grading:

- homework 10%
- midterm 30%
- endterm 60%

textbook:*Linear Algebra*, Hoffman & Kunze

代数：

- 集合上的结构
- 集合间保持结构的映射

基本研究对象: 线性空间和线性映射
公理：在数学对象的定义中，被假定满足的性质

目录

1 Fields

Definition 1.1 (Field)**.** A **field** is a triple $(F, +, \cdot)$ consisting of a set of elements F, and two binary operations $+, \cdot$ satisfying

- addition commutativity
- addition associativity
- additive identity (denoted by 0_F)
- additive inverse (denoted by $-a$)
- multiplication commutativity
- multiplication associativity
- multiplicative identity (denoted by 1_F)
- multiplicative inverse (denoted by a^{-1})
- multiplication distributes over addition
- $1_F \neq 0_F$

Example 1.2

Examples of fields: $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Q}[\sqrt{2}], \mathbb{F}_p$, the set of rational fuctions (with usual addtion and multiplication).

Definition 1.3 (Subfield)**.** A subset F' of a field F is a **subfield** if F' is closed under addtion and multiplication, i.e. if $a, b \in F$, $a +_F b, a \cdot_F b, -a, a^{-1} \in F$ as well.

Definition 1.4. The smallest positive integer k satisfying $\underbrace{1_F + \cdots + 1_F}_{k} = 0_F$ is called the **characteristic** of F (denoted by $\mathrm{char}F$). If no such k exists, we define the characteristic to be 0.

For $n \in \mathbb{N}$, define $n_F := \underbrace{1_F + \cdots + 1_F}_{n}$, and $nx := n_F \cdot x$, $\dfrac{x}{n} := n_F^{-1} \cdot x$ $(n_F \neq 0_F)$.

Some basic properties of $\mathrm{char}F$:

- $\mathrm{char}F$ is either 0 or a prime number.
- $\mathrm{char}F \mid n \iff n_F = 0_F$.
- Let F' be a subfield of F, we have $\mathrm{char}F' = \mathrm{char}F$.

2 Vector spaces

Definition 2.1 (Vector space)**.** Given a field F, A set V with two operations(addition, scaling) is a **vector space** if:

- $(V,+)$ is an abelian group, i.e. addition is commutative, associative, and there exist additive identity and inverse (denoted by 0_V and $-\alpha$);
- scaling is distributive and associative, i.e. $\forall c_1, c_2 \in F, \alpha, \beta \in V$,

  $c_1(c_2\alpha) = (c_1c_2)\alpha, c(\alpha+\beta) = c\alpha + c\beta, (c_1+c_2)\alpha = c_1\alpha + c_2\alpha$;
- $1_F \cdot \alpha = \alpha$.

Example 2.2

$\mathbb{R}^n, \mathbb{R}[x]$ are vector spaces over $\mathbb{R}$.

Let $F^S := \{f | f : S \to F\}$, F^S is a vector space as well.

2.1 Subspaces

Definition 2.3 (Subspace)**.** $W \subset V$ is a (linear) subspace of V if:

- $0_V \in W$
- W is closed under addition and scaling

Remark 2.4 — Alternative definition of subspaces: $W \subset V$ and W is a vector space.

Example 2.5

$\{0_V\}, V$ are (trivial) subspaces of V, $\mathbb{R}$ is a subspace of $\mathbb{C}$.

Proposition 2.6

Let W be a subset of V, the following are equivalent:

1. W is a subspace of V.
2. $W \neq \emptyset$, and the linear combinations of vectors in W are also in W.
3. $0 \in W$, and $\forall \alpha, \beta \in W, c \in F \Rightarrow c\alpha + \beta \in W$.

For $S \subset V$, denoted by $\text{span}(S)$ the subspace spanned by S, and define $\text{span}(\emptyset) = \{0\}$.

Remark 2.7 — $\operatorname{span}(S)$ is the smallest subspace containing S. If $\operatorname{span}(S) = V$, we say that S is **spanning**.

Definition 2.8 (Sum of subspaces)**.** Let $W_1, \dots, W_k \subset V$ be subspaces of V, then

$$\sum_{i=1}^{k} W_i := \left\{ \sum_{i=1}^{k} \alpha_i \middle| \alpha_i \in W_i \right\}$$

is a subspace of V, called the sum of $W_1, \dots, W_k$. moreover,

$$\sum_{i=1}^{k} W_i = \operatorname{span}\left(\bigcup_{i=1}^{k} W_i \right).$$

2.2 Linear independence & Basis

Definition 2.9 (Basis)**.** A subset $S \subset V$ is a **basis** if it's both linear independent and spanning.

Theorem 2.10

Let $S, T \subset V$ be finite subsets of vector space V. If S is linear independent and T is spanning, then $|S| \le |T|$.

Proof. Let $T = \{v_1, v_2, \dots, v_n\}, S = \{w_1, w_2, \dots, w_m\}$. Since T is spanning, we have

$$w_1 = c_1 v_1 + \cdots + c_n v_n.$$

There must be some nonzero coefficient, say c_n. Thus

$$v_n = -\frac{c_1}{c_n} v_1 - \frac{c_2}{c_n} v_2 - \cdots - \frac{c_{n-1}}{c_n} v_{n-1} + \frac{1}{c_n} w_1.$$

We can "throw in" w_1: $T_1 := \{v_1, v_2, \dots, v_{n-1}, w_1\}$ is also spanning.

Do the same thing for $w_2, w_3, \dots, w_m$, the linear independce of S ensures that there's always some v_i whose coefficient is nonzero. Since we can eventually get to T_m, we have $n \ge m$. □

Corollary 2.11

Every basis of a finite dimensional vector space has the same number of vectors (Here "finite dimensional" means it has a finite basis). This number is called the **dimension** of the space, denoted by $\dim V$.

Theorem 2.12

Let $W \subset V$ be a subspace of finite dimensional vector space V. Suppose that $S_1 \subset S_2 \subset W$, S_1 is linear independent and $\operatorname{span}(S_2) = W$. Then there exists a basis S_0 of W such that $S_1 \subset S_0 \subset S_2$.

Proof. Add vectors from S_2 to S_1: if there exists $v \in S_2$ such that $S_1 \cup \{v\}$ is linear independent, then add v to S_1.

Since S_1 is linear independent, $|S_1| \leq \dim V$, the above process must stop at some time.

Now $\forall v \in S_2$, $\{v\} \cup S_1$ is not linear independent, hence $S_2 \subset \operatorname{span}(S_1)$. This tells us $W = \operatorname{span}(S_1)$, meaning S_1 is a basis of W. □

Remark 2.13 — The above theorem implies the following useful facts:

- Given a linear independent subset of V, we can complete it into a basis.
- Given a spanning subset of V, we can extract a basis from it.
- Subspaces of a finite dimensional vector space are also finite dimensional.

Definition 2.14 (Direct sum for finite dimensional spaces)**.** Let $W_1, W_2 \subset V$ be finite dimensional subspaces of V, if every vector in $W_1 + W_2$ can be represent uniquely as a sum of two vectors from W_1 and W_2, then $W_1 + W_2$ is called the **direct sum** of W_1, W_2, written as $W_1 \oplus W_2$.

Theorem 2.15

Let $W_1, W_2 \subset V$ be finite dimensional subspaces of V, we have $W_1 + W_2$ is finite dimensional, and

$$\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2).$$

Proof. Let $\{\alpha_1, \ldots, \alpha_d\}$ be a basis of $W_1 \cap W_2$, extend it to a basis of W_1: $\{\alpha_1, \ldots, \alpha_d, \beta_1, \ldots, \beta_m\}$, and a basis of W_2: $\{\alpha_1, \ldots, \alpha_d, \gamma_1, \ldots, \gamma_n\}$.

We claim that $\{\alpha_1, \ldots, \alpha_d, \beta_1, \ldots, \beta_m, \gamma_1, \ldots, \gamma_n\}$ is a basis of $W_1 + W_2$.

- linear independence:

  If $\sum x_i\alpha_i + \sum y_i\beta_i + \sum z_i\gamma_i = 0$, then $\sum z_i\gamma_i \in W_1 \Rightarrow \sum z_i\gamma_i \in W_1 \cap W_2$, we get $z_i = 0$. Hence $x_i = y_i = 0$.

- spanning: Trival.

□

Remark 2.16 — Alternative definition of direct sums: If

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2,$$

then we write the sum as $W_1 \oplus W_2$.

3 Linear equations

Definition 3.1 (Systems of linear equations)**.** Given a field F, a system of m linear equations in n unknowns $x_1, \ldots, x_n$ is of the form

$$\begin{cases} A_{11}x_1 + \cdots + A_{1n}x_n = y_1 \\ \qquad\qquad \vdots \\ A_{m1}x_1 + \cdots + A_{mn}x_n = y_m \end{cases}$$

where $A_{ij}, y_i \in F$ are constants, A_{ij}'s are called coefficients, y_i's are called constant terms.

The solution $(x_1, x_2, \ldots, x_n)$ is often written as a column vector.

Definition 3.2. A matrix is an array of elements in field F, written as $A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix}$.

Here A is a $m \times n$ matrix in field F, denoted by $A \in F^{m\times n}$

> **Remark 3.3 —** Arrays of numbers are evil. In fact, a matrix should be defined as a way to represent linear maps.

The system of linear equations can be rewritten as

$$\sum_{i=1}^{n} x_i \alpha_i = \beta,$$

where $\alpha_i, \beta \in F^{m\times 1}$ are column vectors $\begin{pmatrix} A_{1i} \\ \vdots \\ A_{mi} \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$.

Let the **coefficient matrix** be $(\alpha_1, \ldots, \alpha_n) \in F^{m\times n}$, the **augmented matrix** be $(\alpha_1, \ldots, \alpha_n, \beta) \in F^{m\times(n+1)}$

Thus the equations have a solution $\iff \beta \in \operatorname{span}\{\alpha_1 \ldots, \alpha_n\}$.

If all the constant terms are zero, we call the system of equations **homogeneous**.

A homogeneous system of equations has only trivial solution(i.e. $x_i = 0$ for all i) $\iff \alpha_1 \ldots, \alpha_n$ are linear independent.

3.1 Linear maps & linear equations

Definition 3.4 (Linear maps)**.** Let V, W be vector spaces, a map $f : V \to W$ is a **linear map** if:

- $f(\alpha + \beta) = f(\alpha) + f(\beta), \forall \alpha, \beta \in V$.
- $f(c\alpha) = cf(\alpha), \forall \alpha \in V, c \in F$.

We can easily deduce that $f(0_V) = 0_W$ and $\mathrm{Im}(f)$ is a subspace of W.

Now we can view the linear equations as linear maps:

Consider a map $T : F^{n\times 1} \to F^{m\times 1} : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} A_{11}x_1 + \cdots + A_{1n}x_n \\ \vdots \\ A_{m1}x_1 + \cdots + A_{mn}x_n \end{pmatrix}$.

It's obvious that T is linear, and for $\beta \in F^{m\times 1}$, the equations has solutions $\iff$ $\beta \in \mathrm{Im}(T)$.

Definition 3.5. The **kernel** of a linear map $f : V \to W$ is the set of vectors whose image is 0_W, that is,

$$\ker(f) = \{\alpha \in V \mid f(\alpha) = 0\}.$$

Clearly $\ker(f)$ is a subspace of V.

Observe that $\ker(T) = T^{-1}(0)$ is the solution set of homogeneous equations.

For $\beta \in \mathrm{Im}(T)$, we want to identify $T^{-1}(\beta)$, which is the solution set of equation $\sum_{i=1}^{m} x_i\alpha_i = \beta$.

Note that if $T(\alpha) = \beta$, then $T(\alpha') = \beta \iff \alpha - \alpha' \in \ker(T)$. Hence $T^{-1}(\beta) = \alpha + \ker(T) := \{\alpha + \gamma \mid \gamma \in \ker(T)\}$, these sets are called **affine subspaces**.

3.2 Elementary operations & row-reduced matrices

Definition 3.6. Elementary row(column) operations on a matrix is one of the following:

- multiply a row by a non-zero scalar $c \in F$;
- replace the s-th row by the sum of s-th row and c times r-th row;
- interchange two rows

Two matrix is **row-equivalent** if one can be obtained by the other by a finite number of elementary row operations. It's clear that this is indeed an equivalence relation.

Corollary 3.7

Two matrices (of the same size) are row-equivalent iff their row spaces are equal.

Remark 3.8 — The row space of a $m \times n$ matrix is the subspace of F^n spanned by its row vectors, denoted by $row(A)$. Similarly we define the column space $col(A) \subset F^m$.

Corollary 3.9

Two systems of equations are equivalent iff their augmented matrices are row-equivalent.

Definition 3.10. A matrix is **row-reduced** if:

- the first non-zero entry(pivot) in each row is 1;
- each column containing a pivot has all its other entries zero.

A matrix is **row-reduced elchelon** if:

- it is row-ruduced;
- the zero rows are at the bottom;
- if the non-zero rows are $1, 2, \ldots, r$, and the i-th row's pivot is in the k_i-th column, then $k_1 < k_2 < \cdots < k_r$.

