v]_{\mathcal{B}}$$

$[L]_{\mathcal{B},\mathcal{B}'}[v]_{\mathcal{B}}$

for any $v \in V$.

Proof. Let $v = c_1v_1 + \dots + c_nv_n$ for $c_1, \dots, c_n \in F$ and $[L]_{\mathcal{B},\mathcal{B}'} = (a_{ij})$. Then

$$L(v) = L\left(\sum_{j=1}^{n} c_jv_j\right) = \sum_{j=1}^{n} c_jL(v_j) = \sum_{j=1}^{n} c_j\left(\sum_{i=1}^{m} a_{ij}w_i\right) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}c_j\right)w_i.$$

Thus the ith component of $[L(v)]_{\mathcal{B}'}$ equals $\sum_{j=1}^{n} a_{ij}c_j$, the same for the product of $[L]_{\mathcal{B},\mathcal{B}'}$ and $[v]_{\mathcal{B}}$. □

Exercise 6.5. *Let $V = \mathrm{M}_2(\mathbb{R}), W = \{A \in \mathrm{M}_2(\mathbb{R}) \mid A = A^T\}$ and*

$$\mathcal{B} = \left\{\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\right\}, \mathcal{B}' = \left\{\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right\}, v = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}.$$

For $L(A) = A + A^T$, find $[v]_{\mathcal{B}}$, $[L(v)]_{\mathcal{B}'}$, $[L]_{\mathcal{B},\mathcal{B}'}$ and check the identity from Theorem 6.4.

Change of basis

Definition. Let $\mathcal{B} = \{u_1, \dots, u_n\}$, $\mathcal{B}' = \{v_1, \dots, v_n\}$ be bases of a vector space V over F. Then $v_j = \sum_{i=1}^{n} c_{ij}u_i$ for some $c_{ij} \in F$. The matrix $C = (c_{ij})_{i,j=1}^{n}$ is called the **transition matrix** (过渡矩阵) from the basis $\mathcal{B}$ to the basis $\mathcal{B}'$ and is denoted by $M_{\mathcal{B}\rightsquigarrow\mathcal{B}'}$. In other words,

$$M_{\mathcal{B}\rightsquigarrow\mathcal{B}'} = \begin{pmatrix}[v_1]_{\mathcal{B}} & [v_2]_{\mathcal{B}} & \dots & [v_n]_{\mathcal{B}}\end{pmatrix}$$

Symbolically, one can write

$$\begin{pmatrix}v_1 & v_2 & \dots & v_n\end{pmatrix}^T = M^T_{\mathcal{B}\rightsquigarrow\mathcal{B}'}\begin{pmatrix}u_1 & u_2 & \dots & u_n\end{pmatrix}^T.$$

Proposition 6.5. *Let $\mathcal{B} = \{u_1, \dots, u_n\}$, $\mathcal{B}' = \{v_1, \dots, v_n\}$, $\mathcal{B}'' = \{w_1, \dots, w_n\}$ be bases of a vector space V. Then*

1. *$M_{\mathcal{B}\rightsquigarrow\mathcal{B}} = E_n$;*
2. *$M_{\mathcal{B}\rightsquigarrow\mathcal{B}''} = M_{\mathcal{B}\rightsquigarrow\mathcal{B}'}M_{\mathcal{B}'\rightsquigarrow\mathcal{B}''}$;*
3. *$M_{\mathcal{B}\rightsquigarrow\mathcal{B}'}$ is invertible and $M^{-1}_{\mathcal{B}\rightsquigarrow\mathcal{B}'} = M_{\mathcal{B}'\rightsquigarrow\mathcal{B}}$.*

Proof. 1. Obvious

2. Since $(v_1, \dots, v_n)^T = M^T_{\mathcal{B}\rightsquigarrow\mathcal{B}'}(u_1, \dots, u_n)^T$ and $(w_1, \dots, w_n)^T = M^T_{\mathcal{B}'\rightsquigarrow\mathcal{B}''}(v_1, \dots, v_n)^T$, one has

$$(w_1, \dots, w_n)^T = M^T_{\mathcal{B}'\rightsquigarrow\mathcal{B}''}(v_1, \dots, v_n)^T = M^T_{\mathcal{B}'\rightsquigarrow\mathcal{B}''}M^T_{\mathcal{B}\rightsquigarrow\mathcal{B}'}(u_1, \dots, u_n)^T.$$

Then $(w_1, \dots, w_n)^T = M^T_{\mathcal{B} \rightsquigarrow \mathcal{B}''}(u_1, \dots, u_n)^T$ implies

$$M^T_{\mathcal{B} \rightsquigarrow \mathcal{B}''}(u_1, \dots, u_n)^T = M^T_{\mathcal{B}' \rightsquigarrow \mathcal{B}''} M^T_{\mathcal{B} \rightsquigarrow \mathcal{B}'}(u_1, \dots, u_n)^T.$$

Since $u_1, \dots, u_n$ is a basis, the equality of its linear combinations with coefficients from each row of both matrices implies the equality of the corresponding coefficients and thus $M^T_{\mathcal{B} \rightsquigarrow \mathcal{B}''} = M^T_{\mathcal{B}' \rightsquigarrow \mathcal{B}''} M^T_{\mathcal{B} \rightsquigarrow \mathcal{B}'}$, whence $M_{\mathcal{B} \rightsquigarrow \mathcal{B}''} = M_{\mathcal{B} \rightsquigarrow \mathcal{B}'} M_{\mathcal{B}' \rightsquigarrow \mathcal{B}''}$.

3. The above properties for $\mathcal{B}'' = \mathcal{B}$ imply $M_{\mathcal{B} \rightsquigarrow \mathcal{B}'} M_{\mathcal{B}' \rightsquigarrow \mathcal{B}} = M_{\mathcal{B} \rightsquigarrow \mathcal{B}} = E_n$. □

Theorem 6.6. *Let V be a vector space with bases $\mathcal{B}$, $\mathcal{B}'$. Then for any $v \in V$*

$$[v]_{\mathcal{B}} = M_{\mathcal{B} \rightsquigarrow \mathcal{B}'}[v]_{\mathcal{B}'}$$

Proof. Let $\mathcal{B} = \{u_1, \dots, u_n\}$, $\mathcal{B}' = \{v_1, \dots, v_n\}$ and $[v]_{\mathcal{B}} = (x_1 \dots x_n)^T$, $[v]_{\mathcal{B}'} = (y_1 \dots y_n)^T$. Then $v = x_1 u_1 + \dots + x_n u_n = y_1 v_1 + \dots + y_n v_n$, that is

$$(x_1 \dots x_n)(u_1 \dots u_n)^T = v = (y_1 \dots y_n)(v_1 \dots v_n)^T.$$

On the other hand, $(v_1 \dots v_n)^T = M^T_{\mathcal{B} \rightsquigarrow \mathcal{B}'}(u_1 \dots u_n)^T$, whence

$$(x_1 \dots x_n)(u_1 \dots u_n)^T = (y_1 \dots y_n) M^T_{\mathcal{B} \rightsquigarrow \mathcal{B}'}(u_1 \dots u_n)^T.$$

Since $u_1, \dots, u_n$ is a basis, the equality of its linear combinations implies the equality of the corresponding coefficients and thus

$$(x_1 \dots x_n) = (y_1 \dots y_n) M^T_{\mathcal{B} \rightsquigarrow \mathcal{B}'}.$$

Therefore

$$[v]_{\mathcal{B}} = (x_1 \dots x_n)^T = M_{\mathcal{B} \rightsquigarrow \mathcal{B}'}(y_1 \dots y_n)^T = M_{\mathcal{B} \rightsquigarrow \mathcal{B}'}[v]_{\mathcal{B}'}.$$

□

Theorem 6.7. *Let U, V be vector spaces over F with bases $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{B}'$, $\mathcal{C}'$, respectively. Then for $L \in \mathcal{L}(U, V)$*

$$[L]_{\mathcal{C},\mathcal{C}'} = M^{-1}_{\mathcal{B}' \rightsquigarrow \mathcal{C}'}[L]_{\mathcal{B},\mathcal{B}'} M_{\mathcal{B} \rightsquigarrow \mathcal{C}}$$

Proof. For $u \in U$

$$[L(u)]_{\mathcal{B}'} = [L]_{\mathcal{B},\mathcal{B}'}[u]_{\mathcal{B}}, \quad [L(u)]_{\mathcal{C}'} = [L]_{\mathcal{C},\mathcal{C}'}[u]_{\mathcal{C}}$$

and

$$[u]_{\mathcal{B}} = M_{\mathcal{B} \rightsquigarrow \mathcal{C}}[u]_{\mathcal{C}}, \quad [L(u)]_{\mathcal{C}'} = M_{\mathcal{C}' \rightsquigarrow \mathcal{B}'}[L(u)]_{\mathcal{B}'}.$$

Therefore,

$$[L]_{\mathcal{C},\mathcal{C}'}[u]_{\mathcal{C}} = [L(u)]_{\mathcal{C}'} = M_{\mathcal{C}' \rightsquigarrow \mathcal{B}'}[L(u)]_{\mathcal{B}'} = M_{\mathcal{C}' \rightsquigarrow \mathcal{B}'}[L]_{\mathcal{B},\mathcal{B}'}[u]_{\mathcal{B}} = M_{\mathcal{C}' \rightsquigarrow \mathcal{B}'}[L]_{\mathcal{B},\mathcal{B}'} M_{\mathcal{B} \rightsquigarrow \mathcal{C}}[u]_{\mathcal{C}}.$$

If u is the ith element of $\mathcal{C}$, then the ith columns of $[L]_{\mathcal{C},\mathcal{C}'}$ and $M_{\mathcal{C}' \rightsquigarrow \mathcal{B}'}[L]_{\mathcal{B},\mathcal{B}'} M_{\mathcal{B} \rightsquigarrow \mathcal{C}}$ are equal, hence the matrices themselves are equal. □

Exercise 6.6. *Let* $U = \mathbb{R}[t]_1 \times \mathbb{R}$, $V = \{f \in \mathbb{R}[t]_3 \mid f(1) = 0\}$, $w = (1,1) \in U$ *and* $L \in \mathcal{L}(U,V), L((\varphi, a)) = a(t^3 - t) + \varphi(t)t^2 + \varphi(-t)t - 2\varphi(0)$. *Let* $\mathcal{F} = \{f_1, f_2, f_3\}$,$\mathcal{F}' = \{f_1', f_2', f_3'\}$ *be two bases of* U *and* $\mathcal{G} = \{g_1, g_2, g_3\}$ *be a basis of* V, *where* $f_1 = (t-1, -1)$, $f_2 = (t, 1)$, $f_3 = (t+2, 1)$; $f_1' = (t+2, -1)$, $f_2' = (2, -1)$, $f_3' = (t-1, 1)$ *and* $g_1 = t^3 + t - 2$, $g_2 = 2t^2 - t - 1$, $g_3 = t^3 + t^2 - 2t$.

i) *Find the matrix of* L *relative to the bases* $\mathcal{F}$ u $\mathcal{G}$;

ii) *Find the matrix of* L *relative to the bases* $\mathcal{F}'$ u $\mathcal{G}$;

iii) *Find the transition matrix from* $\mathcal{F}$ *to* $\mathcal{F}'$ *or from* $\mathcal{F}'$ *to* $\mathcal{F}$;

iv) *Apply Theorem 6.7 and verify the equality obtained;*

v) *Find the coordinates of* w *relative to the bases* $\mathcal{F}$ *and* $\mathcal{F}'$;

vi) *Apply Theorem 6.6 and verify the equality obtained.*

Exercise 6.7. *Let* L_φ *be the counterclockwise rotation of the geometric vector plane about the origin by* φ. *Let* p *be a line passing thought the origin and* R_p *be the reflection across* p. *Prove that the composition* $T = R_p L_\varphi$ *is a reflection using the following method:*

i) *Choose a basis* $\mathcal{B}$ *of two perpendicular unit vectors, one of which lies in* p;

ii) *Find* $[R_p]_\mathcal{B}, [L_\varphi]_\mathcal{B}$ *and their product* $[T]_\mathcal{B}$;

iii) *Find a non-zero vector* v *such that* $Tv = v$ *by solving* $([T]_\mathcal{B} - E_2)[v]_\mathcal{B} = 0$;

iv) *Let* v' *be a non-zero vector perpendicular to* v *and* $\mathcal{C} = \{v, v'\}$. *Find the transition matrix between* $\mathcal{B}$ *and* $\mathcal{C}$.

v) *Apply Theorem 6.7 to find* $[T]_\mathcal{C}$ *and conclude that* T *is a reflection.*

Lemma 6.8. *Let* $\mathcal{C}$ *be a basis of an* n*-dimensional vector space* V *over* F *and* $A \in \mathrm{M}_n(F)$ *be invertible. Then there exists a basis* $\mathcal{B}$ *in* V *such that* $M_{\mathcal{C} \rightsquigarrow \mathcal{B}} = A$.

Proof. Let $A = (a_{ij})$ and $\mathcal{C} = \{v_1, \ldots, v_n\}$. Define $u_j = \sum_{i=1}^{n} a_{ij} v_i$. Since A is invertible, $\mathcal{B} = \{u_1, \ldots, u_n\}$ is a basis by Proposition 4.4. By definition, $M_{\mathcal{C} \rightsquigarrow \mathcal{B}} = A$. □

Theorem 6.9. *Let* U, V *be vector spaces over* F *and* $L \in \mathcal{L}(U,V)$. *Then there are a basis* $\mathcal{B}$ *in* U *and a basis* $\mathcal{B}'$ *in* V *such that*

$$[L]_{\mathcal{B},\mathcal{B}'} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

with $r = \dim \mathrm{Im}(L)$.

