高等代数 (Polynomial).

Date

Chapter 1 polynomial (多项式)

1. Fields (数域).

$\mathbb{Q} = \{\frac{a}{b} \mid a, b \in \mathbb{Z} \wedge b \neq 0\}$ (rational field). $\mathbb{Z}$ $\mathbb{R}$ $\mathbb{N}$ $\mathbb{F}$

$\mathbb{C}$ 复数 (complex number).

Def1 Let F be a subset of $\mathbb{C}$. If $0, 1 \in F$, and if F is closed under addition, subtraction, multiplication and division, namely, for any a and b, there hold $a+b, a-b, a \times b, a \div b$ (when $b \neq 0$) $\in F$. then F is called a number field. 环：整数环、多项式环 (对除法不封闭)

Theorem If P is a field. $\mathbb{Q} \subseteq P$

检测 1) $a, b \in F \Rightarrow ab \in F$, $a+b \in F$
2) $a \in F \Rightarrow a^{-1} \in F$
3) $a \in F \Rightarrow -a \in F$

proof: since P is a field. $1 \in P$.

$n = \underbrace{1+1+\cdots+1}_{n \text{ times}} \in P$. $\mathbb{N}_+ \subseteq P$.

further we have

$0 - n = -n$. $\mathbb{N}_- \subseteq P$.

thus $\mathbb{Z} \subseteq P$.

For any $a, b \in \mathbb{Z}$. $a, b \in P$. with $b \neq 0$. we have $\frac{a}{b} \in P$.

Hence $\mathbb{Q} \subseteq P$. □

2. The Polynomial.

In the section we always use $\mathbb{F}$ to denote a fixed field, and use x to denote a indeterminate (不确定的)

Def2 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = \sum_{i=0}^{n} a_i x^i$

a polynomial over $\mathbb{F}$, where $a_n, a_{n-1}, \cdots a_0$ called coefficient of it.

a_0. constant (不变的). term.

Def3 If n is the largest positive integer such than $a_n \neq 0$.

n is called the degree of $f(x)$. $n = \deg f$

$f(x) = 0$. no degree (assignment)

$a_n x^n$ leading term a_n leading coefficient.

Monic Polynomials. leading coefficient $= 1$. (Scalar) Zone Polynomial $f(x) = 0$.

Zone Polynomial belongs to the Constant Polynomials

linear Polynomials. $\deg = 1$

Polynomial Operations.

proposition 1 (性质)　　　　　　　　　　　　proposition 命题.

1) addition is commutative (交换) and associative

2) multiplication ↗

3) multiplication is distributive with respect to addition

4) cancellation law:

If $f(x)\cdot g(x) = f(x)\cdot h(x)$ with $f(x)\neq 0 \iff g(x)=h(x)$

Denote by $F[x]$ the set of all polynomial over F. $F[x]$ is a ring of one indetermination

$\boxed{\text{Def 3}}$ A polynomial $f(x)$ (in $F[x]$) is called a unit if there exists $g(x)\in F[x]$. $f(x)\cdot g(x)=1$.

Notice that every unit in $F[x]$ is a nonzero constant

3. Divisibility (不随数域扩大而改变).

Def 5 $f(x), g(x)$ in $F[x]$, $\exists\, h(x)$. $f(x) = g(x)\cdot h(x) \Rightarrow g(x) \mid f(x)$

(1) The zero polynomial is divisible by any nonzero polynomial.

(2) Any polynomial is divisible by itself

(3) Any polynomial is divisible by any 0 degree polynomial (nonzero constant).

(4) If $f(x) \mid g(x) \wedge g(x) \mid f(x)$. then $f(x) = c\,g(x)$. ~~with~~ c is a nonzero constant. $f(x), g(x)$ 相伴 (associates of each)

(5) If $f(x) \mid g(x)$. $g(x) \mid h(x)$. then $f(x) \mid h(x)$

(6) If $f(x) \mid a(x)$. $f(x) \mid b(x)$. Then $f(x) \mid a(x) \pm b(x)$.

More generally, for any $s(x), t(x) \in F[x]$. $f(x) \mid \underline{s(x)a(x)+t(x)b(x)}$ ⇑ a combination of $a(x)$ and $b(x)$

余式 remainder 商 quotient 因式 factor

$\boxed{\text{Thm 1}}$ $f(x) = q(x)g(x) + r(x)$

a unique pair of $q(x)$. $r(x)$ where $\deg r(x) < \deg g(x)$ or $r(x)=0$

$\boxed{\text{proof}}$ i) If $f(x)=0$. then $q(x)=r(x)=0$

ii) now assume $\deg f = n$. $\deg g = m \geq 0$.

$f(x) = \sum_{i=0}^{n} a_i x^i \quad g(x) = \sum_{j=0}^{m} b_j x^j$

If $n<m$. then let $q(x)=0$. $r(x)=f(x)$

If $n\geq m$. then let $f_1(x) = f(x) - a_n b_m^{-1} x^{n-m} g(x)$

归纳起点是 $n=m$

If $f_1(x)=0$. $f(x) = a_n b_m^{-1} x^{n-m} g(x)$ (1)

If $f_1(x)\neq 0$. $\deg f_1 < \deg f$.

By the induction assumption there exist $q_1(x) \in F[x]$, $r(x)\in F(x)$ that.

$f_1(x) = q_1(x)g(x) + r(x)$ (2)

combining (1) ∧ (2) $f(x) = [q_1(x) + a_n b_m^{-1} x^{n-m}]\, g(x) + r(x)$

that's the desire form.

高等代数 (Polynomials)

Date . .

proof of uniqueness

we assume that.

$f(x) = q_1(x)g(x) + r_1(x)$.

$\quad = q_2(x)g(x) + r_2(x)$.

$\Rightarrow [q_1(x) - q_2(x)]g(x) = r_2(x) - r_1(x)$

if $q_1(x) - q_2(x) \neq 0$.

then $\partial\{[q_1(x) - q_2(x)]g(x)\} \geq \partial\{g(x)\}$

if $r_2(x) - r_1(x) \neq 0$

then $\partial(r_2(x) - r_1(x)) < \partial(g(x))$.

that causes an contradiction.

hence $\begin{cases} q_1(x) - q_2(x) = 0 \\ r_1(x) - r_2(x) = 0. \end{cases}$

$\therefore q_1(x) = q_2(x)$. $r_1(x) = r_2(x)$

$q(x). r(x)$ is a unique pair.

Thm 2. complete induction

1) ~~Since~~ base case $P(n_0)$ is true.

2). for all $m \geq n$, if $P(k)$ is true for $n_0 \leq k < m$ then $P(m)$ is true.

4. Greadest common divisor

Def 6. Let $f(x)$ and $g(x)$ be in $F[x]$ and let $\varphi(x) \in F[x]$.

If $\varphi(x) | f(x)$. $\varphi(x) | g(x)$. then $\varphi(x)$ is called a common divisor

Furthermore, if a polynomial $d(x) \in F[x]$, is a common divisor of f and g and every common divisor of f and g divides $d(x)$, then $d(x)$ is called a greatest common divisor (g.c.d). of f and g

Lemma 1 Suppose $f(x) = q(x)g(x) + r(x)$. with $g(x) \neq 0$.

Then the list of common divisor of f and g coincides with the list of common divisors of g and r. ($f \cdot r$ 不能换 $g \cdot r$ / $f \cdot g$).

proof: Suppose $\varphi(x) | f(x)$ and $\varphi(x) | g(x)$ then $\varphi(x) | f(x) - q(x)g(x)$.

i.e $\varphi(x) | r(x)$ Thus $\varphi(x)$ is a common divisor of g and r

Eudidean Algorithm. (带余除法). ~~Giv~~ Given 2 polynomials $f(x)$ and $g(x)$, $g(x) \neq 0$. divide g into f. then the remainder g, then that remainder into the previous remainder, etc. This process end with 0 remainder. Symbolically

$f(x) = q_1(x)g(x) + r_1(x)$

$g(x) = q_2(x)r_1(x) + r_2(x)$

...

$r_{n-1}(x) = q_{n+1}(x)r_n(x) + 0$.

Theorem 1 In Euclidean Algorithm, the last nonzero remainder $r_n(x)$ is a g.c.d. of $f(x)$ and $g(x)$

Fact Two common divisors of $f(x)$ and $g(x)$ differ with a nonzero constant

we use (f,g) or $(f(x),g(x))$ to denote the g.c.d with monic polynomial

Thm2 (Bézout) For any polynomials $f(x)$ and $g(x)$ in $F[x]$, there exists a greatest common divisor $d(x)$ in $F[x]$, and $d(x)$ can be express as a linear combination of $f(x)$ and $g(x)$ namely, there exists $u(x)$ and $v(x)$ in $F[x]$ such that $d(x) = u(x)f(x) + v(x)g(x)$. (Bézout's identity)

佰祖恒等式.

$r_s = (1 + q_s(x)\, q_{s-1}(x))\, r_{s-2}(x) - q_s(x)\, r_{s-3}(x)$.

$r_s = r_{s-2}(x) - q_s(x) * r_{s-1}(x)$

逐个消去. $r_s(x)$, $r_{s-1}(x)$, $r_{s-2}(x)$ ... $r_1(x)$, 再并项.

Def 2. If any g.c.d. of $f(x)$ and $g(x)$ is a nonzero constant, then $f(x)$ and $g(x)$ are called coprime or relatively prime

Remark: $f(x)$ and $g(x)$ are coprime $\overset{\text{if and only if}}{\Longleftrightarrow}$ $(f(x), g(x)) = 1$.

Thm 3 If $(f,g)=1$, $\exists\, u(x), v(x)$, such that $u(x)f(x) + v(x)g(x) = 1$.

推论

Corollary 1 If $(f,g)=1$, and $f(x) \mid g(x) \cdot h(x)$, then $f(x) \mid h(x)$

2 If $f_1(x) \mid g(x)$, $f_2(x) \mid g(x)$ and $(f_1, f_2) = 1$, then $f_1(x) f_2(x) \mid g(x)$

proof: By $f_1(x) \mid g(x)$, $\exists\, h_1(x)$ s.t. $q(x) = f_1(x) h_1(x)$ (1)

Now $f_2(x) \mid f_1(x) h(x)$ and $(f_1(x), f_2(x)) = 1$, thus $f_2(x) \mid h(x)$

There exists $h_2(x)$ s.t. $h(x) = f_2(x) h_2(x)$

This gives $g(x) = f_1(x) f_2(x) h_2(x)$ and thus $f_1(x) f_2(x) \mid g(x)$

Thm4. If $(f(x), g(x)) = 1$, and $f(x) \mid g(x)h(x)$, then $f(x) \mid h(x)$.

$u(x)f(x) + v(x)g(x) = 1$.

$h(x)u(x)f(x) + h(x)v(x)g(x) = h(x)$

$f(x) \mid LHS$

Generalized to more than two polynomials,

$\boxed{\text{Def}}$ Let $f_i(x)$ $(i\in[1,s]$, $s\geq 2)$. $d(x)\in F[x]$, if (1) $d(x)\mid f_i(x)$, for each i (2) Each common divisor of f_i's divides $d(x)$. then $d(x)$ is a g.c.d of f_i's.

use $(f_1, f_2, \dots f_s)$ = monic g.c.d.

$(f_1, f_2, \dots f_s) = ((f_1, f_2, \cdots, f_{s-1}), f_s)$.

If $(f_1, f_2, \dots f_s)=1$. then $f_1, f_2 \dots f_s$ are called coprime

$\exists u_i(x)$, $i\in[1,s]$. s.t. $u_1 f_1 + u_2 f_2 + \cdots u_s f_s = 1$.

5. Factorization

$\boxed{\text{Def 8}}$ Let $p(x)$ be a polynomial over $F[x]$ of degree ≥ 1. If $p(x)$ can't be factored a product of two polynomials of degree ≥ 1, then $p(x)$ is called irreducible polynomial

reducible: $p(x) = u(x)\cdot v(x)$ $\deg p, u, v \geq 1$.

$\deg u < \deg p$ $\deg v < \deg p$.

Remark (1) Whether ~~the~~ a polynomial is irreducible depends on the field.

(2) An irreducible polynomial ~~are~~ is only divisible by units or associates of itself.

(3). Let $P(x)$ and $f(x)$ be in $F[x]$, $p(x)$ is irreducible, then $(p(x), f(x))=1$ or $p(x)\mid f(x)$

Thm 5. Let $p(x), f(x), g(x)$ be in $F[x]$, $p(x)$ is irreducible. If $p(x)\mid f(x)\cdot g(x)$, then $p(x)\mid f(x)$ or $p(x)\mid g(x)$

Remark If $p(x)$ is irreducible, and $p(x)\mid f_1(x) f_2(x) \dots f_s(x)$, $s\geq 2$, then $p(x)$ divides some $f_i(x)$

Thm. Every polynomial of degree ≥ 1 over F is irreducible or factor into a product of irreducible polynomials.

proof Let $\deg f = n$. Use induction on n (Second induction.)

For $n=1$, $f(x)$ is irreducible. The theorem holds for $n=1$. Now we may assume that we have proved the theorem for $1\leq k<n$

Consider the case n. If $f(x)$ is irreducible there is nothing to say. If not $f(x)$ can be factored into $f(x)=a(x)b(x)$ where $1\leq \deg a<n$, $1\leq \deg b<n$

By the induction assumption, $a(x)$ and $b(x)$ can factor into a product of irreducible polynomials. Hence $f(x)$ factors into a product of irreducible polynomials So the thm is for n. □

Thm (the uniqueness of factorization).

In $f(x) \in F[x]$. if $f(x)=p_1(x)p_2(x)\cdots p_s(x) = q_1(x)q_2(x)\cdots q_t(x)$ (1) are two factorizations of the polynomial $f(x)$, then $s=t$. and there is a one to one correspondence between the factors $p_1(x), \dots p_s(x)$ and $q_1(x) \dots q_t(x)$. where if $p_i(x)$ correspond ~~ence~~ with $q_j(x)$. then $p_i(x)$ and $q_j(x)$ are associates.

proof. Induction on s. If $s=1$ then $f(x)=p_1(x)$ is irreducible and so $t=1$.

and $f=p_1(x)=q_1(x)$

suppose that we have proved the theorem for $s-1$

Now consider s. By (1). $p_1(x) \mid q_1(x)q_2(x)\cdots q_t(x)$.

Hence, $p_1(x)$ divides some $q_i(x)$ $(1 \le i \le t)$

W.l.g. we assume $p_1(x) \mid q_1(x)$.

As $q_1(x)$. $p_1(x)$ are both irreducible, then $p_1(x) = c_1 q_1(x)$ For $c \in F(x) \neq 0$

subsitating this in (1). by cancellation law, we have

$c_1 p_2(x)p_3(x)\cdots p_s(x) = q_2(x)q_3(x)\cdots q_t(x)$

since the number of irreducible polynomials on the left-hand side is $s-1$

our inductive assumption applies. and thus $s-1=t-1$, namely, $s=t$.

and there is one to one correspondence between the factors $p_2 \sim p_s$

and $q_2 \sim q_s$. This together with $p_1 = c_1 q_1$ shows that $s=t$

Remark2. We can write the irreducible factorization of $f(x)$ of degree ≥ 1

in $F[x]$ in exponential form. $f(x)=p_1^{e_1}p_2^{e_2}\cdots p_r^{e_r}$ (2)

where $p_1, p_2, \dots p_r$ are dinstinct irreducible polynomials. $e_i \geq 1$.

(1) If any $e_i > 1$. we shall say p_i is <u>a multiple factor</u> of f.