Theorem 3.11

Every matrix is row-equivalent to a unique row-reduced elchelon matrix.

The proof of existence is but some computation, the proof of uniqueness can be deduced from the fact that a row-reduced elchelon matrix can be identified by its row space, so I won't include them here.

Now consider a system of linear equations $AX = Y, A \in F^{m\times n}, X \in F^{n\times 1}, Y \in F^{m\times 1}$. We can apply row operarions to the augmented matrix (A, Y), and obtain a row-reduced elchelon matrix (R, Z). We only need to solve the equation $RX = Z$.

Assume the pivots are in columns $k_1, \ldots, k_r$, $J = \{1, \ldots, n\}\backslash\{k_1, \ldots, k_r\}$, and let $Z = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix}$. We can rewrite the equations as

$$\begin{cases} x_{k_1} + \sum_{j\in J} c_{1j}x_j = z_1 \\ \qquad\qquad\vdots \\ x_{k_r} + \sum_{j\in J} c_{rj}x_j = z_r \\ \qquad\qquad 0 = z_{r+1} \\ \qquad\qquad\vdots \\ \qquad\qquad 0 = z_m \end{cases}$$

There are 3 cases:

1. If $r < m$ and $z_{r+1}, \ldots, z_m$ are not all zero, then the equation has no solutions.
2. Otherwise if $r = n$, then $J = \emptyset$, $k_i = i$, the equation has unique solution $(z_1, \ldots, z_n)$.
3. Otherwise $r < n$, the equation has more than one solutions, and $x_j (j \in J)$ are free variables.

Proposition 3.12

Let $F \subset K$ be fields. If a system of linear equations on F has solutions on K, it has solutions on F as well.

When the equations are homogeneous, we won't come to case 1. For each $j_0 \in J$, let $x_{j_0} = 1, x_j = 0 (j \in J, j \neq j_0)$, we obtain a solution $\alpha_{j_0} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, where $x_{k_i} = -c_{ij_0}$.

Proposition 3.13

$\{\alpha_j \mid j \in J\}$ is a basis of the solution space $\ker(T)$.

Proof. Linear independence:

note that only α_{j_0} has non-zero j_0-th entry$(\forall j_0 \in J)$.

Spanning:

let $\alpha = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \ker(T)$, we have $\alpha = \sum_{j \in J} x_j \alpha_j$. □

Corollary 3.14

if $m < n$ then $AX = 0$ has non-trivial solutions.

Proof. $r \leq m < n \implies n - r > 0 \implies J \neq \emptyset$. □

Remark 3.15 — This is equivalent to any linear independent subset has fewer elements than a spanning set.

4 Matrices

4.1 Matrix multiplication

Definition 4.1. Let $A, B \in F^{m \times n}$ be matrices, define $(A + B)_{ij} = A_{ij} + B_{ij}, (cA)_{ij} = cA_{ij}$. Thus the matrices form a vector space, and $\dim F^{m \times n} = mn$.

Example 4.2

Matrices as vector space:

- The symmetirc matrices is a subspace of $F^{n\times n}$
- We call a matrix $A \in \mathbb{C}^{n\times n}$ is Hermite if $A_{ij} = \overline{A_{ji}}$. Note that $\{A \mid A \text{ Hermite}\}$ is not a subspace over $\mathbb{C}$, but it is a subspace over $\mathbb{R}$, and $\dim_{\mathbb{R}}\{A \mid A \text{ Hermite}\} = n^2$.

Recall that we can view a system of linear equations as a linear map, we can do the same thing for matrices.

Let $A \in F^{m\times n}$, $L_A : F^n \to F^m : L_A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} A_{11}x_1 + \cdots + A_{1n}x_n \\ \vdots \\ A_{m1}x_1 + \cdots + A_{mn}x_n \end{pmatrix}$. We can check that L_A is indeed a linear map.

In fact, the above linear maps are all linear maps from F^n to F^m.

Proposition 4.3

For every linear map $T : F^n \to F^m$, there exists a unique matrix $A \in F^{m\times n}$ such that $T = L_A$.

Proof. Trivial. □

Remark 4.4 — We can regard matrices as a way to write down linear maps between F^n and F^m.

Note that the difinitions of addition and scaling is the same: $L_A + L_B = L_{A+B}, cL_A = L_{cA}$. But for linear maps there's one more operation: the composition of maps. Hence we need a coressponding operation for matrices, the solution is matrix multiplication.

Definition 4.5 (Matrix multiplication)**.** For $A \in F^{m\times n}, B \in F^{n\times p}$, define $AB \in F^{m\times p}$ such that $L_{AB} = L_A \circ L_B$. Note that the multiplication is not commutative.

And now we can compute AB in terms of A and B:

Proposition 4.6 (Alternative definition of matrix multiplication)

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj}$$

Proof. $(AB)_{ij} = (L_{AB}(\varepsilon_j))_i = (L_A(L_B(\varepsilon_j)))_i = \left(L_A \begin{pmatrix} B_{1j} \\ \vdots \\ B_{nj} \end{pmatrix} \right)_i = \sum_{k=1}^{n} A_{ik}B_{kj}$ □

Proposition 4.7 (Why we use column vectors instead of row vectors)

Let $A \in F^{m\times n}, X \in F^{n\times 1}$, then $L_A(X) = AX$.

Proof. We view both A and X as linear maps. It's clear that $L_X(c) = cX$, so we can write X as $L_X(1)$. Hence $L_A(X) = L_A(L_X(1)) = L_{AX}(1) = AX$. □

Proposition 4.8 (Properties of matrices multiplication)

Let A, B, C be matrices.

1. (Associativity) $(AB)C = A(BC)$;
2. (Distribution over addtion) $(A+B)C = AC + BC, A(B+C) = AB + AC$;
3. $\forall c \in F, c(AB) = (cA)B = A(cB)$;
4. Let I_n denote the $n \times n$ identity matrix. Then $I_m A = AI_n = A(A \in F^{m\times n})$.
5. Let $B = (B_1, \dots, B_n)$, where B_i's are column vectors, then $AB = (AB_1, \dots, AB_n)$.

4.2 Elementary matrices

Let $A \in F^{m\times n}, B \in F^{n\times p}, AB \in F^{m\times p}$, if we write $B = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$, then $AB = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix}$,

where $\gamma_i = \sum_{k=1}^{n} A_{ik}\beta_k$.

This can be viewed as a row operation on B. Therefore we want to find the matrices which correspond to elementary row operations, those are the elementary matrices.

It's easy to verify the three kinds of elementary matrices:

$$\begin{pmatrix} 1 & & & & & & \\ \vdots & \ddots & & & & & \\ 0 & & 1 & & & & \\ 0 & & & c & & & \\ 0 & & & & 1 & & \\ \vdots & & & & & \ddots & \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}, \begin{pmatrix} 1 & & & & & & \\ \vdots & \ddots & & & & & \\ 0 & & 1 & & & & \\ \vdots & & & \ddots & & & \\ 0 & & c & & 1 & & \\ \vdots & & & & & \ddots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}, \begin{pmatrix} 1 & & & & & & & & & \\ \vdots & \ddots & & & & & & & & \\ 0 & & 1 & & & & & & & \\ 0 & & & 0 & & & 0 & 1 & & \\ 0 & & & & 1 & & & 0 & & \\ \vdots & & & & & \ddots & & & & \\ 0 & & & 0 & & & 1 & & & \\ 0 & & & 1 & 0 & & & 0 & & \\ \vdots & & & & & & & & \ddots & \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{pmatrix}$$

Corollary 4.9

Let A, B be matrices, the following are equivalent:

- $A \sim B$(row equivalent);
- There exists elementary matrices $E_1, \ldots, E_k \in F^{m\times m}$, such that $B = E_1 \cdots E_k A$.

4.3 Invertible matrices

Definition 4.10. If $AB = BA = I_n$, we say A is **invertible**, and B is the **inverse** of A.

$AB = I_n \iff L_A L_B = \mathrm{id}, BA = I_n \iff L_B L_A = \mathrm{id}$, so the inverse of A is unique, since $L_B = L_A^{-1}$. We write A^{-1} for the inverse of A.

If we only have $AB = I_n$, can we imply that A is invertible? The answer is yes in finite dimensional vector space. To prove this, we need the following theorem:

Theorem 4.11

Let V, W be vector spaces, $T : V \to W$ is a linear map, then

$$\dim \ker(T) + \dim \mathrm{Im}(T) = \dim V.$$

Sketch of the proof. Take a basis of $\ker(T)$, say $\{\alpha_1, \ldots, \alpha_k\}$.

Extend it to a basis of V: $\{\alpha_1, \ldots, \alpha_n\}$.

It's clear that $\{T(\alpha_{k+1}), \ldots, T(\alpha_n)\}$ is a basis of $\mathrm{Im}(T)$. □

Corollary 4.12

Let V be a finite dimensional vector space, $T : V \to V$ a linear map.

We have T injective $\iff$ T surjective.

Proof. T injective $\iff \ker(T) = \{0\} \iff \dim \mathrm{Im}(T) = \dim V \iff T$ surjective. □

Corollary 4.13

Let V be a finite dimensional vector space, T_1, T_2 are linear maps from V to itself, if $T_1 T_2 = \mathrm{id}$, then $T_2 T_1 = \mathrm{id}$.

Therefore $AB = I_n \implies BA = I_n \implies B = A^{-1}$.

Some properties of invertible matrices:

- A invertible $\implies (A^{-1})^{-1} = A$;
- A, B invertible $\implies (AB)^{-1} = B^{-1}A^{-1}$;

- AB invertible $\implies$ A, B invertible.

It's clear that elementary matrices are invertible, as elementary row operations are invertible. Hence we can deduce:

Proposition 4.14

The following are equivalent:

1. A is invertible;
2. A and I_n are row equivalent;
3. A can be written as a product of elementary matrices.

Proof. It's obvious that (2) $\implies$ (3) $\implies$ (1).

If A is invertible, let R be the row-reduced elchelon matrix such that $A \sim R$. If $R \neq I_n$, then R has a zero row, hence L_R is not injective. This is a contradiction since R is invertible.

Thus (1) $\implies$ (2) and we're done. □

Define the **general linear group** $\mathrm{GL}_n(F)$ to be the set of invertible $n \times n$ matrices.

Corollary 4.15

For $A, B \in F^{m\times n}$, $\mathrm{row}(A) = \mathrm{row}(B) \iff \exists P \in \mathrm{GL}_m(F), B = PA$.

Proof. $\mathrm{row}(A) = \mathrm{row}(B) \iff A \sim B \iff B = PA$ for some invertible matrix P. □

Proposition 4.16 (How to compute the inverse of a matrix)

Suppose $A \in \mathrm{GL}_n(F)$, $e_1 \ldots, e_k$ are the elementary row operations such that $e_1 \cdots e_k(A) = I_n$, then $e_1 \cdots e_k(I_n) = A^{-1}$.

Proof. Trivial. □

Remark 4.17 — To compute the inverse of A, we can apply elementary row operations to the $n \times 2n$ matrix (A, I_n). If we get I_n on the left half, then the right half is precisely A^{-1}, i.e. the matirx becomes (I_n, A^{-1}).

The same method can be used to compute $A^{-1}B$. To compute BA^{-1}, we can use column operations instead.

Proposition 4.18

Let $A = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = (A_1, \dots, A_n)$ be a matrix. TFAE:

- A is invertible;
- $\{\alpha_1, \dots, \alpha_n\}$ is a basis of F^n;
- $\{A_1, \dots, A_n\}$ is a basis of F^n.

Proof. A is invertible $\iff$ $A \sim I_n$ $\iff$ $\mathrm{row}(A) = \mathrm{row}(I_n) = F^n$ $\iff$ $\{\alpha_1, \dots, \alpha_n\}$ is a basis of F^n. □

4.4 Coordinates in vector space

In the previous sections, we only discuss the linear spaces F^n and $F^{n\times 1}$. What if we run into a general vector space in which vectors aren't a list of numbers?

Definition 4.19. Given a vector space V and an ordered basis $B = (\alpha_1, \dots, \alpha_n)$, for every vector $\alpha \in V$ there's a unique n-tuple of numbers $(x_1, \dots, x_n)$ such that $\sum_{i=1}^{n} x_i\alpha_i = \alpha$. The x_i's are called the **coordinate** of α under basis B.

Let $[\alpha]_B := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, we can write $\alpha = (\alpha_1, \dots, \alpha_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Consider a map $\Gamma_B : V \to F^n : \alpha \mapsto [\alpha]_B$, we can see that this is a linear isomorphism. Thus any finite dimensional vector space is completely determined by its dimension.

Remark 4.20 — In practice, we won't regard all the spaces as F^n, because usually there's no natural basis for a general vector space.

Take n vectors $\alpha'_1, \dots, \alpha'_n \in V$, there exists an $n \times n$ matrix P such that $(\alpha'_1, \dots, \alpha'_n) = (\alpha_1, \dots, \alpha_n) \cdot P$.

Proposition 4.21

$\{\alpha'_1, \dots, \alpha'_n\}$ is a basis $\iff$ P is invertible.