Proof. Let $\dim U = m, \dim V = n$. Choose arbitrary bases $\mathcal{C}$ in U and $\mathcal{C}'$ in V. Then

$$[L]_{\mathcal{C},\mathcal{C}'} = P_t P_{t-1} \cdots P_1 D Q_1 \cdots Q_{s-1} Q_s = PDQ, \quad D = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

where $P_1, \dots, P_t; Q_1, \dots, Q_s$ are elementary matrices, $P = P_t P_{t-1} \cdots P_1 \in \mathrm{M}_n(F)$, $Q = Q_1 \cdots Q_s \in \mathrm{M}_m(F)$ are invertible and $r = \operatorname{rank}[L]_{\mathcal{C},\mathcal{C}'}$. By Lemma 6.8 there are a basis $\mathcal{B}$ in U and a basis $\mathcal{B}'$ in V such that $M_{\mathcal{C}\rightsquigarrow\mathcal{B}} = Q^{-1}$, $M_{\mathcal{C}'\rightsquigarrow\mathcal{B}'} = P$. Then $M_{\mathcal{B}\rightsquigarrow\mathcal{C}} = Q$, $M^{-1}_{\mathcal{B}'\rightsquigarrow\mathcal{C}'} = P$ and Theorem 6.7 implies $[L]_{\mathcal{B},\mathcal{B}'} = D$.

If $\mathcal{B} = \{u_1, \dots, u_m\}, \mathcal{B}' = \{v_1, \dots, v_n\}$, then $L(u_1) = v_1, \dots, L(u_r) = v_r, L(u_{r+1}) = \cdots = L(u_m) = 0$, whence $\dim \operatorname{Im}(L) = \dim \operatorname{Span}(L(u_1), \dots, L(u_m)) = \dim \operatorname{Span}(v_1, \dots, v_r) = r$. □

Exercise 6.8. *Let $U = V = \mathbb{R}[t]_2$ and $L \in \mathcal{L}(U, V), L(f) = tf'(t) - f(t)$. Find bases in U, V such that the matrix of L relative to these bases has a form as in Theorem 6.9.*

Hint. Find $\operatorname{Ker}(L)$ and choose its basis, then extend this basis to a basis of U.

Corollary 6.10. *Let U, V be vector spaces over F with bases $\mathcal{C}, \mathcal{C}'$, respectively, and $L \in \mathcal{L}(U, V)$. Then* $\dim \operatorname{Im}(L) = \operatorname{rank}[L]_{\mathcal{C},\mathcal{C}'}$.

Proof. Theorem 6.9 allows one to choose a basis $\mathcal{B}$ in U and a basis $\mathcal{B}'$ in V such that $\dim Im(L) = \operatorname{rank}[L]_{\mathcal{B},\mathcal{B}'}$. But $[L]_{\mathcal{C},\mathcal{C}'} = M^{-1}_{\mathcal{B}'\rightsquigarrow\mathcal{C}'}[L]_{\mathcal{B},\mathcal{B}'} M_{\mathcal{B}\rightsquigarrow\mathcal{C}}$ and $M^{-1}_{\mathcal{B}'\rightsquigarrow\mathcal{C}'}, M_{\mathcal{B}\rightsquigarrow\mathcal{C}}$ are invertible, therefore $\operatorname{rank}[L]_{\mathcal{C},\mathcal{C}'} = \operatorname{rank}[L]_{\mathcal{B},\mathcal{B}'}$. □

7 Isomorphism

Proposition 7.1. *If $L \in \mathcal{L}(V, W)$ is invertible, then the inverse map is also a linear transformation.*

Proof. Let $L' \colon W \to V$ be the inverse map of L. For $w_1, w_2 \in W$, let $L'(w_1) = v_1$, $L'(w_2) = v_2$, whence $L(v_1) = w_1$, $L(v_2) = w_2$. Then $L(v_1 + v_2) = L(v_1) + L(v_2) = w_1 + w_2$, and hence, $L'(w_1 + w_2) = v_1 + v_2 = L'(w_1) + L'(w_2)$.

Similarly, for $w \in W$ and $a \in F$, let $L'(w) = v$, whence $L(v) = w$. Then $L(av) = aL(v) = aw$, and hence, $L'(aw) = av = aL'(w)$. □

Definition. An invertible linear transformation is called an **isomorphism** (同构). Vector spaces V and W over F are **isomorphic** (同构的) if there is an isomorphism from V to W.

Examples. 1. id_V is an isomorphism from V to V

2. $L(f) = (f(0), f(1), f(-1))$ is an isomorphism from $\mathbb{R}[t]_2$ to $\mathbb{R}^3$

3. $\ln(x)$ is an isomorphism from the exotic vector space to $\mathbb{R}$

Exercise 7.1. *Show that* $L(A) = A\begin{pmatrix}1\\1\end{pmatrix}$ *is an isomorphism from*

$$\left\{A \in \mathrm{M}_2(\mathbb{R}) \mid A\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}\right\}$$

to $\mathrm{M}_{2,1}(\mathbb{R})$

Exercise 7.2. *Let* U, W *be subspaces of a vector space* V *and* $V = U \oplus W$*. Show that*

$$L: U \times W \to V, \quad L((u, w)) = u + w$$

is an isomorphism.

Exercise 7.3. *Let* $L \in \mathcal{L}(V, W)$ *be an isomorphism and* $v_1, \ldots, v_n$ *be a basis of* V*. Prove that* $L(v_1), \ldots, L(v_n)$ *is a basis of* W*.* dim keeps same only need to proof l.i or spanning

Proposition 7.2. *Isomorphism has the following properties:* (equivalent relation)

1. *any vector space is isomorphic to itself*

2. *if* V *is isomorphic to* W *then* W *is isomorphic to* V

3. *if* V *is isomorphic to* W *and* W *is isomorphic to* U *then* V *is isomorphic to* U

In other worlds, the relation on the set of vector spaces over F *given by isomorphism is an equivalence relation.*

Proof. 1. Obvious

2. There is an isomorphism $L: V \to W$ and then $L^{-1}: W to V$ is an isomorphism by Proposition 7.1.

3. There are an isomorphism $L: V \to W$ and an isomorphism $L': W \to U$. Then $L'L: V \to U$ is a linear transformation as the composition of linear transformations and is invertible as the composition of invertible maps. □

Theorem 7.3. *Two vector spaces over* F *are isomorphic if and only if their dimensions are equal.*

Proof. Suppose V and W are isomorphic, that is, there is a invertible $L \in \mathcal{L}(V, W)$. Then L is bijective. Since L is injective, Proposition 5.4 implies that $\mathrm{Ker}(L) = 0$. Since L is surjective, $\mathrm{Im}(L) = W$. Now $\dim V = \dim W$ by Theorem 5.6.

Conversely, suppose $\dim V = \dim W = n$. Choose a basis $v_1, \ldots, v_n$ in V and a basis $w_1, \ldots, w_n$ in W. By Theorem 5.1 there is $L \in \mathcal{L}(V, W)$ such that $L(v_i) = w_i, 1 \leq i \leq n$. We show that L is bijective.

For any $v \in \mathrm{Ker}(L)$ one has $v = a_1v_1 + \cdots + a_nv_n$. Then $0 = L(v) = a_1L(v_1) + \cdots + a_nL(v_n) = a_1w_1 + \cdots + a_nw_n$. The vectors $w_1, \ldots, w_n \in W$ form a basis, whence $a_1 = \cdots = a_n = 0$. It yields $v = 0$ whence L is injective by Proposition 5.4.

For any $w \in W$ one has $w = b_1w_1 + \cdots + b_nw_n = b_1L(v_1) + \cdots + b_nL(v_n) = L(b_1v_1 + \cdots + b_nv_n)$ whence L is surjective. □

Corollary 7.4. *Any vector space V over F is isomorphic to the coordinate space F^n and the column space $\mathrm{M}_{n,1}(F)$, where $n = \dim V$. Moreover, if $\mathcal{B}$ is a basis of space V, then the map $\varphi\colon V \to \mathrm{M}_{n,1}(F), \varphi(v) = [v]_{\mathcal{B}}$ is an isomorphism.*

Proof. Since $\dim F^n = \dim \mathrm{M}_{n,1}(F) = \dim V = n$, these vector spaces are isomorphic.

Let $\mathcal{B} = \{v_1, \ldots, v_n\}$. Theorem 6.3 implies that φ is a linear transformation, it remains to check that φ is bijective. If $v \in \mathrm{Ker}(\varphi)$ then the coordinates of v are all zero, whence $v = 0 \cdot v_1 + \cdots + 0 \cdot v_n = 0$, i.e. $\mathrm{Ker}(\varphi) = \{0\}$. Further, for $w = (a_1 \ldots a_n)^T \in \mathrm{M}_{n,1}(F)$ consider $v = a_1v_1 + \cdots + a_nv_n \in V$. Clearly $[v]_{\mathcal{B}} = w$, which proves that φ is surjective. □

Theorem 7.5. *Let V, W be vector spaces over F. Then $\mathcal{L}(V, W)$ is isomorphic $\mathrm{M}_{m,n}(F)$, where $m = \dim W$, $n = \dim V$. Moreover, if $\mathcal{B}, \mathcal{B}'$ are bases of V, W, respectively, the map $\varphi\colon \mathcal{L}(V, W) \to \mathrm{M}_{m,n}(F), \varphi(L) = [L]_{\mathcal{B},\mathcal{B}'}$ is an isomorphism.*

Proof. Theorem 6.1 implies that φ is a linear transformation. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$, $\mathcal{B}' = \{w_1, \ldots, w_m\}$. If $L \in \mathrm{Ker}(\varphi)$ and $v \in V$, the coordinates of $L(v)$ in the basis $\mathcal{B}'$ are zero by Theorem 6.4 and therefore $L(v) = 0$. Since this is true for any $v \in V$, one has $L = 0$, whence $\mathrm{Ker}(\varphi) = \{0\}$.

Finally, let $A = (a_{ij}) \in \mathrm{M}_{m,n}(F)$. Theorem 5.1 implies that there exists $L \in \mathcal{L}(V, W)$ such that $L(v_j) = a_{1j}w_1 + \cdots + a_{mj}w_m, 1 \le j \le n$. Then $[L]_{\mathcal{B},\mathcal{B}'} = A$, whence $\mathrm{Im}(\varphi) = \mathrm{M}_{m,n}(F)$. □

Corollary 7.6. *If V, W are vector spaces over F, then $\dim \mathcal{L}(V, W) = \dim V \cdot \dim W$.*

Proof. Follows from Theorems 7.5, 7.3 and the fact that $\dim \mathrm{M}_{m,n}(F) = mn$. □

Definition. Let V be a vector space. A linear transformation $L\colon V \to V$ is called a **linear operator** (线性算子) on V. The set of linear operators on V is denoted by $\mathcal{L}(V)$.

Proposition 7.7. *Let V be a vector space and $L \in \mathcal{L}(V)$. The following statements are equivalent:*

1. *L is bijective*
2. *L is injective*
3. *L is surjective.*

Proof. Obviously, (1) implies (2) and (3). Let us show that (1) follows from (2). Since L is injective, $\mathrm{Ker}(L) = 0$ by Proposition 5.4. Then $\dim \mathrm{Im}(L) = \dim V$ by Theorem 5.6. Finally Proposition 4.2 implies that $\mathrm{Im}(L) = V$, therefore L is bijective.

It remains to show that (3) implies (1). Since L is surjective, $\mathrm{Im}(L) = V$ and Theorem 5.6 implies $\dim \mathrm{Ker}(L) = \{0\}$. Then L is injective by Proposition 5.4 and thus bijective. □

Exercise 7.4. *Suppose V and W are vector spaces, $v \in V, v \neq 0$ and $\dim V = n, \dim W = m$. Let*

$$K = \{L \in \mathcal{L}(V, W) \mid L(v) = 0\}.$$

Show that K is a subspace of $\mathcal{L}(V, W)$ and find its dimension.

Im(T) is a subspace W
we know nothing about W.
dim Im(T) = dim W or 0 (idea, not strict prove)

Hint. Consider $T \in \mathcal{L}(\mathcal{L}(V, W), W), T(L) = L(v)$ and apply Theorem 5.6.

8 Eigenvectors and eigenvalues

Definition of eigenvectors and eigenvalues

Definition. Let V be a vector space over F and $L \in \mathcal{L}(V)$. An **eigenvalue** (特征值) of L is $\lambda \in F$ such that $L(u) = \lambda u$ for some non-zero $u \in V$. An **eigenvector** (特征向量) of L is $u \in V$ such that $L(u) = \lambda u$ for some $\lambda \in F$. Thus, if $L(u) = \lambda u$, the eigenvector u is associated to the eigenvalue λ, and the eigenvalue λ is associated to the eigenvector u.

Examples.

1. If V is the geometric plane, L is the reflection across OX, then L has two eigenvalues 1 and -1 which correspond to eigenvectors lying on OX and OY, respectively.
2. If V is the geometric plane, L is the clockwise rotation around O by 90°, then L has no eigenvalues and eigenvectors.
3. If $V = \mathrm{M}_2(F), L(A) = A^T$, then L has two eigenvalues 1 and -1, and the associated eigenvectors are the symmetric matrices and the matrices of the form $\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$, respectively.
4. If $V = \mathbb{R}[x], L(f) = f'$, then L has only one eigenvalue 0, and the associated eigenvectors are the constant polynomials.

Exercise 8.1. *Find the eigenvalues and the eigenvectors of* $L \in \mathcal{L}(\mathbb{R}[t]_3), L(f) = f(2t)$.

Exercise 8.2. *For* $L \in \mathcal{L}(V)$ *prove that* 0 *is its eigenvalue if and only if* $\mathrm{Ker}(L) \neq \{0\}$.

Exercise 8.3. *Suppose* $L \in \mathcal{L}(V)$ *is invertible.* first we need $\lambda \neq 0$ (by Ex 8.2 $\lambda = 0 \Leftrightarrow \ker(L) \neq \{0\}$. not invertible

i) *Prove that* $\lambda \in F$ *is an eigenvalue of* L *if and only if* λ^{-1} *is an eigenvalue of* L^{-1}.

ii) *Prove that* $v \in V$ *is an eigenvector of* L *associated to* λ *if and only if* v *is an eigenvector of* L^{-1} *associated to* λ^{-1}.