If $e_i = 1$. we shall say p_i is <u>a simple factor</u> of f

(2) We write (2) as. $f(x) = a\, q_1^{e_1} q_2^{e_2} \cdots q_r^{e_r}$.

where $q_1 \sim q_r$ is monic irreducible polynomials. a is nonzero constant. → leading coefficient.

such a factorization is called a <u>normalized factorization</u> of $f(x)$

高等代数 (Polynomials)

Date

Def 9. An irreducible polynomial $p(x)$ is called a factor of the polynomial $f(x)$ of multiplicity k. if $p^k(x) \mid f(x)$. and $p^{k+1}(x) \nmid f(x)$ where k is a positive integer. If $k=1$. we say $p(x)$ is a simple factor of $f(x)$. If $k>1$, we say $p(x)$ is a multiple factor of $f(x)$.

Def. The formal derivate (形式导数) of the polynomial $f(x)=\sum_{i=0}^{n} a_i x^i$ is the polynomial $f'(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \cdots + a_1$

We also use the notation $Df = f'$ We have the higher formal dirivatives.

$$Df' = f'' \qquad f^{(3)} = Df'' = D^3 f.$$

(1) $(c f(x))' = c f'(x)$ (2) $(f(x)+g(x))' = f'(x) + g'(x)$

(3) $(f(x)\cdot g(x))' = f'(x)g(x) + g'(x)f(x)$ (4) $(f^m(x))' = m f^{m-1}(x) f'(x)$

Fact. If $\deg f(x) = n \geq 1$. then

(1) $\deg f' = n-1$ (2) $f^{(n)}(x) = n!\, a_n$ (3) $f^{(n+1)}(x) = 0$

Thm 6 Let $p(x)$ be an irreducible factor of $f(x)$ of multiplicity k with $k \geq 1$.

The $p(x)$. is a factor of $f'(x)$ of multiplicity $k-1$. (反推不成立. $f(x)$可加任意常数)

In particular. if $k=1$. then $p(x) \nmid f'(x)$.

proof: $\Rightarrow f'(x) = p^{k-1}(\cdots)$ $\Rightarrow p^{k}(x) \nmid (\cdots)$

Corollay 1 If $p(x)$ is a multiple irreducible factor of $f(x)$ of ~~multipdication~~ multiplication k $(k \geq 1)$ then $p(x)$ is a factor of $f(x)$. $f'(x)$, $f^{(k-1)}(x)$, but $p(x)$ is not a factor of $f^{(k)}(x)$

corollay 2 An ~~re~~irreducible polynomial $p(x)$ is a multiple factor of $f(x)$ if and only if $p(x)$ is a common factor of $f(x)$ and $f'(x)$.

Corollay 3 $f(x)$ has no multiple factor if and only if $f(x)$ and $f'(x)$ are coprime.

Remark Consider the normalized factorization of $f(x)$: $f(x) = a p_1^{r_1}(x) p_2^{r_2}(x) \cdots p_s^{r_s}(x)$ where $p_1(x), p_2(x), \ldots p_s(x)$. are distinct irreducible factors of $f(x)$. Then for each i the highest power of $p_i(x)$ that divides $f'(x)$ is $p_i^{r_i-1}(x)$ and so the greatest common divisor of $f(x)$ and $f'(x)$ is $(f(x), f'(x)) = p_1^{r_1-1}(x) p_2^{r_2-1}(x) \cdots p_s^{r_s-1}(x)$.

$$\frac{f(x)}{(f(x), f'(x))} = a p_1(x) p_2(x) \cdots p_s(x) \rightarrow \text{the square-free part of } f(x) \text{ (无重因式)}$$

求重根 $(f(x), f'(x)) = (f'(x), f(x) - f'(x))$.

Let $f(x) = \sum_{i=0}^{n} a_i x_i$, $\alpha \in E$. $f(\alpha) = a_n\alpha^n + \cdots + a_1\alpha + a_0 \in E$.

which is called the value of $f(x)$ at $x=\alpha$. Hence. we may define a polynomial function $\mathfrak{f} = f = \sum_{i=0}^{n} a_i x^i$; $\mathfrak{f}: E \to E$ by letting for every $\alpha \in E$

Therefore. $\mathfrak{f}$ is determined by the polynomial $f(x)$. By abuse of notation, we also denote $\mathfrak{f}$ by f sometimes.

Def 1. If $f(\alpha) = 0$ then α is said to be a root of $f(x)$

Thm 1 (Remain Theorem). When dividing $f(x)$ by $x-\alpha$. then the remainder is $f(\alpha)$.

proof: By Euclidean algorithm, there exist $q(x)$ and a number r.

s.t. $f(x) = q(x)(x-\alpha) + r \Rightarrow f(\alpha) = r$

Corol α is a root of $f(x)$ if and only if $x-\alpha \mid f(x)$

Thm 2. (D'Alembert's) Let $f(x)$ be a polynomial of degree n, $n \geq 1$.

Then $f(x)$ has at most n roots

$f(x) = (x-\alpha_1)^{k_1}(x-\alpha_2)^{k_2}\cdots(x-x_n)^{k_n} g(x)$ $\deg f n = \sum_{k=1}^{n} k_i + \deg g$ $\therefore$

Def (多项式相等). $f(x) = \sum_{i=0}^{n} a_i x^i$ $g(x) = \sum_{j=0}^{m} a_j x^j$ are equal if and only if the coefficients of each power of x are equal.

But two functions $f(x)$ and $g(x)$ defined on F are equal if and only if $f(\alpha) = g(\alpha)$, $\forall \alpha \in F$. ($f(x) \neq g(x)$ as polynomials $\Rightarrow$ $f(x) \neq g(x)$ as functions. in $F[x]$)

Thm 3 Let $f(x)$ and $g(x)$ be in $F[x]$ of degree not exceeding n. If there exist $n+1$ distinct numbers. $\alpha_1, \ldots, \alpha_{n+1}$ in $F[x]$. s.t. $f(\alpha_i) = g(\alpha_i)$, $1 \leq i \leq n+1$

Then $f(x) = g(x)$ as polynomials.

proof: Let $h(x) = f(x) - g(x)$. Then $h(x) = 0$ or $\deg h \leq n$

By assumption. $h(\alpha_i) = f(\alpha_i) - g(\alpha_i) = 0$, $1 \leq i \leq n+1$

That is, $h(x)$ has $n+1$ distinct roots. This is impossible by D'Alembert Theorem

Hence, $h(x) = 0$ and such that $f(x) = g(x)$ (同时说明不同多项式定义不同多项式函数.

8. Factorization for ~~com~~ Complex polynomials.

(A) over $\mathbb{C}$ (coefficient)

Fundamental Theorem of Algebra: Every polynomial in $\mathbb{C}[x]$ of degree ≥ 1 has a root in $\mathbb{C}$.

—复系数多项式域分解.

Thm2 Every polynomial $f(x)$ in $F[x]$ of degree ≥ 1 can factor into a product of polynomials of degree 1.

Fact: A polynomial $f(x)$ in $\mathbb{C}[x]$ of degree $n \geq 1$ has exactly n roots in $\mathbb{C}$

(B) over $\mathbb{R}$. (coefficient)

proposition 1. Let $f(x)$ be a polynomial over $\mathbb{R}$. If ~~z~~ $=a+bi$ is a root of $f(x)$

Then $\bar{z}=a-bi$ is a root of $f(x)$.

($\alpha, \beta \in \mathbb{C}$ $\overline{\alpha+\beta}=\bar{\alpha}+\bar{\beta}$, $\overline{\alpha\beta}=\bar{\alpha}\cdot\bar{\beta}$, $\overline{\alpha^n}=(\bar{\alpha})^n$)

proposition 2. If $f(x)=x^2+bx+c$ is a real polynomial of degree 2 then $f(x)$ is irreducible if and only of $b^2-4c<0$

—实系数多项式实数域分解

Thm3. No irreducible polynomial over $\mathbb{R}$ of degree >2.

Let $f(x)$ be in $\mathbb{R}[x]$ of degree >2.

Assume that $f(x)$ has no real roots. Let $\alpha=a+bi$ $(b\neq 0)$ be a nonreal complex root of $f(x)$. Thus $b\neq 0$. Let $p(x)=(x-\alpha)(x-\bar{\alpha})=x^2-2ax+a^2+b^2$, $p(x)\in\mathbb{R}[x]$

We have $f(x)=q(x)p(x)+r(x)$ for some $q(x), r(x)\in\mathbb{R}[x]$. $r(x)=0$ or $\deg r<\deg p$

write $r(x)=u+vx$. $(u, v\in\mathbb{R})$

Then $0=f(\alpha)=q(\alpha)p(\alpha)+r(\alpha)$

i.e. $r(\alpha)=u+v\alpha=u+v(a+bi)=(u+va)+(bv)i=0. \overset{b\neq 0}{\Rightarrow}\begin{cases}v=0\\u=0\end{cases}\Rightarrow r(x)=0.$

Thm3' Every polynomial over $\mathbb{R}$ factors into a ~~poly~~ product of polynomials of degree 1 and irreducible polynomials of degree 2. (in real numbers)

Remark Every polynomial $f(x)$ in $\mathbb{R}[x]$ of degree ≥ 1 has a normalized factorization

$$f(x)=a(x-\alpha_1)^{k_1}\cdots(x-\alpha_n)^{k_n}(x^2+a_1x+b_1)^{l_1}\cdots(x^2+a_mx+b_m)^{l_m}.$$

where $\alpha_1, \alpha_2 \cdots \alpha_n$ are distinct roots of $f(x)$ ~~b~~ $a_i, a, b_i\in\mathbb{R}$ and ~~b~~ $a_i^2-4b_i<0$.

9. Rational Polynomials.

Def. (本原多项式). Let $f(x)$ be in $\mathbb{Z}[x]$. If all coefficients of $f(x)$ are relatively coprime, then $f(x)$ is <u>primitive</u>. (Primitive doesn't include the zero polynomial)

Remark. Every polynomial in $\mathbb{Q}[x]$ is an associate of a primitive polynomial.

Lemma If $g(x)$ is primitive, $f(x)$ is in $\mathbb{Z}[x]$, $f(x) = a g(x)$. for $a \in \mathbb{Q}$, then $a \in \mathbb{Z}$.

If moveover $f(x)$ is primitive. then $a = \pm 1$.

proof. Let $f(x) = \sum_{i=0}^{n} a_i x^i$ $g(x) = \sum_{i=0}^{n} b_i x^i$ $a = \frac{r}{s}$, with $(r,s)=1$

Then $s a_i = r b_i$ $(0 \le i \le n)$

Thus $s \mid b_i$. since $g(x)$ is primitive, $s = \pm 1$. so $a = \pm r \in \mathbb{Z}$

If $f(x)$ is also primitive. then by the same argument, we have $s = \pm 1$ and thus $a = \pm 1$

Thm 10 (Gauss's Lemma) The product of two primitive polynomials in primitive

Proof. Let $f(x) = \sum_{i=0}^{n} a_i x^i$ $g(x) = \sum_{j=0}^{m} b_j x^j$ are primitives.

Assume that their product $h(x) = f(x) g(x)$ is not primitive

There exists a prime number p divides every coefficient of $h(x)$

Let s, t be the least indices (pl. index) s.t. $p \nmid a_s$. $p \nmid b_t$

Consider the coefficient of x^{s+t} in $h(x)$

$C_{s+t} = a_s b_t + (a_{s-1} b_{t+1} + \cdots + a_0 b_{s+t}) + (a_{s+1} b_{t-1} + \cdots + a_{s+t} b_0)$

$p \mid a_{s-1} b_{t+1} + \cdots + a_0 b_{s+t}$ $p \mid a_{s+1} b_{t-1} + \cdots + a_{s+t} b_0$. $p \mid C_{s+t}$.

$\Rightarrow p \mid a_s b_t$. which is contradicts to

Thm 11. (Gauss's Lemma) A polynomial in $\mathbb{Z}[x]$ is reducible over $\mathbb{Q}$, then it is reducible over $\mathbb{Z}$

Assume that $f(x) = g(x) h(x)$, where $g(x), h(x) \in \mathbb{Q}[x]$ $1 \le \deg g < \deg f$, $1 \le \deg h < \deg f$

Then $f(x) = a f_1(x)$ $g(x) = b g_1(x)$ $h(x) = c h_1(x)$

where $f_1(x)$. $g_1(x)$ are primitives, $a \in \mathbb{Z}$ and $b, c \in \mathbb{Q}$

Then $a f_1(x) = bc\, g_1(x) h_1(x)$

By Lemma of Gauss. $a = \pm bc$. Thus $b.c \in \mathbb{Z}$, we obtain $f(x) = (bc\, g_1(x)) h_1(x)$

That is. $f(x)$ is reducible over $\mathbb{Z}$

coro 1 If $f(x)$ is in $\mathbb{Z}[x]$. $f(x) = g(x) h(x)$, $g(x)$ is primitive $h(x) \in \mathbb{Q}[x]$, then $h(x) \in \mathbb{Z}[x]$

高等代数 (Polynomials)

Date

Thm12. (Descarte's Rational Root Theorem)

Let $f(x) = a_n x^n + \cdots + a_0$ be in $\mathbb{Z}[x]$. suppose $\frac{r}{s}$ with $(r,s)=1$. is a rational root of $f(x)$. Then $s \mid a_n$. $r \mid a_0$. If, in particular, $f(x)$ is monic, then every rational root of $f(x)$ is an integer, and it divides a_0.

Proof: By assumption, $x - \frac{r}{s} \mid f(x)$

thus $sx - r \mid f(x)$.

Then $f(x) = (sx-r)(b_{n-1}x^{n-1} + \cdots + b_0)$. (1)

for $b_{n-1}, \ldots, b_0 \in \mathbb{Q}$.

By the corollary of Gauss Theorem, $b_i \in \mathbb{Z}$ for $0 \le i \le n-1$

By (1). we have $s b_{n-1} = a_n$, $-r b_0 = a_0$,

which gives $(s \mid a_n) \wedge (r \mid a_0)$

Thm 13 (Eisenstein irreducible criterion).

Suppose $f(x) = \sum_{i=0}^{n} a_i x^i$. be in $\mathbb{Z}[x]$. and there exists a prime p. s.t.

(1) $p \nmid a_n$. (2) $p \mid a_i$. $0 \le i \le n-1$ (3) $p^2 \nmid a_0$.

then $f(x)$ is irreducible in $\mathbb{Q}[x]$.

proof: suppose that, on the contrary, there exists a factorization $f(x) = g(x)h(x)$,

where $g(x) = \sum_{j=0}^{s} b_j x^j$ $h(x) = \sum_{k=0}^{t} c_k x^k$. $b_s \neq 0$. $c_t \neq 0$. $s>0, t>0$ $s+t=n$

We have $a_0 = b_0 \cdot c_0$. Since. $p^2 \nmid a_0$,

We concludes that only one of b_0 and c_0 is divisible by P.

We may suppose $P \mid b_0$, $P \nmid c_0$.

Note that not all b_i are divisible by p. Let k denote the least indix s.t. $p \nmid b_k$. ($p \mid b_{k-1}$)

By the fact $a_k = b_k c_0 + b_{k-1} c_1 + \cdots + b_0 c_k$.