Proof. α'_i linear independent $\iff$ "$\forall X \in F^{n\times 1}$, $(\alpha'_i)X = 0 \implies X = 0$"

$$\iff \text{"}(\alpha_i)PX = 0 \implies X = 0\text{"}$$

Notice that $(\alpha_i)PX = 0 \iff PX = 0$, hence P invertible is equivalent to α'_i linear independent. □

Proposition 4.22

Let $B' = (\alpha'_1, \dots, \alpha'_n)$ be a basis, then

$$\forall \alpha \in V,\ [\alpha]_B = P[\alpha]_{B'}.$$

Proof. $\alpha = (\alpha_1, \dots, \alpha_n)[\alpha]_B = (\alpha'_1, \dots, \alpha'_n)[\alpha]_{B'} = (\alpha_1, \dots, \alpha_n)P[\alpha]_{B'}$. □

4.5 Row spaces

Let A, B be $m \times n$ matrices, P be a $m \times m$ matrix. We know that

- If $B = PA$, then $\mathrm{row}(B) \subset \mathrm{row}(A)$.
- Moreover if P is invertible, then $\mathrm{row}(B) = \mathrm{row}(A)$.

Definition 4.23. The row rank of a matrix A $\mathrm{rowrank}(A) := \dim \mathrm{row}(A)$. Simlilarly define the column rank.

Theorem 4.24

For all matrix $A \in F^{m\times n}$, we have $\mathrm{rowrank}(A) = \mathrm{colrank}(A) =: \mathrm{rank}(A)$.

Proof. $\dim \ker A + \dim \mathrm{im}\, A = n$, $\mathrm{col}(A) = \mathrm{im}\, A$.

Let R be the row reduced matrix which is row equivalent to A. We have $\ker A = \ker R$, $\mathrm{row}(A) = \mathrm{row}(R)$.

Suppose there are r nonzero rows in R, it's clear that $\ker R = n - r$, $\mathrm{rowrank}(R) = r$. □

Corollary 4.25

$\mathrm{rank}(A) + \dim \ker A = n$.

Remark 4.26 (Porperties of ranks) — A, B, C are matrices with suitable size,

- $\mathrm{rank}(AB) \le \min\{\mathrm{rank}(A), \mathrm{rank}(B)\}$
- $\mathrm{rank}(A + B) \le \mathrm{rank}(A, B) \le \mathrm{rank}(A) + \mathrm{rank}(B)$
- $\mathrm{rank}(AB) + \mathrm{rank}(BC) \le \mathrm{rank}(ABC) + \mathrm{rank}(B)$

The last one can be proved using the linear map $L_A|_{\mathrm{row}(BC)}$.

Proposition 4.27

Given two integers $m \le n$. Let W be a subspace of F^n, and $\dim W \le m$, then there exists a unique row reduced elchelon matrix $R \in F^{m\times n}$ such that $\mathrm{row}(R) = W$.

5 Linear maps

We have already defined linear maps back in Definition 3.4. Now we'll have a closer look at linear maps.

5.1 The algebraic structure of linear maps

To describe a linear map $T : V \to W$, we don't need to tell what $T(v)$ is for every $v \in V$. In fact, all the inforamtion we need is the image of a basis of V.

Proposition 5.1

Let V be a finite dimensional vector space, and $\{\alpha_1, \dots, \alpha_n\}$ be a basis of V. $\forall \beta_1, \dots, \beta_n \in W$, there exists a unique linear map $T : V \to W$, such that $T(\alpha_i) = \beta_i, i = 1, 2, \dots, n$.

Proof. Note that if $\alpha = \sum x_i \alpha_i$, then $T(\alpha) = \sum x_i T(\alpha_i)$. □

Definition 5.2. Given two vector spaces V, W over a field F, all the linear maps from V to W forms a vector space, called $\mathrm{Hom}_F(V, W)$.

The operations are given by $(T_1 + T_2)(v) = T_1(v) + T_2(v)$ and $(c \cdot T)(v) = c(T(v))$.

Sometimes when F is understood, we'll just write $\mathrm{Hom}(V, W)$.

Remark 5.3 — Combined with the multiplication structure of linear maps, we find that $\mathrm{Hom}(V, V)$ is in fact an **F-algebra** (that is, a vector space with multiplication having properties it should have).

Proposition 5.4

$T \in \mathrm{Hom}(V, W)$, TFAE:

- T injective;
- $\forall S \subset V$ linear independent, $T(S)$ is linear independent;
- $\exists$ a basis $S \subset V$, $T(S)$ is linear independent.

Proposition 5.5

$T \in \mathrm{Hom}(V, W)$, TFAE:

- T surjective;
- $\forall S \subset V$ spaning, $T(S)$ is spanning;
- $\exists S \subset V$ (S don't have to be a basis), $T(S)$ is spanning.

The proof is just a bunch of abstract nonsense, so I won't include them here.

5.2 Represent linear maps as matrices

Recall that matrices and linear maps are somehow the same thing:

Proposition 5.6

Let $V = F^n, W = F^m$, then for every linear map $T : V \to W$, there exists a unique $m \times n$ matrix A s.t. $T(\alpha) = A\alpha$. For the right multiplication we have the same result.

Now we'll do this for linear maps on general vector spaces instead of F^n, F^m.

Given two vector space V, W of dimension n, m, respectively. Fix a basis $B = \{\alpha_1, \dots, \alpha_n\}$ of V, and $B' = \{\beta_1, \dots, \beta_m\}$ of W.

Now for a linear map $T : V \to W$, there's a unique matrix such that $A[v]_B = [T(v)]_{B'}, \forall v \in V$ (In fact A is just putting the coordinates of $T(\alpha_i)$ into a matrix, satisfying the following commutative diagram). We say A is the matrix of T under basis B and B'.

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle \Gamma_B}\downarrow & & \downarrow{\scriptstyle \Gamma_{B'}} \\ F^{n\times 1} & \xrightarrow{L_A} & F^{m\times 1} \end{array}$$

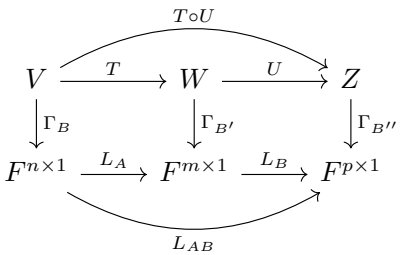
Remark 5.7 — Always keep in mind that the matrix of a linear map depends on the bases chosen for V and W.

This induces a bijection between linear maps and matrices, it's clear that this map is a linear isomorphism,

$$\mathrm{Hom}(V, W) \cong F^{m\times n}.$$

Remark 5.8 — This bijection is not canonical, it depends on the bases chosen for V and W. To understand what exactly $\mathrm{Hom}(V, W)$ is, we need the dual space, which will be discussed later.

If $T : V \to W$ and $U : W \to Z$, fix a basis for V, W, Z each. We have the matrix of $T \circ U$ is just the product of the matrices of T and U.



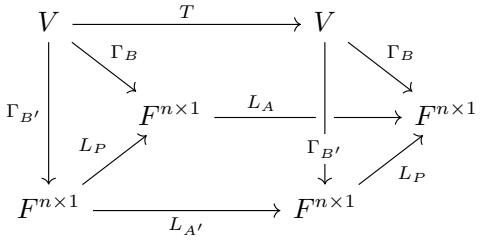
Now let's look at linear maps in $\mathrm{Hom}(V, V)$. If we fix a basis B, for any T, we denote its matrix (under B and B) as $[T]_B$.

Then we could write $[T(\alpha)]_B = [\alpha]_B, [T \circ U]_B = [T]_B[U]_B$. Hence sometimes we just write $T(\alpha)$ as $T\alpha$ and $T \circ U$ as $T \cdot U$.

For a linear map $T \in \mathrm{Hom}(V, V)$, what will happen to its matrix if we take different basis of V?

Take two basis B and B' of V, there's an invertible matrix P such that $[\alpha]_B = P[\alpha]_{B'}$. Suppose $[T]_B = A, [T]_{B'} = A'$, we have

$$A'[\alpha]_{B'} = [T\alpha]_{B'} = P^{-1}[T\alpha]_B = P^{-1}A[\alpha]_B = P^{-1}AP[\alpha]_{B'}.$$



We deduce

$$A' = P^{-1}AP.$$

Definition 5.9. Two $n \times n$ matrices A and A' is **similar**, if there exists $P \in \mathrm{GL}_n(F)$, s.t. $A' = P^{-1}AP$.

Now we know that similar matrices are just the same linear map under different basis.

midterm exam

5.3 Dual spaces & pairings

Definition 5.10. Let $V^* := \mathrm{Hom}(V, F)$ be the **dual space** of V. It consists of linear functions from V to F. Note that $\dim V^* = \dim V$ (This is only true when $\dim V < +\infty$).

There's a natural map $\phi : V^* \times V \to F$ called "evaluation". This map ϕ is **bilinear**, which means it is linear for both V and V^*. This shows why V^* is called the "dual" of V.

Definition 5.11. We say a function φ is **bilinear** (also called **pairing**), if

- $\forall \alpha \in V$, $f_\alpha : W \to F, f_\alpha(\beta) = \varphi(\alpha, \beta)$ is linear.
- $\forall \beta \in W$, $g_\beta : V \to F, g_\beta(\alpha) = \varphi(\alpha, \beta)$ is linear.

And we define

$$\begin{cases} L_\varphi : V \to W^* : \alpha \mapsto f_\alpha \\ R_\varphi : W \to V^* : \beta \mapsto g_\beta \end{cases}$$

We say φ is **non-degenerate**, if $\ker(L_\varphi) = \{0\} = \ker(R_\varphi)$.

Example 5.12

Examples of non-degenerate bilinear function:

- $\mathrm{ev} : V^* \times V \to F$, $\mathrm{ev}(f, \alpha) = f(\alpha)$.
- Let $V = F^n$, $W = F^{n\times 1}$, $\varphi(\alpha, \beta) = \alpha\beta$.
- If $A \in F^{m\times n}$, $V = F^m$, $W = F^{n\times 1}$. $\varphi(\alpha, \beta) = \alpha A\beta$. φ is non-degenerate iff $m = n$ and A invertible.

Next we give a different understanding of why the row rank of a matrix is equal to its column rank.

Theorem 5.13

Let V, W be finite dimensional vector spaces, if there exists a non-degenerate pairing $\varphi : V \times W \to F$, then $\dim V = \dim W$.

Proof. L_φ is injective $\implies \dim V \le \dim W^* = \dim W$;
R_φ is injective $\implies \dim W \le \dim V^* = \dim V$. □

Remark 5.14 — More generally, $\dim V - \dim \ker(L_\varphi) = \dim W - \dim \ker(R_\varphi)$.
This induces a non-degenerate pairing $\widetilde{\varphi} : V/\ker(L_\varphi) \times W/\ker(R_\varphi) \to F$.

We can use this theorem to give an alternate proof of Theorem 4.24.

Let $A \in F^{m\times n}$, $V = \mathrm{row}(A)$, $W = \mathrm{col}(A)$. Define $\varphi : V \times W \to F$:

For $\alpha \in V$, take $\alpha' \in F^m$ s.t. $\alpha = \alpha' A$.

For $\beta \in W$, take $\beta' \in F^{n\times 1}$ s.t. $\beta = A\beta'$.

$\varphi(\alpha, \beta) := \alpha' A\beta' = \alpha\beta' = \alpha'\beta$. We can check φ is a well-defined pairing.

Now we claim that φ is non-degenerate.

Indeed, if $\alpha \in \ker(L_\varphi)$, we have $\alpha\beta' = 0, \forall \beta' \in F^{n\times 1} \implies \alpha = 0$. The same process implies that $\ker(R_\varphi) = \{0\}$.

Thus from the above theorem, we get $\dim V = \dim W$. **In this proof, the rows and columns are totally symmetric.**

This also gives a hint of what the row vectors are. In fact for $\alpha \in F^k$, define $f_\alpha \in (F^{k\times 1})^*$: $f_\alpha(X) = \alpha X$, thus a row vector can be regarded as an element in the dual space of column vectors. The idea of "rows and columns are symmetric" appears again.

Proposition 5.15

Let $\dim V = n$, and $B = \{\alpha_1, \dots, \alpha_n\}$ be a basis. Then there exists a unique basis $\{f_1, \dots, f_n\}$ of V^*, such that $f_i(\alpha_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

Proof. Trivial. □

We call this $\{f_1, \ldots, f_n\}$ the **dual basis** of $\{\alpha_1, \ldots, \alpha_n\}$, sometimes we write f_i as α_i^*.

Observe that $f_i(\alpha)$ is the coefficient of α_i when we write α as linear combination of $\alpha_!, \ldots, \alpha_n$, and $\{\alpha_1, \ldots, \alpha_n\}$ is the dual basis of $\{f_1, \ldots, f_n\}$.