Proposition 8.1. *Let* $L \in \mathcal{L}(V), \lambda \in F$. *Then* $V_\lambda(L) = \{u \in V \mid Lu = \lambda u\}$ *is a subspace of* V.

Proof. If $L(u) = \lambda u, L(u') = \lambda u'$, then $L(u + u') = \lambda(u + u')$. Similarly, if $\mu \in F, L(u) = \lambda u$, then $L(\mu u) = \mu L(u) = \mu(\lambda u) = \lambda(\mu u)$. □

Definition. The subspace from Proposition 8.1 is called the **eigenspace** (特征空间) of L, associated to λ. Note that $V_\lambda(L) \neq \{0\}$ if and only if λ is an eigenvalue of L.

Definition. Let $A \in \mathrm{M}_n(F)$. An **eigenvalue** (特征值) of A is $\lambda \in F$ such that $AX = \lambda X$ for a nonzero column vector $X \in \mathrm{M}_{n,1}(F)$.

A **column eigenvector** (特征列向量) of A is $X \in \mathrm{M}_{n,1}(F)$ such that $AX = \lambda X$ for some $\lambda \in F$.

Proposition 8.2. *Let V be an n-dimensional vector space over F.*

1. *If $\lambda \in F$ is an eigenvalue of $L \in \mathcal{L}(V)$, then λ is an eigenvalue of $[L]_{\mathcal{B}}$ for any basis $\mathcal{B}$ of V.*

2. *If $\lambda \in F$ is an eigenvalue of $[L]_{\mathcal{B}}$ for a basis $\mathcal{B}$ of V, then λ is an eigenvalue of L.*

3. *If $u \in V$ is an eigenvector of $L \in \mathcal{L}(V)$, then $[u]_{\mathcal{B}}$ is a column eigenvector of $[L]_{\mathcal{B}}$ for any basis $\mathcal{B}$ of V.*

4. *If $X \in \mathrm{M}_{n,1}(F)$ is a column eigenvector of $[L]_{\mathcal{B}}$ for a basis $\mathcal{B}$ of V, then $X = [u]_{\mathcal{B}}$, where $u \in V$ is an eigenvector of L.*

Proof. 1. If $\lambda \in F$ is an eigenvalue of $L \in \mathcal{L}(V)$, there is $u \in V, u \neq 0$ such that $L(u) = \lambda u$. Then

$$[L]_{\mathcal{B}}[u]_{\mathcal{B}} = [L(u)]_{\mathcal{B}} = [\lambda u]_{\mathcal{B}} = \lambda [u]_{\mathcal{B}}.$$

It remains to note that $[u]_{\mathcal{B}} \neq 0$, since $u \neq 0$.

2. If $\lambda \in F$ is an eigenvalue of $[L]_{\mathcal{B}}$, there is $X \in \mathrm{M}_{n,1}(F), X \neq 0$ such that $[L]_{\mathcal{B}} X = \lambda X$. If $X = [u]_{\mathcal{B}}$ for $u \in V$ then

$$[L(u)]_{\mathcal{B}} = [L]_{\mathcal{B}}[u]_{\mathcal{B}} = [L]_{\mathcal{B}} X = \lambda X = \lambda [u]_{\mathcal{B}} = [\lambda u]_{\mathcal{B}},$$

whence $L(u) = \lambda u$. It remains to note that $u \neq 0$ since $[u]_{\mathcal{B}} = X \neq 0$.

The proof of (3) and (4) is similar. □

Exercise 8.4. *Prove (3) and (4) from Proposition 8.2.*

Properties of eigenvalues and eigenvectors

Theorem 8.3. *Let $L \in \mathcal{L}(V)$ and $v_1, \ldots, v_n \in V$ be nonzero eigenvectors associated to distinct eigenvalues $\lambda_1, \ldots, \lambda_n \in F$. Then $v_1, \ldots, v_n$ are linearly independent.*

Proof. Assume $v_1, \ldots, v_n$ are linearly dependent. v_1 is nonzero Lemma 3.2 implies that there is an index j such that v_j can be expressed via $v_1, \ldots, v_{j-1}$. Let j be the smallest index satisfying the property. Then

$k>j$. $a_1 \ldots a_k$ are l.i.

$$v_j = a_1 v_1 + \cdots + a_{j-1} v_{j-1}.$$

Since $v_j \neq 0$, let $1 \leq k \leq j-1$ be the maximum index such that $a_k \neq 0$. Then

$$v_j = a_1 v_1 + \cdots + a_k v_k,$$

and one has

$$\lambda_j v_j = L(v_j) = a_1 L(v_1) + \cdots + a_k L(v_k) = \lambda_1 a_1 v_1 + \cdots + \lambda_k a_k v_k.$$

These equalities imply

$$0 = (\lambda_1 - \lambda_j) a_1 v_1 + \cdots + (\lambda_k - \lambda_j) a_k v_k,$$

one can express v_k as a linear combination of $v_1, \ldots, v_{k-1}$, which is a contradiction. □

Proposition 8.4. *Let $L \in \mathcal{L}(V)$ and $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of L. Then the sum $V_{\lambda_1}(L) + \cdots + V_{\lambda_m}(L)$ is direct. Moreover, $\dim V_{\lambda_1}(L) + \cdots + \dim V_{\lambda_m}(L) \le \dim V$.*

Proof. Applying Proposition 4.6 assume $u_1 + \cdots + u_m = 0$, where $u_j \in V_{\lambda_j}(L)$. Then Theorem 8.3 implies $u_1 = \cdots = u_m = 0$.

Proposition 4.6 also gives

$$\dim V_{\lambda_1}(L) + \cdots + \dim V_{\lambda_m}(L) = \dim V_{\lambda_1}(L) \oplus \cdots \oplus V_{\lambda_m}(L) \le \dim V.$$

□

Invariant subspaces

Definition. Let $L \in \mathcal{L}(V)$. A subspace U of V is **invariant** under L (在线性变换L下不变) or L**-invariant** (L-不变的) if $L(u) \in U$ for any $u \in U$.

Example. 1. $\{0\}$ and V are trivial L-invariant subspaces

2. If $L \in \mathcal{L}(F[t]_n), L(f) = f'$ then $U_1 = F[t]_m$ is L-invariant for $m \le n$ and $U_2 = \{f \in F[t]_n \mid f(1) = 0\}$ is not.

3. If L is the counterclockwise rotation about the origin by φ, then L has no non-trivial invariant subspaces unless $\varphi = 0, \pi$.

4. If L is the reflection across OX, then OX and OY are the only non-trivial invariant subspaces.

Proposition 8.5. *$V_\lambda(L)$ is L-invariant.*

Proof. If $v \in V_\lambda(L)$ then $L(v) = \lambda v \in V_\lambda(L)$ since $V_\lambda(L)$ is a subspace. □

Definition. Let $L \in \mathcal{L}(V)$, and U be an L-invariant subspace of V. The map $L|_U \colon U \to U$ given by $(L|_U)(u) = L(u)$, is called the **restriction** (限制) of L to U.

Proposition 8.6. *The restriction $L|_U$ is correctly defined and belongs to $\mathcal{L}(U)$.*

Proof. If $u \in U$, then $(L|_U)(u) = L(u) \in U$, thus $L|_U \colon U \to U$. Finally, $L|_U$ is a linear transformation since L is a linear transformation. □

Exercise 8.5. *Let $\mathcal{B} = \{v_1, v_2, v_3, v_4\}$ be a basis of V and $L \in \mathcal{L}(V)$. If*

$$[L]_{\mathcal{B}} = \begin{pmatrix} 1 & -1 & 2 & 0 \\ -2 & 2 & 1 & 1 \\ 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & -1 \end{pmatrix},$$

prove that $U = \mathrm{Span}(v_1 + v_2, v_3 + v_4)$ is L-invariant and find a matrix of $L|_U$.

Proposition 8.7. *Let V be a vector space and $L \in \mathcal{L}(V)$. Let $V_1, \dots, V_m$ be L-invariant subspaces of a vector space V with bases $\mathcal{B}_1, \dots, \mathcal{B}_m$, respectively, and $V = V_1 \oplus \dots \oplus V_m$. Then*

$$[L]_{\mathcal{B}} = \operatorname{diag}([L|_{V_1}]_{\mathcal{B}_1}, \dots, [L|_{V_m}]_{\mathcal{B}_m}),$$

where $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_m$ is a basis of V (see Proposition 4.6).

Proof. Let $v_i \in \mathcal{B}_i$. Since V_i is L-invariant, $L(v_i) \in V_i$. Hence all the coordinates of $L(v_i)$ relative to $\mathcal{B}$ are zero except perhaps the coordinates corresponding to $\mathcal{B}_i$. These coordinates are exactly the coordinates of $L|_{V_i}(v_i) = L(v_i)$ relative to $\mathcal{B}_i$. □

9 Characteristic polynomial

Roots or characteristic polynomial

Definition. Let $A \in \mathrm{M}_n(F)$. The polynomial $\chi_A(t) = |tE_n - A|$ is the **characteristic polynomial** (特征多项式) of A.

Example. If $A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$, then

$$\chi_A(t) = \begin{vmatrix} t-1 & -1 \\ 0 & t+1 \end{vmatrix} = t^2 - 1.$$

Definition. The **trace** (迹) of $A = (a_{ij}) \in \mathrm{M}_n(F)$ is defined by $\operatorname{Tr} A = \sum_{i=1}^n a_{ii}$.

Proposition 9.1. *If $A \in \mathrm{M}_n(F)$ then $\chi_A(t) = t^n - \operatorname{Tr} A \cdot t^{n-1} + \dots + (-1)^n |A|$.*

Proof. For $A = (a_{ij})$ one has

$$\chi_A(t) = \begin{vmatrix} t - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & t - a_{22} & & -a_{2n} \\ \vdots & & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & t - a_{nn} \end{vmatrix}.$$

In the sum defining the determinant, the highest degree of t is given by the term

$$(t - a_{11})(t - a_{22}) \cdots (t - a_{nn}).$$

Therefore $\deg \chi_A = n$ and its highest-degree coefficient is 1. Clearly, no other term can produce ct^{n-1}, so the coefficient at t^{n-1} in the determinant coincides with the corresponding coefficient of this product, i.e. $-\operatorname{Tr} A$. Finally, the constant term of $\chi_A(t)$ is equal to its value at 0, i.e. $|-A| = (-1)^n |A|$. □

Lemma 9.2. *If $A, B \in \mathrm{M}_n(F)$ are such that $A = UBU^{-1}$ for an invertible matrix $U \in \mathrm{M}_n(F)$, then $\chi_A = \chi_B$.*

Proof. One has

$$\chi_A(t) = |tE_n - A| = |tE_n - UBU^{-1}| = |U(tE_n - B)U^{-1}| = |U|\cdot|tE_n - B|\cdot|U|^{-1} = |tE_n - B| = \chi_B(t).$$

□

Lemma 9.3. *Let $L \in \mathcal{L}(V)$ and $\mathcal{B}, \mathcal{B}'$ be bases of V. Then $\chi_{[L]_\mathcal{B}} = \chi_{[L]_{\mathcal{B}'}}$.*

Proof. If $U = M_{\mathcal{B}\rightsquigarrow\mathcal{B}'}$, then $[L]_{\mathcal{B}'} = U^{-1}[L]_\mathcal{B}U$ and Lemma 9.2 completes the proof. □

Definition. Let $L \in \mathcal{L}(V)$. Lemma 9.3 implies that $\chi_{[L]_\mathcal{B}}(t)$ does not depend on the choice of $\mathcal{B}$. Thus this polynomial is called the **characteristic polynomial** (特征多项式) of L.

Similarly, Proposition 9.1 implies that $\mathrm{Tr}[L]_\mathcal{B}$ and $|[L]_\mathcal{B}|$ do not depend on the choice of $\mathcal{B}$. They are called the **trace** (迹) of L and the **determinant** (行列式) of L, respectively.

Proposition 9.4. *The eigenvalues of an operator/a matrix coincide with the roots of its characteristic polynomial.*

$|\lambda E_n - A| = 0$

Proof. Let $A \in \mathrm{M}_n(F)$. If $\lambda \in F$ is an eigenvalue of A, then the system of linear equations $(\lambda E_n - A)X = 0$ has a non-zero solution, whence $\lambda E_n - A$ is singular. Therefore $\chi_A(\lambda) = |\lambda E_n - A| = 0$. Note that all the implications here are reversible.

Let $L \in \mathcal{L}(V)$ and $\mathcal{B}$ be a basis of V. If $\lambda \in F$ is an eigenvalue of L, then λ is an eigenvalue of $A = [L]_\mathcal{B}$ whence λ is a root of $\chi_A = \chi_L$. All the implications here are reversible. □

Corollary 9.5. *Let V be a nonzero vector space over $\mathbb{C}$, $L \in \mathcal{L}(V)$. Then L has an eigenvalue.*

An algorithm for computing the eigenvalues λ and the associated eigenvectors u of $L \in \mathcal{L}(V)$ based on Propositions 8.2 and 9.4:

I. Choose a basis $\mathcal{B}$ in V;

II. Calculate $A = [L]_\mathcal{B}$ and $\chi_A(t)$;

III. Find the roots of $\chi_A(t)$;

IV. For each root λ solve the system $(\lambda E_n - A)X = 0$;

V. For each solution X find $u \in V$ such that $[u]_\mathcal{B} = X$.

Problem 9.6. *For $V = \{f \in \mathbb{R}[t]_3 \mid f(1) = 0\}$, find eigenvalues and eigenvectors of $L \in \mathcal{L}(V), Lf = f(2-t)$.*

Solution. Choose a basis of V, let $\mathcal{B} = \{f_1, f_2, f_3\}$ and $f_1 = t-1, f_2 = t(t-1), f_2 = t^2(t-1)$. Then

$$\begin{aligned}
L(f_1) &= (2-t) - 1 = 1 - t = -f_1 \\
L(f_2) &= (2-t)((2-t)-1) = (t-2)(t-1) = -2f_1 + f_2 \\
L(f_3) &= (2-t)^2((2-t)-1) = (-t^2+4t-4)(t-1) = -4f_1 + 4f_2 - f_3
\end{aligned}$$

whence

$$A = [L]_{\mathcal{B}} = \begin{pmatrix} -1 & -2 & -4 \\ 0 & 1 & 4 \\ 0 & 0 & -1 \end{pmatrix}.$$

Then

$$\chi_A(\lambda) = \begin{vmatrix} \lambda+1 & 2 & 4 \\ 0 & \lambda-1 & -4 \\ 0 & 0 & \lambda+1 \end{vmatrix} = 0,$$

whence $\lambda_1 = 1, \lambda_2 = -1$.