We have $p \nmid a_k$ $(k<n)$ This contradicts the hypothesis that $p \mid a_k$. for $k<n$

$\varepsilon = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$. $\varepsilon^n = 1$. $(\varepsilon^2)^n = 1$ $(\varepsilon^{n-1})^n = 1$

$x^n - 1 = (x-1)(x-\varepsilon)(x-\varepsilon^2)\cdots(x-\varepsilon^{n-1})$

case 1 n is odd $x^n - 1 = (x-1)(x^2 - (\varepsilon + \varepsilon^{n-1})x + 1)(x^2 - (\varepsilon^2 + \varepsilon^{n-2})x + 1)\cdots[x^2 - (\varepsilon^{\frac{n-1}{2}} + \varepsilon^{\frac{n+1}{2}})x + 1]$

n is even $x^n - 1 = (x+1)(x-1)(x^2 - (\varepsilon + \varepsilon^{n-1})x + 1)$ $[x^2 - (\varepsilon^{\frac{n-1}{2}} + \varepsilon^{\frac{n+1}{2}})x + 1]$

Summary

1. field (closed under $+-\times\div$) ring (closed under $+-\times$).
2. polynomial. concept: term. degree. monic. coefficient.

operations of polynomial. (similar as number. * cancellation law)

3. (1) Division. (Eulicdean Algorithm)

$f(x) = g(x) \cdot q(x) + r(x)$ — $q(x)$ quotient, $r(x)$ remainder

$\boxed{\text{Thm1}}$ $r(x) = 0$ or $\deg r < \deg q$.

$\boxed{\text{Lemma}}$ g, f and r, g have same list of common divisor.

(2) Divisibility. ① transitivity.

② $f(x) \mid a(x)$. $f(x) \mid b(x)$. $f(x) \mid u(x)a(x) + v(x)b(x)$ ← a combination of $a(x)$ and $b(x)$

(3) g.c.d

definition: $d(x)$. any common divisor $\varphi(x)$. $\varphi(x) \mid d(x)$

① co-prime. $(f, g) = 1$.

② Bezout's Theorem. $d(x) = u(x)f(x) + v(x)g(x)$

③ $f_1(x) \mid g(x)$. $f_2(x) \mid g(x)$ $(f_1, f_2) = 1$.
$f_1(x) \cdot f_2(x) \mid g(x)$

④ $(f, g) = 1$, $f(x) \mid g(x)h(x) \Rightarrow f(x) \mid h(x)$.

⑤ Generalization.
$(f_1, f_2 \ldots f_s) = 1 \Rightarrow (f_1, f_2 \ldots f_s$ are relatively prime.

4. Factorization. depend on field.

Thm Every polynomial is irreducible or can be factored into a product of irreducible polynomial.

① Complex field. → factor into a product of polynomials of deg 1.

② a polynomial over $\mathbb{R}$ → have root $a+bi$ → have root $a-bi$ (conjugate)
→ can product into polynomials of deg 1. and irreducible polynomials of deg 2.

③ primitive polynomial. (Gauss Lemma).

Descarte's Rational Root Theorem. / Eisenstein irreducible criterion.

$\frac{r}{s}$. $s \mid a_n$. $r \mid a_0 \Rightarrow x - \frac{r}{s}$ $\qquad$ $p \nmid a_n$. $p \mid a_0 \ldots a_{n-1}$ $p^2 \nmid a_0 \Rightarrow$ irreducible.

5. Multiple factor and root.

derivate (formal) $f(x)$ and $f'(x)$ have common root $\Leftrightarrow f(x)$ have multiple root.

$\frac{f(x)}{(f(x), f'(x))} = a p_1(x) p_2(x) \ldots p_s(x)$. → the square-free part.

~~pol~~ D'Alembert's Theorem a polynomial of deg n. has at most n roots.

Chapter 2 Determinants.

§1. Determinants of lower ~~od~~ orders

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}$$

When $a_{11}a_{22} - a_{12}a_{21} \neq 0$.

$$x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{12}a_{21}} \qquad x_2 = \frac{b_2a_{11} - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}}$$

We call $a_{11}a_{22} - a_{12}a_{21}$ the determinant of order 2.

denoted by $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = D$ major diagonal

$\to$ (2,2) entry

$\to$ column index (列标)

$\to$ row index (行标)

"Diagonal rule" (don't fit the order ≥ 4). 行标 自然排列 列标 $p_1 \cdots p_n$

$$D = \sum_{p_1p_2} (-1)^{\tau(p_1p_2)} a_{1p_1}a_{2p_2}$$

similarly $b_1a_{22} - a_{12}b_2 = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} = D_1$ Hence, $x_1 = \frac{D_1}{D}$

$x_2 = \frac{D_2}{D}$.

$b_2a_{11} - a_{21}b_1 = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} = D_2$.

Determinant of order 3.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

$$D = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$= \sum_{p_1p_2p_3} (-1)^{\tau(p_1p_2p_3)} a_{1p_1}a_{2p_2}a_{3p_3}$$

When $D \neq 0$ $\quad x_1 = \dfrac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{D}. \quad x_2 = \dfrac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{D} \quad x_3 = \dfrac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{D}$

§2. Permutations.

Def 1 A permutation of the natural numbers $1, 2, \ldots, n$ is a linear arrangement of then in some order.

Remark (1) The number of all permutations of $1, 2, \ldots n$ is $n!$

(2) The permutation $123 \cdots n$ is called natural permutation. (自然顺序).

(3) For any permutation $p_1 \cdots p_n$. $p_i \neq p_j$ if $i \neq j$.

Def 2. Given a permutation $p_1p_2 \ldots p_n$, the number $p_1, p_2, \ldots p_{i-1}$ greater than p_i is called inversion number (逆序数) of p_i, denoted by t_i.

The sum $t_1 + t_2 + \ldots + t_n$ is called the inversion number of $p_1p_2 \ldots p_n$ denoted by $\tau(p_1p_2 \ldots p_n)$

Def 3. For the permutation $p_1p_2 \ldots p_n$. if $\tau(p_1p_2 \ldots p_n)$ is an odd number, then $p_1p_2 \ldots p_n$ is said to odd. Otherwise, it is said to even.

Thm 1 If two numbers of a permutation $p_1p_2 \ldots p_n$ are interchanged, then the parity of the new permutation is opposite to $p_1p_2 \ldots p_n$

coro 1. interchanged odd times $\to$ parity is changed.

even times $\to$ parity not change.

$$\tau(x_1x_2 \cdots x_n) + \tau(x_nx_{n-1} \cdots x_1) = \frac{(n-1)n}{2}$$

coro 2. An odd permutation $\xrightarrow{\text{interchange odd}}$ natural permutation

coro 3. For all permutation of $1, 2, \ldots n$. the number of odd permutation is equal to the even one.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{i_1 i_2 \cdots i_n} (-1)^{\tau(i_1 i_2 \cdots i_n)} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}$$

(for row index and column index are symmetric).

Remark: (1) D is unique. sum of $n!$ terms.

(2) Each term of D contains exactly one element from each row (resp. column) of D. (同行/同列不得重复取).

Special determinant

$$\begin{vmatrix} a_{11} & 0 & 0 \cdots 0 & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11} \cdot a_{22} \cdots a_{nn}$$

lower triangular (upper triangular similarly)

especially, we have diagonal determinant. $|{}^{\backslash}{}_{0}^{0}{}^{\backslash}|$

$$\begin{vmatrix} 0 & \cdots & \cdots & a_{1n} \\ \vdots & & a_{2,n-1} & a_{2n} \\ & a_{n-1,2} & \cdots & a_{n-1,n} \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} = (-1)^{\overset{\frac{n(n-1)}{2}}{\tau(n \cdots 321)}} a_{1n} a_{2(n-1)} \cdots a_{(n-1)2} a_{n1}$$

Equivalent Definition

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{p_1 p_2 \cdots p_n} (-1)^{\tau(p_1 \cdots p_n)} a_{1p_1} a_{2p_2} \cdots a_{np_n} \quad \text{(行展开)}$$

$$= \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{j_1 1} a_{j_2 2} \cdots a_{j_n n} \quad \text{(列展开)}$$

§4 Properties of Determinant

$$D = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}, \quad D' = \begin{vmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{nn} \end{vmatrix}.$$

(可写作 $D = |a_{ij}|_{n\times n}$)

(1) Pro 1: $D = D'$. D' is called the transpose of D. (转置行列式). (行的性质列也有)

(2) If two rows or two columns of D are interchanged then D changes sign (变号) 换法

Coro 1: Especially, if there are two rows or columns have same elements (are identical). $D = -D = 0$

(3) If D_1 is obtained from D by multiplying every element in one row or one column by k, then $D_1 = kD$. (倍法)

coro 1: especially, if two rows or two columns of D are proportional, then $D = 0$

(4)
$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_1+b_1 & a_2+b_2 & \cdots & a_n+b_n \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_1 & a_2 & \cdots & a_n \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_1 & b_2 & \cdots & a_n \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

(行列拆=)

(5) If D_1 is obtained from D by adding a row or a column of multiple k of (D a) another row or another column of D, then $\underline{D_1 = D}$. (消法)

一般解法: 找第一列非零元(尽可能小)

e.g.1 $D = \begin{vmatrix} a & b & b & b & \cdots & b \\ b & a & b & b & & \vdots \\ b & b & a & b & & \\ b & b & b & a & \cdots & b \\ \vdots & & & & \ddots & \\ b & & & & & a \end{vmatrix} = \begin{vmatrix} a+(n-1)b & b & \cdots & b \\ a+(n-1)b & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ a+(n-1)b & b & \cdots & a \end{vmatrix} = a+(n-1)b \begin{vmatrix} 1 & b & b & \cdots & b \\ 1 & a & b & \cdots & b \\ 1 & b & a & \cdots & b \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & b & b & \cdots & a \end{vmatrix}$

(行等和)

$= a+(n-1)b \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & a-b & & & \vdots \\ \vdots & 0 & a-b & & 0 \\ 1 & 0 & \cdots & 0 & a-b \end{vmatrix} = [a+(n-1)b](a-b)^{n-1}$

Solution 2.

$= \begin{vmatrix} a & b & \cdots & b \\ b-a & a-b & 0 & \cdots 0 \\ b-a & 0 & \ddots & \vdots \\ \vdots & \vdots & & \\ b-a & 0 & \cdots & a-b \end{vmatrix} \underset{(爪形)}{=} \begin{vmatrix} a+(n-1)b & b & \cdots & b \\ 0 & a-b & & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & & \\ 0 & 0 & & a-b \end{vmatrix} = [a+(n-1)b]\cdot(a-b)^{n-1}$

e.g.2 $D_1 = |a_{ij}|_{k\times k}$ $D_2 = |b_{ij}|_{s\times s}$

$D = \begin{vmatrix} a_{ij\,k\times k} & 0 \cdots \\ c_{ij\,k\times s} & b_{ij\,s\times s} \end{vmatrix}$ $\boxed{\text{Then } D = D_1 \times D_2}$

(Theorem: Any determinant can be change into triangular determinant by these properties above)

§5. Cofactor expansion (余子式展开)

$$|a_{ij}|_{3\times3} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix},$$

Def 7. Let $D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$, detete the i th row and j column, the remaining determinant of order $n-1$ is called the minor (余子式) of a_{ij}, denote by M_{ij}.

(Each element in determinant has a unique minor).

Def 8. Then the cofactor of a_{ij}, written as A_{ij}. is give by $A_{ij} = (-1)^{i+j} M_{ij}$.

Thm1 (cofaction theorem). Let $D = |a_{ij}|_{n\times n}$. Then.

(1) D can be expanded in terms of the elements of the ith row $D = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots a_{in}A_{in}$ (can also be a definition of nth determinant)

(2) column. similarly (jth column)

Lemma 1. $D = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n-1\,1} & \cdots & a_{n-1\,n} \\ 0 & 0 \cdots & a_{nn} \end{vmatrix} = a_{nn}A_{nn}$ (expanded by last column

Thm1 + Thm2 $\sum_{j=1}^{n} a_{kj}A_{ij} = \begin{cases} D, & k=i \\ 0, & k\neq i \end{cases}$

Thm2. keep notations as above.

(1) $a_{k1}A_{i1} + a_{k2}A_{i2} + \cdots + a_{kn}A_{in} = 0$ ($i\neq k$, 非本行对应)

(2) ~~a_{ij}~~ $a_{1k}A_{1j} + a_{2k}A_{2j} + \cdots + a_{nk}A_{nj} = 0$ ($j \neq k$)

(1) proof: $D_1 = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \begin{matrix} \\ \to i \\ \to k \\ \\ \end{matrix}$ $D_1 = 0$

Denote $D_1 = |b_{ij}|_{n\times n}$, $b_{i1} = a_{k1} \cdots b_{in} = a_{kn}$

Let B_{ij} be the cofactor of b_{ij}

Then $B_{ij} = A_{ij}$, expand D_1 by ith row.

e.g. the Vandermonde determinant.

$$D=\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \cdots & a_n^2 \\ \vdots \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \cdots & a_n^{n-1} \end{vmatrix} = \underset{\text{连乘}}{\prod}(a_i-a_j) \quad (1\le j<i\le n)$$

$$= (a_n-a_{n-1})(a_n-a_{n-2})\cdots(a_n-a_1)(a_{n-1}-a_{n-2})\cdots(a_{n-1}-a_1)\cdots(a_2-a_1)$$

(共 $\frac{n(n-1)}{2}$ 项).

$0=\prod_{1\le j<i\le n}(a_i-a_j) \Leftrightarrow a_1, a_2 \ldots a_n$ have at least 2 equal terms

proof: Induction. 倒消(从最后一行开始) (a_1倍).

§6. Cramer's rule

Consider a system of n linear equation with n unknowns:

$$(1)\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n = b_1 & \leftarrow \text{constant term.} \\ a_{21}x_1 + a_{22}x_2 + \cdots a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots a_{nn}x_n = b_n \end{cases}$$

We can $D=|a_{ij}|_{n\times n}$ the coefficient determinant of (1)

When $b_1=b_2=\cdots=b_n=0$. (1) is called homogeneous. (must have solution $x_1=x_2\cdots=x_n=0$.) Otherwise, (1) is called nonhomogeneous

Thm4 (Cramer's Rule). If $D\neq 0$. then the linear system (1) has a unique solution. and. $x_1=\frac{D_1}{D}$. $x_2=\frac{D_2}{D}$, $\cdots$ $x_n=\frac{D_n}{D}$, where D_j is obtained from D by replacing the jth column of D by $b_1, b_2, \ldots b_n$

代入方程组验证 ✓ (存在性)

uniqueness: Assume $x_i=d_i$ $(1\le i\le n)$ is a solution of (1).

$$d_iD=\begin{vmatrix} a_{11} & \cdots & d_ia_{1i} & \cdots & a_{1n} \\ a_{21} & \cdots & d_ia_{2i} & \cdots & a_{2n} \\ \vdots \\ a_{n1} & \cdots & d_ia_{ni} & \cdots & a_{nn} \end{vmatrix}$$

做 n-1 次列变换. $a_{11}\times d_1$, $a_{12}\times d_2\cdots$ $a_{1n}\times d_n$ 加到第i行
第i列变成了 $b_1, b_2, \ldots b_i$.

$d_iD=D_i \Rightarrow d_i=\frac{D_i}{D}$ (unique)

linear system 2个解 → 无穷多个解.

Coro 1 If (1) has no solution or has infinitely many solutions, then $D=0$.

Coro 2. If $b_1=b_2\cdots=b_n=0$. and if $D\neq 0$, then (1) has only zero solution (trivial).

conversely, if a homogeneous (齐次) system has nontrivial solutions then $D=0$

existence: we should proof $a_{11}\frac{D_1}{D}+a_{12}\frac{D_2}{D}+\cdots+a_{1n}\frac{D_n}{D}=b_1$

$$\Leftrightarrow a_{11}(b_1A_{11}+b_2A_{21}+\cdots b_nA_{n1})+a_{12}(b_1A_{12}+b_2A_{22}+\cdots b_nA_{n2})+\cdots a_{1n}(b_1A_{1n}+\cdots+b_nA_{nn})$$

$$=b_1(a_{11}A_{11}+a_{12}A_{12}+\cdots+a_{1n}A_{1n})+\cdots+b_n(a_{11}A_{11}+\cdots a_{1n}A_{nn})=$$

$$=b_1(a_{11}A_{11}+a_{12}A_{12}+\cdots+a_{1n}A_{1n})=b_1D.$$

* step determinant (matrix). $\frac{n^3+2n-3}{3}$ times transformation.