We can think of the dual basis is taking the coefficient of each term when writting a vector as a linear combination of the original basis. That is,

$$\alpha = \sum_{i=1}^{n} f_i(\alpha)\alpha_i, \quad f = \sum_{i=1}^{n} f(\alpha_i)f_i.$$

In terms of bilinear functions, the above proposition can be state as:

Proposition 5.16

If there's a non-degenerate pairing $\varphi : V \times W \to F$, for any basis $\{\alpha_1, \ldots, \alpha_n\}$ of V, there exists a unique basis $\{\beta_1, \ldots, \beta_n\}$ of W, such that $\varphi(\alpha_i, \beta_j) = \delta_{ij}$.

Where $\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$.

Proof. Take $\beta_i = R_\varphi^{-1}(f_i)$, where $\{f_1, \ldots, f_n\}$ is the dual basis of $\{\alpha_1, \ldots, \alpha_n\}$. □

5.4 Double dual

If we apply this to V^* and V^{**}, we can see that elements in V^{**} are functions on V^*, and there's a natural map between V and V^{**}, in which $\alpha \in V$ maps to "evaluate at α"(denoted by L_α in this handout), which is a function on V^*. Observe that this map is an injective linear map:

When $\dim V < \infty$, we can take a basis containing α, thus $\ker L_\varphi = \{0\}$. For infinte dimensional spaces, the same thing can also be done, but it requires Zorn's lemma.

So when $\dim V < \infty$, $V \cong V^{**}$. This intensifies the idea that we should view $\alpha \in V$ as a function on V^*, and $\alpha(f) = f(\alpha) = \mathrm{ev}(f, \alpha)$, this is what "dual" means.

Definition 5.17. For $S \subset V$, let $S^0 := \{\beta \in W \mid \varphi(\alpha, \beta) = 0, \forall \alpha \in S\}$ be the **annihilator** of S. Similarly we can define this for subsets of W.

Note that annihilators are subspaces and $S^0 = \mathrm{span}(S)^0$.

Remark 5.18 — Equivalent definitions:

$$g_\beta(\alpha) = 0, \forall \alpha \in S \iff S \subset \ker g_\beta \iff g_\beta|_S = 0 \iff g_\beta(S) = \{0\}$$

In non-degenerate pairings, $R_\varphi : W \to V^*$ is a natural isomorphism, so we can view W and V^* as the same space. In what follows, if the space W is not mentioned, we'll assume that $W = V^*$.

Theorem 5.19

Let Z be a subspace of V, then

$$\dim Z + \dim Z^0 = \dim V.$$

Proof. Take a basis of Z, complete it to a basis of V and take the dual basis in V^*. The rest is some trivial computation. □

In fact we can prove $Z = \bigcap_{i=k+1}^{n} \ker f_i$, where $Z^0 = \operatorname{span}\{f_{k+1}, \dots, f_n\}$. Thus if $\dim Z = n-1$, there exists a unique $f \in V^* \backslash \{0\}$ such that $Z = \ker f$. In this case Z is called a hyperplane.

Remark 5.20 — Now we come back to matrices again.

Let $A \in F^{m\times n}$, consider the pairing $F^n \times F^{n\times 1} \to F$.

$\operatorname{row}(A)^0 = \{X \in F^{n\times 1} \mid \alpha X = 0, \forall \alpha \in \operatorname{row}(A)\} = \{X \in F^{n\times 1} \mid AX = 0\} = \ker A$.

So we get $\dim \ker A + \dim \operatorname{row}(A) = n$.

Recall $\dim \operatorname{col}(A) = \dim \operatorname{im} A = n - \dim \ker A$, thus $\dim \operatorname{row}(A) = \dim \operatorname{col}(A)$.

Corollary 5.21

Note that $(Z^0)^0 \subset V$, we claim that $(Z^0)^0 = Z$.

Proof. Clearly $Z \subset (Z^0)^0$, and equality follows from dimensional reasons. □

This tells us that if S is a subset of V, $(S^0)^0 = \operatorname{span}(S)$.

5.5 Quotient spaces

Definition 5.22. Let W be a subspace of V. We say a **coset**(or **affine subspace**) of W is of the form $\alpha + W := \{\alpha + \beta \mid \beta \in W\}$.

Notice that $\alpha_1 + W = \alpha_2 + W \iff \alpha_1 - \alpha_2 \in W$. So this gives a equivalent relation on V(kind of like "taking modulo W").

Define the **quotient space** V/W to be the set of all cosets of W(i.e. all the equivalency classes), with operations given by

$$(\alpha + W) + (\beta + W) = (\alpha + \beta) + W \text{ and } c(\alpha + W) = c\alpha + W.$$

We check that this definition is well-defined: if $\alpha' + W = \alpha + W, \beta' + W = \beta + W$, then $(\alpha + \beta) + W = (\alpha' + \beta') + W$ and $c\alpha + W = c\alpha' + W$.

Now there's a natural projection map $Q : V \to V/W : \alpha \mapsto \alpha + W$, it's a surjective linear map with kernel W.

Corollary 5.23

Let W be a subspace of a finite dimensional vector space V. We have

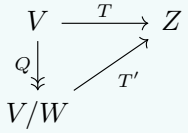
$$\dim V/W = \dim V - \dim W.$$

Proof. Just take a basis $\{\alpha_1, \dots, \alpha_n\}$ of V such that $\{\alpha_1, \dots, \alpha_k\}$ is a basis of W. Clearly $\{\alpha_{k+1} + W, \dots, \alpha_n + W\}$ is a basis of V/W.

(Alternatively, $\dim V/W = \dim \operatorname{im} Q = \dim V - \dim \ker Q = \dim V - \dim W$.) □

Proposition 5.24

Let $T \in \operatorname{Hom}(V, Z)$ such that $W \subset \ker T$. Then there's a unique $T' \in \operatorname{Hom}(V/W, Z)$ such that the following diagram commutes, and $\ker T' = (\ker T)/W, \operatorname{im} T' = \operatorname{im} T$.



In particular, if $\ker T = W$ and T is surjective, then T' is an isomorphism. This is to say $V/\ker T \cong \operatorname{im} T$.

Proof. Let $T'(\alpha + W) = T(\alpha)$, easy to verify T' is unique, well-defined and linear.

$\alpha + W \in \ker T' \iff \alpha \in \ker T \iff \alpha + W \in (\ker T)/W$ so $\ker T' = (\ker T)/W$. □

Some facts about quotient spaces:

There's a bijection between subspaces of V containing W and subspaces of V/W which perserves inclusion:

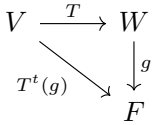
$$\{\text{subspaces of } V \text{ containing } W\} \xleftrightarrow{1:1} \{\text{subspaces of } V/W\}$$

$$N \longmapsto N/W$$

$$Q^{-1}(\overline{N}) \longmapsfrom \overline{N}$$

5.6 Transpose of linear maps

Definition 5.25. Let V, W be vector spaces and $T \in \operatorname{Hom}(V, W)$. Define $T^t : W^* \to V^*$ to be the **transpose** of T, given by $T^t(g) = g \circ T$.



Since $T^t(cg_1 + g_2) = (cg_1 + g_2) \circ T = cT^t(g_1) + T^t(g_2)$, we get T^t is linear.

Those who are familiar with transpose of matrices might guess that $(T^t)^t = T$ and some other properties same as matrices.

Indeed, $(T^t)^t : V^{**} \to W^{**}$, for any $g \in V^*$,

$$(T^t)^t(L_\alpha)(g) = (L_\alpha \circ T^t)(g) = (g \circ T)(\alpha) = g(T(\alpha)) = L_{T(\alpha)}(g),$$

which means $(T^t)^t$ sends $L_\alpha \in V^{**}$ to $L_{T(\alpha)}$, and that's what we expected.

Consider $\begin{array}{ccccc} V & \xrightarrow{T} & W & \xrightarrow{U} & Z \\ V^* & \xleftarrow[T^t]{} & W^* & \xleftarrow[U^t]{} & Z^* \end{array}$. We want to prove $(U \circ T)^t = T^t \circ U^t$.

$$(UT)^t(g) = g(UT) = T^t(gU) = T^tU^tg$$

Proposition 5.26

Let V, W be finite dimensional vector spaces, $T \in \mathrm{Hom}(V, W)$. B, B' are bases of V, W, and their dual bases are B^*, B'^*. Then $[T^t]_{B'^*,B^*} = [T]^t_{B,B'}$

Proof. Let $B = \{\alpha_1, \ldots, \alpha_n\}, B' = \{\beta_1, \ldots, \beta_m\}$. From the definition of matrices, we have

$$(T\alpha_1, \ldots, T\alpha_n) = (\beta_1, \ldots, \beta_m)[T]$$

$$(T^t\beta_1^*, \ldots, T^t\beta_m^*) = (\alpha_1^*, \ldots, \alpha_n^*)[T^t]$$

Some annoying computation:

$$T^t(\beta_j^*)(\alpha_i) = \beta_j^*(T\alpha_i) = \sum_{k=1}^{m} \beta_j^*([T]_{ik}\beta_k) = [T]_{ij}$$

and

$$T^t(\beta_j^*)(\alpha_i) = \sum_{k=1}^{n} [T^t]_{jk}\alpha_k^*(\alpha_i) = [T^t]_{ji}$$

□

Proposition 5.27

Let V, W be finite dimensional vector spaces, $T \in \mathrm{Hom}(V, W)$. Then

$$\ker T^t = (\mathrm{Im}\, T)^0, \quad \mathrm{Im}\, T^t = (\ker T)^0.$$

Proof. $\ker T^t = \{g \in W^* \mid T^tg = gT = 0\}$.

But $gT = 0 \iff g\big|_{\mathrm{Im}\, T} = 0 \iff g \in (\mathrm{Im}\, T)^0$.

The other equality follows similarly (just take the annihilators of both). □

Corollary 5.28

$\mathrm{rank}(T^t) = \mathrm{rank}(T)$. (Same as row rank = column rank)

Proof.

$$\dim \ker T^t + \dim \operatorname{im} T^t = \dim W^*,$$

and from previous proposition,

$$\dim \ker T^t + \dim \operatorname{im} T = \dim W.$$

So the conclusion follows. □

6 Determinants

Let V be a vector space over field $\mathbb{R}$ and T a linear map from V to itself.

We want to find out whether T is invertible. Note that for a non-degenerate "figure" in $\mathbb{R}^n$, T is invertible if and only if the image is also non-degenerate (meaning it cannot be included in a proper subspace of $\mathbb{R}^n$).

Hence we want to find a way to decide whether a figure is non-degenerate, from the experience in $\mathbb{R}^3$, we come to define a general "volume" for n-dimensional figures and see how it changes under T.

To do this, we need some prerequisites.

6.1 Symmetry groups

(Skip this section if you're familiar with S_n and A_n)

Definition 6.1. We say a bijection σ from $\{1, 2, \ldots, n\}$ to itself is a **permutation**.

Define the n-th **symmetry group** $S_n := \{\text{all permutations of } \{1, 2, \ldots, n\}\}$, the group operation is the composition of maps.

If a permutation σ has exactly $n-2$ fixed points, i.e. only exchange two elements s, t, we call σ a **transposition**, written as (s, t).

Proposition 6.2

S_n can be generated by transpositions. Moreover, it can be generated by adjacent transpositions (exchanging two adjacent elements).

Definition 6.3 (reversed numbers)**.** Let σ be a permutation of $\{1, 2, \ldots, n\}$. If $i < j$ but $\sigma(i) > \sigma(j)$, we say (i, j) is a reversed pair of σ.

Let $l(\sigma)$ be the number of reversed pairs of σ, $\operatorname{sgn}(\sigma) = (-1)^{l(\sigma)}$ be its sign.

Proposition 6.4

Let $k_1, \ldots, k_n$ be a permutation of $1, 2, \ldots, n$.

$$\operatorname{sgn}(\sigma) = \prod_{1 \le i < j \le n} \frac{\sigma(k_i) - \sigma(k_j)}{k_i - k_j}$$

Proposition 6.5

The map sgn: $S_n \to \{\pm 1\}$ is a group homomorphism.

In particular, if σ is wirtten as a product of transpositions, the sign of σ indicates the pairity of the number of transpositions.

Proof. Use the above proposition. □

Definition 6.6. Let $A_n := \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\}$ be the n-th **alternating group**. Then A_n is a subgroup of S_n and $S_n/A_n \cong \mathbb{Z}_2$.

6.2 Alternating functions

Let V be a vector space over $\mathbb{R}$ with dimension n. Take $\alpha_1, \dots, \alpha_n \in V$.

Let $P(\alpha_1, \dots, \alpha_n) = \{\sum_{i=1}^n c_i\alpha_i \mid 0 \le c_i \le 1\}$ be the "parallelepiped".

Instead of general diagrams, we'll only define a volume for those parallelepipeds for simplicity. Of course we would like some properties that is the same as 2-dimensional parallelograms and 3-dimensional parallelepipeds:

- If $\alpha_i = \alpha_j$ for some $i \neq j$, then its volume should be 0.
- The volume should be linear with respect to each α_i.

  This leads to a problem that the volume might be negative. Here we allow it to be negative, in the sense of "oriented volume", and it's also easier for us to operate on them.

We'll abstract these ideas in the following.

The second condition is called multilinearity:

Definition 6.7. Given a vector space V over F.

A map $L : V^r \to F$ is an **r-multilinear function** if it's linear with respect to each entry.