For $\lambda_1 = 1$, the corresponding system of linear equations is

$$\begin{cases} 2x + 2y + 4z = 0 \\ \qquad\qquad - 4z = 0\,, \\ \qquad\qquad\quad 2z = 0 \end{cases}$$

whence $x = -\alpha, y = \alpha, z = 0$, where α is a parameter. Then $v_1 = -\alpha f_1 + \alpha f_2 = \alpha(t-1)^2$ is the associated eigenvector.

For $\lambda_2 = -1$, the corresponding system of linear equations is

$$\begin{cases} 2y + 4z = 0 \\ -2y - 4z = 0 \end{cases},$$

whence $x = \alpha, z = \beta, y = -2\beta$, where α, β are parameters. Then $v_2 = \alpha f_1 - 2\beta f_2 + \beta f_3 = \alpha(t-1) + \beta t(t-1)(t-2)$ is the associated eigenvector.

Indeed, $L((t-1)^2) = (t-1)^2$, $L(t-1) = -(t-1)$ and $L(t(t-1)(t-2)) = -t(t-1)(t-2)$. $\square$

Exercise 9.1. *For* $V = \{f(t) = a_0 + a_1 \sin t + a_2 \cos t + a_3 \sin 2t + a_4 \cos 2t \mid a_i \in \mathbb{R}, f(\pi/4) = 0\}$, *find the eigenvalues and the associated eigenvectors of* $L \in \mathcal{L}(V), L(f) = f(\pi/2 - t)$.

Exercise 9.2. *For* $V = \mathrm{M}_{2,1}(\mathbb{C})$ *over* $\mathbb{R}$, *find the eigenvalues and the associated eigenvectors of* $L \in \mathcal{L}(V)$,

$$L(u) = \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} u.$$

Exercise 9.3. *Let* $A \in \mathrm{M}_n(F)$ *and*

$$A = \begin{pmatrix} B & * \\ 0 & C \end{pmatrix}, B \in \mathrm{M}_m(F), C \in \mathrm{M}_{n-m}(F).$$

Show that $\lambda \in F$ *is an eigenvalue of* A *if and only if* λ *is an eigenvalue of* B *or an eigenvalue of* C.

Polynomials applied to matrices and operators

Definition. Let $f \in F[t]$, $f = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$ and $A \in \mathrm{M}_m(F)$, $L \in \mathcal{L}(V)$. Then $f(A) \in \mathrm{M}_m(F)$ and $f(L) \in \mathcal{L}(V)$ are defined as follows:

(new operation. all the operation rule need to be checked)

$$f(A) = a_0 E_n + a_1 A + a_2 A^2 + \cdots + a_n A^n,$$

$$f(L) = a_0 \operatorname{id}_V + a_1 L + a_2 L^2 + \cdots + a_n L^n, \quad \text{where } L^i = \underbrace{L \circ \cdots \circ L}_{i}.$$

Proposition 9.7. *If $f, g \in F[t]$, $a \in F$ and $A \in \mathrm{M}_m(F)$, $L \in \mathcal{L}(V)$. Then $(f+g)(A) = f(A) + g(A)$, $(fg)(A) = f(A)g(A)$, $(af)(A) = af(A)$, $(f \circ g)(A) = f(g(A))$ and $(f+g)(L) = f(L) + g(L)$, $(fg)(L) = f(L)g(L)$, $(\lambda f)(L) = \lambda f(L)$, $(f \circ g)(L) = f(g(L))$.*

Proof. Let $f = a_0 + a_1 t + \cdots + a_m x^m$, $g = b_0 + b_1 t + \cdots + b_s t^s$. Then $fg = \sum_\ell \left(\sum_{i+j=\ell} a_i b_j \right) t^\ell$ and

$$f(L) = a_0 \operatorname{id}_V + a_1 L + \cdots + a_m L^m, g(L) = b_0 \operatorname{id}_V + b_1 L + \cdots + b_s L^s,$$

therefore

$$f(L)g(L) = \sum_i a_i L^i \left(\sum_j b_j L^j \right) = \sum_{i,j} a_i b_j L^{i+j} = \sum_\ell \left(\sum_{i+j=\ell} a_i b_j \right) L^\ell = (fg)(L).$$

Similarly $f(L) + g(L) = \sum (a_i + b_i) L^i = (f+g)(L)$ and $af(L) = \sum (a a_i) L^i = (af)(L)$.

The verification of the properties of polynomials applied to matrices is completely analogous.

□

Corollary 9.8. *If $f, g \in F[t]$, $A \in \mathrm{M}_m(F)$, $L \in \mathcal{L}(V)$ then $f(A)g(A) = g(A)f(A)$ and $f(L)g(L) = g(L)f(L)$.* (看似是矩阵可交换、其实展开后均为同一种运算)

Lemma 9.9. *Let V be a vector space over F and $\mathcal{B}$ be a basis of V. If $L \in \mathcal{L}(V)$ and $f \in F[t]$, then $[f(L)]_\mathcal{B} = f([L]_\mathcal{B})$.*

Proof. Let $f = a_0 + a_1 t + \cdots + a_m t^m$. Then by Theorems 6.1 and 6.2

$$[f(L)]_\mathcal{B} = [a_0 \operatorname{id}_V + a_1 L + \cdots + a_m L^m]_\mathcal{B} = a_0 E_n + a_1 [L]_\mathcal{B} + \cdots + a_m [L]_\mathcal{B}^m = f([L]_\mathcal{B}).$$

Coordinates (column vectors) w.r.t $\mathcal{B}$

□

Lemma 9.10. *Let $L \in \mathcal{L}(V)$ and $p \in F[t]$. Then $\operatorname{Ker}(p(L))$ and $\operatorname{Im}(p(L))$ are L-invariant.*

Proof. If $v \in \operatorname{Ker}(p(L))$, then $p(L)(v) = 0$ and

$$p(L)(L(v)) = (p(L)L)(v) = (Lp(L))(v) = L(p(L)(v)) = L(0) = 0.$$

If $v \in \operatorname{Im}(p(L))$ then $v = p(L)(u)$ for some $u \in V$. Then

$$L(v) = L(p(L)(u)) = (Lp(L))(u) = (p(L)L)(u) = p(L)(L(u)) \in \operatorname{Im}(p(L)).$$

□

Hamilton-Cayley theorem

Theorem 9.11 (Hamilton-Cayley). *1. Let $A \in \mathrm{M}_n(F)$. Then $\chi_A(A) = 0$.*

2. Let $L \in \mathcal{L}(V)$. Then $\chi_L(L) = 0$.

$$B = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \vdots & & & \\ A_{n1} & & \cdots & A_{nn} \end{bmatrix}^T$$

$AA^* = A^*A = |A| I$

adjoint matrix

Proof. Let $\chi_A(t) = c_0 + c_1 t + \cdots + c_n t^n$, $c_n = 1$. Let $B = \mathrm{cof}(tE_n - A)$ be the cofactor matrix of $tE_n - A$, i.e. $(tE_n - A)B = \chi_A(t)E_n$. Since the entries of B are cofactors, $B \in \mathrm{M}_n(F[t]_{n-1})$. Therefore $B = B_0 + tB_1 + \cdots + t^{n-1}B_{n-1}$, $B_i \in \mathrm{M}_n(F)$ and the equality

$$(tE_n - A)(B_0 + tB_1 + \cdots + t^{n-1}B_{n-1}) = (c_0 + c_1 t + \cdots + c_n t^n)E_n$$

implies

$$\begin{aligned} -AB_0 &= c_0 E_n \\ B_0 - AB_1 &= c_1 E_n \\ &\vdots \\ B_{n-2} - AB_{n-1} &= c_{n-1}E_n \\ B_{n-1} &= c_n E_n \end{aligned}$$

Multiply the second equality on the left by A, the third equality by A^2, and so on, and then add all them up to get $0 = c_0 E_n + c_1 A + \cdots + c_n A^n$, as required.

For $L \in \mathcal{L}(V)$, choose a basis $\mathcal{B}$ in V. Then $\chi_L = \chi_{[L]_\mathcal{B}}$ and

$$[\chi_L(L)]_\mathcal{B} = [\chi_{[L]_\mathcal{B}}(L)]_\mathcal{B} = \chi_{[L]_\mathcal{B}}([L]_\mathcal{B}) = 0$$

by Lemma 9.9 and (1). Finally, if the matrix of an operator relative to a certain basis is zero, then the operator itself is zero. □

Exercise 9.4. *Using the Hamilton-Cayley theorem, compute the inverse of the matrix*

$$A = \begin{pmatrix} -1 & 2 & 0 \\ 1 & 0 & -1 \\ 2 & 1 & -1 \end{pmatrix}$$

Hint. Using the identity from the Hamilton-Cayley theorem, express A^{-1} as a polynomial in A.

Exercise 9.5. *Let $V = \mathrm{M}_2(F)$ and $L \in \mathcal{L}(V)$, $L(X) = X\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Write the identity from the Hamilton-Cayley theorem and verify it directly.*

10 Special types of operators and matrices

Diagonalizable operators and matrices

Definition. An operator $L \in \mathcal{L}(V)$ is called **diagonalizable** (可对角化的) if its matrix relative to some basis of V is diagonal.

Theorem 10.1. *Let $L \in \mathcal{L}(V)$, $\lambda_1, \ldots, \lambda_m \in F$ be the distinct eigenvalues of L. Then the following conditions are equivalent:*

1. *L is diagonalizable;*

2. *V has a basis consisting of the eigenvectors of L;*

3. *$V = V_{\lambda_1}(L) \oplus \cdots \oplus V_{\lambda_m}(L)$;*

4. *$\dim V = \dim V_{\lambda_1}(L) + \cdots + \dim V_{\lambda_m}(L)$.*

Proof. 1 $\Leftrightarrow$ 2 The matrix of L relative to the basis $v_1, \ldots, v_n$ has the form

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

if and only if $L(v_j) = \lambda_j v_j$ for all $j = 1, \ldots, n$.

2 $\Rightarrow$ 3 If V has a basis consisting of of eigenvectors, then any vector of V is a linear combination combination of these eigenvectors, and thus $V = V_{\lambda_1}(L) + \cdots + V_{\lambda_m}(L)$. Proposition 8.4 shows that this sum is direct.

3 $\Rightarrow$ 4 Follows from Proposition 4.6.

4 $\Rightarrow$ 2 Let $\dim V = n$. Choose a basis in each subspace $V_{\lambda_j}(L)$. Join them together and obtain a set $v_1, \ldots, v_n$, consisting of n eigenvectors of L. Show that it is linearly independent. Assume $a_1 v_1 + \cdots + a_n v_n = 0$ for some $a_1, \ldots, a_n \in F$. Let u_j be the sum of the terms $a_i v_i, v_i \in V_{\lambda_j}(L)$. Then $u_j \in V_{\lambda_j}(L)$, and $u_1 + \cdots + u_m = 0$. Then Theorem 8.3 implies $u_1 = \cdots = u_m = 0$. Since each u_j is the linear combination of the basis vectors of $V_{\lambda_j}(L)$, all the coefficients a_i are zero. A linearly independent set of n vectors in n-dimensional vector space is a basis.

□

Corollary 10.2. *Let $\dim V = n$ and $L \in \mathcal{L}(V)$ have n distinct eigenvalues. Then L is diagonalizable.*

Proof. Let $v_1, \ldots, v_n$ be eigenvectors of L associated to the distinct eigenvalues which are linearly independent by Theorem 8.3. A linearly independent set of n vectors in n-dimensional vector space is a basis. □

Examples. 1. Let $\mathcal{B} = \{e_1, e_2\}$ be a basis in V and $L \in \mathcal{L}(V), [L]_{\mathcal{B}} = A$, where

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Since $\chi_L(t) = \chi_A(t) = t^2$, L has only one eigenvalue 0 with associated eigenvector $u = \alpha e_1$. Therefore, V has no basis consisting of eigenvectors, and hence L is not diagonalizable.

2. Let $V = \mathrm{M}_2(F)$ and $L \in \mathcal{L}(V), L(A) = A^T$. Then

$$A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, A_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, A_4 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

is a basis of V consisting of eigenvectors of L, whence L is diagonalizable.

Exercise 10.1. *Which of the following linear operators $L \in \mathcal{L}(V)$ are diagonalizable:*

a. $V = \mathbb{R}[t]_n, L(f) = f'(t)$

b. $V = \mathbb{R}[t]_n, L(f) = tf'(t)$

c. V is the geometric plane, L is the reflection across a line passing through the origin

Exercise 10.2. *Let $\mathcal{B}$ be a basis in V and $L \in \mathcal{L}(V), [L]_\mathcal{B} = A$, where*

$$A = \begin{pmatrix} 0 & -1 & 2 & -1 \\ -2 & 1 & -4 & 2 \\ -2 & 2 & -5 & 2 \\ -1 & 1 & -2 & 0 \end{pmatrix}$$

Determine if L is diagonalizable.