* Laplace Theorem.

Chapter 3. Linear System(s)

(A) An $m\times n$ linear system of equations (or simply an $m\times n$ LS in variable $x_1 \dots x_n$).

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \quad (1)$$ a solution $(\alpha_1 \dots \alpha_n)$

The set of all possible solutions of (1) is called the solutions of (1)

Def: If a LS has a solution, it is called consistent: if it has no solution, it's called nonconsistent (consistent 对应有解, 可以是 infinite solutions).

Def: Two linear systems S_1 and S_2 is said to be equivalent if they have the same solution set.

Gauss-Jorden eliminate method (消元法).

(Def: echelon form (row echelon from): the first nonzero coeffient in each equation is to the right of the first nonzero coefficient in the preceding equation). 先前的, 前面的

[Elementary Operations on LS].

(I) Interchange two equation: $r_i \leftrightarrow r_j$

(II) Multipling any equation by a nonzero number k: kr_i

(III) Adding to the jth equation k times the ith equation: $r_j + kr_i$

Fact: Linear system under elementary operations have the solutions and thus are equivalent systems.

After transforming to the echelon form

(If we find $0 \neq b_i$ finally. the system has no solution

If we find $c_i x_{n-1} + c_{i+1} x_n = d_i$. in the last row, the system has infinite solutions).

(1) If $r = n$. has a unique solution. * (此处r是转化为阶梯形矩阵之后的x_r).

(2) If $r < n$. we can represent $x_1, x_2 \dots x_m$ by $x_{r+1} \dots x_n$. For $x_{r+1} = k_{r+1} \dots x_n = k_n$. then $x_1 \dots x_m$ has a unique solution. $x_{r+1} \dots x_n$ ~~are~~ is called free variables of the system 有多个, 取前个. $x_1, \dots x_r$ is called the principle/basic variables

Theorem 1: Every homogeneous $m\times n$ LS has nontrivial solutions of if $m<n$. (not only if)

e.g. $\begin{cases} 2x_1 - x_2 + 3x_3 = 1 \\ 4x_1 - 2x_2 + 5x_3 = 4 \\ 2x_1 - x_2 + 4x_3 = -1 \end{cases} \Rightarrow \begin{cases} x_1 = \frac{1}{2}(7 + x_2) \\ x_2 \text{ is free} \\ x_3 = -2 \end{cases}$ (general solution)

For S $\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{s1}x_1 + \cdots + a_{sn}x_n = b_s \end{cases}$

We denote $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{s1} & a_{s2} & \cdots & a_{sn} & b_s \end{bmatrix}$ an Augmented Matrix of S (增广矩阵)

§2. Introduction to Matrices and Matrix representation of a Linear System.

(A) Def. An array of numbers in m rows and n columns is called an $m\times n$ matrix denoted by $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ The numbers a_{ij} are called elements or entries. the subscripts s and j identify the row and column. Especially, if $m=n$. the matrix is called n-square matrix.

↳ We will write $A = (a_{ij})_{m\times n}$ for short.

(B) Matrix Representation of an LS

In general, with the $m\times n$ LS (1).

We associate two matrices. The coefficient matrix for (1). is the $m\times n$ matrix. A.

The augmented matrix for (1) is the $m\times(n+1)$ matrix. $B=[A|b]$

(column vector) $b=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$

(C) Elementary row operation on a matrix (only! when related ~~wit~~ to LS).

Def3. The elementary row operation on a matrix are ~~one~~ of three types.

(I) interchange (II) Multiply scalar k (III) r_j+kr_j $(i\neq j)$

Def4. Two $m\times n$ matrices are called row equivalent if one can be obtained from the other by sequence of elementary row operations.

Def5. A matrix is said to be in (row) echelon from or simply an elechon matrix. if.

(1) the zero rows. if any, are below all nonzero rows.

(2) the leading entry (the left-most nonzero, entry). of a nonzero row lies in a column to the right of the leading entry of the row above it.

(3) All entries in a column below a leading entry are zeros.

Def6. A matrix in row echelon-form is said to be in reduce row echelon form, if. (最简)

(1) the leading entry in each row is 1.

(2) each leading 1 is the only nonzero entry in its column.

e.g $\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{bmatrix}$ $\begin{bmatrix} 1 & 2 & 0 & 1 & -1 \\ 0 & 0 & 1 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Thm1 (uniqueness of the reduced echelon from of a matrix).

Each matrix is row equivalent. to one and only one reduced echelon matrix.

(D) The row Reduction Algorithm (we use "→" instead of "=")

Step1. Find the left-most nonzero column

Step2. Bring a nonzero entry in the left-most Column to the top (by interchanging row).

Step3. making the leading entry of the 1st row into 1. (Multiple).

Step4. Eliminate. nonzero entries below leading 1.

Step 5. Temporarily ignore the first row of this matrix and repeat step 1-4 on the remaining submatrix.

When we obtain the ~~colu~~ echelon matrix. make it reduced. (right→left).

Def.6. A pivot position in a matrix A is a location in A that corresponds to a leading 1 in the reduced echelon form of A → pivot position corresponds the basic variable others are free. (其实选择不唯一)

(pivot position. the column ~~of~~ contains a pivot position)

§3. Using row reduction to solve a Linear System

(A) Given LS → Augmented Matrix → Reduced Matrix → Reduced LS → Solution

(B) Existence and uniqueness Question.

Thm 1 (1) A linear system is consistent (有解) if and only if an echelon form of the augmented matrix has no row of the form $[0 \cdots 0\ 0\ b]$ with b nonzero

(2) If an LS is consistent. then the solution set contains either

(i) a unique solution, when there is no free variable(s).

or (ii) infinitely many solutions when there is at least one free variable

§4. n-dimensional vector space.

(A) Definition of n-dimensional vectors

Let F denote the field of real number or the field of complex number

Def 1. An $n\times 1$ matrix $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ is called an n-dimensional (维) column vector or a vector, simply.

A $1\times n$ matrix $[a_1, a_2, \ldots a_n]$ is called an n-dimensional row vector (Nature is the same)

(B) Operations of Vectors.

Let. $u = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$. $v = \begin{bmatrix} b_1 \\ \vdots \\ b_2 \end{bmatrix} \in F^n$ (the set of all n-dimensional vector) /* either column or row. can't be both / mixture. */

(1) u and v are equal. $(u=v) \Leftrightarrow a_i = b_i\ (i=1,2,\cdots n)$

(2) The sum. substraction $u \pm v = \begin{bmatrix} a_1 \pm b_1 \\ a_2 \pm b_2 \\ \vdots \\ a_n \pm b_n \end{bmatrix}$

(3) Given a scalar $k \in F$, $k\cdot u = \begin{bmatrix} ka_1 \\ \vdots \\ ka_n \end{bmatrix}$

(4) The negative of u is $(-1)\cdot u$

(5) The zero vector. $0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

properties: addition (1) commutative law (2) associative law, (3) $u-u=0$ (4) $u+0=u$

multiplication (1) $1\cdot u = u$ (2) $c\cdot(du) = (c\cdot d)\cdot u$ (3) $(c+d)u = cu + du$ (4) $c(u+v) = cu + cv$.

(For, c, d are scalars u, v are vectors)

we call F^n the n-dimensional vector space

(C) Linear Combination

Def 2. Given vectors $\alpha_1, \alpha_2 \ldots \alpha_s \in F^n$ and given scalars $k_1, k_2 \ldots k_s \in F$.

the ~~β~~ vector $\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s$. is a linear combination of $\alpha_1, \alpha_2 \ldots \alpha_s$.

with weights (权,系数组合) $k_1, k_2 \ldots k_s$. (the weight in a linear combination can be any numbers. including 0)

The set of all linear combinations of $\alpha_1, \alpha_2, \cdots \alpha_p$ is called the linear span of $\alpha_1, \alpha_2 \ldots \alpha_p$ denoted by span $\{\alpha_1, \alpha_2 \ldots \alpha_p\}$ or $\langle \alpha_1, \alpha_2 \ldots \alpha_p \rangle$. That is $\text{span}\{\alpha_1, \alpha_2 \ldots \alpha_p\} = \{k_1\alpha_1 + k_2\alpha_2 + \cdots k_p\alpha_p \mid k_1, \ldots k_p \in F\}$

In general, an LS $\begin{cases} a_{11}x_1 + \cdots a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{m1}x_1 + \cdots a_{mn}x_n = b_m \end{cases}$ whose augmented matrix is $[A|b] = [\alpha_1, \alpha_2, \cdots \alpha_n, \beta]$

then we can obtain the vector equation.

$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \beta$. (1)

Thm 1 A vector equation (1) has the same solution set as the LS whose augment matrix is $[\alpha_1, \alpha_2 \ldots \beta]$ (2). In particular. β can be written as a linear combination of $\alpha_1 \ldots \alpha_n$ $\Leftrightarrow$ there is a solution to the LS corresponding to the matrix (2)

(D) The matrix equation $AX=\beta$. (矩阵×列向量). 矩阵方程

Def 3. Let A be an $m\times n$ matrix with columns $\alpha_1, \alpha_2, \dots \alpha_n$, and let $X=[x_1, x_2, \cdots x_n]$ the product of A and X, denoted by AX, is the linear combination of the columns of A using the corresponding entries in X as weights.

i.e. $AX = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n$.

(the dimensional of X. should be equal to the amout of column vector(s). of A).

Warning: AX is defined only if the number of colums of A equals to the number of entires in X.

Thm 2 If A is an $m\times n$ matrix with columns $\alpha_1, \alpha_2 \dots \alpha_n$ and if β is in $\mathbb{R}^n$, the matrix equation $AX=\beta$. has the same solution as the vector equation $x_1\alpha_1 + \cdots + x_n\alpha_n = \beta$. which has the same solution as the LS.

row Vector rule: (for computing AX).
If the product AX is defined. then the the ith entry in AX is the sum of the products of corresponding entries from ith row of A and from the vector X.

Thm 3. If A is an $m\times n$ matrix, α and β are in $\mathbb{R}^n$ and c is in $\mathbb{R}$. then

1° $A(\alpha+\beta) = A\alpha + A\beta$. 2° $A(c\alpha) = c(A\alpha)$ [linear property].

Corollay 1: Let A be an $m\times n$ matrix.
If X_1 and X_2 are solution of $AX=0$.
then ① X_1+X_2 is a solution of $AX=0$
② c is a scarlar. $c\in\mathbb{R}$. cX_1 is a solution of $AX=0$.

§5. Solution Sets of LS

(A) Homogeneous LS. $AX=0$.

$\boxed{\text{Fact}}$. $AX=0$ has a nontrivial solution $\Leftrightarrow$ the equation has at least one free variable.

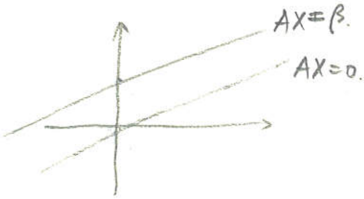
If the solution set has one free variable
$X = x_n\begin{bmatrix}\vdots\\\vdots\end{bmatrix} = tv$ (t is free variable) Namely. the solution set is span$\{v\}$

n variables. $X = t_1v_1 + t_2v_2 + \cdots t_2v_n$ span$\{v_1, v_2, \cdots v_n\}$

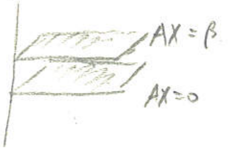
(B) Nonhomogeneous LS: $AX=\beta$. $\beta\neq 0$.

$X = p + \sum_{i=1}^{n} t_iv_i$ p is called particular solution

when the number of free variable $=1$. $X = p+tv$



2 dimensional



§6. Linear Independence

(A) Definition (Both vectors and the set of ~~the~~ vectors can be the subject)

<u>Def 1</u>: An indexed set of vectors $\{v_1, v_2 \dots v_p\}$ in F^n is called **linear independence** if the vector equation $x_1v_1 + x_2v_2 + \dots + x_pv_p = 0$. has only trivial solution. The set is called **linear dependence** if there exist weights $c_1, \dots c_p$ not all zero s.t. $c_1v_1 + \dots c_pv_p = 0$. (whether has free variable or not).

Remark: linear dependence relation among $v_1, v_2, \dots v_p$.

(B) Linear Independence of Matrix columns

Given an $m \times n$ matrix $A = [\alpha_1\ \alpha_2 \cdots \alpha_n]$, $AX = 0 \Leftrightarrow x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n = 0$.

<u>Fact</u>. The set of column vectors is linearly independent $\Leftrightarrow$ $AX=0$ has only zero solution
$\Leftrightarrow$ The LS has no free variables

(C) set of one or two vectors.

one vector. ① $v \neq 0 \Leftrightarrow$ linear independent ($x=0$)
② $v = 0 \Leftrightarrow$ linear dependent

two vector: $\{v_1, v_2\}$ is linearly dependent $\Leftrightarrow$ at least one vector is a <u>multiple of the other</u> (存在倍数关系)
linearly independent $\Leftrightarrow$ Neither of them is a multiple of the other

(D) set of two or more vectors

<u>Thm1</u> [Characterization of linearly dependent sets]

An indexed set $S = \{v_1 \dots v_p\}$ with $p \geq 2$ is linearly dependent. $\Leftrightarrow$ <u>At least</u> (not all) one of the vector in S is a linear combination of the others.

"$\Rightarrow$"

proof: There exist weights $c_1 \dots c_p$ not all zero. such that

$$c_1v_1 + c_2v_2 + \dots + c_pv_p = 0.$$

Since $c_1 \dots c_p$ are not all zero. We may assume $c_1 \neq 0$.

Then $v_1 = -\frac{c_2}{c_1}v_2 - \frac{c_3}{c_1}v_3 - \dots \frac{c_p}{c_1}v_p$. $\quad -\frac{c_i}{c_1} = d_i$

$\Leftrightarrow \quad v_1 + d_2v_3 + \dots + d_pv_p = 0$.

<u>Thm2</u>. If $\vec{0} \in S = \{v_1, \dots v_p\}$. then S is linearly dependent.

<u>Thm3</u>. Any set $\{v_1, \dots v_p\}$ in F^n is linearly dependent if $p > n$ (p is number of variable, n is number of equation).

<u>Thm4</u>

(a) If any subset of $\{v_1 \dots v_p\}$ is linearly dependent, the entire set of vectors is linear dependent

(b) Any subset of a linearly independent set of vectors $\{v_1 \dots v_p\}$ is linearly independent.

<u>(c)</u>. If $v_1 \dots v_p$ are linearly independent. $v_1, \dots v_p, v$. are linearly dependent then v is a <u>unique</u> linear combination of $v_1, v_2 \dots v_p$.

uniqueness: ① $a_1v_1 + a_2v_3 + \dots a_pv_p = v$
② $b_1v_1 + b_2v_2 + \dots b_pv_p = v$ $\quad$ ① − ②. $a_i - b_i = 0$.

(d) If $v_1 \dots v_p$ are linearly independent. and v_{p+1} can not express as linear combination of $v_1, \dots v_p$ then $v_1, \dots v_p, v_{p+1}$ are linearly independent (converse-negative proposition of (c)).

(e). Let $\alpha_1 \dots \alpha_p$ be in F^s, $\beta_1 \dots \beta_p$ be in F^t. If $\alpha_1 \dots \alpha_p$ are ~~li~~ linearly independent, then the vector

$v_1 = \begin{bmatrix}\alpha_1\\ \beta_1\end{bmatrix}$, $v_2 = \begin{bmatrix}\alpha_2\\ \beta_2\end{bmatrix}$, $\dots$ $v_p = \begin{bmatrix}\alpha_p\\ \beta_p\end{bmatrix}$ are linearly independent.