We can define operations on $M^r(V) := \{L : V^r \to F, L \text{ } r\text{-multilinear}\}$ in the obvious way, so $M^r(V)$ is also a vector space over F.

We write $f_1 \otimes f_2 \otimes \cdots \otimes f_r \in M^r(V) : (\alpha_1, \dots, \alpha_r) \mapsto f_1(\alpha_1) \cdots f_r(\alpha_r)$ for $f_1, \dots, f_r \in V^*$.

Remark 6.8 — The symbol $\otimes$ is called the tensor product, but here we won't talk much about it.

The first condition can be interpreted as follows:

Definition 6.9. We say a multilinear function $L \in M^r(V)$ is **alternating** if

$$\forall s \neq t, \quad \alpha_s = \alpha_t \implies L(\alpha_1, \dots, \alpha_n) = 0.$$

Let $\Lambda^r(V) = \{L \in M^r(V) \mid L \text{ alternating}\}$, clearly it's a subspace of $M^r(V)$.

It's easy to discover that: If L is alternating,

$$L(\alpha_1, \alpha_2, \ldots) = -L(\alpha_1 + \alpha_2, -\alpha_2, \ldots) = -L(\alpha_1 + \alpha_2, \alpha_1, \ldots) = -L(\alpha_2, \alpha_1, \ldots)$$

This means if we interchange two entries of L, its value differs by a change of sign, that's where the name "alternating" comes from.

Example 6.10

When $r = 2$, let $f, g \in V^*$, define the wedge product $f \wedge g := f \otimes g - g \otimes f$. Then $f \wedge g \in \Lambda^2(V)$ is a alternating bilinear function.

Let $V = F^{2n}$, $\{f_1, \ldots, f_{2n}\}$ is a basis of V^*. Let $\omega_0 = \sum_{i=1}^{n} f_i \wedge f_{i+n} \in \Lambda^2(V)$ be the standard symplectic form. It's also an example of alternating function.

The above construction can be generalized to arbitary r:

Definition 6.11. Let $f_1, \ldots, f_r \in V^*$, define

$$f_1 \wedge f_2 \wedge \cdots \wedge f_r = \sum_{\sigma \in S_r} \operatorname{sgn}(\sigma) f_{\sigma(1)} \otimes \cdots \otimes f_{\sigma(r)}$$

It's easy to check $f_1 \wedge \cdots \wedge f_s \wedge \cdots \wedge f_t \wedge \cdots \wedge f_r = -f_1 \wedge \cdots \wedge f_t \wedge \cdots \wedge f_s \wedge \cdots \wedge f_r$ (check it yourself!).

So it's indeed an alternating function.

Corollary 6.12

Let $\sigma \in S_r$, $f_1 \wedge \cdots \wedge f_r = \operatorname{sgn}(\sigma) f_{\sigma(1)} \wedge \cdots \wedge f_{\sigma(r)}$.

Corollary 6.13

Let $f_1, \ldots, f_r \in V^*$, then

$$f_1 \wedge \cdots \wedge f_r \neq 0 \iff f_1, \ldots, f_r \text{ are linear independent}$$

Moreover, from this we can compute the dimension of $\Lambda^r(V)$:

Take a basis of V we discover that $\dim \Lambda^r(V) = \binom{n}{r}$.

Remark 6.14 — In fact we need to prove all the element in $\Lambda^r(V)$ can be generated by $f_1 \wedge \cdots \wedge f_r$ first.

Proof. Let $L \in \Lambda^r(V)$.

Let $\{f_1, \ldots, f_n\}$ be a basis of V^* and $\{\alpha_1, \ldots, \alpha_n\}$ be the dual basis.

We claim that:

$$L = \sum_{i_1 < \cdots < i_r} L(\alpha_{i_1}, \ldots, \alpha_{i_r}) f_{i_1} \wedge \cdots \wedge f_{i_r}.$$

We check manually they're equal:

$\forall \beta_1, \ldots, \beta_r \in V$, expand $L(\beta_1, \ldots, \beta_r)$ into linear combinations of $L(\alpha_{i_1}, \ldots, \alpha_{i_r})$ and we'll get the result.

The computation is boring so it's omitted. □

Since we only care about n-dimensional volumes, let's look at the case where $r = n$. Surprisingly we see that $\dim \Lambda^n(V) = 1$, which means there's only one way to define the volume! (up to a scalar)

Proposition 6.15

Let $L \in \Lambda^n(V) \backslash \{0\}$, $\alpha_1, \ldots, \alpha_n \in V$, then $\alpha_1, \ldots, \alpha_n$ form a basis iff $L(\alpha_1, \ldots, \alpha_n) \neq 0$.

Proof. "$\Longrightarrow$": Take a dual basis, from dimensional reasons we get $L = cf_1 \wedge \cdots \wedge f_n \implies L(\alpha_1, \ldots, \alpha_n) = c$.

"$\Longleftarrow$": Assume they're linearly dependent, then

$$L(\alpha_1, \ldots, \alpha_n) = \sum_{i=2}^{n} c_i L(\alpha_i, \alpha_2, \ldots, \alpha_n) = 0$$

□

This means it makes sense when we interpret L as a volume function.

6.3 Determinants of linear maps

For $T \in \mathrm{Hom}(V, V)$, we'll define the determinant of T such that

$$L(T\alpha_1, \ldots, T\alpha_n) = \det(T) L(\alpha_1, \ldots, \alpha_n).$$

We need to check this is well-defined:

- It's independent with respect to the choice of L, since $\dim \Lambda^n(V) = 1$;
- It's independent with respect to α_i's:

  Choose $L' \in \Lambda^n(V)$ such that

  $$L'(\alpha_1, \ldots, \alpha_n) = L(T\alpha_1, \ldots, T\alpha_n)$$

  There exists a constant $c \in F$ s.t. $L' = cL$, hence $\det(T) = c$.

But the above definition is still not so good, since it need to be checked it's well-defined. We'll give a more intrinsic definition.

Definition 6.16. Let V, W be F-vector space, $T \in \mathrm{Hom}(V, W)$. Let $T^{(r)} : \Lambda^r(W) \to \Lambda^r(V)$ to be

$$T^{(r)}(L)(\alpha_1, \ldots, \alpha_n) = L(T\alpha_1, \ldots, T\alpha_n), \quad \forall L \in \Lambda^r(W)$$

Remark 6.17 — Note that $T^{(1)} = T^t$ and $(UT)^{(r)} = T^{(r)} \circ U^{(r)}$.

We can see that when $T \in \mathrm{Hom}(V, V)$, $T^{(n)} : \Lambda^n(V) \to \Lambda^n(V)$ must be a "multiply by a scalar" map.

Definition 6.18. Let $\det(T)$ be the scalar satisfying $T^{(n)} = \det(T)\mathrm{id}_{\Lambda^n(V)}$.

Remark 6.19 — It's easy to see this is the same definition as above.

Example 6.20

$\det(\mathrm{id}_V) = 1$ as $\mathrm{id}_V^{(n)} = \mathrm{id}_{\Lambda^n(V)}$.

Now some well-known properties of determinant become trivial:

Proposition 6.21

Let $T, U \in \mathrm{Hom}(V, V)$, then $\det(TU) = \det(T)\det(U)$.

Proposition 6.22

Let $T \in \mathrm{Hom}(V, V)$, then T invertible $\iff \det(T) \neq 0$, and $\det(T^{-1}) = \det(T)^{-1}$.

Proof. Let $\alpha_1, \ldots, \alpha_n$ be a basis of V.

$$\det(T) \neq 0 \iff L(T\alpha_1, \ldots, T\alpha_n) = \det(T)Ł(\alpha_1, \ldots, \alpha_n) \neq 0$$

$$\iff T\alpha_1, \ldots, T\alpha_n \text{ is a basis of } V.$$

□

Proposition 6.23

Let $T \in \mathrm{Hom}(V, V)$, $\phi : V \to W$ is an linear isomorphism, $\det(\phi T \phi^{-1}) = \det(T)$.

Proof.

$$(\phi T \phi^{-1})^{(n)} = (\phi^{-1})^{(n)} \det(T)\phi^{(n)} = \det(T)(\phi \circ \phi^{-1})^{(n)} = \det(T)\mathrm{id}_{\Lambda^n(W)}$$

□

Proposition 6.24 (Computing the determinant)

Let α_i and f_i be a pair of dual bases of V and V^*.

$$\det(T) = f_1 \wedge \cdots \wedge f_n(T\alpha_1, \ldots, T\alpha_n)$$

Define the **special linear group** $\mathrm{SL}(V)$:

$$\mathrm{SL}(V) = \{T \in \mathrm{Hom}(V, V) \mid \det(T) = 1\}.$$

6.4 Determinants of square matrices

Lemma 6.25

Let $A \in F^{n\times n}$, then $\det(L_A) = \det(R_A)$.

Proof. Computation. □

Remark 6.26 — More generally we have $\det(T) = \det(T^t)$.

Definition 6.27. Let $A \in F^{n\times n}$, define the determinant of A to be $\det(A) := \det(L_A)$.

We write $\det(A) = \begin{vmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{vmatrix}$. And define $\mathrm{SL}_n(F) = \{A \in F^{n\times n} \mid \det(A) = 1\}$.

Proposition 6.28

Let T be a linear map $V \to V$, for any basis B, $\det([T]_B) = \det(B)$.

Proof. The coordinate map $\Gamma : \alpha \mapsto [\alpha]_B$ is a linear homomorphism, and $L_{[T]_B} \circ \Gamma = \Gamma \circ T$, by Proposition 6.23,

$$\det(T) = \det(L_{[T]_B}) = \det([T]_B).$$

□

So we see that the determinant of a matrix is essentially the same thing as the determinant of a linear map, so all the properties in the previous section carries over.

Corollary 6.29 (Computing the determinant)

$$\det(A) = \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{n\sigma(n)}$$

Proof. Apply Proposition 6.24 and expand. □

Proposition 6.30

For $A_1, A_2, \ldots, A_n \in F^{n\times 1}$, the map $(A_1, \ldots, A_n) \mapsto \det(A_1, \ldots, A_n)$ is an alternating function.

For $\alpha_1, \ldots, \alpha_n \in F^n$, the map $(\alpha_1, \ldots, \alpha_n) \mapsto \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ is also an alternating function.

Proof. From Proposition 6.24 $\det(A_1, \ldots, A_n) = f_1 \wedge \cdots \wedge f_n(A_1, \ldots, A_n)$.
Where the f_i's form the dual basis of A_i. □

Some well-known properties of determinants: (see to it that you can prove them)

1. If a row/column is zero, then the determinant is 0.
2. If we interchange two rows/columns, the determinant changes its sign.
3. Adding a multiple of some row/column to another row/column doesn't change the determinant.
4. If a row/column is multiplied by $c \in F$, the determinant multiplied by c as well.

Definition 6.31. We say a matrix $A \in F^{n\times n}$ is **upper triangular** if $A_{ij} = 0, \forall i > j$ and **lower triangular** if $A_{ij} = 0, \forall i < j$.

Proposition 6.32

If A is an upper/lower triangular matrix, then its determinant is equal to the product of all the diagonal entries.

For block upper/lower triangular matrices we have the same result.

Lemma 6.33

If $A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$, where B, D are square matrices, then $\det(A) = \det(B)\det(D)$.

Proof. Computation. □

Remark 6.34 — In this section most of the proofs are computational because they're the proofs my teacher gave. But I'm not pleased with them and I wonder whether there's another approach.

Theorem 6.35 (Expand the determinant with respect to a row/column)

Let $A = (a_{ij})$, and denoted by M_{ij} the $(n-1) \times (n-1)$ matrix obtained by crossing out the i-th row and j-th column of A.

For all $i \in \{1, 2, \ldots, n\}$,

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(M_{ij})$$

Proof. Computation again! □

We call $\det(M_{ij})$ the complement minor of a_{ij}, and $C_{ij} := (-1)^{i+j} \det(M_{ij})$ the **cofactor** of a_{ij}.

Corollary 6.36

For an $n \times n$ matrix A:

$$\sum_{k=1}^{n} a_{ik} C_{jk} = \sum_{k=1}^{n} a_{ki} C_{kj} = \delta_{ij} \det(A)$$

where δ_{ij} is the dirac function.

Definition 6.37 (Adjoint matrices)**.** The (classic) **adjoint matrix** of A (denoted by $\operatorname{adj}(A)$) is the transpose of the matrix (C_{ij}). i.e. $\operatorname{adj}(A)_{ij} = C_{ji}$.

Corollary 6.38

Let $A \in F^{n \times n}$, $A \operatorname{adj}(A) = \operatorname{adj}(A) A = \det(A) I_n$.

At last we come to the famous theorem in linear equations:

Theorem 6.39 (Cramer's rule)

Let $A \in \mathrm{GL}_n(F)$, $Y \in F^{n \times 1}$, we already know that $AX = Y$ has a unique solution. Let B_j be the matrix obtained by repacing the j-th column with Y, then the solution $X = (x_1, \ldots, x_n)^t$ can be written explicitly:

$$x_j = \frac{\det(B_j)}{\det(A)}.$$

Proof. Assume $n \geq 2$. $X = A^{-1} Y = \det(A)^{-1} \operatorname{adj}(A) Y$.