Definition. A matrix $A \in \mathrm{M}_n(F)$ is called **diagonalizable** (可对角化的) if there are a diagonal matrix $D \in \mathrm{M}_n(F)$ and an invertible matrix $U \in \mathrm{M}_n(F)$ such that $A = UDU^{-1}$.

Proposition 10.3. *1. Let $\mathcal{B}$ be a basis in a vector space V and $L \in \mathcal{L}(V)$. Then L is diagonalizable if and only if $[L]_\mathcal{B}$ is diagonalizable.*

2. Let $A \in \mathrm{M}_n(F)$. Then $L \in \mathcal{L}(\mathrm{M}_{n,1}(F)), L(X) = AX$ is diagonalizable if and only if A is diagonalizable.

Proof. 1. Assume L is diagonalizable and $\mathcal{C}$ is a basis such that $[L]_\mathcal{C}$ is diagonal. Put $U = M_{\mathcal{B}\rightsquigarrow\mathcal{C}}$, then $[L]_\mathcal{B} = U[L]_\mathcal{C}U^{-1}$ by Theorem 6.7.

Conversely, let $[L]_\mathcal{B} = UDU^{-1}$ for a diagonal matrix $D \in \mathrm{M}_n(F)$ and an invertible matrix $U \in \mathrm{M}_n(F)$. Lemma 6.8 asserts that there is a basis C of V such that $M_{\mathcal{B}\rightsquigarrow\mathcal{C}} = U$. Then

$$[L]_\mathcal{C} = M_{\mathcal{B}\rightsquigarrow\mathcal{C}}^{-1}[L]_\mathcal{B}M_{\mathcal{B}\rightsquigarrow\mathcal{C}} = U^{-1}[L]_\mathcal{B}U = D$$

is diagonal.

2. Let $\mathcal{F}$ be the standard basis of $\mathrm{M}_{n,1}(F)$, i.e., $\mathcal{F} = \{e_1, \dots, e_n\}$, where $e_i = (0 \cdots 0\,1\,0 \cdots 0)^T$ has 1 at ith position. Then $[L]_\mathcal{F} = A$ and one can use (1). □

Exercise 10.3. *For the matrix*

$$A = \begin{pmatrix} -7 & -4 & -4 \\ 8 & 5 & 4 \\ 4 & 2 & 3 \end{pmatrix},$$

find a diagonal matrix $D \in \mathrm{M}_3(\mathbb{R})$ and an invertible matrix $U \in \mathrm{M}_3(\mathbb{R})$ such that $A = UDU^{-1}$. Verify the equality obtained.

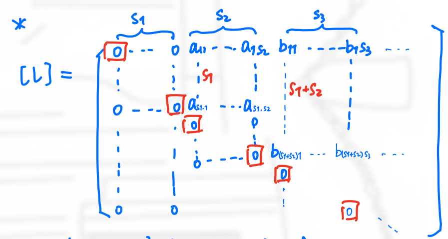
Hint. Consider $L \in \mathcal{L}(\mathrm{M}_{n,1}(F)), L(X) = AX$. If $\mathcal{C}$ is a basis such that $D = [L]_{\mathcal{C}}$ is diagonal, denote by U the matrix which has $\mathcal{C}$ as its columns. Then $M_{\mathcal{F} \rightsquigarrow \mathcal{C}} = U$ if $\mathcal{F}$ is the standard basis of $\mathrm{M}_{n,1}(F)$.

Nilpotent operators and matrices

Definition. An operator $L \in \mathcal{L}(V)$ is called **nilpotent** (幂零的) if its matrix relative to some basis of V is strictly upper triangular, i.e. only the elements above the diagonal may be non-zero.

Proposition 10.4. *Let $N \in \mathcal{L}(V)$ and $\dim(V) = n$. The following properties are equivalent*

*



let $\{e_{11} \ldots e_{s1}\}$ form a basis of ker N.

$\{e_{11} \ldots e_{s_1}, e_{12} \ldots e_{s_2}\}$ a basis of ker N.

$N(e_{11}) = \cdots = N(e_{s1}) = 0$ since $e_{11} \ldots e_{s1} \in \ker N$

$N^2(e_{12}) = 0 \Rightarrow N(N(e_{12})) = 0$. $N(e_{12}) \in \ker N$. $\to$ write as l.c of basis

1. *N is nilpotent*
2. *$\chi_N(t) = t^n$*
3. *$N^j = 0$ for some $j \in \mathbb{N}$*
4. *For any $v \in V$ there is $j \in \mathbb{N}$ such that $N^j(v) = 0$*

Proof. $1 \Rightarrow 2$ Obvious

$2 \Rightarrow 3$ Follows from the Hamilton-Cayley theorem

$3 \Rightarrow 4$ Obvious

$4 \Rightarrow 3$ Let $v_1, \ldots, v_n$ be a basis of V. Then there are $j_1, \ldots, j_n$ such that $N^{j_k}(v_k) = 0$ for $1 \leq k \leq n$. If $j = \max(j_i, \ldots, j_n)$ then $N^j(v_k) = 0$ for $1 \leq k \leq n$. It implies that $N^j(v) = 0$ for any $v \in V$ since v can be represented as a linear combination of $v_1, \ldots, v_n$.

$3 \Rightarrow 1$ First notice that

$$\operatorname{Ker}(N) \subset \operatorname{Ker}(N^2) \subset \cdots \subset \operatorname{Ker}(N^j) = V.$$

Choose a basis of Ker N. Then extend this to a basis of Ker N^2. Then extend to a basis of Ker N^3. Continue in this fashion, eventually getting a basis of V. Then the matrix N* relative to this basis is strictly upper triangular. Indeed, first columns, corresponding to basis vectors from Ker N, consist of all 0's, because the images of these vectors is 0. The next set of columns comes from basis vectors in Ker N^2. Their images lie in Ker N and thus are linear combinations of the previous basis vectors. Thus all nonzero entries in these columns lie above the diagonal etc.

□

Corollary 10.5. *If $N \in \mathcal{L}(V)$ is nilpotent then $N^n = 0$, where $n = \dim(V)$.*

Example. The differentiation operator on $F[t]_n$ is nilpotent

Exercise 10.4. *Show that there is no non-zero operator which is both diagonalizable and nilpotent.*

Definition. A matrix $A \in \mathrm{M}_n(F)$ is called **nilpotent** (幂零的) if $A^j = 0$ for some $j \in \mathbb{N}$.

Exercise 10.5. *Let $\mathcal{B}$ be a basis in a vector space V and $N \in \mathcal{L}(V)$. Then N is nilpotent if and only if $[N]_{\mathcal{B}}$ is nilpotent.*

Proposition 10.6. *A matrix $A \in \mathrm{M}_n(F)$ is nilpotent if and only if $A = UQU^{-1}$ for a strict upper triangular matrix $Q \in \mathrm{M}_n(F)$ and an invertible matrix $U \in \mathrm{M}_n(F)$.*

Proof. Consider $L \in \mathcal{L}(\mathrm{M}_{n,1}(F)), L(X) = AX$. Remind that $[L]_{\mathcal{F}} = A$, where $\mathcal{F}$ is the standard basis of $\mathrm{M}_{n,1}(F)$. Clearly L is nilpotent if and only if A is nilpotent.

If L is nilpotent and $\mathcal{C}$ is a basis in $\mathrm{M}_{n,1}(F)$ such that $Q = [L]_{\mathcal{C}}$ is strict upper triangular. Denote by U the matrix which has $\mathcal{C}$ as its columns, then $M_{\mathcal{F}\rightsquigarrow\mathcal{C}} = U$. Now $A = UQU^{-1}$ by Theorem 6.7.

For the reverse implication, remark that $Q^n = 0$ for any strict upper triangular matrix $Q \in \mathrm{M}_n(F)$ and thus $A^n = (UQU^{-1})^n = UQ^nU^{-1} = 0$. □

Exercise 10.6. *For the matrix*

$$A = \begin{pmatrix} 3 & 1 & 3 \\ -4 & -2 & -4 \\ -1 & 0 & -1 \end{pmatrix},$$

find a strict upper triangular matrix $Q \in \mathrm{M}_3(\mathbb{R})$ and an invertible matrix $U \in \mathrm{M}_3(\mathbb{R})$ such that $A = UQU^{-1}$. Verify the equality obtained.

Hint. Follow the proof of Propositions 10.6 and 10.4.

11 Generalized eigenvectors

Generalized eigenspace decomposition

Definition. Let $L \in \mathcal{L}(V)$ and $\lambda \in F$ be its eigenvalue. A nonzero vector $v \in V$ is called a **generalized eigenvector** (广义特征向量) of L associated to λ if $(L - \lambda\,\mathrm{id}_V)^j(v) = 0$ for some $j \in \mathbb{N}$.

The set of all generalized eigenvectors of L associated to λ, together with 0, is called the **generalized eigenspace** (广义特征空间) and is denoted by $V(\lambda, L)$.

Remark. Any eigenvector is a generalized eigenvector vector and $V_\lambda(L) \subset V(\lambda, L)$.

Proposition 11.1. *Let $L \in \mathcal{L}(V)$ and $\lambda \in F$. Then*

1. *$V(\lambda, L) \neq \{0\}$ if and only if λ is an eigenvalue of L*
2. *$V(\lambda, L)$ is a subspace of V*

Proof. 1. If $v \in V(\lambda, L)$ and $v \neq 0$, then there is $m \geq 0$ such that $(L - \lambda\,\mathrm{id}_V)^m(v) \neq 0, (L - \lambda\,\mathrm{id}_V)^{m+1}(v) = 0$. Then $L(u) = \lambda u, u \neq 0$ for $u = (L - \lambda\,\mathrm{id}_V)^m(v)$, whence λ is an eigenvalue of L. The inverse implication is obvious since $V_\lambda(L) \subset V(\lambda, L)$.

2. If $(L-\lambda\,\mathrm{id}_V)^j(v_1)=0$ and $(L-\lambda\,\mathrm{id}_V)^i(v_2)=0$, then $(L-\lambda\,\mathrm{id}_V)^{\max(i,j)}(v_1+v_2)=0$ and $(L-\lambda\,\mathrm{id}_V)^j(av_1)=0$ for any $a\in F$. $[\, L^{-j}(L-\lambda id_V)^j = (L^{-1})^j(L-\lambda id_V)^j = (L^{-1}(L-id_V))^j = (id_V - L^{-1}\lambda)^j \,]$ □

$(id_V - \lambda L^{-1})^j(v) = L^{-j}(L-\lambda id_V)^j(v) = L^{-j}(0)=0 \Rightarrow \lambda^{-j}(id_V-\lambda L^{-1})^j(v) = \lambda^{-j}\cdot 0 = 0. \Rightarrow [\lambda^{-1}(id_V-\lambda L^{-1})]^j(v) = (\lambda^{-1}id_V - L^{-1})^j(v)=0$
$\Rightarrow (L^{-1}-\lambda^{-1}id_V)^j(v)=0$

Exercise 11.1. *Suppose $L\in\mathcal{L}(V)$ is invertible. Prove that $V(\lambda,L)=V(\lambda^{-1},L^{-1})$ (compare with Exercise 8.3).*

Lemma 11.2. *Let $N\in\mathcal{L}(V)$ and $\mathrm{Ker}(N^j)=\mathrm{Ker}(N^{j+1})$ for $j>0$ then $\mathrm{Ker}(N^j)=\mathrm{Ker}(N^{j+m})$ for any $m\in\mathbb{N}$. In other words, if two consecutive terms in the chain*

$$\mathrm{Ker}(N)\subset\mathrm{Ker}(N^2)\subset\cdots\mathrm{Ker}(N^j)\subset\mathrm{Ker}(N^{j+1})\subset\cdots$$

are equal then all the subsequent terms are equal.

Proof. It suffices to prove that $\mathrm{Ker}(N^{j+k})=\mathrm{Ker}(N^{j+k+1})$ for any $k\in\mathbb{N}$. Let $v\in\mathrm{Ker}(N^{j+k+1})$ and $N^{j+k+1}(v)=N^{j+1}(N^k(v))=0$. Then $N^k(v)\in\mathrm{Ker}(N^{j+1})=\mathrm{Ker}(N^j)$, whence $N^{j+k}(v)=N^j(N^k(v))=0$ as required. □

Problem 11.3. *Let $V=\mathbb{R}[t]_2$ and $L\in\mathcal{L}(V), L(f)=f(1-t)-f(0)t$. Find the eigenvalues, the associated eigenspaces and the associated generalized eigenspaces of L.*

Solution. Choose a basis of V, let $\mathcal{B}=\{f_1,f_2,f_3\}$ and $f_1=1,f_2=t,f_3=t^2$. Then

$$\begin{aligned}L(f_1)&=1-t=f_1-f_2\\L(f_2)&=1-t=f_1-f_2\\L(f_3)&=1-2t+t^2=f_1-2f_2+f_3,\end{aligned}$$

whence

$$A=[L]_{\mathcal{B}}=\begin{pmatrix}1&1&1\\-1&-1&-2\\0&0&1\end{pmatrix}.$$

Then $\chi_A(\lambda)=\lambda^2(\lambda-1)=0$, whence $\lambda_1=0,\lambda_2=1$ are the eigenvalues of L.

First solve $B_1X=0$, where $B_1=A-\lambda_1E_3=A$, which gives $V_{\lambda_1}(L)=\{\alpha-\alpha t\mid\alpha\in\mathbb{R}\}$. In order to find $v\in V$ satisfying $(L-\mathrm{id}_V)^j(v)=0$, solve the equation $B_1^jX=0$. Since

$$B_1^2=B_1^3=\begin{pmatrix}0&0&0\\0&0&-1\\0&0&1\end{pmatrix},$$

Lemma 11.2 implies that $V(\lambda_1,L)=\mathrm{Ker}(L-\lambda_1\,\mathrm{id}_V)^2=\{\alpha+\beta t\mid\alpha,\beta\in\mathbb{R}\}$.