$A = [\alpha_1 \dots \alpha_p]_{s \times p}$. $\quad B = [v_1 \dots v_p]_{(s+t) \times p}$

(I) $AX = 0$ S_1 (II) $BX = 0$. S_2.

$S_2 \subseteq S_1 = \{0\}$.

vector set equivalent $\underset{\not\Leftarrow}{\Rightarrow}$ same rank
$\Leftrightarrow$ can be written as a linear combination of each other

§ 7. The Rank of a Matrix. (矩阵的秩).

(A). Bases and Dimension

$\text{span}\{v_1, \dots v_p\} = \{c_1v_1 + c_2v_2 + \cdots + c_pv_p \mid c_1, \dots c_p \in \mathbb{R}\}$.

Def 1. Let V be a nonempty subset of $\mathbb{R}^n$,
If (a) $0 \in V$
(b) $\forall \alpha, \beta \in V$, $\alpha + \beta \in V$
(c) $\forall \alpha \in V$, $\forall c \in \mathbb{R}$, $c\alpha \in V$
then V is called *a subspace* of $\mathbb{R}^n$

Remark: the solution set of homogenous LS. is a subspace of $\mathbb{R}^n$.
$A_{m\times n}$, $X \in \mathbb{R}^n$. $AX = 0$.
$N(A) = \{\xi \mid A\xi = 0, \xi \in \mathbb{R}^n\}$.

e.g. $\text{span}\{v_1, \dots v_p\}$ is a subspace of $\mathbb{R}^n$

Def 2. A basis for a subspace V of $\mathbb{R}^n$ is a linearly independent set in V that spans V.
(用有限个表示无穷多个).
(任意添一个向量，即线性相关)

极大无关组&基：
线形空间(极大无关组的全部线性组合)，则极大无关组才是基.
极大无关组：一定从给定的向量组中产生
(maximal linearly independent array)
(行向量做完变换，得出基，对应不了极大无关组.
列向量即可，行向量会换行).

(1) In $\mathbb{R}^n$, consider the unit column-vectors

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots \quad e_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

$\forall X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$, then $x = e_1x_1 + e_2x_2 + \cdots + e_nx_n$ ($e_1, e_2, \cdots e_n$ is a basis of $\mathbb{R}^n$. the standard basis)

($\varepsilon_1 = e_1$, $\varepsilon_2 = e_1 + e_2$, $\varepsilon_3 = e_1 + e_2 + e_3$, … also make up a basis of $\mathbb{R}^n$)

Lemma 1. *Let V be a subspace of $\mathbb{R}^n$ with basis $\alpha_1 \dots \alpha_r$ and let $\beta_1, \beta_2 \dots \beta_s$ be a linearly independent set of V. Then $s \le r$* (极大无关组为基).

Thm: $\alpha_1, \dots \alpha_r$ can be represented $\beta_1 \dots \beta_s$.
若 if $r > s$. then $\alpha_1 \dots \alpha_r$ is linearly dependence.

proof: Let:
$$\begin{cases} \beta_1 = a_{11}\alpha_1 + \cdots + a_{r1}\alpha_r \\ \vdots \\ \beta_s = a_{1s}\alpha_1 + \cdots + a_{rs}\alpha_r \end{cases}$$

Where a_{ij} are scalars. Suppose $s > r$. we form a linear combination of the vectors β_j with weights x_j: $x_1\beta_1 + x_2\beta_2 + \cdots x_s\beta_s = (x_1a_{11} + x_2a_{12} + \cdots + x_sa_{1s})\alpha_1 + (\dots)\alpha_2$

consider
$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1s}x_s = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2s}x_s = 0 \\ \vdots \\ a_{r1}x_1 + \cdots a_{r2}x_2 + \cdots + a_{rs}x_s = 0. \end{cases}$$
if $s > r$. the system must have free variable(s)
$\Rightarrow$ must have nonzero solution $X = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} \neq 0$

But this gives also $k_1\beta_1 + k_2\beta_2 + \cdots k_s\beta_s = 0$. contradicts.

Theorem 1: *Every nonzero subspace $V \subseteq \mathbb{R}^n$ has a finite basis. All basis of V have the same number $r \le n$ of vector (The number r is called the dimension of V denoted by $\dim V = r$ or $\dim_{\mathbb{R}} V = r$)*

proof: $V \neq \{0\}$ Let $0 \neq \alpha_1 \in V$ suppose that we have found k linearly independent vectors in V: $\alpha_1 \dots \alpha_k$ If $\text{span}\{\alpha_1 \dots \alpha_k\} \neq V$. then we choose $\alpha_{k+1} \in V$ which is not in $\text{span}\{\alpha_1 \dots \alpha_k\}$ Then the set $\alpha_1, \dots \alpha_k, \alpha_{k+1}$ is linearly independent. This process can not go on infinitely. Hence, for some $r \le n$. the linearly independent set $\alpha_1 \dots \alpha_r \in V$. becomes maximal (i.e. no matter what vector $X \in V$ we add to the set $\{\alpha_1 \dots \alpha_r, X\}$ is linearly independent This means that $X \in \text{span}\{\alpha_1, \dots \alpha_r\}$. Hence $V = \text{span}\{\alpha_1 \dots \alpha_r\}$ $\square$

Remark: If $\{\alpha_1 \dots\}$ is set of vector in $\mathbb{R}^n$, $\dim(\text{span}\{\alpha_1, \alpha_2, \dots\}) \le n$. which is called the rank of $\alpha_1, \alpha_2, \dots$

$\text{rank}\{\alpha_1, \alpha_2, \dots\} = \dim \text{span}\{\alpha_1, \alpha_2, \dots\}$.

If $V = \{0\}$, then $\dim V = 0$.

* pivot column → linearly independent (reduced echelon matrix).

If $A = (a_{ij})_{m\times n}$ Let $A = [\alpha_1, \alpha_2 \dots \alpha_n]$

$A = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix}$ Then column space $\text{col}(A) = \text{span}\{\alpha_1, \alpha_2, \dots \alpha_n\}$ → (α_1-α_n中的某些个可组合 col(A) 的基)

$\text{row}(A) = \text{span}\{\beta_1 \cdots \beta_m\}$.

column rank of $A = \dim(\text{col}(A))$

row rank of $A = \dim(\text{row}(A))$

Fact: If B is obtained from A by using a finite sequence of row elementary operations. then the columns of $A = [a_1, \dots a_n]$ have exactly the same linear dependence relation as the column of $B = [b_1 \dots b_n]$.

~~since~~ $x_1 a_1 + \dots + x_n a_n = 0$ (1) (1) and (2) have same solution.
$x_1 b_1 + \dots + x_n b_n = 0$ (2).

A.B have same dimension. rank. linearly dependence/independence. basis (indices are same)

(注. basis 中只是位置对应 如 $a_1, a_2, a_5 \to b_1, b_2, b_5$ 不是 b_1, b_2, b_5 是A的基】空间是不同的.

Theorem2. The pivot columns of a matrix A from a basis for $\text{col}(A)$

Theorem3. Let A be an $m\times n$ matrix and suppose that A is row equivalent to the $m\times n$ matrix B Then $\text{Row}(A) = \text{Row}(B)$ (空间完全一样).

简述
sketch proof: $A = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \xrightarrow[\text{interchange}]{\text{row}} B = \begin{bmatrix} \beta_2 \\ \beta_1 \\ \beta_3 \end{bmatrix}$

$\xrightarrow[\text{multipling}]{\text{adding}}$ same.

Theorem4 If the nonzero matrix A is row equivalent to B in echelon form then the nonzero rows of B form a basis for $\text{row}(A)$

rank of A = the number of nonzero rows in B.

corollary2: If A is an $m\times n$ matrix. then the row rank of A equals the column rank of A.

(Thus the common value of the column rank and value of the row rank is called the rank of matrix)

Determinant rank of a matrix

If $A = [a_{ij}]$ is $n\times n$ matrix, then we call $\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$ the determinant of A. denoted by $|A|$.

Def. Let A be $m\times n$ matrix, a matrix obtain from A by deleting some (but not all) of its row and column is called submatrix of A. The determinant of an $r\times r$ submatrix of A is called a minor (子式) of order r (determinant). /the highest order of $A = \min\{m, n\}$/

Def: If a matrix A has a minor of r not equal to 0. but no nonzero minors of order $r+1$. the number r is called a determinant rank of A. (rank: 非零子式的最高阶数).

Theorem5. 3 ranks are equal

Let A be an $m\times n$ matrix. consider the LS.

$AX=\beta$ (1)

$\Leftrightarrow x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n=\beta$ (2) $\alpha_1,\alpha_2,\dots\alpha_n$ are column of A.

(If $x_1=k_1$. $x_2=k_2$ .... $x_n=k_n$ is a solution of (1))

$\Leftrightarrow k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n=\beta$. $\beta\in \mathrm{Span}\{\alpha_1\dots\alpha_n\}=\mathrm{Col}(A)$

$\Leftrightarrow \mathrm{span}\{\beta,\alpha_1,\dots\alpha_n\}=\mathrm{span}\{\alpha_1\dots\alpha_n\}$ rank/dimension equal.

$\Leftrightarrow \mathrm{rank}(A|\beta)=\mathrm{rank}(A)$

Thm 6. (Kronecker – Capelli) 克罗内克-卡佩利

An $m\times n$ nonhomogenous (i.e. $\beta\neq 0$) LS $AX=\beta$ is consistent if and only if $\mathrm{rank}(A)=\mathrm{rank}(A|\beta)$. (系数矩阵.增广矩阵等秩).

Thm 7. Given an $m\times n$ matrix A with columns $\alpha_1,\alpha_2\dots\alpha_n$ in R^n, Then.

$\alpha_1,\dots\alpha_n$ are linearly independent $\Leftrightarrow \mathrm{rank}(A)=n$.

Summary

☆☆☆ Structure of Solution set 解集合是子空间

(A1) Homogenous LS $AX=0$ (A is an $m\times n$ matrix)

1° $AX=0$ has only the trivial solution.

$\Leftrightarrow AX=0$ has no free variable $\Leftrightarrow$ A has n pivot columns (basic variable)

$\Leftrightarrow \mathrm{rank}(A)=n$ $\Leftrightarrow$ 解空间 Null space of A, i.e. $N(A)=\{X\in R^n\,|\,AX=0\}=\{0\}$

i.e. $\dim N(A)=0$.

2° $AX=0$ has nontrivial solution (infinite many)

$\Leftrightarrow \mathrm{rank}(A)=r<n$ $\Leftrightarrow$ $AX=0$ has $(n-r)$ free variable

$\Leftrightarrow \dim(N(A))=n-r$. In this case. say $\xi_1,\xi_2,\dots\xi_{n-r}$ are a basis of $N(A)$.

general solutions (基础解系) Then for any $X\in N(A)$, we have $X=t_1\xi_1+t_2\xi_2+\cdots+t_{n-r}\xi_{n-r}$

where $t_1,t_2,\dots t_{n-r}$ are scalars 基础解系所含解的个数 $n-r$

(A2) Nonhomogenous LS $AX=\beta$ ($\beta\in R^n$. $\beta\neq 0$) $AX=\beta$ (1) $AX=0$ (2) [(2)是(1)导出组]

解集合不是子空间

1° if X_1 and X_2 are solution of (1). then X_1-X_2 is solution of (2)

2° if X is a solution of (1). Y is a solution of (2) $\Rightarrow X+Y$ is a solution of (1).

(a) The LS (1) has unique solution $\Leftrightarrow R(A)\underset{\text{有解}}{=}R(A|\beta)\underset{\text{唯一}}{=}n \leftrightarrow$ (列向量组线性无关)

(b) The LS (1) has infinitely solution $\Leftrightarrow R(A)=R(A|\beta)=r<n$

Let η^* be a particular (特解) solution of (1) and let $\xi_1,\xi_2,\dots\xi_{n-r}$, be a basis of $N(A)$.

Then the general solution of $AX=\beta$ is $X=\eta^*+t_1\xi_1+t_2\xi_2+\cdots+t_{n-r}\xi_{n-r}$ ($t_1\dots t_{n-r}$ are scalars)

proof: is solution substitute into. $A\eta^*+A\cdot t_1\xi_1+At_2\xi_2+\cdots At_{n-r}\xi_{n-r}$

any solution form $X-\eta^*\in N(A)$ $X-\eta^*=t_1\xi_1+t_2\xi_2+\cdots+t_{n-r}\xi_{n-r}$

Chapter 4. Algebra of Matrices

§1. Operations.

(A) Linear operations.

Let $A=(a_{ij})_{m\times n}$, $B(b_{ij})_{p\times q}$ we say $A=B \Leftrightarrow m=p,\ q=n.\ a_{ij}=b_{ij}.\ (1\le i\le m,\ 1\le j\le n)$.

<u>Def. 1</u>. 数乘 Let $A=(a_{ij})_{m\times n}$, k a scalar, define $kA=(ka_{ij})_{m\times n}$.

properties (1) $1\cdot A=A$ (2) $(k+l)A=kA+lA$ (3) $(kl)A=k(lA)$ (4) $kA=Ak$.

<u>Def 2</u>. 加法 Let $A=(a_{ij})_{m\times n}$, $B=(b_{ij})_{m\times n}$. define $A+B=(a_{ij}+b_{ij})_{m\times n}$. <u>同型矩阵才可相加</u>

<u>warning</u> $A+B$ is defined only if they have the same size.

properties: (1) commutive. (2) associative

(3) $k(A+B)=kA+kB$. (distributive with scalar).

(4) <u>Def 3</u>. The zero matrix whose entries are 0. is called the zero matrix denoted by $O_{m\times n}$ or O. ($\boxed{!}$ $O_{3\times 2}\neq O_{2\times 2}$. size).

$O+A=A$.

<u>Def 4</u>. 负矩阵 Let $A=(a_{ij})_{m\times n}$. $-A=(-1)A=(-a_{ij})_{m\times n}$ negative matrix of A.

property: $A+(-A)=0$.

(B) Multiplication of Matrices.

row-vector rule $AX=\beta$. row A = column X.

<u>Def 5</u>. Let $A=(a_{ij})_{m\times n}$, $B=(b_{ij})_{n\times p}$

Define $AB=C=(C_{ij})_{m\times p}$ where.

$$C_{ij}=[a_{i1},a_{i2},\cdots a_{in}]\begin{bmatrix}b_{1j}\\ b_{2j}\\ \vdots\\ b_{nj}\end{bmatrix}$$ A 第i行 × B 第j列.

$= a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots a_{in}b_{nj}.\ (1\le i\le m,\ 1\le j\le p)$.

<u>warning</u>. (1) AB (no sign between them) is defined $\Leftrightarrow$ the number of columns of A equals the number of rows of B.

(2) There are divisors of zero $AB=0$. $(A\neq 0)\wedge(B\neq 0)$ can be held.

(3). In general, $AB=AC \not\Rightarrow B=C$ (cancellation law not exists)

<u>properties</u>. Let $A.B.C$ be matrices.

(1) $A(BC)=(AB)C$ (both are multiplicative)

(2) $A(B+C)=AB+AC$

(3) $(A+B)C=AC+BC$

(4) $k(AB)=(kA)B=A(kB)$

proof (1): Let $A=(a_{ij})_{m\times n}$, $B=(b_{ij})_{n\times p}$, $C=(C_{ij})_{p\times q}$.