Consider the j-th row of both sides, we get

$$x_j = \det(A)^{-1} \sum_{i=1}^{n} \operatorname{adj}(A)_{ji} y_i = \det(A)^{-1} \sum_{i=1}^{n} y_i C_{ij} = \det(A)^{-1} \det(B_j).$$

□

Alternative proof of Cramer's rule. Let the columns of A be $A_1, \dots, A_n$. Then

$$Y = AX = (A_1, \dots, A_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} x_i A_i.$$

Note that

$$\begin{aligned} \det(B_j) &= \det(A_1, \dots, A_{j-1}, \sum_{i=1}^{n} x_i A_i, A_{j+1}, \dots, A_n) \\ &= \det(A_1, \dots, A_{j-1}, x_j A_j, A_{j+1}, \dots, A_n) = x_j \det(A) \end{aligned}$$

So we're done. □

6.5 Tensor product & wedge product for functions

Definition 6.40 (Tensor product)**.** Let $L \in M^r(V), M \in M^s(V)$.

Define their tensor product $L \otimes M \in M^{r+s}(V)$ to be

$$L \otimes M(\alpha_1, \dots, \alpha_{r+s}) = L(\alpha_1, \dots, \alpha_r) M(\alpha_{r+1}, \dots, \alpha_{r+s}).$$

Some properties of tensor product:

- The map $M^r(V) \times M^s(V) \to M^{r+s}(V)$, $(L, M) \mapsto L \otimes M$ is bilinear.
- Tensor product has associativity, i.e. $L \otimes (M \otimes N) = (L \otimes M) \otimes N$, $\forall L \in M^r(V), M \in M^s(V), N \in M^t(V)$.

Theorem 6.41

Let V be a vector space with dimension n. Let $\{f_1, \dots, f_n\}$ be a basis of V^*.

(1) $\{f_{i_1} \otimes \cdots \otimes f_{i_r}\}$ is a basis of $M^r(V)$, so $\dim M^r(V) = n^r$;

(2) Let $\{\alpha_1, \dots, \alpha_n\}$ be the dual basis, then $\forall L \in M^r(V)$,

$$L = \sum_{i_1, \dots, i_r} L(\alpha_{i_1}, \dots, \alpha_{i_r}) f_{i_1} \otimes \cdots \otimes f_{i_r}.$$

Proof. Some computation will give (2) (much in the sense of Remark 6.14), so $\{f_{i_1} \otimes \cdots \otimes f_{i_r}\}$ is spanning.

As for linear independece: Plug $(\alpha_{j_1}, \dots, \alpha_{j_r})$ into

$$\sum_{i_1, \dots, i_r} c_{i_1, \dots, i_r} f_{i_1} \otimes \cdots \otimes f_{i_r} = 0$$

we get $c_{j_1, \dots, j_r} = 0$. □

Now we come back to alternating functions again:

Our goal is to define an "alternating product" of altenating functions. Here's an idea from earlier constructions:

Definition 6.42. Let $L \in M^r(V)$. The alternate of L is defined as

$$\mathrm{Alt}(L)(\alpha_1, \dots, \alpha_r) = \sum_{\sigma \in S_r} \mathrm{sgn}(\sigma) L(\alpha_{\sigma(1)}, \dots, \alpha_{\sigma(r)})$$

It's obvious $\mathrm{Alt}(L)$ is an alternating function.

With this we have:

Definition 6.43. Let $L \in \Lambda^r(V), M \in \Lambda^s(V)$, when $\mathrm{char}(F) = 0$ define their wedge product to be $L \wedge M = \frac{1}{r!s!}\mathrm{Alt}(L \otimes M)$, i.e.

$$L \wedge M(\alpha_1, \dots, \alpha_{r+s}) = \frac{1}{r!s!} \sum_{\sigma \in S_{r+s}} \mathrm{sgn}(\sigma) L(\alpha_{\sigma(1)}, \dots, \alpha_{\sigma(r)}) M(\alpha_{\sigma(r+1)}, \dots, \alpha_{\sigma(r+s)})$$

Remark 6.44 — The constant here grants the associativity. It also can be $\frac{1}{(r+s)!}$.

In general cases, observe that in $\mathrm{Alt}(L \otimes M)$, each term appears $r!s!$ times, this means $\frac{1}{r!s!}$ makes more senses than $\frac{1}{(r+s)!}$, as it can be generalized to fields with chararteristic p.

Definition 6.45 (Wedge product)**.** Let $L \in \Lambda^r(V), M \in \Lambda^s(V)$, let

$$\mathrm{Sh}(r, s) = \{\sigma \in S_{r+s} \mid \sigma(1) < \cdots < \sigma(r), \sigma(r+1) < \cdots < \sigma(r+s)\}.$$

define their wedge product to be

$$L \wedge M(\alpha_1, \dots, \alpha_{r+s}) = \sum_{\sigma \in \mathrm{Sh}(r,s)} \mathrm{sgn}(\sigma) L(\alpha_{\sigma(1)}, \dots, \alpha_{\sigma(r)}) M(\alpha_{\sigma(r+1)}, \dots, \alpha_{\sigma(r+s)})$$

Remark 6.46 — The "Sh" here comes from the word "shuffle". Can you see why this makes sense?

To check $L \wedge M$ is indeed alternating, we only need to check when $\alpha_1 = \alpha_2$, $L \wedge M(\alpha_1, \dots, \alpha_{r+s}) = 0$.

From the alternativity of L, M, we only need to consider the terms with $\sigma^{-1}(1) \leq r, \sigma^{-1}(2) > r$ or $\sigma^{-1}(1) > r$, $\sigma^{-1}(2) \leq r$.

Let $\tau = (12) \in S_{r+s}$, $P = \{\sigma \in \mathrm{Sh}(r, s) \mid \sigma^{-1}(1) \leq r, \sigma^{-1}(2) > r\}$, we have

$$\begin{aligned} L \wedge M(\alpha_1, \dots, \alpha_{r+s}) &= \sum_{\sigma \in \mathrm{Sh}(r,s)} \mathrm{sgn}(\sigma) L(\alpha_{\sigma(1)}, \dots, \alpha_{\sigma(r)}) M(\alpha_{\sigma(r+1)}, \dots, \alpha_{\sigma(r+s)}) \\ &= \sum_{\sigma \in P} \mathrm{sgn}(\sigma) L(\alpha_{\sigma(1)}, \dots, \alpha_{\sigma(r)}) M(\alpha_{\sigma(r+1)}, \dots, \alpha_{\sigma(r+s)}) \\ &\quad - \sum_{\sigma \in P} \mathrm{sgn}(\sigma) L(\alpha_{\tau\sigma(1)}, \dots, \alpha_{\tau\sigma(r)}) M(\alpha_{\tau\sigma(r+1)}, \dots, \alpha_{\tau\sigma(r+s)}) \\ &= 0 \end{aligned}$$

Thus if $\alpha_i = \alpha_{i+1}$, $L \wedge M(\alpha_1, \dots, \alpha_{r+s}) = 0$, which proves it's alternating.

Example 6.47

When $r = s = 1$, $f, g \in V^*$, $f \wedge g(\alpha, \beta) = f(\alpha)g(\beta) - f(\beta)g(\alpha)$.

We can see that $f \wedge f = 0$ for any $f \in V^*$.

When $r = 2, s = 1$,

$$L \wedge M(\alpha_1, \alpha_2, \alpha_3) = L(\alpha_1, \alpha_2)M(\alpha_3) + L(\alpha_2, \alpha_3)M(\alpha_1) - L(\alpha_1, \alpha_3)M(\alpha_2).$$

Some more examples:

Example 6.48

Let $1 \leq k < n$, and $L \in \Lambda^k(F^{n\times 1})$ be

$$L\left(\begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix}, \ldots, \begin{pmatrix} x_{1k} \\ \vdots \\ x_{nk} \end{pmatrix}\right) = \begin{vmatrix} x_{11} & \cdots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{k1} & \cdots & x_{kk} \end{vmatrix}.$$

Let $M \in \Lambda^{n-k}(F^{n\times 1})$ be

$$M\left(\begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix}, \ldots, \begin{pmatrix} x_{1,n-k} \\ \vdots \\ x_{n,n-k} \end{pmatrix}\right) = \begin{vmatrix} x_{k+1,1} & \cdots & x_{k+1,n-k} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,n-k} \end{vmatrix}.$$

Consider $L \wedge M \in \Lambda^n(F^{n\times 1})$. Note that $\dim \Lambda^n(F^{n\times 1}) = 1$, so we get $L \wedge M$ is always equal to the determinant multiplying a scalar c.

Subtitute I_n into $L \wedge M$ we get $c = 1$, and expanding $L \wedge M$ we get the formula of expanding the determinant with respect to k rows.

Proposition 6.49 (Laplace, expanding determinant with respect to k rows)

Let $A \in F^{n\times n}$, $k < n$. For index sets $I = \{i_1, i_2, \ldots, i_k\}, J = \{j_1, j_2, \ldots, j_k\}$, let A_{IJ} be the determinant of the $k \times k$ matrix obtained by taking the $i_1, \ldots, i_k$-th rows and $j_1, \ldots, j_k$-th columns of A.

For a fixed set I, we have

$$\det(A) = \sum_{|J|=k} (-1)^{\sum_J + \sum_I} \det(A_{IJ}) \det(A_{I^cJ^c})$$

Where I^c, J^c are complement sets of I, J, $\sum_I$ is the short hand of $\sum_{i\in I} i$.

Remark 6.50 — The computation is really boring (involving computing the sign of some weird permutation), and this proposition won't be of much use.

The A_{IJ}'s are called a **minor determinant** (or simply a **minor**) of A, and $(-1)^{\sum_I + \sum_J} A_{I^cJ^c}$

is called the **cofactor** of A_{IJ}.

Now we come back to wedge produts. Here are some properties of wedge product:

- $(L, M) \mapsto L \wedge M$ is a bilinear map.
- Associativity, $(L \wedge M) \wedge N = L \wedge (M \wedge N)$.
- If $L \in \Lambda^r(V), M \in \Lambda^s(V)$, $L \wedge M = (-1)^{rs} M \wedge L$.

Proposition 6.51

Let $f_1, \dots, f_n \in V^*$, $\alpha_1, \dots, \alpha_n \in V$, then

$$f_1 \wedge \cdots \wedge f_n(\alpha_1, \dots, \alpha_n) = \det(f_i(\alpha_j)).$$

Corollary 6.52

Let $f_1, \dots, f_n \in V^*$, $\sigma \in S_n$. We have

$$f_1 \wedge f_2 \wedge \cdots \wedge f_n = \operatorname{sgn}(\sigma) f_{\sigma(1)} \wedge \cdots \wedge f_{\sigma(n)}.$$

Theorem 6.53

Let $\dim V = n$, $\{f_1, \dots, f_n\}$ is a basis of V^*,

(1) If $1 \le r \le n$,

$$\{f_{i_1} \wedge \cdots \wedge f_{i_r} \mid i_1 < \cdots < i_r\}$$

is a basis for $\Lambda^r(V)$, so $\dim \Lambda^r(V) = \binom{n}{r}$;

(2) If $r > n$, $\Lambda^r(V) = \{0\}$.

Proof. For the spanning part, see Remark 6.14.

As for linear independence, just plug $(\alpha_{i_1}, \dots, \alpha_{i_r})$ in to see that each coefficient is zero. □

Proposition 6.54

$f_1 \wedge \cdots \wedge f_r \neq 0 \iff \{f_1, \dots, f_r\}$ linearly independent. (Also true for infintely dimensional space)

Proof. " $\Longrightarrow$ " is obvious by multilinearity.

For the other direction, we claim that $T : V \to F^r$ which sends $\alpha \mapsto (f_1(\alpha), \dots, f_r(\alpha))$ is surjective.

To prove it we only need to check $\mathrm{Im}(T)^0 = \{0\}$:

$$\mathrm{Im}(T)^0 = \{g \in (F^r)^* \mid g \circ T = 0\} = \left\{ \sum_{i=1}^{r} c_i e_i^* \mid \sum_{i=1}^{r} c_i f_i(\alpha) = 0, \forall \alpha \in V \right\} = \{0\}$$

Hence $\exists \alpha_i$ s.t. $T(\alpha_i) = e_i$, i.e. $f_j(\alpha_1) = \delta_{ij}$. We have

$$f_1 \wedge \cdots \wedge f_r(\alpha_1, \ldots, \alpha_r) = \sum_{\sigma \in S_r} \mathrm{sgn}(\sigma) f_1(\alpha_{\sigma(1)}) \cdots f_r(\alpha_{\sigma(r)}) = 1.$$

□

7 Polynomials

7.1 Polynomial algebra

Definition 7.1 (Algebras)**.** Let $\mathcal{A}$ be an F-vector space. If there's a binary operation "multiplication" which satisfies associativity, disturbution and compatible with scaling in F, then we say $\mathcal{A}$ is an F**-algebra**.