For $\lambda_2=1$, one has

$$B_2=A-\lambda_2E_3=A-E_3=\begin{pmatrix}0&1&1\\-1&-2&-2\\0&0&0\end{pmatrix},$$

which gives $V_{\lambda_2}(L) = \{\alpha t - \alpha t^2 \mid \alpha \in \mathbb{R}\}$. Since

$$B_2^2 = \begin{pmatrix} -1 & -2 & -2 \\ 2 & 3 & 3 \\ 0 & 0 & 0 \end{pmatrix},$$

the solutions of $B_2^2 X = 0$ and $B_2 X = 0$ coincide. Therefore $\operatorname{Ker}(L - \lambda_2 \operatorname{id}_V)^2 = \operatorname{Ker}(L - \lambda_2 \operatorname{id}_V)$ and Lemma 11.2 implies $V(\lambda_2, L) = V_{\lambda_2}(L)$. □

Theorem 11.4. *Let V be a vector space over $\mathbb{C}$ and $L \in \mathcal{L}(V)$. If $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of L, then*

1. *$V = V(\lambda_1, L) \oplus \cdots \oplus V(\lambda_m, L)$;* Divide and Conquer
2. *$V(\lambda_1, L), \ldots, V(\lambda_m, L)$ are L-invariant;*
3. *$(L - \lambda_j \operatorname{id}_V)|_{V(\lambda_j, L)}$ is nilpotent.*

Proof. Let λ be one of the eigenvalues of L. For any $v \in V(\lambda, L)$ there is $j \in \mathbb{N}$ such that $(L - \lambda \operatorname{id}_V)^j(v) = 0$. Then $(L - \lambda \operatorname{id}_V)^{j-1}((L - \lambda \operatorname{id}_V)(v)) = 0$, whence $(L - \lambda \operatorname{id}_V)(v) \in V(\lambda, L)$ and $V(\lambda, L)$ is $(L - \lambda \operatorname{id}_V)$-invariant. Then one can consider $(L - \lambda \operatorname{id}_V)|_{V(\lambda, L)}$ which is nilpotent by Proposition 10.4. The fact that $V(\lambda, L)$ is $(L - \lambda \operatorname{id}_V)$-invariant easily implies that $V(\lambda, L)$ is L-invariant.

Over $\mathbb{C}$ one has a factorization

$$\chi_L(t) = (t - \lambda_1)^{k_1} \cdots (t - \lambda_m)^{k_m}.$$

Denote

$$\chi_j(t) = (t - \lambda_1)^{k_1} \cdots (t - \lambda_{j-1})^{k_{j-1}} (t - \lambda_{j+1})^{k_{j+1}} \cdots (t - \lambda_m)^{k_m}.$$

Then the greatest common divisor of $\chi_1, \ldots, \chi_m$ equals 1, hence $f_1 \chi_1 + \cdots + f_m \chi_m = 1$ for some $f_1, \ldots, f_m \in \mathbb{C}[t]$,

Consider $W_j = \operatorname{Im} f_j(L) \chi_j(L)$ which is L-invariant by Lemma 9.10. Since

$$\begin{aligned} &(L - \lambda_j \operatorname{id}_V)^{k_j} f_j(L) \chi_j(L) \\ &= (L - \lambda_j \operatorname{id}_V)^{k_j} f_j(L)(L - \lambda_1 \operatorname{id}_V)^{k_1} \cdots (L - \lambda_{j-1} \operatorname{id}_V)^{k_{j-1}} (L - \lambda_{j+1} \operatorname{id}_V)^{k_{j+1}} \cdots (L - \lambda_m \operatorname{id}_V)^{k_m} \\ &= f_j(L) \chi_L(L) = 0 \end{aligned}$$

by the Hamilton-Cayley theorem, $W_j \subset V(\lambda_j, L)$ for any $1 \leq j \leq m$.

The equality $f_1 \chi_1 + \cdots + f_m \chi_m = 1$ implies

$$f_1(L) \chi_1(L) + \cdots + f_m(L) \chi_m(L) = \operatorname{id}_V,$$

whence $W_1 + \cdots + W_m = V$. Since $W_j \subset V(\lambda_j, L)$, one has $V(\lambda_1, L) + \cdots + V(\lambda_m, L) = V$.

Assume $v_j \in V(\lambda_j, L) \cap U_j$, where

$$U_j = V(\lambda_1, L) + \cdots + V(\lambda_{j-1}, L) + V(\lambda_{j+1}, L) + \cdots + V(\lambda_m, L).$$

Thus $v_j = v_1 + \cdots v_{j-1} + v_{j+1} + \cdots v_m$, where $v_i \in V(\lambda_i, L), 1 \leq i \leq m$. Let $n = \dim V$. Then $(L - \lambda_i \operatorname{id}_V)^{k_i'}(v_i) = 0$, where $k_i' = \dim V(\lambda_i, L) \leq n$, whence $(L - \lambda_i \operatorname{id}_V)^n(v_i) = 0$ for all $1 \leq i \leq m$. Therefore $(L - \lambda_j \operatorname{id}_V)^n(v_j) = 0$ and

$$(L - \lambda_1 \operatorname{id}_V)^n \cdots (L - \lambda_{j-1} \operatorname{id}_V)^n (L - \lambda_{j+1} \operatorname{id}_V)^n \cdots (L - \lambda_m \operatorname{id}_V)^n (v_j) = 0.$$

The polynomials $(t - \lambda_j)^n$ and

$$\rho_j(t) = (t - \lambda_1)^n \cdots (t - \lambda_{j-1})^n (t - \lambda_{j+1})^n \cdots (t - \lambda_m)^n$$

are coprime, whence $h_j(t)(t - \lambda_j)^n + g_j(t)\rho_j(t) = 1$ for some $h_j, g_j \in \mathbb{C}[t]$. Then

$$h_j(L)(L - \lambda_j \operatorname{id}_V)^n + g_j(L)\rho_j(L) = \operatorname{id}_V,$$

whence

$$v_j = h_j(L)(L - \lambda_j \operatorname{id}_V)^n(v_j) + g_j(L)\rho_j(L)(v_j) = 0$$

which proves (1). □

Corollary 11.5. *Let V be a vector space over $\mathbb{C}$ and $L \in \mathcal{L}(V)$. Then there is a basis of V consisting of generalized eigenvectors of L.*

Exercise 11.2. *Let*

$$V = \mathbb{R}[t, s]_2 = \{a_{0,0} + a_{1,0}t + a_{0,1}s + a_{2,0}t^2 + a_{1,1}ts + a_{0,2}s^2 \mid a_{0,0}, a_{1,0}, a_{0,1}, a_{2,0}, a_{1,1}, a_{0,2} \in \mathbb{R}\}$$

and $L \in \mathcal{L}(V), L(f) = f(t + 1, s + 1)$.

i) *Find the eigenvalues, the associated eigenspaces and the associated generalized eigenspaces of L.*

ii) *Choose bases of the generalized eigenspaces $V(\lambda, L)$ and find the matrices of the restrictions $L|_{V(\lambda,L)}$ relative to them.*

iii) *Verify that $L|_{V(\lambda,L)} - \lambda \operatorname{id}_{V(\lambda,L)}$ is nilpotent*

Hint. Follow the solution of Problem 11.3.

Jordan basis for a nilpotent operator

Theorem 11.6. *Let V be a vector space over F, $N \in \mathcal{L}(V)$ be a nilpotent operator. Then there exist $v_1, \ldots, v_s \in V$ and $m_1, \ldots, m_s \in \mathbb{N}$ such that the vectors*

$$\begin{gathered} N^{m_1}(v_1), \ldots, N(v_1), v_1, \\ N^{m_2}(v_2), \ldots, N(v_2), v_2, \\ \vdots \\ N^{m_s}(v_s), \ldots, N(v_s), v_s \end{gathered}$$

form a basis of V and $N^{m_1+1}(v_1) = \cdots = N^{m_s+1}(v_s) = 0$.

$V = \operatorname{span}(e)$

$N(e) = \lambda e \quad N^2(e) = N(\lambda e) = \lambda^2 e$

$\vdots$

$N^k(e) = \lambda^k e = 0 \Rightarrow \lambda^k = 0 \Rightarrow \lambda = 0 \Rightarrow N = 0$

$N^J = 0 \Rightarrow N^{J-1} = N^J \cdot N^{-1} = 0. \Rightarrow \cdots \Rightarrow N = 0$ (not invertible)

or N bijective. N^J bijective ($N^J = 0$. impossible)

Proof. We use induction on $\dim V$. The case $\dim V = 1$ is trivial since a nilpotent operator on one-dimensional space must be zero, and we can take $s = 1$, any nonzero $v_1 \in V$ and $m_1 = 0$.

Suppose $\dim V > 1$. If $\mathrm{Im}(N) = V$, then N is invertible by Proposition 7.7, which is impossible for the nilpotent operator. Therefore $\dim \mathrm{Im}(N) < \dim V$. If $\mathrm{Im}(N) = \{0\}$, then $N = 0$, and hence one can take an arbitrary basis $v_1, \ldots, v_s$ in V and put $m_1 = \cdots = m_s = 0$.

Imp(L) is L-invariant

n.假设

Thus $\mathrm{Im}(N)$ is N-invariant by Lemma 9.10 and our induction hypothesis applied to $L|_{\mathrm{Im}(N)}$ gives $v_1, \ldots, v_s \in \mathrm{Im}(N)$ and $m_1, \ldots, m_s \in \mathbb{N}$ such that the conclusion of the theorem holds for $\mathrm{Im}(N)$). If $v_i = Nu_i$ for some $u_i \in V$, it means that

$$\begin{aligned}
&N^{m_1+1}(u_1), \ldots, N^2(u_1), N(u_1),\\
&N^{m_2+1}(u_2), \ldots, N^2(u_2), N(u_2),\\
&\vdots\\
&N^{m_s+1}(u_s), \ldots, N^2(u_s), N(u_s)
\end{aligned}$$

form a basis of $\mathrm{Im}(N)$, and $N^{m_1+2}(u_1) = \cdots = N^{m_s+2}(u_s) = 0$.

Extend this linearly independent set with $u_1, \ldots, u_s$ to a bigger linearly independent set

$$\begin{aligned}
&N^{m_1+1}(u_1), \ldots, N^2(u_1), N(u_1), u_1,\\
&N^{m_2+1}(u_2), \ldots, N^2(u_2), N(u_2), u_2,\\
&\vdots\\
&N^{m_s+1}(u_s), \ldots, N^2(u_s), N(u_s), u_s.
\end{aligned}$$

Indeed, consider a zero linear combination of these vectors and apply N to it. It gives a linear combination of the vectors

$$\begin{aligned}
&N^{m_1+2}(u_1) = 0, N^{m_1+1}(u_1), \ldots, N^2(u_1), N(u_1),\\
&N^{m_2+2}(u_2) = 0, N^{m_2+1}(u_1), \ldots, N^2(u_2), N(u_2),\\
&\vdots\\
&N^{m_s+2}(u_s) = 0, N^{m_s+1}(u_1), \ldots, N^2(u_s), N(u_s).
\end{aligned}$$

Therefore all its coefficients except possibly the coefficients $c_1, \ldots, c_s$ corresponding to the vectors $N^{m_1+1}(u_1), \ldots, N^{m_s+1}(u_s)$ in the original linear combination are zero. Then $c_1 N^{m_1+1}(u_1) + \cdots + c_s N^{m_s+1}(u_s) = 0$, whence $c_1 = \cdots = c_s = 0$ because $N^{m_1+1}(u_1), \ldots, N^{m_s+1}(u_s)$ are linearly independent.

We have shown that the vectors

$$\begin{aligned}
&N^{m_1+1}(u_1), \ldots, N^2(u_1), N(u_1), u_1,\\
&N^{m_2+1}(u_2), \ldots, N^2(u_2), N(u_2), u_2,\\
&\vdots\\
&N^{m_s+1}(u_s), \ldots, N^2(u_s), N(u_s), u_s
\end{aligned}$$

are linearly independent. Extend them with $w_1, \dots, w_t \in V$ to a basis of V. Since $N(w_j) \in \operatorname{Im}(N), 1 \le j \le t$, the vector Nw_j is a linear combination of the vectors

$$\begin{aligned}
&N^{m_1+1}(u_1), \dots, N^2(u_1), N(u_1),\\
&N^{m_2+1}(u_2), \dots, N^2(u_2), N(u_2),\\
&\vdots\\
&N^{m_s+1}(u_s), \dots, N^2(u_s), N(u_s).
\end{aligned}$$

Then $N(w_j) = N(x_j)$, where x_j is a linear combination of the vectors

$$\begin{aligned}
&N^{m_1}(u_1), \dots, N(u_1), u_1,\\
&N^{m_2}(u_2), \dots, N(u_2), u_2,\\
&\vdots\\
&N^{m_s}(u_s), \dots, N(u_s), u_s.
\end{aligned}$$

Now put $u_{s+j} = w_j - x_j$. Then the vectors

$$\begin{aligned}
&N^{m_1+1}u_1, \dots, N^2u_1, Nu_1, u_1,\\
&\vdots\\
&N^{m_s+1}u_s, \dots, N^2u_s, Nu_s, u_s,\\
&u_{s+1},\\
&\vdots\\
&u_{s+t}
\end{aligned}$$

form a basis of V since it is obtained by subtracting from certain basis vectors linear combinations of the others. Moreover,

$$N^{m_1+2}(u_1) = \dots = N^{m_s+2}(u_s) = N(u_{s+1}) = \dots = N(u_{s+t}) = 0,$$

since $N(u_{s+j}) = N(w_j - x_j) = N(w_j) - N(x_j) = 0, 1 \le j \le t$. $\square$

12 Jordan normal form

Jordan normal form for operators and matrices

Definition. An $n \times n$ matrix of the form

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0\\ 0 & \lambda & 1 & \dots & 0 & 0\\ 0 & 0 & \lambda & \dots & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & \dots & \lambda & 1\\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

is called a **Jordan block** (若尔当块). A block diagonal matrix in which each block is a Jordan block, is called a **Jordan matrix** (若尔当矩阵)

$$\begin{pmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & J_{n_s}(\lambda_s) \end{pmatrix}.$$

For $L \in \mathcal{L}(V)$ a basis of V is a **Jordan basis** if the matrix of L relative to this basis is a Jordan matrix. This matrix is then called a **Jordan normal form** (若尔当标准型) of L.

