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.*

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

Linear combinations and linear span

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

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

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

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

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

1. $\mathrm{Span}(v_1, \ldots, v_n)$ *is a subspace of V;*

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

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

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 $U_1 + \cdots + U_m \subset U$.*

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 $u_i \in U_i$.

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 $U \cap W = \{0\}$.*

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 $a_1 v_1 + \cdots + a_n v_n = 0$ (**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).

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.*

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

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

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

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

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 .*

Finite-dimensional vector spaces

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

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

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 $u_1, \ldots, u_n$ 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 $u_1, \ldots, u_n$.*

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

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

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

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

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 $v = a_1 v_1 + \cdots + a_n v_n$. 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* $t-1, t^2+1, t^2-t$.

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.*

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 $\dim V$.

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* $U = V$.

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.*

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.*

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

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

Direct sum revisited

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

1. $V = U_1 \oplus \cdots \oplus U_m$;

2. $V = U_1 + \cdots + U_m$ *and* $U_j \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_m) = \{0\}$ *for any* $1 \leq j \leq m$.

3. $V = U_1 + \cdots + U_m$ *and if* $u_1 + \cdots + u_m = 0$ *for* $u_i \in U_i$, *then* $u_1 = \cdots = u_m = 0$;

4. *the union of bases for* U_i *is a basis for* V;

5. $V = U_1 + \cdots + U_m$ *and* $\dim V = \dim U_1 + \cdots + \dim U_m$;

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 $L(0) = 0$.

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 $i = 1, \ldots, n$.*

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. $(L'_1 + L'_2)L = L'_1 L + L'_2 L$ *and* $L'(L_1 + L_2) = L'L_1 + L'L_2$ *for* $L, L_1, L_2 \in \mathcal{L}(U,V)$ *and* $L', L'_1, L'_2 \in \mathcal{L}(V,W)$.

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 $((LL')(f)) = (tf(t))' = tf'(t) + f(t)$.

Kernel and image

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

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

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

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

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

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

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

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

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

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

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 $(a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$ 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 $\mathcal{B} = \mathcal{B}'$.

Theorem 6.1. *Let V, W be vector spaces over F with bases $\mathcal{B} = \{u_1, \ldots, u_l\}$,$\mathcal{B}' = \{v_1, \ldots, v_m\}$, 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.

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

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

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

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

is a linear transformation.

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*

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

for any $v \in V$.

Change of basis

Definition. Let $\mathcal{B} = \{u_1, \ldots, u_n\}$, $\mathcal{B}' = \{v_1, \ldots, v_n\}$ 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,

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

Symbolically, one can write

$$\begin{pmatrix} v_1 & v_2 & \ldots & v_n \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & \ldots & u_n \end{pmatrix} M_{\mathcal{B} \rightsquigarrow \mathcal{B}'}.$$

Proposition 6.5. *Let $\mathcal{B} = \{u_1, \ldots, u_n\}$, $\mathcal{B}' = \{v_1, \ldots, v_n\}$, $\mathcal{B}'' = \{w_1, \ldots, w_n\}$ 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* $M^{-1}_{\mathcal{B} \rightsquigarrow \mathcal{B}'} = M_{\mathcal{B}' \rightsquigarrow \mathcal{B}}$.

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

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

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 $L \in \mathcal{L}(U, V)$*

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

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 $M_{\mathcal{C} \rightsquigarrow \mathcal{B}} = A$.*

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

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* $\dim \operatorname{Im}(L) = \operatorname{rank}[L]_{\mathcal{C},\mathcal{C}'}$.

7 Isomorphism

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

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.

Proposition 7.2. *Isomorphism has the following properties:*

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.

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

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.*

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.*

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

Definition. Let V be a vector space. A linear transformation $L \colon V \to V$ is called a **linear operator** 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* $V \in \mathcal{L}(V)$*. The following statements are equivalent:*

1. *L is bijective*

2. *L is injective*

3. *L is surjective.*

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.

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.*

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.*

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.*

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

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 $u \in U$.

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

Definition. Let $L \in \mathcal{L}(V)$, and U be an L-invariant subspace of V. The map $L|_U : U \to U$ 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 $\mathcal{L}(U)$.*

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

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

where $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_m$ is a basis of V (see Proposition4.6).