Definition 7.2. Some special algebras:

- We say an F-algebra has a **unit**, if there exists $1_A \in \mathcal{A}$ s.t. $\forall \alpha \in \mathcal{A}$, $1_A \alpha = \alpha 1_A = \alpha$.
- We say it's **commutative** if the multiplication is commutative.

Example 7.3

Examples of algebras:

- Let $\mathcal{A} = F^{n \times n}$, $\mathcal{A}$ is an F-algebra with matrix multiplication, whose unit is the unit matrix I_n.
- Similarly, $\mathrm{Hom}(V, V)$ is also an algebra with unit id_V.
- Given two fields $F' \subset F$, we can see F is a commutative F'-algebra with a unit.

We'll construct a more complicated algebra, the formal power series:

Definition 7.4 (Formal power series)**.** Let $F[[x]] := \{(a_0, a_1, \ldots) \mid a_i \in F\} = F^{\mathbb{N} \cup \{0\}}$ be a vector space.

For $f = (a_0, a_1, \ldots), g = (b_0, b_1, \ldots)$, define the multiplication to be:

$$f \cdot g = (c_0, c_1, \ldots), \quad c_k = \sum_{i=0}^{k} a_i b_{k-i}.$$

$F[[x]]$ then becomes a commutative F-algebra. Note that $(1, 0, 0, \ldots)$ is the unit of $F[[x]]$.

Remark 7.5 — The map $c \mapsto (c, 0, 0 \dots)$ is an embedding of F into $F[[x]]$. So we won't distinguish c from $(c, 0, 0, \dots)$ later. (Clearly $c(a_0, a_1, \dots) = (c, 0, \dots) \cdot (a_0, a_1, \dots)$)

You might wonder where is the x in $F[[x]]$. Let $x := (0, 1, 0, 0, \dots)$ be the **indeterminated**. You may recognize the multiplcation as "polynomial multiplication" in this sense.

We can check that $x^k = (0, \dots, 0, 1, 0, \dots)$ with 1 at the kth entry. Let x^0 be $1 = (1, 0, 0, \dots)$.

Hence if $f = (a_0, a_1, \dots, a_n, 0, \dots)$, we can recognize it as $\sum_{k=0}^{n} a_k x^k$.

For a general f, we denote it by $\sum_{k=0}^{\infty} a_k x^k$. This notation has no relation with the acutal sum or anything about limits. It's merely a way of writing them. (As the name "formal" suggests)

Note that $\{x^0, x^1, x^2, \dots\}$ is linearly independent but not spanning, as there can be infinte linear combinations of $\{x^k\}$, which are not in its span.

Definition 7.6. Let

$$F[x] = \operatorname{span}\{1, x, x^2, \dots\} = \left\{ \sum_{k=0}^{\infty} a_k x^k \mid a_k \text{ has only finitely many nonzero terms} \right\}$$

This is known as the **polynomial ring** over F. Indeed we can check $F[x]$ is closed under multiplication easily.

Clearly $F[x]$ is a subalgebra of $F[[x]]$.

In high school mathematics, polynomials are always viewed as "polynomial function". However, in modern algebra these are different concepts.

Example 7.7

In $\mathbb{F}_p[x]$, x^p and x correspond to the same *function*, but as *polynomials* they are not the same.

Definition 7.8. Let $f = \sum_{k=0}^{\infty} a_k x^k \in F[x]$, the **degree** of f is

$$\deg f := \max\{k \geq 0 \mid a_k \neq 0\}.$$

If $f = 0$, define $\deg f = -\infty$.

If $f = \sum_{k=0}^{n} a_k x^k \in F[x], a_n \neq 0$. a_n is called the **leading coefficient**, we say f is **monic** if $a_n = 1$.

Proposition 7.9

Let $f, g \in F[x]$, then

$$\deg(fg) = \deg f + \deg g.$$

If f, g are monic, so is fg.

Proof. Trivial. Note this is also true when $f = 0$ or $g = 0$. □

Proposition 7.10

Let $f, g, h \in F[x]$, $f \neq 0$, then $fg = fh \implies g = h$.

Proof. Trivial. In fact this is saying $F[x]$ is an *integral domain*. □

Definition 7.11 (Polynomial map)**.** Let $\mathcal{A}$ be an F-algebra with unit, $f = \sum_{k=0}^{n} a_k x^k \in F[x]$. The map

$$f_{\mathcal{A}} : \mathcal{A} \to \mathcal{A}, \quad f_{\mathcal{A}}(\alpha) = \sum_{k=0}^{n} a_k \alpha^k,$$

where $\alpha^0 = 1$, is called the **polynomial map**.

When $\mathcal{A}$ is understood, we write f rather than $f_{\mathcal{A}}$ for simplicity.

Example 7.12

When $\mathcal{A} = F$, $f_F : F \to F$ is the polynomial function we learned in high school.

When $\mathcal{A} = F[x]$, $f_{F[x]}(g) = \sum_{k=0}^{n} a_k g^k$. In particular, $f_{F[x]}(x) = f$.

Some other choices of $\mathcal{A}$ includes $F^{n\times n}, \mathrm{Hom}(V, V)$.

Definition 7.13. Let $\mathcal{A}, \mathcal{A}'$ be F-algebras with unit. A map $\varphi : \mathcal{A} \to \mathcal{A}'$ is a **homomorphism** if

- φ is an F-linear map.
- $\varphi(\alpha\beta) = \varphi(\alpha)\varphi(\beta)$.
- $\varphi(1_{\mathcal{A}}) = 1'_{\mathcal{A}}$

If φ is also a bijection, we say φ is an **isomorphism** of algebras. In this case $\mathcal{A}$ and $\mathcal{A}'$ is **isomorphic**, written as $\mathcal{A} \cong \mathcal{A}'$.

Proposition 7.14

Let $\mathcal{A}$ be an F-algebra, the map

$$\phi_\alpha : F[x] \to \mathcal{A}, \quad f \mapsto f(\alpha)$$

is a homomorphism of algebras.

Proof. Check it yourself. ϕ_α is just "evaluate at α". In fact definition is chosen so that this property holds. □

7.2 Ideals of the polynomial ring

Definition 7.15. A subspace $M \subset F[x]$ is called an **ideal**, if $\forall f \in F[x]$, $g \in M$, $fg \in M$, i.e. $F[x]M \subset M$, or M *absorbs* multiplication.

Here $VW := \{\sum v_i w_i \mid v_i \in V, w_i \in W\}$.

Proposition 7.16

Let $M_1, M_2, \ldots, M_k \subset F[x]$ be ideals. Then $\sum_{i=1}^{k} M_i$ and $\bigcap_{i=1}^{k} M_i$ are ideals as well.

Example 7.17

Let $d_1, d_2, \ldots, d_k \in f[x]$. The ideal

$$(d_1, \ldots, d_k) := \left\{ \sum_{i=1}^{k} f_i d_i \mid f_i \in F[x] \right\}$$

is the ideal generated by $d_1, \ldots, d_k$. An ideal generated by one element is called a **principal ideal**.

Theorem 7.18

The polynomial ring $F[x]$ is a Principal Ideal Domain (PID), i.e. every ideal of $F[x]$ is a principal ideal. Moreover, the generator is unique up to a scalar. (i.e. there is a unique monic polynomial d s.t. $M = (d)$.)

Proof. This can be proved using division algorithm. □

Definition 7.19. Let $p_1, \ldots, p_k \in F[x]$ (not all zero), define their **greatest common divisor(gcd)** to be the monic generator of $(p_1, \ldots, p_k)$. (Here you get a hint why sometimes gcd is just written as $(a_1, \ldots, a_n)$)

Let $p_1, \ldots, p_k \in F[x] \backslash \{0\}$, their **least common multiple(lcm)** is the monic generator of the ideal $\bigcap_{i=1}^{k}(p_i)$, denoted by $\operatorname{lcm}(p_1, \ldots, p_k)$.

To see why this makes sense, just take a look at the definition of divisiblility in polynomial rings:

Definition 7.20. Let $f, d \in F[x]$, if $(f) \subset (d)$, we say f is divisible by d, denoted by $d \mid f$.

Remark 7.21 — You can check these definitions coincide with the elementary definitions in number theory. In fact, the above definitions work for any PID, you might learn it in abstract algebra.

Corollary 7.22 (Bezout's theorem)

Let $p_1, \ldots, p_k \in F[x]$ (not all zero), the followings are equivalent:

- $\gcd(p_1, \ldots, p_k) = 1$;
- $\exists u_1, \ldots, u_k \in F[x]$, such that $\sum_{i=1}^{k} u_i p_i = 1$.

We say $p_1, \ldots, p_k$ are **coprime** when they satisfy one of the above conditions.

Proof. Using the language of ideals, this is obvious. □

7.3 Unique factorization of polynomials

Definition 7.23. A polynomial $f \in F[x]$ is called **reducible**, if there exists $g, h \in F[x] \backslash F$ such that $f = gh$. Otherwise f is called **irreducible**.

A polynomial $p \in F[x]$ is called **prime** if for any $f, g \in F[x]$, $p \mid fg \implies p \mid f$ or $p \mid g$.

Proposition 7.24

In polynomial rings, prime $\iff$ irreducible.

Proof. Obviously prime $\implies$ irreducible.

For the other direction, suppose p irreducible and $p \mid fg$, WLOG $p \nmid f$.

Let $d = \gcd(p, f)$, then either $d = 1$ or $d = p$. Since $p \nmid f$, $d = 1$.

Thus $\exists u, v \in F[x]$, $up + vf = 1 \implies upg + vfg = g$, which is divisible by p. □

Remark 7.25 — This is saying PID $\implies$ UFD(Unique Factorization Domain). The conclusion is just the definition of UFD.

This property implies the following well-known theorem:

Theorem 7.26 (Unique factorization)

Let $f \in F[x]$ be a non-zero polynomial. Then there exists prime elements $p_1, \ldots, p_r \in F[x]$ and $e_1, \ldots, e_r \in \mathbb{N}, c \in F$ such that

$$f = c \prod_{i=1}^{r} p_i^{e_i}.$$

Moreover, this factorization is unique up to permutation.

Proof. The proof is essentially the same as the proof of Fundamental Theorem of Arithmetics, where the polynomial ring $F[x]$ are replaced with the integer ring $\mathbb{Z}$. □

Remark 7.27 — As the name suggests, this is true for any UFD.

The factorization provides another way of computing the gcd's and lcm's. (In exactly the same way you do with integers)

7.4 Roots of polynomials

Definition 7.28. Let $f \in F[x]$, $c \in F$, if $f(c) = 0$, we say c is a **root** or **zero** of f.

Proposition 7.29

c is a root of $f \iff (x-c) \mid f$.

Proof. Division algorithm gives $f = (x-c)q + r$, where $r \in F, q \in F[x]$. Since $f \mapsto f(c)$ is a homomorphism of F-algebras, so $f(c) = (c-c)q(c) + r = r$. □

Proposition 7.30

For any field F, the followings are equivalent:

- Every $f \in F[x] \backslash F$ has a root.
- Every prime polynomial $p \in F[x]$ has degree 1.

In this case, F is called an **algebraically closed** field.

Proof. If every polynomial of degree ≥ 1 has a root, take a monic prime polynomial $p \in F[x]$. p has a root $\iff (x-c) \mid p \implies p = x - c \implies \deg p = 1$.

If every prime has degree 1, for any $f \in F[x] \backslash F$, factor f into product of prime polynomials. It's clear that any polynomial of degree 1 has a root, so f must have a root. □

Theorem 7.31 (Fundamental Theorem of Algebra)

The complex number field $\mathbb{C}$ is algebraically closed.

Remark 7.32 — Here we won't prove this theorem, as it cannot be proved without analysis method, which is not the focus of this course.

We also commit the following result:

Theorem 7.33

Let F be any field.

- There exists an algebraically closed field $\overline{F}$ such that $F \subset \overline{F}$ and every $c \in \overline{F}$ is a root of some polynomials in $F[x]$.
- Such $\overline{F}$ is unique under isomorphisms of fields.

This field $\overline{F}$ is called the **algbraic closure** of F.

Example 7.34

$\overline{\mathbb{R}} = \mathbb{C}$, $\overline{\mathbb{Q}}$ is the set of algebraic numbers

Proposition 7.35

Let $f \in F[x]\backslash\{0\}$, then f has at most $\deg f$ roots.

Proof. Induction by n. The case $n = 0$ is obvious.

If f has a root c, then $f = (x - c)q$ for some $q \in F[x]$, $\deg q = \deg f - 1$. By induction hypothesis the conclusion follows. □

Corollary 7.36

Let $t_0, t_1, \ldots, t_n$ be distinct elements in F. The *Vandermonde* matrix

$$A = \begin{pmatrix} 1 & t_0 & \cdots & t_0^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & \cdots & t_n^n \end{pmatrix}$$

is invertible.

Proof. If $\alpha = (a_0, \ldots, a_n)^t \in F^{(n+1)\times 1}$ satisfies $A\alpha = 0$, then $t_0, \ldots, t_n$ are all roots of the polynomial $a_0 + a_1 x + \cdots + a_n x^n$, which implies $a_0 = \cdots = a_n$. □

Recall that a polynomial function is a map $\phi : F \to F$ such that exists $f \in F[x]$, $\phi = f_F$.