Remark. In a Jordan matrix $\operatorname{diag}(J_{n_1}(\lambda_1), \ldots, J_{n_s}(\lambda_s))$ the elements $\lambda_1, \ldots, \lambda_s$ may not be all distinct.

Theorem 12.1. *Let V be a vector space over $\mathbb{C}$ and $L \in \mathcal{L}(V)$. Then there exists a Jordan basis for L.*

Moreover, a Jordan normal form of the operator is unique up to the order of the Jordan blocks (without proof).

Proof. By Theorem 11.4, the space V is a direct sum of the generalized eigenspaces of L. Moreover, if $\lambda_i \in F$ is its eigenvalue, the restriction of $L - \lambda_i \operatorname{id}_V$ on $V(\lambda_i, L)$ is nilpotent. Then Theorem 11.6 implies that the matrix of $(L - \lambda_i \operatorname{id}_V)|_{V(\lambda_i, L)}$ relative to a certain basis has a form

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}.$$

Then the matrix of $L|_{V(\lambda_i, L)}$ relative to this basis is $J_{n_j}(\lambda_i)$, where $n_j = \dim V(\lambda_i, L)$. Proposition 8.7 completes the proof. □

Remark. The basis vectors $u_1, \ldots u_m$ corresponding to a Jordan block $J_m(\lambda)$ are characterized by the following identities:

$$\begin{aligned} L(u_1) &= \lambda u_1 \quad \to \text{eigenvector} \\ L(u_2) &= u_1 + \lambda u_2 \quad \to \text{generalized eigenvector.} \\ &\vdots \\ L(u_m) &= u_m + \lambda u_{m-1}. \end{aligned}$$

and are called a **Jordan chain** (若尔当链). A Jordan chain starts from an eigenvector. If $\mathcal{J}$ is a Jordan basis and $C = [L]_{\mathcal{J}} - \lambda E_n$, it satisfies $CX_1 = 0, CX_2 = X_1, \ldots, CX_m = X_{m-1}$, where $X_j = [u_j]_{\mathcal{J}}, 1 \leq j \leq m$.

Corollary 12.2. *Let V be a vector space over $\mathbb{C}$ and $L \in \mathcal{L}(V)$. The sum of the orders of the Jordan blocks associated with λ is equal to its multiplicity as a root of the characteristic polynomial. The number of the Jordan blocks associated with λ is equal to* $\dim \operatorname{Ker}(L - \lambda \operatorname{id}_V)$.

Proof. Let J be a Jordan normal form of L and $\dim V = n$. Then $tE_n - J$ is an upper triangular matrix, and its determinant equals $\prod_{i=1}^{n}(t - c_i)$, where $c_1, \dots, c_n$ are the diagonal entries of J.

重根数.

The multiplicity of λ as a root of $\chi_L(t) = \det(tE_n - J)$ equals the number of occurrences of λ among $c_1, \dots, c_n$ which in turn equals the sum of the orders of the Jordan blocks, associated with λ.

Let s be the number of the Jordan blocks associated with λ and $u_1, \dots, u_s$ be the starting vectors for the Jordan chains. The form of the Jordan matrix implies that $v \in V$ is an eigenvector of L associated with λ if and only its coordinates relative to all the vectors of the Jordan basis other than $u_1, \dots, u_s$ are all zero. Therefore $u_1, \dots, u_s$ form a basis of $\operatorname{Ker}(L - \lambda \operatorname{id}_V)$. □

if not, the image will have terms like $\lambda v_j + v_{j+1}$ ...

Example. If $\chi_L(t) = (t-1)^2(t+1)$, then a Jordan normal form of L can take one of two forms up to the order of the blocks:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

They correspond to the cases where $\dim \operatorname{Ker}(L - \operatorname{id}_V) = 2$ and $\dim \operatorname{Ker}(L - \operatorname{id}_V) = 1$, respectively.

Problem 12.3. *Let $V = \mathrm{M}_2(\mathbb{C})$ and $L \in \mathcal{L}(V)$,*

$$L(A) = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} A \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

$$\begin{pmatrix} 1 & 0 & -1 & 0 \\ -2 & 1 & 1 & -1 \\ 1 & 0 & -1 & 0 \\ -1 & 1 & 0 & -1 \end{pmatrix}$$

Find a Jordan normal form and a Jordan basis for L.

$$\begin{pmatrix} 1 & -2 & 1 & -1 \\ 0 & 1 & 0 & 1 \\ -1 & 1 & -1 & 0 \\ 0 & -1 & 0 & -1 \end{pmatrix}\begin{pmatrix} \ \end{pmatrix}.$$

Solution. Choose the standard basis $\mathcal{B} = \{E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2}\}$ for $\mathrm{M}_2(\mathbb{C})$, then

$$[L]_{\mathcal{B}} = \begin{pmatrix} 2 & 0 & -1 & 0 \\ -2 & 2 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{pmatrix}.$$

$$\begin{cases} x_2 + x_4 = 0 \\ x_2 = x_1 + x_3 \end{cases}$$

$(1 \ 1 \ 0 \ -1)$

L has the unique eigenvalue 1 and two linearly independent eigenvectors $[A_1]_{\mathcal{B}} = (1, 1, 1, 0)^T$ and $[A_2]_{\mathcal{B}} = (1, 0, 1, -1)^T$. Denote

$(0 \ 1 \ 1 \ -1)$

$$C = [L]_{\mathcal{B}} - E_4 = \begin{pmatrix} 1 & 0 & -1 & 0 \\ -2 & 1 & 1 & -1 \\ 1 & 0 & -1 & 0 \\ -1 & 1 & 0 & -1 \end{pmatrix}.$$

Note that the equations $CX = [A_1]_{\mathcal{B}}$ and $CX = [A_2]_{\mathcal{B}}$ are inconsistent, thus either A_1 or A_2 is not suitable for building a Jordan chain. Why these equations are inconsistent? Notice that the sum of the first and the second rows of C equals its fourth row. Therefore the equation $CX = (y_1, y_2, y_3, y_4)^T$ is solvable only if $y_1 + y_2 = y_4$. Thus one needs to find a linear combination of the eigenvectors A_1 and A_2 satisfying this condition. Clearly, $A_2' = A_1 - A_2$ satisfies it and $[A_3]_{\mathcal{B}} = (-\frac{1}{2}, 0, -\frac{1}{2}, -\frac{1}{2})^T$ is a solution of the system of linear equations

$$\begin{cases} x_1 - x_3 = 0 \\ -2x_1 + x_2 + x_3 - x_4 = 1 \\ x_1 - x_3 = 0 \\ -x_1 + x_2 - x_4 = 1 \\ x_1 + x_2 - x_4 = 0 \end{cases} .$$

Here the first four equations are from the condition $CX = [A_2']^T$ and the fifth one guarantees that the next element of the Jordan chain can be found. Finally, $[A_4]_{\mathcal{B}} = (0, 0, \frac{1}{2}, \frac{1}{2})^T$ is a solution of $CX = [A_3]^T$ and the Jordan chain is completed. This gives the Jordan basis $\mathcal{J} = \{A_1, A_2', A_3, A_4\}$, where

$$A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, A_2' = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, A_3 = -\begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, A_4 = \begin{pmatrix} 0 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

with $L(A_1) = A_1, L(A_2') = A_2', L(A_3) = A_2' + A_3, L(A_3) = A_3 + A_4$ and

$$[L]_{\mathcal{J}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} .$$

□

Exercise 12.1. *Find a Jordan normal form and a Jordan basis for the linear operator from Exercise 11.2.*

Theorem 12.4. *Let $A \in \mathrm{M}_n(\mathbb{C})$. Then there exist an invertible matrix $U \in \mathrm{M}_n(\mathbb{C})$ and a Jordan matrix $J \in \mathrm{M}_n(\mathbb{C})$ such that $A = UJU^{-1}$. Moreover, such the Jordan matrix J is unique up to the order of the Jordan blocks.*

Proof. Consider $L \in \mathcal{L}(\mathrm{M}_{n,1}(\mathbb{C}))$, $L(X) = AX$. Remind that $[L]_{\mathcal{F}} = A$, where $\mathcal{F}$ is the standard basis of $\mathrm{M}_{n,1}(\mathbb{C})$. By Theorem12.1, there exists a basis $\mathcal{J}$ such that $[L]_{\mathcal{J}} = J$ is a Jordan matrix. Let $U \in \mathrm{M}_n(\mathbb{C})$ be the transition matrix from $\mathcal{F}$ to $\mathcal{J}$. Then $A = [L]_{\mathcal{F}} = U[L]_{\mathcal{J}}U^{-1} = UJU^{-1}$ as required.

Let $A = U'J'U'^{-1}$ for an invertible matrix U' and a Jordan matrix J'. By Lemma 6.8 there exists a basis $\mathcal{J}'$ of $\mathrm{M}_{n,1}(\mathbb{C})$ such that $M_{\mathcal{B} \rightsquigarrow \mathcal{J}'} = U'$. Then $[L]_{\mathcal{J}'} = U'^{-1}[L]_{\mathcal{B}}U' = U'^{-1}AU' = J'$, i.e. $\mathcal{J}'$ is a Jordan basis. Theorem 12.1 implies that J and J' are equal up to the order of the blocks. □

Definition. The Jordan matrix J from the above theorem is called a **Jordan normal form** (若尔当标准型) of A.

An algorithm for computing for a given $A \in \mathrm{M}_n(\mathbb{C})$ an invertible matrix $U \in \mathrm{M}_n(\mathbb{C})$ and a Jordan matrix $J \in \mathrm{M}_n(\mathbb{C})$ such that $A = UJU^{-1}$ based on the proof of Theorems 12.1 and 12.4:

$N = L - \lambda \mathrm{id}|_{V(\lambda, L)}$

the row of C: $(\ \vdots\ )$ $(\ker N)^\perp$ solution of $CX=0$ denote the eigenspace $\ker N$

I. Find the eigenvalues of A and for every eigenvalue λ put $C = A - \lambda E_n$;

II. Find independent non-trivial zero linear combinations of the rows of C. A formal way to find them is to solve the system $C^T Z = 0$;

find $\ker N \cap \operatorname{Im} N$

$\operatorname{Im} C \cap V(\lambda, L) = \operatorname{Im} N$ $\operatorname{Im}(C|_{V(\lambda,L)})$

$\operatorname{Im} C = \operatorname{span}$ columns of C $(z_1, \dots, z_n)\big(\ |\ |\ |\ |\ |\ \big) = 0$

we need to find $(\operatorname{Im} N)^\perp$, $(\operatorname{Im} C)^\perp$ (to find orthogonal complement.

III. Extend C by the rows of the coefficients of these linear combinations to form $\hat{C}$;

$\ker N \cap \operatorname{Im} N = \{x : \hat{C}x = 0\}$

$\to \ker N \cap \operatorname{Im} N$

IV. Find a basis $X_1^{(1)}, \dots, X_{s-t}^{(1)}$ of the solutions of the system $\hat{C}X = 0$ and its extension $Y_{s-t+1}, \dots, Y_s$ to a basis of the solutions of the system $CX = 0$. The latter column vectors are Jordan chains of length 1, and the former column vectors are the first elements of Jordan chains of length greater than 1;

$\to \ker N$

$X_1^{(1)} \dots X_{s-t}^{(1)}$ Y_{s-t+1}, Y_s

$u_1 \in \ker N \cap \operatorname{Im} N^2$ $N^2 x = u_1$ has solution

$\hat{C}x = \left(\frac{C}{D}\right) x = \begin{pmatrix} X_1^{(1)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

V. For any $1 \leq j \leq s - t$, consider the system $\hat{C}X = \overline{X}_j^{(1)}$, where $\overline{X}_j^{(1)}$ is obtained by appending a suitable number of zeros to the end of $X_j^{(1)}$. If it is inconsistent then $X_j^{(1)}$ and Y_j, a solution of the system $CX = X_j^{(1)}$, is a Jordan chain of length 2. If it is solvable, its solution $X_j^{(2)}$ is the second element of a Jordan chain starting from $X_j^{(1)}$;

VI. Proceed with the remaining elements of the form $X_j^{(2)}$ and consider the systems $\hat{C}X = \overline{X}_j^{(2)}$. If it is inconsistent then $X_j^{(1)}, X_j^{(2)}$ and Y_j, a solution of the system $CX = X_j^{(2)}$, is a Jordan chain of length 3. If it is solvable, its solution $X_j^{(3)}$ is the third element of a Jordan chain starting from $X_j^{(1)}, X_j^{(2)}$;

VII. Proceed in a similar manner until the number of the column vectors obtained equals n;

VIII. Let $X_j^{(1)}, \dots, X_j^{(k_j - 1)}, Y_j$ for $1 \leq j \leq s$ be the jth Jordan chain. Put

$$U(\lambda) = \left(X_1^{(1)} \mid \cdots \mid X_1^{(k_1-1)} \mid Y_1 \mid \cdots\cdots \mid X_s^{(1)} \mid \cdots \mid X_s^{(k_s-1)} \mid Y_s \right),$$

$$J(\lambda) = \operatorname{diag}(J_{k_1}(\lambda), \dots, J_{k_s}(\lambda)).$$

IX. Finally, put

$$U = (U(\lambda_1) \mid \cdots \mid U(\lambda_q)),$$

$$J = \operatorname{diag}(J(\lambda_1), \dots, J(\lambda_q)),$$

where $\lambda_1, \dots, \lambda_q$ are the eigenvalues of A.