$AB=(d_{ij})_{m\times p}$ $(AB)C=(e_{ij})_{m\times q}$. $BC=(f_{ij})_{n\times q}$ $A(BC)=(g_{ij})_{m\times q}$

$$e_{ij}=[d_{i1}\cdots d_{ip}]\begin{bmatrix}C_{1j}\\ \vdots\\ C_{pj}\end{bmatrix}=d_{i1}C_{1j}+\cdots+d_{ip}C_{pj}=(a_{i1}b_{11}+a_{i2}b_{21}+\cdots+a_{in}b_{n1})C_{1j}+\cdots+(a_{i1}b_{1p}+\cdots a_{in}b_{np})\cdot C_{pj}$$
$$=a_{i1}(b_{11}C_{1j}+b_{12}C_{2j}+\cdots+b_{1p}C_{pj})+\cdots+a_{in}(b_{n1}C_{1j}+\cdots b_{np}C_{pj})$$
$$=a_{i1}f_{1j}+\cdots a_{in}f_{nj}=(g_{ij})_{m\times q}$$

(C) Special Matrices.

Let A be $m\times n$ matrices.

(a) $A = \begin{bmatrix} a_{11} & & 0 \\ & a_{22} & \\ 0 & & \ddots \\ & & & a_{nn} \end{bmatrix}$ diagonal matrix.

$= diag(a_{11}\ a_{22} \dots a_{nn})$

(b) scalar matrix.

$A = diag(c\ c \cdots c)_{n\times n}$

(c) Identity Matrix ($c=1$. in (b))

denote by I_n , $diag(c\,c\cdots c)_{n\times n} = cI_n$.

$I_n \cdot A_{n\times s} = A_{n\times s}$

$A_{s\times n} \cdot I_n = A_{s\times n}$.

For any $n\times n$ matrix A, $\underline{I_nA = AI_n = A}$.

(d) upper triangular Matrix. (lower. similarly).

$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & & \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$

Def. Transpose of a matrix.

Let. $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{bmatrix}_{m\times n}$. The $m\times n$ matrix. A^T is called the transpose of A

$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}_{n\times m}$

properties. (1) $(A^T)^T = A$ (2) $(A+B)^T = A^T + B^T$ (3) $(AB)^T = B^TA^T$ (4) $(KA)^T = kA^T$

(3) in general $(A_1A_2 \dots A_s)^T = A_s^T A_{s-1}^T \cdots A_2^T A_1^T$

§2. Inverse of a matrix. (可逆矩阵)

Def. Let A be an $n\times n$ matrix. If there exists an $n\times n$ matrix B. s.t. $AB = I_n$, $BA = I_n$. then B is an inverse of A. and A is said to be invertible
nonsingular 非退化的

e.g. invertible: I_n, $cI_n\,(c\neq 0)$.

noninvertible. 0. (proof by contradiction).

Fact. The inverse of A (if it exists) is unique. denote by A^{-1}

proof: If $AB = I_n = BA$, $AC = I_n = CA$. ~~and~~

$B = BI_n = B(AC) = (BA)C = I_nC = C$

Elementary Matrices.

There are also three types of elementary column operations of an $m\times n$ matrix A:

Type 1 $c_i \leftrightarrow c_j$ interchange

Type 2 $kc_i\,(k\neq 0)$ multiply.

Type 3. $c_j + kc_i$ Add k times the ith column to the jth column

Def. An elementary matrix E is one obtained by a single elementary operation on an identity matrix.

e.g. $E_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

$E_1 A = \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{bmatrix} \xrightarrow{r_1 \leftrightarrow r_3} A.$ $(E_1 \xleftrightarrow{r_1 \leftrightarrow r_3} E)$. 做行变换. 左乘相应初等矩阵
做列变换. 右乘相应初等矩阵.

Fact 2. (1) An elementary row (resp. column) operation on a matrix A is equivalent to premultiplying (resp. post-) by an elementary matrix E, which is obtained by performing the same row (resp. column) operation on I.

postmultiply 右乘 premultiply 左乘.

Properties of Elementary Matrixes.

(1) Every elementary Matrix E is invertible.

Three types. Type 1. $E(i,j)$ = interchange $c_i \leftrightarrow c_j$. / $r_i \leftrightarrow r_j$
Type 2. $E(i(k))$ = multiple ith row/column by scalar k (nonzero)
☆ Type 3. $E(i,j(k))$ = jth row k times to ith row
i column k times to jth column

the inverse of elementary matrix is elementary matrix.

(2) $E(i,j)^{-1} = E(i,j)$.

$E(i(k))^{-1} = E(i(\frac{1}{k}))$ $(k \neq 0)$

$E(i,j(k))^{-1} = E(i,j(-k))$. $(i \neq j)$

(3) The determinant of every elementary matrix is not equal to zero

$|E(i,j)| = -1$.

$|E(i(k))| = k$.

$|E(i,j(k))| = 1$.

Theorem An $n \times n$ matrix A is invertible $\Leftrightarrow$ A is row equivalent to I_n. and in the case, any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1}

proof: "$\Rightarrow$" Suppose A is invertible
Consider $AX = b$. then we have $(A^{-1}A)X = A^{-1}b$, i.e. $X = A^{-1}b$.
Since $AX = b$ has a solution for each $b \in \mathbb{R}^n$. A has a pivot position in every row. Because A is square, then n pivot position must be on the diagonal. which implies, that the reduced row echelon form of A is I_n. That is, A is row equivalent to I_n

"$\Leftarrow$" If A is row equivalent to I_n,
Then, since each step of the row reduction of A corresponding to premultiplication by an elementary matrix, there exists elementary matrixes $E_1, E_2, \dots E_p$
s.t. $A \Rightarrow E_1A \Rightarrow E_2E_1A \to \cdots \to E_pE_{p-1} \cdots E_2E_1A = I_n$ (more: 一个可逆矩阵是一些初等矩阵乘积)
i.e. $\underbrace{E_pE_{p-1} \cdots E_2E_1}_{A^{-1}}A = I_n$. $\Rightarrow A = (E_p \cdots E_2E_1)^{-1} I_n = E_1^{-1}E_2^{-1}E_3^{-1} \cdots E_p^{-1}$

* ! If A, B are invertible. matrices then AB is invertible $(AB)^{-1} = B^{-1}A^{-1}$.

A is invertible $\Leftrightarrow A^T$ is invertible. $(A^{-1})^T = (A^T)^{-1}$.

Coro 1. An $n \times n$ matrix A is invertible $\Leftrightarrow$ A is a product of a sequence of elementary matrices.

Algorithm for finding A^{-1}.

Row reduce the argumented matrix $[A|I]$. If A is row equivalent to I. then $[A|I]$ is row equivalent to $[I|A^{-1}]$

§3. Product Theorem of Determinant

Thm2. If A is an matrix and E is an elementary matrix of size $n\times n$. then

(1) $|EA| = |E||A|$ (2) $|AE| = |A||E|$

(proof: (1) discuss respectively 3 kinds of row operation. (2). column operation).

Coro1. Let A be $n\times n$ matrix. $E_1 \dots E_p$ are $n\times n$ elementary matrices.

① $|E_1E_2\cdots E_pA| = |E_1||E_2|\cdots|E_p||A|$.

② $|AE_1E_2\cdots E_p| = |A||E_1|\cdots|E_p|$

Coro2. Let A be $n\times n$ matrix. Then

(1) If A is nonsingular, then $|A|\neq 0$.

(2) If A is singular, then $|A| = 0$.

Thm3. (Product of determinant).

If A and B are $n\times n$ matrices, then $|AB| = |A||B| = |B|\cdot|A| = |BA|$.

不可逆矩阵的乘积不可能是可逆阵.

proof: 1°. A is singular, then $|A| = 0$.

If AB is nonsingular, then there exists a matrix C. s.t. $(AB)C = I_n$.

i.e. $A(BC) = I_n$. $\Rightarrow$ BC is a inverse of A. contradicts.

Hence AB is ~~no~~ singular $|AB| = 0$

2° A is nonsingular. Then A is a product of elementary matrices.

$A = E_1\dots E_p$ so $|AB| = |E_1E_2\cdots E_pB| = |E_1|\cdot|E_2|\cdots|E_p||B| = |E_1E_2\cdots E_pI_n|\cdot|B| = |A|\cdot|B|$

§4. Adjoint of A Matrix

Def. Let $A = (a_{ij})$ be $n\times n$ matrix. Let A_{ij} be the cafactor of a_{ij} in $|A|$. for $1\le i,j\le n$.

Then the $n\times n$ matrix $A^* = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}^T = \begin{bmatrix} A_{11} & \cdots & A_{n1} \\ A_{12} & & \vdots \\ \vdots & & \vdots \\ A_{1n} & \cdots & A_{nn} \end{bmatrix}$ is called the adjoint (matrix) of A.

we write as A^* or adj.(A)

$a_{11}A_{11} + a_{12}A_{12} + \cdots a_{1n}A_{1n} = A$.

$a_{11}A_{21} + a_{12}A_{22} + \cdots a_{1n}A_{2n} = 0$.

Fact $AA^* = A^*A = |A|I_n$

proof: $AA^* = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix} = \begin{bmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & |A| \end{bmatrix}$

Thm4. Let A be $n\times n$ matrix Then A is invertible $\Leftrightarrow |A|\neq 0$.

proof: $\Leftarrow$ $AA^* = A^*A = |A|I_n$

$\therefore A\frac{A^*}{|A|} = \frac{A^*}{|A|}A = I_n$. $\therefore A$ is invertible. $A^{-1} = \frac{1}{|A|}A^*$

$\Rightarrow$ Thm2. Coro2.

$r(A^*) = \begin{cases} n, & r(A) = n \\ 1, & r(A) = n-1 \\ 0, & r(A) < n-1 \end{cases}$

Def: Let A and B be matrices. If B is obtained from A by a sequence of elementary operations then A and B are equivalent, denote by $A \sim B$.

Fact. $\operatorname{rank}(A) = \operatorname{rank}(B)$.

Thm: Let A, B be $m \times n$ matrices. Then

$A \sim B \Leftrightarrow \exists$ elementary matrices $P_1, P_2, \dots P_s$. $Q_1, Q_2 \dots Q_t$.

~~结~~ s.t. $B = (P_s \cdots P_1) A (Q_1 \cdots Q_t)$.

$\Leftrightarrow$ i.e. $\exists$ invertible matrices P and Q s.t. $B = PAQ$.

Thm: Let A be an $m \times n$ matrix. If $\operatorname{rank}(A) = r$, then A is equivalent to unique (可用列变换). matrix of the form.

$$D = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

normalized form (matrix.

§5. Partitioned Matrices.

(A) Def. A matrix can be splited into several parts (submatrices / blocks) by means of horizontal and/or vertical.

partitioned by rows $A = [a_{ij}]_{m \times n} = [\alpha_i]_m$

by columns $A = [a_{ij}]_{m \times n} = [\beta_j \cdots]_{1 \times n}$

(B) Operations.

(i) Addition. $A = (a_{ij})_{m \times n}$ $B = (b_{ij})_{m \times n}$. $A + B = (A_{ij} + B_{ij})_{s \times t}$

$= (A_{ij})_{s \times t}$ $= (B_{ij})_{s \times t}$. (A_{ij}. B_{ij} are the same size)

(ii) scalar multiplication.

(iii) Multiplication Let A be an $m \times n$ matrix, B be an $n \times p$ matrix.
If the column partition of A matches the row partition of B.
Then A and B can be multiplied by the usual row-column rules

e.g. if A is $m \times n$ B is $n \times p$. $B = [\beta_1\ \beta_2 \cdots \beta_p]$. $A = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix}$

then $AB = [A\beta_1\ A\beta_2\ \cdots\ A\beta_p]$

$$AB = \begin{bmatrix} \alpha_1\beta_1 & \cdots & \alpha_1\beta_p \\ \alpha_2\beta_2 & \cdots & \alpha_2\beta_p \\ \vdots & & \vdots \\ \alpha_m\beta_n & \cdots & \alpha_m\beta_p \end{bmatrix}$$

(c) Transpose: All blocks should change (transpose

e.g. $A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$ $A^T = \begin{bmatrix} A_{11}^T & A_{21}^T \\ A_{12}^T & A_{22}^T \\ A_{13}^T & A_{23}^T \end{bmatrix}$

(d) Special partitioned Matrices.

(1) Block diagonal matrix.

$A = \begin{bmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_s \end{bmatrix}$. A_i are square matrice. $1 \le i \le s$ $A^{-1} = \begin{bmatrix} A_1^{-1} & & & 0 \\ & A_2^{-1} & & \\ & & \ddots & \\ 0 & & & A_s^{-1} \end{bmatrix}$

$|A| = |A_1|\,|A_2| \cdots |A_s|$ A is invertible $\Leftrightarrow$ each block A_i is invertible ($\Leftrightarrow |A| \neq 0$).

Fact: if A and B are $n \times n$ matrices. $AB = I_n \Rightarrow A$ is invertible (无需验证 $BA = I_n$.

proof: $|AB| = |I_n| \Rightarrow |A| \cdot |B| = 1 \Rightarrow |A| \neq 0$.

(2) Block up (lower) triangular Matrix.

$$A=\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ & A_{22} & \cdots & A_{2n} \\ & & \ddots & \vdots \\ 0 & & & A_{nn} \end{bmatrix} \qquad \underline{A_{ii} \text{ is a square}}$$

$$|A| = |A_{11}|\cdot|A_{22}|\cdots|A_{nn}|.$$

(D) Elementary operations on a partitioned matrix. (乘初阵!).

(i) Interchange.

(ii) premultiply the ith row by an $\underset{\text{可逆矩阵}}{\underline{\text{invertible matrix } P}}$. $\quad Pr_i$

postmultiply the jth column by an invertible matrix $P \quad c_jP$.

(iii) premultiply the ith row by a $\underline{\text{matrix } K}$ and then add to the jth row $(i\neq j)$ $\quad r_j+Kr_i$

or postmultiply. $K \quad c_j + c_iK$.

<u>Def 2.</u> A partitioned elementary matrix is obtained by a single elementary operation on block identity matrix.

e.g. $\begin{bmatrix} I_m & 0 \\ 0 & I_n\end{bmatrix} \to \begin{bmatrix} 0 & I_n \\ I_m & 0\end{bmatrix}$

$\qquad \to \begin{bmatrix} P & 0 \\ 0 & I_n\end{bmatrix} \begin{bmatrix} I_m & 0 \\ 0 & P\end{bmatrix}$ (P is invertible)

$\qquad \to \begin{bmatrix} I_n & K \\ 0 & I_n\end{bmatrix} \begin{bmatrix} I_m & 0 \\ K & I_n\end{bmatrix}$

<u>Fact.</u> (1) Partition elementary matrice are invertible

(2) If an elementary row (column) operation is performed on $M=\begin{bmatrix} A & B \\ C & D\end{bmatrix}$. the resulting matrix can be written as EM (ME). Where E is created by performing the same operation on $\begin{bmatrix} I_m & 0 \\ 0 & I_n\end{bmatrix}$

e.g. $A=(a_{ij})_{m\times n}$. Suppose $\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk}\end{vmatrix} \neq 0$. $\quad$ ($1\le k\le n$. 顺序主子式), 共n个).

such that there exists a lower triangular matrix B. s.t. BA is upper triangular matrix.

proof: (induction). suppose ... holds for $n-1$

$$A_1 = \begin{bmatrix} a_{11} & \cdots & a_{1n-1} \\ \vdots & & \vdots \\ a_{n-1\,1} & \cdots & a_{n-1\,n-1}\end{bmatrix}$$

there exist an $(n-1)\times(n-1)$ lower triangular matrix B_1 such that B_1A_1 is $(n-1)\times(n-1)$ upper ...