Proposition 7.37

The map $F[x] \to \{\text{polynomial function on } F\}, f \mapsto f_F$ is a surjective homomorphism of F-algebras. Moreover if F is infinite, this homomorphism is an isomorphism.

Proof. Clearly it's a surjective homomorphism.

When F is infinite, if $f_F = 0$, it has infinitely many zeros, so f must be the zero polynomial. So the map has a trivial kernel $\implies$ injectivity. □

Remark 7.38 — When F is finite, the above homomorphism is always not an isomorphism, as there are only finitely many maps $F \to F$, but $F[x]$ is always infinte.

Definition 7.39. Let $c \in F$ be a root of $f \in F[x]$. Let $m(c, f) := \max\{k \geq 1 \mid (x - c)^k \mid f\}$ be the **multiplicity** of c in f.

If $m(c, f) = 1$, we say c is a **simple root**, otherwise c is a **multiple root**.

Definition 7.40. Let $f = \sum_{k=0}^{n} a_k x^k \in F[x]$, define its **formal derivative** to be $f' := Df := \sum_{k=1}^{n} k a_k x^{k-1}$.

We can easily check that some properties in calculus carry over to formal derivatives, like the product rule and chain rule.

Proposition 7.41

Let $c \in F$ be a root of f. c is a multiple root $\iff f'(c) = 0$.

Proof. Since c is a root, $f = (x-c)g$ for some $g \in F[x]$.

So $f' = g + (x-c)g'$ has the root $c \iff (x-c) \mid g \iff (x-c)^2 \mid f$. □

Proposition 7.42

Let $f \in F[x] \backslash F$.

- If f, f' are coprime, then f has no multiple roots. (Note that coprime and formal derivatives do not depend on the base field!)
- If F is algebraically closed, the reverse also holds.

Proof. Let $d = \gcd(f, f')$.

If f has a multiple root c, by the previous proposition, $x - c \mid d \implies d \neq 1$.

On the other hand, when $d \neq 1$, d has a root c in F. Hence $f(c) = 0, f'(c) = 0 \implies c$ is a multiple root of f. □

Proposition 7.43

Let F be a field with $\mathrm{char}(F) = 0$. Let $f \in F[x] \backslash F$, $c \in F$. We have

$$m(c, f) = \min\{k \geq 0 \mid (D^k f)(c) \neq 0\}.$$

Proof. We can check $\deg f \in \{k \geq 0 \mid (D^k f)(c) \neq 0\}$, as

$$D^{\deg f} f = (\deg f)! c,$$

where c is the leading coefficient of f. So the minimum is well-defined.

Remark 7.44 — For $\mathrm{char}(F) = p$, this set can be empty, for example when $f = x^p$.

Apply the chain rule repeatedly, we get $\forall k \leq m := m(c, f)$, $(x-c)^{m-k} \| D^k f$. □

Proposition 7.45

Let $f \in F[x] \backslash F$,

- If f and f' are coprime, then f has no multiple factors in $F[x]$(i.e. doesn't exist prime polynomial $p \in F[x]$ s.t. $p^2 \mid f$).
- If $\mathrm{char}(F) = 0$, the inverse also holds.

Remark 7.46 — When $\mathrm{char}(F) = p$, $F = \mathbb{F}_p(t), f = x^p - t$ is a counter example for (2).

Proof. If f has a multiple factor p, suppose $f = p^2 g$. Then $f' = 2pp'g + p^2 g'$ has a prime factor $p \implies p \mid \gcd(f, f')$.

On the other hand, if $\gcd(f, f') \neq 1$, take one of its prime factor p. Suppose $f = pg$, $f' = p'g + pg' \implies p \mid p'g$. Since $\mathrm{char}(F) = 0$, $p' \neq 0$, hence $p \nmid p'$.

Thus $p \mid g \implies p^2 \mid f$. □

7.5 Lagrange interpolation

Theorem 7.47

Let $n \geq 1$, $t_0, t_1, \ldots, t_n \in F$ are distinct elements. For any $c_0, c_1, \ldots, c_n \in F$, there exists a unique polynomial $f \in F[x]$ such that

$$\deg f \leq n, \quad f(t_i) = c_i, i = 0, 1, \ldots, n.$$

Proof. As we learned from high school olympiads,

$$f = \sum_{i=0}^{n} c_i \prod_{j \neq i} \frac{x - t_j}{t_i - t_j}$$

satisfies the above condition. This is known as **Lagrange interpolation formula**.

For the uniqueness part, just notice the fact that a nonzero polynomial with degree $\leq n$ has at most n roots. □

Well, this is just the same proof you might learned in high school. It seems the mysterious formula somehow jumps out of thin air. To motivate the formula, let's see another proof first.

Alternate proof using vector spaces. Let $V = \{f \in F[x] \mid \deg f \leq n\}$ be a vector space.

Note that $\dim V = n + 1$, for $\{1, x, \ldots, x^n\}$ is a basis of V.

Consider the linear map $T : V \to F^{n+1}, f \mapsto (f(t_0), \ldots, f(t_n))$. We claim that T is injective, as $f \in \ker T \implies f$ has $n + 1$ distinct roots $\implies f = 0$.

By dimensional reasons, T must be a linear isomorphism, which gives the desired result. □

Let $L_i \in V^*$ such that $L_i(f) = f(t_i), i = 0, 1, \ldots, n$. Then $T = (L_0, L_1, \ldots, L_n)$.

T isomorphism $\implies \{L_i\}$ is a basis of V^*. In fact,

$$\ker T = \bigcap_{i=0}^{n} \ker L_i.$$

If we take the dual basis of $\{L_0, \ldots, L_n\}$, say $\{p_0, \ldots, p_n\}$, then $p_i(t_j) = \delta_{ij}$. From this we can compute $p_i = \prod_{j \neq i} \frac{x - t_j}{t_i - t_j}$.

Thus for any $f \in V$, we know that $T(f) = (c_0, \ldots, c_n)$, so we must have $f = \sum_{i=0}^{n} c_i p_i$.

Moreover, if we take $f = x^j$ in the above formula, we discover that the matrix between $\{1, x, \ldots, x^n\}$ and $\{p_0, \ldots, p_n\}$ is precisely Vandermonde matrix:

$$(1, x, \ldots, x^n) = (p_0, p_1, \ldots, p_n) \begin{pmatrix} 1 & t_0 & \cdots & t_0^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & \cdots & t_n^n \end{pmatrix}$$

7.6 Bonus section

7.6.1 Rational function fields

The goal of this section is to construct the rational function field over F in details, namely $F(x)$. The following process applies for any integral domain R, and $F(x)$ corresponds to the fraction field $\mathrm{Frac}(R)$.

Consider the set

$$X := F[x] \times (F[x] \backslash \{0\}) = \{(f, g) \mid f, g \in F[x], g \neq 0\}.$$

Next we set an equivalence relation on X: $(f_1, g_1) \sim (f_2, g_2) \iff f_1 g_2 = f_2 g_1$. (Check this is indeed an equivalence relation)

Let $F(x) := X/\sim := \{\text{all the equivalence classes}\}$, and denote the class in which (f, g) lies by $\frac{f}{g}$, i.e. $\frac{f}{g} = \{(p, q) \in X \mid fq = gp\}$.

Now we define addition and multiplication on $F(x)$:

$$\frac{f_1}{g_1} + \frac{f_2}{g_2} := \frac{f_1 g_2 + f_2 g_1}{g_1 g_2}, \quad \frac{f_1}{g_1} \cdot \frac{f_2}{g_2} := \frac{f_1 f_2}{g_1 g_2}.$$

We can check that they are well-defined, and $F(x)$ forms a field under these two operations. (The computation is again left out as it's really tiring) This field is called the **rational function field over** F. (Don't be confused with its name! The definition has nothing to do with functions!)

There's a natural embedding $F[x] \hookrightarrow F(x)$ by $f \mapsto \frac{f}{1}$, this is clearly a ring homomorphism. So $F[x]$ is a subring of $F(x)$, F is a proper subfield of $F(x)$. This tells us any field is a proper subfield of a larger field.

7.6.2 Perfect fields

Recall that for $f \in F[x] \backslash F$,

$$f \text{ has no multiple roots in } \overline{F} \iff \gcd(f, f') = 1 \underset{\mathrm{char} F = 0}{\rightleftharpoons} f \text{ has no multiple factors in} F[x]$$

Proposition 7.48

Let F be a field, the followings are equivalent:

- For any $f \in F[x]$, if f has no multiple factors, then f has no multiple roots in $\overline{F}$.
- Any prime $p \in F[x]$ has no multiple roots in $\overline{F}$.

In this case, F is called a **perfect field**.

Proof. Trivial by the unique factorization of polynomial rings. □

Proposition 7.49

Some sufficient conditions for perfect fields:

- Fields with characteristic zero;
- Algebraically closed fields;
- Finite fields.

Proof. Here we only prove $F = \mathbb{F}_p$ is perfect. Let $q = \sum_{k=0}^{n} a_k x^k$ be a prime polynomial in $F[x]$, if q has multiple roots in $\overline{F}$,

$$\gcd(q, q') \neq 1 \implies q \mid q' \implies q' = 0.$$

This gives $a_k = 0$ for all $p \nmid k$. Hence there exists $g \in F[x]$ such that $q(x) = g(x^p) = (g(x))^p$, contradicts the assumption that q is a prime. □

Example 7.50

Example of imperfect fields: $\mathbb{F}_p(t)$.

Note that $x^p - t \in \mathbb{F}_p(t)[x]$ is a prime polynomial (by Eisenstein criterion). Let $c \in \overline{\mathbb{F}_p(t)}$ be a root of $x^p - t$, we claim that $x^p - t = (x - c)^p$:

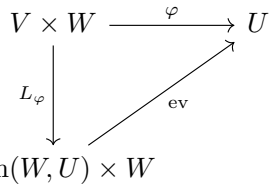
$$(x - c)^p = \sum_{k=0}^{p} \binom{p}{k} x^k (-c)^{p-k} = x^p - c^p = x^p - t.$$

Thus c is a multiple root of $x^p - t$.

A Where does tensor product come from?

After studying bilinear functions, we'll naturally come to bilinear maps, that is $\varphi : V \times W \to U$, where U, V, W are vector spaces.

We want to classify all the bilinear maps, notices that φ can be determined by $L_\varphi : V \to \mathrm{Hom}(W, U)$.



But this is not good enough, because it depends on U, and you couldn't tell whether φ is right-degenerated from just L_φ. We want to find a way to describe bilinear maps which is symmetric about V, W and independent of U.

Now let's think about what φ really is. First it sends every pair $(v, w) \in V \times W$ to a vector in U. It also satisfies the condition that it's linear for both V and W. And that's all!

So we're going to construct a space with above structure:

Let $S = V \times W$, and we're not interest in the structure of vector space in S, so we *forget* the structure in S, i.e. view S as a *set* instead of a vector space.

Next, we add the structure we want into the set S. Because $\mathrm{im}\,\varphi$ is a linear space, and we can write something like $\varphi(v_1, w_1) + \varphi(v_2, w_2)$ or $c\varphi(v, w)$, so we'll consider the vector space freely generated by S, namely F^S(to view S as a basis, and generate other elements by linear combinations).

According to the bilinear conditions, we want to make $(v_1, w)+(v_2, w) = (v_1+v_2, w), c(v, w) = (cv, w)$ and so on. This is actually forcing elements like $(v_1, w) + (v_2, w) - (v_1 + v_2, w)$ equal to 0. To achieve this, we just make use of quotient spaces:

Let

$$S_0 = \mathrm{span}\{(v_1, w)+(v_2, w)-(v_1+v_2, w), c(v, w)-(cv, w), (v, w_1)+(v, w_2)-(v, w_1+w_2), c(v, w)-(v, cw)\}$$

where v_i, w_i can be any elements in V, W, respectively.

Finally we come to the space S/S_0, which has all the properties we want. We call this **tensor prodouct** of V and W, written as $V \otimes_F W$.

In other words,

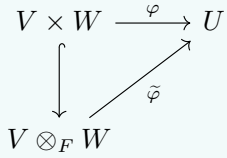
Definition A.1 (Tensor product)**.** The tensor product $V \otimes_F W$ is a vector space whose *generators* are $v \otimes w$ for all $v \in V, w \in W$, with the relations:

- $v_1 \otimes w + v_2 \otimes w = (v_1 + v_2) \otimes w$;
- $v \otimes w_1 + v \otimes w_2 = v \otimes (w_1 + w_2)$;
- $\forall c \in F, c \cdot (v \otimes w) = (cv) \otimes w = v \otimes (cw)$.

With this we get a description of bilinear maps:

Proposition A.2

For every bilinear map $\varphi : V \times W \to U$, there exists a map $\widetilde{\varphi} : V \otimes_F W \to U$ such that the following diagram commutes.



Where $V \times W \hookrightarrow V \otimes_F W$ is given by $(v, w) \mapsto v \otimes w$.

Proof. Just let $\widetilde{\varphi}(v \otimes w) = \varphi(v, w)$. Obviously this is a linear map. □