Lemma 8.8. *Let $L \in \mathcal{L}(V)$ and $p \in F[x]$. Then $\operatorname{Ker}(p(L))$ and $\operatorname{Im}(p(L))$ are L-invariant.*

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.

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 $\chi_A(t) = t^n - \operatorname{Tr} A \cdot t^{n-1} + \cdots + (-1)^n |A|$.*

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 $\chi_A = \chi_B$.*

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

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 $\operatorname{Tr}[L]_\mathcal{B}$ 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.*

Corollary 9.5. *Let F be algebraically closed, V be a nonzero vector space over F, $L \in \mathcal{L}(V)$. 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 $[u]_{\mathcal{B}} = X$.

Hamilton-Cayley theorem

Definition. Let $f \in F[x]$, $f = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ 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:

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

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

Proposition 9.7. *If $f, g \in F[x]$, $a \in F$ 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 $(f+g)(L) = f(L) + g(L)$, $(fg)(L) = f(L)g(L)$, $(\lambda f)(L) = \lambda f(L)$, $(f \circ g)(L) = f(g(L))$.*

Corollary 9.8. *If $f, g \in F[x]$, $A \in \mathrm{M}_m(F)$, $L \in \mathcal{L}(V)$ then $f(A)g(A) = g(A)f(A)$ and $f(L)g(L) = g(L)f(L)$.*

Theorem 9.9 (Hamilton-Cayley)**.**

1. *Let $A \in \mathrm{M}_n(F)$. Then $\chi_A(A) = 0$.*
2. *Let $L \in \mathcal{L}(V)$. Then $\chi_L(L) = 0$.*

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 $L \in \mathcal{L}(V)$, $\lambda_1, \dots, \lambda_m \in F$ 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. $\dim V = \dim V_{\lambda_1}(L) + \cdots + \dim V_{\lambda_m}(L)$.

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

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

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.*

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*

1. *N is nilpotent*
2. $\chi_N(t) = t^n$
3. *$N^j = 0$ for some $j \in \mathbb{N}$*

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

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

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

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*

Theorem 11.3. *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. $V = V(\lambda_1, L) \oplus \cdots \oplus V(\lambda_m, L)$;
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.*

Corollary 11.4. *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.*

Jordan basis for a nilpotent operator

Theorem 11.5. *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*

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

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

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). 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

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

the elements $\lambda_1, \dots, \lambda_s$ 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).

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* $\dim \operatorname{Ker}(L - \lambda \operatorname{id}_V)$.

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.*

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:

I. Find the eigenvalues of A and for every eigenvalue λ put $C = A - \lambda A$;

II. Find independent non-trivial zero linear combinations of the rows of C. A formal way to find them is to solve the system $C^T Z = 0$;

III. Extend C by the rows of the coefficients of these linear combinations to form $\hat{C}$;

IV. Find a basis $X_1^{(1)}, \ldots, X_{s-t}^{(1)}$ of the solutions of the system $\hat{C}X = 0$ and its extension $Y_{s-t+1}^{(1)}, \ldots, Y_s^{(1)}$ 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;

V. For any $1 \leq j \leq s-t$, consider the system $\hat{C}X = X_j^{(1)}$. If it is inconsistent then $X_j^{(1)}$ and $Y_j^{(2)}$, 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 $\hat{C}X = X_j^{(2)}$. If it is inconsistent then $X_j^{(1)}, X_j^{(2)}$ and $Y_j^{(2)}$, 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 $X_1^{(j)}, \dots, X_{k_j-1}^{(j)}, Y^{(j)}$ for $1 \le j \le s$ be the jth Jordan chain. Put

$$U(\lambda) = \left(X_1^{(1)} \mid \cdots \mid X_{k_1-1}^{(1)} \mid Y^{(1)} \mid \cdots\cdots \mid X_1^{(s)} \mid \cdots \mid X_{k_s-1}^{(s)} \mid Y^{(s)}\right),$$

$$J(\lambda) = \operatorname{diag}(J_{k_1}(\lambda), \dots, J_{k_s}(\lambda)).$$

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.

Applications of Jordan normal form

An alternative form of Jordan blocks has 1 on the subdiagonal:

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

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 $J_n'(\lambda)$.*

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* $N < m$.

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 $S^2 = E_n + Q$.*

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 $A = B^2$.*