Now consider $A=\begin{bmatrix} A_1 & \beta \\ \alpha & a_{nn}\end{bmatrix} \xrightarrow{r_2-\alpha A_1^{-1}r_1} \begin{bmatrix} A_1 & \beta \\ 0 & a_{nn}-\alpha A_1^{-1}\beta\end{bmatrix} \xrightarrow{B_1r_1} \begin{bmatrix} B_1A_1 & B_1\beta \\ 0 & a_{nn}-\alpha A_1^{-1}\beta\end{bmatrix}$

$$\begin{bmatrix} I_{n-1} & 0 \\ 0 & 1\end{bmatrix} \xrightarrow{r_2-\alpha A_1^{-1}r_1} \begin{bmatrix} I_{n-1} & 0 \\ -\alpha A_1^{-1} & 1\end{bmatrix} = E_1.$$

$$\xrightarrow{B_1r_1} \begin{bmatrix} B_1 & 0 \\ 0 & 1\end{bmatrix} = E_2. \qquad B=E_2E_1$$

e.g. Let $A_{m\times n}$. $B_{n\times m}$. λ a scalar.

*分块换法,结果乘$(-1)^{\cdots}$

$$\begin{vmatrix} \lambda I_n & B \\ A & I_m\end{vmatrix} = |\lambda I_n - BA| = \lambda^{n-m}|\lambda I_m - AB|. \quad \text{(行列式降阶公式.)}$$

proof: $\begin{vmatrix} \lambda I_n & B \\ A & I_m\end{vmatrix} \overset{r_1-Br_2}{=\!=\!=} \begin{vmatrix} \lambda I_n - BA & 0 \\ A & I_m\end{vmatrix} = |\lambda I_n - BA|$

$\overset{c_2+c_1(-\frac{1}{\lambda}B)}{=\!=\!=\!=} \begin{vmatrix} \lambda I_n & 0 \\ A & I_m-\frac{1}{\lambda}AB\end{vmatrix} = |\lambda I_n|\cdot|I_m-\frac{1}{\lambda}AB| = \lambda^n\cdot|I_m-\frac{1}{\lambda}AB| = \lambda^{n-m}|\lambda I_m - BA$

每行一个λ!

e.g. find ranks of matrices.

Recall properties. $A_{m\times n}$.

(1) $0 \le r(A) \le \min\{m,n\}$.

(2) $\text{rank}(A^T) = \text{rank}(A)$.

(3) If A_1 is a submatrix of A. then $r(A_1) \le r(A)$

(4) Let k be a scalar then $\text{rank}(kA) = \begin{cases} 0, & k=0 \\ \text{rank}(A). & k\ne 0. \end{cases}$

(5) $\text{rank}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = r(A) + r(B)$.

proof: suppose $\text{rank}(A) = r$. $\text{rank}(B) = s$.

$$A \sim \begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix}. \quad B \sim \begin{bmatrix} E_s & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \sim \begin{bmatrix} E_r & 0 & 0 & \\ 0 & 0 & & \\ 0 & & E_s & 0 \\ & & 0 & 0 \end{bmatrix} \sim \left[\begin{array}{cc|c} E_r & & 0 \\ & E_s & \\ \hline 0 & & 0 \end{array}\right]. \quad r\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = r+s$$

(6). $\text{rank}\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \ge \text{rank}(A) + \text{rank}(B)$.

proof: similar as (5). $\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \sim \left[\begin{array}{cc|cc} E_r & 0 & D_{11} & D_{12} \\ 0 & 0 & D_{21} & D_{22} \\ \hline 0 & & E_s & 0 \\ & & 0 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} E_r & & & \\ & E_s & & 0 \\ & & D_{22} & \\ \hline & 0 & & 0 \end{array}\right]$

(7). $\text{rank}[A \quad B] \le \text{rank}(A) + \text{rank}(B)$

$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \to \begin{bmatrix} A & B \\ 0 & B \end{bmatrix}$. $[A \ B]$ is a submatrix of $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$.

(8) $\text{rank}(A+B) \le \text{rank}(A) + \text{rank}(B)$

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \to \begin{bmatrix} A & B \\ 0 & B \end{bmatrix} \to \begin{bmatrix} A+B & B \\ B & B \end{bmatrix}.$$

(9). $\text{rank}(AB) \le \min\{r(A). r(B)\}$.

proof: $A_{m\times n}$. $B_{n\times p}$

$$[A \ \ O_{m\times p}] \xrightarrow{C_2 + C_1 B} [A \ \ AB] \qquad \begin{bmatrix} B \\ O_{n\times p} \end{bmatrix} \to \begin{bmatrix} B \\ AB \end{bmatrix}$$

(10). If A is $m\times n$ and B is $n\times p$ then.

$\text{rank}(AB) \ge \text{rank}(A) + \text{rank}(B) - n$.

proof: $\begin{bmatrix} A & 0 \\ I_n & B \end{bmatrix} \to \begin{bmatrix} 0 & -AB \\ I_n & B \end{bmatrix} \to \begin{bmatrix} 0 & -AB \\ I_n & 0 \end{bmatrix} \to \begin{bmatrix} -AB & 0 \\ 0 & I_n \end{bmatrix} \to \begin{bmatrix} AB & 0 \\ 0 & I_n \end{bmatrix}$

$$\text{rank}(AB) + n = \text{rank}\begin{bmatrix} AB & 0 \\ 0 & I_n \end{bmatrix} = \text{rank}\begin{bmatrix} A & 0 \\ I_n & B \end{bmatrix} \ge \text{rank}(A) + \text{rank}(B).$$

* Symmetric Matrix : $A = A^T$ $(n\times n)$. $\Leftrightarrow a_{ij} = a_{ji}$ $1 \le i,j \le n$

Chapter 5 Quadratic Form.

(1) Def. A quadratic form in variable $x_1 \dots x_n$ is an expression of the form

$$f(x_1, x_2, \dots x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} x_i x_j$$ (where the coefficient b_{ij} are scalar in the ground field).

when $F = \mathbb{R}$, we call f is a real quadratic; $\mathbb{C}$, complex ...

$\Downarrow$ polynomial in several variables.

(2) Matrix representation of quadratic form.

Let $A = (b_{ij})_{n\times n}$. $b_{ij} = b_{ji}$. always have $A = A^T$ (symmetric).

Find the quadratic form matrix.

step 1. $a_{ii} = b_{ii}$ $1 \le i \le n$.

step 2. $a_{ij} = \frac{b_{ij} + b_{ji}}{2} = a_{ji}$. $1 \le i, j \le n$. $i \ne j$.

Let $\bar{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ $\quad f = \bar{X}^T A \bar{X} = (x_1, x_2, \cdots x_n) \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

we call the rank(A) the rank of f. i.e. rank(f).

(3) Change of variables in a quadratic form.

Let $f(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j = X^T A X$ (二次型 $\Leftrightarrow$ 对称矩阵).

A change of variables is an equation of the form $\begin{cases} x_1 = c_{11}y_1 + c_{12}y_2 + \cdots + c_{1n}y_n \\ x_2 = c_{21}y_1 + c_{22}y_2 + \cdots c_{2n}y_n \\ \cdots \\ x_n = c_{n1}y_1 + c_{n2}y_2 + \cdots + c_{nn}y_n \end{cases}$

or equivalently. $X = PY$. where $P = [c_{ij}]$ $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ and P is invertible.

In this case, the transformation $X = PY$ is called a nonsingular/invertible linear transformation.

A nonsingular linear transformation $X = PY$ maps a quadratic form $f(x) = X^T A X$ into another one: $f = X^T A X = (PY)^T A (PY) = Y^T P^T A P Y = Y^T (P^T A P) Y$

set $B = P^T A P$. $B^T = P^T A^T P = P^T A P = B$. $Y^T B Y$ is a new quadratic form

(4). Def. congruent matrix

a nonsingular matrix P. s.t. $B = P^T A P$. A, B are said to be congruent. (莫文书: 前提 ~~__~~对称矩阵)

Def. Quadratic forms associated to ~~a~~ congruent matrices are said to be equivalent ($A \sim B$)

$A \sim B$ (二次型等价) $\Leftrightarrow$ $\exists$ elementary matrices $P_1 \dots P_s$, $Q_1 \dots Q_t$ s.t. $B = P_s \cdots P_1 A Q_1 \cdots Q_t$ (不一定要是方阵 但可逆阵是方阵)

i.e. $\exists$ invertible P, Q. $B = PAQ$.

the real symmetric matrix must have a congruent-diagonal matrix (for quadratic form, only have x_i^2 terms).

Remark: A symmetric $n \times n$ matrix can't be congruent to an nonsymmetric matrix.

(5) Standard forms (标准)

Def. A quadratic form f is called a standard (quadratic form) if f is a sum of squares

that is, $f = d_1 x_1^2 + d_2 x_2^2 + \cdots + d_n x_n^2$.

最终都是等价变换. 秩不变

Thm 1. Every quadratic form $f(x)=X^TAX$ over F is equivalent to a standard form (by pairwise elementary operation). $\boxed{P^T}\boxed{A}\boxed{P}$

"把二次型对角化"

the method of Lagrange.

e.g. $f(x_1,x_2,x_3) = x_1^2+2x_2^2+2x_1x_2+x_3^2-4x_2x_3$

$$= (x_1+x_2)^2 + (x_2-2x_3)^2 - 3x_3^2$$

Let. $y_1 = \cdots \quad y_2 = \cdots \quad y_3 = \cdots \qquad f(y_1,y_2,y_3) = y_1^2+y_2^2-3y_3^2$

$z_1 = \cdots \quad z_2 = \cdots \quad z_3 = \cdots \qquad f(z_1,z_2,z_3) = z_1^2+z_2^2-z_3^2$ (非0.平方项个数=.秩=).

Remark: There are infinitely many ways to reduce a quadratic form to standard forms. but the number of nonzero square terms is equal to ...

Normal Forms. (规范型).

(1) Complex quadratic form 等秩⟺合同

Thm 2. Every complex quadratic form $f=X^TAX$ ~~is eq~~ equivalent to a normal form.: $f=y_1^2+y_2^2+\cdots+y_r^2$ where $r=\operatorname{rank}(A)$. and the normal form is unique

Every complex symmetric matrix is congruent to a diagonal matrix of the type $\begin{bmatrix} \overbrace{1 \ \ddots \ 1}^{r} & \\ & 0 \ \ddots \ 0 \end{bmatrix}$ where $r=\operatorname{rank}(A)$.

(2) Real quadratic form

Let f be a real quadratic form. Reduce f to a standard form

$f = d_1y_1^2+d_2y_2^2+\cdots+d_py_p^2-d_{p+1}y_{p+1}^2-d_ry_r^2$. $0<d_i\in\mathbb{R}$. $1\le i\le r$. $r=\operatorname{rank}(f)$.

$\to d_k = \dfrac{\det D_k}{\det D_{k-1}}$, denote $\det D_0 = 1$.

$z_1 = \frac{1}{\sqrt{d_1}}y_1 \ \cdots \ z_r = \frac{1}{\sqrt{d_r}}y_r \ (z_{p+1}=y_{p+1} \cdots z_n = y_n.)$ 令0.

$$\begin{bmatrix} \overbrace{1 \ \ddots \ 1}^{p个} & & \\ & \overbrace{-1 \ \ddots \ -1}^{r-p个} & \\ & & \overbrace{0 \ \ddots \ 0}^{n-r个} \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{d_1}} & & & \\ & \frac{1}{\sqrt{d_2}} & & \\ & & \ddots & \\ & & & \frac{1}{\sqrt{d_r}} \\ & & & & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \Leftrightarrow Y = CZ$$

Then $f = z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots z_r^2$

def: $(p, r-p)$ is called signature.

(Jacobi rule of signs.)

Thm 3. (Sylvester's law of inertia). 惯性定律 Every real quadratic form can be reduced a unique normal form (under nonsingular linear transformations).

proof: (p is unique).

It suffices to prove the uniqueness. Suppose that a real quadratic form $f(x_1,\dots x_n)$ is reduced to a normal form. $f=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots y_r^2$

(by the nonsingular linear transformation $X=BY$.).

Another change of variable $X=CZ$. reduces f to $f=z_1^2+\cdots+z_q^2-z_{q+1}^2-\cdots z_r^2$

We have $Z=C^{-1}X=C^{-1}BY=(C^{-1}B)Y$.

signature of A (quadratic) $\left[\begin{array}{l} p \to n_+(A) \text{ positive signature of interia of } f. \\ r-p \to n_-(A) \text{ negative} \\ n-r \to n_0(A) \end{array}\right.$: $n_+(A)-n_-(A)$ the difference of signature

$$C^{-1}B = G = \begin{bmatrix} g_{11} & \cdots & g_{1n} \\ g_{21} & & \vdots \\ g_{n1} & & g_{nn} \end{bmatrix}$$

Hence, (3) is equivalent to

$$\begin{cases} z_1 = g_{11}y_1+g_{12}y_2+\cdots+g_{1n}y_n \\ z_2 = g_{21}y_1+g_{22}y_2+\cdots+g_{2n}y_n \\ \vdots \\ z_n = g_{n1}y_1+g_{n2}y_2+\cdots+g_{nn}y_n \end{cases}$$

Assume that $p>q$. consider the LS

$$\begin{cases} g_{11}y_1+g_{12}y_2+\cdots+g_{1n}y_n=0. \\ \vdots \\ g_{q1}y_1+g_{q2}y_2+\cdots g_{qn}y_n=0 \\ y_{p+1}=0 \\ \vdots \\ y_n=0. \end{cases} \quad (4)$$

we see that the LS has n variables and has $q+(n-p) = n-(p-q) < n$.

Hence. LS. has at least one nonzero solution say: $y_1=k_1, \cdots y_p=k_p, \cdots y_{p+1}=k_{p+1} \cdots y_n=k_n$ is nonzero solution of LS. (4) $k_{p+1}=\cdots=k_n=0$.

Then $f = y_1^2+\cdots+y_p^2-y_{p+1}^2-y_r^2$

$$= k_1^2+\cdots+k_p^2 > 0.$$

$f = z_1^2+\cdots+z_q^2-z_{q+1}^2-z_r^2$

$$= -z_{q+1}^2-\cdots-z_r^2 < 0. \text{ (Thus. } p\le q\text{. similarly } q\le p\text{).}$$

$\Leftarrow z_1=\cdots z_q=0.$

Thm4. Let A and B be congruent real symmetric matrices: $\Leftrightarrow$ A and B have the same signature. $n_+(A) = n_+(B)$ $(-, 0)$. (normal formal =).

§ 3. Positive Definite Matrices (正定矩阵).

Def1. A real quadratic. $f = X^TAX > 0$. $\forall \bar{X} \in \mathbb{R}^n \setminus 0$. $\Rightarrow$ positive definite.
$\hookrightarrow$ (所有 $x \neq 0$).

(sometimes. ≥ 0. positive semidefinite. when $n > r$).

$f = d_1x_1^2 + \cdots + d_nx_n^2$ is positive definite $\Leftrightarrow$ $d_i > 0$. $(1 \leq i \leq n)$

Fact. If $f = X^TAX = Y^T(C^TAC)Y = Y^TBY$. $\forall 0 \neq Y \in \mathbb{R}^n$. $X = CY \neq 0$. (C is invertible)
and A is Positive definite. B is. (非退化线性变形不改正定性)

Thm1. A real quadratic is positive definite $\Leftrightarrow$ it's positive interia of is n.

Coro1. The normal form of a positive definite quadratic from $f(x) = X^TAX$ is the type: $y_1^2 + y_2^2 + \cdots + y_n^2$

Def2. real symmetric matrix is positive definite. if $X^TAX > 0$. for all $0 \neq \bar{X} \in \mathbb{R}^n$.

coro2. A real symmetric matrix: positive definite $\Leftrightarrow$ congruent to an identify matrix
(moreover, we have. $|A| > 0$. A must be invertible).

proof: As A is positive definite. exists. an invertible matrix C.
s.t.: $A = C^TIC$ i.e. $|A| = |C^T| \cdot |C| = |C|^2 > 0$.