Example. Let us consider how the above algorithm works for $A = [L]_{\mathcal{B}}$ from Problem 12.3. It has the unique eigenvalue 1 and

$$C = A - E_4 = \begin{pmatrix} 1 & 0 & -1 & 0 \\ -2 & 1 & 1 & -1 \\ 1 & 0 & -1 & 0 \\ -1 & 1 & 0 & -1 \end{pmatrix}.$$

$\cancel{x_2} + x_4 = 0$

$x_2 = x_1 + x_3$

$(1, 1, 0, -1)$

Independent zero linear combinations of the rows of C are given by $\underline{y_1 + y_2 - y_4} = 0$ and $y_1 - y_3 = 0$. Then

$(0\ 1\ 1\ -1)$

$\begin{pmatrix} -\frac{1}{2} \\ \\ -\frac{1}{2} \end{pmatrix}$

$X_2 + X_4 = \frac{1}{2}$

$X_1 - X_3 = 0$

$X_1 = -\frac{1}{2}$

$$\hat{C} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ -2 & 1 & 1 & -1 \\ 1 & 0 & -1 & 0 \\ -1 & 1 & 0 & -1 \\ 1 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix}. \quad \begin{matrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{matrix}$$

$X_1 - X_3 = 0.$

$X_4 = X_1 + X_2.$

$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$

The solutions of the system $\hat{C}X = 0$ are spanned by $X_1^{(1)} = (0, 1, 0, 1)^T$; the solution space of the system $CX = 0$ is two-dimensional, it is spanned by $X_1^{(1)}$ and $Y_2 = (1, 1, 1, 0)^T$.

Then consider the system $\hat{C}X = \overline{X}_1^{(1)}$, i.e.,

$$\begin{cases} x_1 \phantom{{}+x_2} - x_3 \phantom{{}-x_4} = 0 \\ -2x_1 + x_2 + x_3 - x_4 = 1 \\ x_1 \phantom{{}+x_2} - x_3 \phantom{{}-x_4} = 0 \\ -x_1 + x_2 \phantom{{}+x_3} - x_4 = 1 \\ x_1 + x_2 \phantom{{}+x_3} - x_4 = 0 \\ x_1 \phantom{{}+x_2} - x_3 \phantom{{}-x_4} = 0 \end{cases}.$$

It is solvable and $X_1^{(2)} = (-\frac{1}{2}, 0, -\frac{1}{2}, -\frac{1}{2})^T$ is one of its solution. Further, the system $\hat{C}X = \overline{X}_1^{(2)}$, i.e.,

$$\begin{cases} x_1 \phantom{{}+x_2} - x_3 \phantom{{}-x_4} = -\frac{1}{2} \\ -2x_1 + x_2 + x_3 - x_4 = 0 \\ x_1 \phantom{{}+x_2} - x_3 \phantom{{}-x_4} = -\frac{1}{2} \\ -x_1 + x_2 \phantom{{}+x_3} - x_4 = -\frac{1}{2} \\ x_1 + x_2 \phantom{{}+x_3} - x_4 = 0 \\ x_1 \phantom{{}+x_2} - x_3 \phantom{{}-x_4} = 0 \end{cases}$$

is inconsistent. It remains to find a solution of the system $CX = X_1^{(2)}$, for example $Y_1 = (0, 0, \frac{1}{2}, \frac{1}{2})^T$. Finally

$$U = \begin{pmatrix} 0 & -\frac{1}{2} & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 1 \\ 1 & -\frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}, \qquad J = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Exercise 12.2. *Find a Jordan normal form of the matrix*

$x_1=0$
$x_2=x_3=x_4$

$$\begin{pmatrix} 2&1&1&1\\ 1&0&0&0\\ 0&0&0&0 \end{pmatrix} \begin{matrix} 1\\0\\0 \end{matrix}$$

$$\begin{matrix} 0&0&0&0&0\\ 0&-1&1&0&0\\ 0&-1&0&1&0 \end{matrix}$$

$$\begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

$\lambda_1=1 \quad \lambda_2=2$

$$\begin{matrix} 0&0&1&2&1\\ 0&0&1&0\\ &0&1&0\\ &&1&0\\ 0&-1&1&0\\ 0&-1&0&1 \end{matrix}$$

$x_1=0$
$x_2+x_3+x_4=0$
$(0,-1,1,0)$
$(0,-1,0,1)$

Exercise 12.3. *Let*

$$A = \begin{pmatrix} 0 & -1 & 1 & 0 \\ 0 & -2 & 1 & 1 \\ -1 & 0 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix}$$

$(0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3})$

$x_1=(1,0,0,0)$
$x_2=(0,0,1,0)$

Find an invertible matrix $U \in \mathrm{M}_n(\mathbb{C})$ and a Jordan matrix $J \in \mathrm{M}_n(\mathbb{C})$ such that $A = UJU^{-1}$. Verify the equality obtained.

Applications of Jordan normal form

An alternative form of Jordan blocks has 1 on the subdiagonal:

$$J'_n(\lambda) = \begin{pmatrix} \lambda & 0 & 0 & \dots & 0 & 0 \\ 1 & \lambda & 0 & \dots & 0 & 0 \\ 0 & 1 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 0 \\ 0 & 0 & 0 & \dots & 1 & \lambda \end{pmatrix}.$$

Accordingly, in this alternative notation, a Jordan matrix J' is block diagonal where each diagonal block equals $J'_n(\lambda)$.

Theorem 12.5. *Let V be a vector space over $\mathbb{C}$ and $L \in \mathcal{L}(V)$. Then there exists a basis $\mathcal{J}'$ in V such that $[L]_{\mathcal{J}'} = J'$. Moreover, J' is unique up to the order of the Jordan blocks $J'_n(\lambda)$.*

Proof. Let

$$\begin{aligned} &u_{m_1}, \dots, u_2, u_1, \\ &v_{m_2}, \dots, v_2, v_1, \\ &\vdots \\ &w_{m_s}, \dots, w_2, w_1 \end{aligned}$$

be a Jordan basis represented as the union of Jordan chains for L, i.e.

$$\begin{aligned} &L(u_{m_1}) = \lambda_1 u_{m_1} + u_{m_1-1}, \dots, L(u_2) = \lambda_1 u_2 + u_1, L(u_1) = \lambda_1 u_1, \\ &L(v_{m_2}) = \lambda_2 v_{m_2} + v_{m_2-1}, \dots, L(v_2) = \lambda_2 v_2 + v_1, L(v_1) = \lambda_2 v_1, \\ &\vdots \\ &L(w_{m_s}) = \lambda_s w_{m_s} + w_{m_s-1}, \dots, L(w_2) = \lambda_s w_2 + w_1, L(w_1) = \lambda_s w_1. \end{aligned}$$

Obviously, the matrix of L relative to the basis

$$\begin{aligned}
&u_1, u_2, \dots, u_{m_1},\\
&v_1, v_2, \dots, v_{m_2},\\
&\vdots\\
&w_1, w_2, \dots, w_{m_s}
\end{aligned}$$

has the required form.

The uniqueness is proved in a similar way. $\square$

Exercise 12.4. *Show that for any $A \in \mathrm{M}_n(\mathbb{C})$ there exist an invertible matrix $U \in \mathrm{M}_n(\mathbb{C})$ and a Jordan matrix $J' \in \mathrm{M}_n(\mathbb{C})$ with 1 on its Jordan blocks' subdiagonals such that $A = UJ'U^{-1}$.*

Hint. Consider $L \in \mathcal{L}(\mathrm{M}_{n,1}(\mathbb{C}))$, $L(X) = AX$.

Lemma 12.6.

$$J_m(\lambda)^N = \begin{pmatrix}
\lambda^N & \binom{N}{1}\lambda^{N-1} & \binom{N}{2}\lambda^{N-2} & \dots & \binom{N}{n-2}\lambda^{N-n+2} & \binom{N}{n-1}\lambda^{N-n+1} \\
0 & \lambda^N & \binom{N}{1}\lambda^{N-1} & \dots & \binom{N}{n-3}\lambda^{N-n+3} & \binom{N}{n-2}\lambda^{N-n+2} \\
0 & 0 & \lambda^N & \dots & \binom{N}{n-4}\lambda^{N-n+4} & \binom{N}{n-3}\lambda^{N-n+3} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \dots & \lambda^N & \binom{N}{1}\lambda^{N-1} \\
0 & 0 & 0 & \dots & 0 & \lambda^N
\end{pmatrix},$$

and it is assumed that $\binom{N}{m} = 0$ for $N < m$.

$J^n = \left[\lambda E_n + \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}\right]^n$

Proof. Induction on N using the identity $\binom{N+1}{m} = \binom{N}{m} + \binom{N}{m-1}$. $\square$

Lemma 12.6 gives a way to compute A^n, where $A \in \mathrm{M}_n(\mathbb{C})$. Indeed, Theorem 12.4 implies that there exist an invertible matrix $U \in \mathrm{M}_n(\mathbb{C})$ and a Jordan matrix $J \in \mathrm{M}_n(\mathbb{C})$ such that $A = UJU^{-1}$. Then $A^N = UJ^NU^{-1}$, and since block diagonal matrices of the same structure multiply block by block, Lemma 12.6 completes the calculation.

Problem 12.7. *Find A^{1001} if*

$$A = \begin{pmatrix} 6 & 2 & 5 \\ -4 & -1 & -4 \\ -5 & -2 & -4 \end{pmatrix}.$$

Solution. One has $A = UJU^{-1}$, where

$$J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} -1 & -1 & -1 \\ 1 & 0 & 2 \\ 1 & 1 & 0 \end{pmatrix}.$$

Lemma 12.6 yields

$$J^{1001} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1001 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus

$$A^{1001} = UJ^{1001}U^{-1} = \begin{pmatrix} 1006 & 2 & 1005 \\ -4 & -1 & -4 \\ -1005 & -2 & -1004 \end{pmatrix}.$$

$\square$

Exercise 12.5. *Find A^{101} if*

$$A = \begin{pmatrix} -1 & 0 & -1 \\ 1 & -2 & 0 \\ 1 & -1 & 0 \end{pmatrix}.$$

Lemma 12.8. *Let $Q \in \mathrm{M}_n(\mathbb{C})$ be nilpotent. Then there is $S \in \mathrm{M}_n(\mathbb{C})$ such that $S^2 = E_n + Q$.*

Proof. First prove that for any $m \geq 0$ there is $f_m \in \mathbb{R}[t]_m$ such that $f_m(0) = 1$ and $f_m^2(t) = 1 + t + g_m(t)t^{m+1}$ for some $g_m \in \mathbb{R}[t]$. Clearly $f_0(t) = 1$ and $f_1(t) = 1 + \frac{1}{2}t$ satisfy these conditions. Assume that there is $f_{m-1} \in \mathbb{R}[t]$ such that $f_{m-1}^2(t) = 1 + t + g_{m-1}(t)t^m$ and put $f_m(t) = f_{m-1}(t) + at^m$. Then for some $h \in \mathbb{R}[t]$ one has

$$f_m^2(t) = (f_{m-1}(t) + at^m)^2 = f_{m-1}^2(t) + 2at^m + h(t)t^{m+1} = 1 + t + g_{m-1}(t)t^m + 2at^m + h(t)t^{m+1}$$

and we can put $a = -\frac{c}{2}$, where $c \in \mathbb{R}$ is the constant term of g_{m-1}.

Now put $S = f_{n-1}(Q)$, then $S^2 = f_{n-1}^2(Q) = E_n + Q + g_{n-1}(Q)Q^n = E_n + Q$ by Corollary 10.5. $\square$

Remark. One can show that f equals the nth Taylor polynomial for $\sqrt{1+t}$:

$$f(t) = 1 + \sum_{j=1}^{n-1} \frac{\frac{1}{2}(\frac{1}{2} - 1)\cdots(\frac{1}{2} - j + 1)}{1 \cdot 2 \cdots j} t^j,$$

whence

$$S = E_n + \sum_{j=1}^{n-1} \frac{\frac{1}{2}(\frac{1}{2} - 1)\cdots(\frac{1}{2} - j + 1)}{1 \cdot 2 \cdots j} Q^j.$$

Proposition 12.9. *Let $A \in \mathrm{M}_n(\mathbb{C})$ be invertible. Then there exists an invertible $B \in \mathrm{M}_n(\mathbb{C})$ such that $A = B^2$.*

Proof. Theorem 12.4 implies that there exist an invertible matrix $U \in \mathrm{M}_n(\mathbb{C})$ and a Jordan matrix $J \in \mathrm{M}_n(\mathbb{C})$ such that $A = UJU^{-1}$. Let $J^{(1)}, \ldots, J^{(s)}$ be the Jordan blocks of J.

For any $1 \leq k \leq s$ one has $J^{(k)} = J_m(\lambda) = \lambda E_m + Q$, where Q is a strictly upper triangular matrix. Since A is invertible, $\lambda \neq 0$ and then $\lambda^{-1}J^{(k)} = E_m + \lambda^{-1}Q$. Lemma 12.8 implies that there exists an invertible $S_k \in \mathrm{M}_m(\mathbb{C})$ such that $S_k^2 = \lambda^{-1}J^{(k)}$. Then $R_k^2 = J^{(k)}$, where $R_k = \sqrt{\lambda}S_k$. Finally $B = URU^{-1}$ is a required matrix, where $R = \mathrm{diag}(R_1, \ldots, R_s)$. $\square$

Exercise 12.6. *Find $B \in \mathrm{M}_2(\mathbb{C})$ such that*

$$B^2 = \begin{pmatrix} 5 & -6 & 2 \\ 4 & -5 & 2 \\ 4 & -6 & 3 \end{pmatrix}.$$

Hint. Use Remark after Lemma 12.8.

Exercise 12.7. *Let $L \in \mathcal{L}(V)$ be invertible. Prove that there exists an invertible $T \in \mathcal{L}(V)$ such that $T^2 = L$.*

Hint. Fix a basis $\mathcal{B}$ in V and find a square root of $[L]_{\mathcal{B}}$.