Fact: The sum of positive definite matrice is positive. (列满秩)

Let A be real matrix. if the columns of A are independent, then $B = A^TA$ is positive definite

since $B^T = (A^TA)^T = A^TA = B$. B is real symmetric
$X^TBX = X^TA^TAX = (X^TA^T)(AX)$ (for every $0 \neq \bar{X} \in R^n$)
$= (AX)^T(AX)$

Charaterization of positive definite matrices

Let $A = (a_{ij})_{n \times n}$. $A_r = \begin{bmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \\ a_{r1} & \cdots & a_{rr} \end{bmatrix}_{r \times r}$ $1 \leq r \leq n$

we call A_r leading principal (顺序主子矩阵) submatrix
$|A_r|$ leading principal minor of order r of A (顺序主子式).

Thm2.
a matrix is positive finite $\Leftrightarrow$ each of it's leading principal minor > 0.

proof: "necessity $\Rightarrow$"
Assume that $f(x) = X^TAX$ is positive definite for $1 \leq r \leq n$. let $X_r = (x_1, x_2, \ldots x_r)^T$ be any nonzero vector in $\mathbb{R}^r$ and let $X = (x_1, x_2, \ldots x_r, 0, \ldots 0)^T \in \mathbb{R}^n$. So $X \neq 0$.

Note $A = \begin{bmatrix} A_r & B \\ B^T & C \end{bmatrix}$. $\bar{X} = \begin{bmatrix} \bar{X}_r \\ 0 \end{bmatrix}$

$0 < X^TAX = [\bar{X}_r^T \ 0^T] \begin{bmatrix} A_r & B \\ B^T & C \end{bmatrix} \begin{bmatrix} \bar{X}_r \\ 0 \end{bmatrix}$

$= [\bar{X}_r^T \ 0^T] \begin{bmatrix} A_r \bar{X}_r \\ B^T \bar{X}_r \end{bmatrix} = \bar{X}_r^T A_r \bar{X}_r$

so $\bar{X}_r A_r \bar{X}_r$ is a positive definite quadratic form, and thus A_r is a positive definite matrix. Hence, $|A_r| > 0$. for $1 \leq r \leq n$

"$\Leftarrow$" Induction. suppose holds for $n-1$.

Let $A = \begin{bmatrix} A_{n-1} & \alpha \\ \alpha^T & a_{nn} \end{bmatrix}$ $\quad \alpha = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{n-1\,n} \end{bmatrix}$

By assumption, the leading principal minors of A are all positive. Thus, the leading principal minors of A_{n-1} are all positive. By induction. there $\exists$ G $(n-1)\times(n-1)$. which is invertible s.t. $G^T A_{n-1} G = I_{n-1}$

Now we reduce A by elementary operations:

$$A = \begin{bmatrix} A_{n-1} & \alpha \\ \alpha^T & a_{nn} \end{bmatrix} \to \begin{bmatrix} I_{n-1} & 0 \\ 0 & a \end{bmatrix} \quad (a = a_{nn} - \alpha^T G G^T \alpha)$$

Let $C_1 = \begin{bmatrix} G & 0 \\ 0 & 1 \end{bmatrix}$ $\quad C_2 = \begin{bmatrix} I_{n-1} & -G^T\alpha \\ 0 & 1 \end{bmatrix}$ $\quad$ set $C = C_1 C_2$

$C^TAC = C_2^T C_1^T A C_1 C_2 = \begin{bmatrix} I_{n-1} & 0 \\ 0 & a \end{bmatrix}$. $\quad |C^TAC| = a > 0$

That is A is congruent to I_n. Hence A is positive definite and the corresponding quadrantic form $f = X^TAX$ is positive definite

Def3. A real quadratic form f is called. (defined analogously).

(a) negative definite if $f(x) < 0$. for all $\bar{X} \neq 0$.

(b) positive semidefinite if $f(x) \geq 0$. for all $\bar{X} \in \mathbb{R}^n$

(c) negative semidefinite if $f(x) \leq 0$ for all $\bar{X} \in \mathbb{R}^n$

(d) indefinite if $f(x)$ assumes both positive and negative values.

For (a). we have $(-f)$ is positive definited. $n_-(A) = n$. A is congruent to $-I_n$

the leading principal minor $\quad |A_r| > 0$. r is even
$\quad |A_r| < 0$ $\quad r$ is odd.

proof: $-f = X^T(-A)X > 0$.

$|-A_r| = (-1)^r |A_r| > 0$.

Thm3. Let $f = x^TAX$ be a real quadratic form. The following statements are equivalent.

(a) f is positive semidefinite (i.e. A is positive semidefinite).

(b). $n_+(A) = \text{rank}(A) = \text{rank}(f)$

(c) there exists. a invertible C. s.t.

$$C^TAC = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}, \text{ with } d_i \geq 0.\ 1 \leq i \leq n.$$

(d) there exists a real matrix P. s.t. $A = P^TP$.

Summary

① cofactor-expansion $\sum_{s=1}^{n} a_{ks} A_{is} = \begin{cases} d, & k=i \\ 0, & k \neq i \end{cases}$

$\sum_{s=1}^{n} a_{sl} A_{sj} = \begin{cases} d, & l=j \\ 0, & l \neq j. \end{cases}$

② Vandermonde - Determinant

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^n & a_2^n & & a_n^n \end{vmatrix} = \prod_{1 \le i < j \le n} (a_j - a_i)$$

③ Eisenstein irreducible criterion.

$f(x) = \sum_{i=1}^{n} a_i x^i$ 1) $p \nmid a_n$ 2) $p \mid a_1, \dots a_n$ 3) $p^2 \nmid a_n$.

④ Descarte's rational root theorem.

$f(x) = \sum_{i=1}^{n} a_i x^i$ in $\mathbb{Z}[x]$, $\frac{r}{s}$ is a rational roots $\Rightarrow s \mid a_n$, $r \mid a_0$.

⑤ Homogeneous Linear Systems (Solution judgement)

1). It's coefficient determinant equals to 0. $\Rightarrow$ has nonzero solution (infinite many).

otherwise ($|A| \neq 0$). $\Rightarrow$ only has zero solution.

2). $\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + \cdots + a_{2n}x_n = 0 \\ \vdots \qquad\qquad \vdots \\ a_{s1}x_1 + \cdots + a_{nn}x_n = 0. \end{cases}$

if $s < n$. must exist nonzero solution (when transforming to the row echelon form. $r \le s < n$)

3). $x_1\alpha_1 + \cdots + x_n\alpha_n = 0$. (column vectors)

$\alpha_1, \dots \alpha_n$ linear dependence $\Leftrightarrow$ nontrivial solutions.

4). LS. unique solution $\Leftrightarrow$ coefficient matrix $|A| \neq 0$.

HLS only trivial $\Leftrightarrow |A| \neq 0$. nontrivial $\Leftrightarrow |A| = 0$.

*LS. (1) have solution. ① augment / coefficient matrix has same rank. ($\bar{A}$ / A)

*HLS. (1). general solutions. → a m.l.i.a of solutions set.

(the system has $(n-r)$ solutions)

= 导出组+特解. (derived group + particular solution).

⑥ Laplace Theorem

Def. minor (子式) M. cofactor 余子式

$(-1)^{(i_1+i_2+\cdots+i_k)+(j_1+j_2+\cdots+j_k)} M'$ (the algebraic minor of M).

Thm. $D = M_1A_1 + M_2A_2 + \cdots + M_tA_t$. $t = C_n^k = \frac{n!}{k!(n-k)!}$ (k行元素组成的一阶切k阶子式).

Thm2. $D_1 = |a_{ij}|_{n\times n}$ $D_2 = |b_{ij}|_{n\times n}|$ $D_3 = |C_{ij}|_{n\times n}$

$D_3 = D_1D_2$, $C_{ij} = a_{i1}b_{1j} + \cdots a_{in}b_{nj}$

⑦ Linear Independence.

1) "represent" ⇒ every vector can be represented ⇒ the set can be represented.

"represent each other" ⇒ equivalent. (reflexivity / symmetry / transitivity).

2) $k_1\alpha_1 + k_2\alpha_2 + \cdots k_s\alpha_s = 0$. ($k_i$ not all equal to 0 → dependence
all equal to 0 → independence).

3) vector set α. is linear dependence ⇔ $\alpha = 0$.

the set of unit vectors ε_n is linear independence.

4). $\{\alpha_1, \dots \alpha_r\}$. can be represented by $\{\beta_1, \beta_2 \dots \beta_s\}$. $(r>s)$. $\{\alpha_1, \dots \alpha_r\}$ linear dependence.

(大组被小组表出.⇒ 大组线性相关) . A组线性无关,可被B组表出. 数目 $A \le B$.

Coro. Any $n+1$ many n-dimensional vectors are linear dependence.

Coro Two equivalent linear independent vector sets must have the equal number of vectors

5). $A = (a_{ij})_{m\times n}$. row (column) vectors linearly independence ⇔ $|A| \neq 0$.

⑧ Maximal linearly independent array

1) equalivalent to the set. 2). has the same number of vectors (rank).

3) 0 vectors. rank = 0. no maximal linearly independent array

row vector → basis ($\neq$ m.l.i.a). column vector → basis ($=$ m.l.i.a)
位置对应

⑨ Rank.

1). Matrix ① Def. the highest order of nonzero minor

②row / column / determinant rank equals.

③ elementary row / column operation can't change it.

④ equivalent → the same rank ($\not\leftarrow$)

⑩ Operation. $(AB)C = A(BC)$

$A(B+C) = AB + AC$ $(A+B)C = AC + BC$

E_n. Identity Matrix.

kE_n Scalar Matrix.

$r(AB) \le \min[r(A), r(B)]$

(AB's vector sets can be represented by A/B.).

11. Inverse. $(n \times n)$

① invertible. / nonsingular (可逆/非退化). A^{-1}. unique. (定义中. $AB=BA=I_n$.
实际只需一边. $|AB|=|I_n|=1$.

* row transformation. $E'A$. (左乘)
column transformation AE'. (右乘)

② $|A| \neq 0$. (nonsingular).

③ Adjoint matrix. $A^* = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}$ (转置后的代数余子式矩阵

$AA^* = A^*A = dE$.

④ invertible $\Leftrightarrow$ row equivalent to E_n. A^{-1} a product of a sequence of elementary matrice.

⑤ $(AB)^{-1} = B^{-1}A^{-1}$ (A. B both. AB is).

⑥ represented as a sequence of elementary matrices 's. product. $\Leftrightarrow$ invertible

12. Partition.

1) Block diagonal. $A = \begin{bmatrix} A_1 & & 0 \\ & A_2 & \\ 0 & & \ddots \ A_n \end{bmatrix}$. A_i are square matrices

A invertible. $A^{-1} = \begin{bmatrix} A_1^{-1} & & 0 \\ & A_2^{-1} & \\ 0 & & \ddots \ A_n^{-1} \end{bmatrix}$ each invertible

2). operation interchange / pre(post) multiply. $\rightarrow$ invertible.

消法. ⇒ 可乘即可.
用于行列式时、换法 注意乘 $(-1)^n$. 倍法注意 λ^n

13. ~~equal~~ equivalent matrices. B can be transformed by a sequence of fundmental transformation by A.

standard form. $\begin{bmatrix} 1 & & & 0 \\ 0 & 1 & & \\ & & 1 & \\ 0 & \cdots & & 0 \\ 0 & \cdots & & 0 \end{bmatrix}$

14. Quadratic Form.

$f(x_1, x_2, \ldots x_n) = X^TAX = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$.

Def congruent matrix. $B = C^TAC$ (C is invertible) (复对称.合同 ⇔ 等秩)

$X = CY$ (nonsingular linear transformation).

15. Standard Form.

~~Def~~. every standard form has equivalent quadratic form.

16. Normal Form.

(1) complex. $f = \sum_{i=1}^{n} y_i^2$ (2) real. coefficient $= \pm 1$.

Thm Sylvester 's law of inertia.

every real quadratic form $\xrightarrow[\text{linear transformation}]{\text{nonsingular}}$ unique normal form.

$n_+(A) = p$ 个 $n_-(A): r-p$ 个 $n_0(A): n-r$ 个.

17. Positive Definite Quadratic

positive ~~interia~~ inertia index = n.

Thm. real symmetric matrix positive definite $\Leftrightarrow$ congrument to identity matrix

$$A_r = \begin{bmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{bmatrix}$$

A_r. Leading principal submatrix

$|A_r|$ leading principal minor of order r of A (顺序主子式).

Thm. positive finite $\Leftrightarrow$ each leading principal minor >0

Vector space

1. Geometric vector plane.

O (a fixed point in the plane → origin

V the set of directed line segments starting at O.

$\vec{OA}$. (A is a point on the plane). the corresponding directed line segment.

Properties:

I. commutation (sum. scalar product).

II. association (sum. scalar product)

III. for any point A there is a point B s.t. $\vec{OA}+\vec{OB}=\vec{OO}$

for any point A. $\vec{OO}+\vec{OA}=\vec{OA}$. $\vec{OA}=\vec{OA}$
(additive identity)

2. Coordinate Space.

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

$(a_1, \dots, a_n)+(b_1, \dots, b_n)=(a_1+b_1, \dots a_n+b_n)$.

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

One can consider F^n as $1\times n$ matrices with standard matrix addition and multiplication by a scalar.

Then the following properties are satisfied.

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

(2) $A+(0, \dots 0)=A$ for any $A \in F^n$

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

(4). $A+B=B+A$. $a(A+B)=aA+aB$ $(a+b)A=aA+bA$.

$abA=a(bA)$. $A=A$ (for any ...)

Def. Vector space.

set V. whose elements are called vectors, together with vector addition and scalar multiplication.

the two operation need to have the properties above[(1)(2)(3)(4)].

Remark: + denotes addition both in V and F ($\mathbb{R}$ or $\mathbb{C}$)

· denotes scalar multiplication in V and multiplication in F

e.g. Zero vector space. $V=\{0\}$ $0+0=0$ $a \cdot 0=0$ for $a \in F$.

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

$M_{m\times n}(F)$ → $m\times n$ matrix over F.

e.g. The exotic vector space

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

$a \cdot v = v^a$ for $v \in V$, $a \in \mathbb{R}$

$1 \in V$. additive identity (加法恒等元)

v^{-1} additive inverse of v.

Right distributivity holds. $a(u+v) = a^{uv} = a^u \cdot a^v = au + av$ (other axioms holds. also)

The product of vector space.

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

scalar multiplication $a(u,v) = (au, av)$

addition $(u,v) + (u',v') = (u+u', v+v')$.

(easy to check the axioms).

Basic properties of vector spaces

1. $\vec{0}$ is unique

proof: Let 0 and $0'$ s.t. $\begin{cases} 0+v = v \\ 0'+v = v \end{cases} \Rightarrow 0 = 0+0' = 0'$.

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

proof: $\begin{cases} v+(-v) = 0 \\ v+(-v') = 0 \end{cases}$ $\quad -v' = 0+(-v') = (-v+v)+(-v') = -v+0 = -v$.

3. $0 \cdot v = 0$ for any $v \in V$

proof: $-0 \cdot v \in V$. s.t. $0 \cdot v + (-0 \cdot v) = 0$

$0 = (0 \cdot v) + (-0 \cdot v) = (0 \cdot v + 0 \cdot v) + (-0 \cdot v) = 0 \cdot v + 0 = 0 \cdot v$

4. $a \cdot 0 = 0$

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

5. $(-1) \cdot v = -v$.

By 3. $\vec{0} = 0 \cdot v = (1+(-1)) \cdot v = v + (-1) \cdot v$.

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

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

if $a \neq 0$. $\exists a^{-1} \in F$.

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

§ Subspace (2)

Def. subspace

Let V be a vector space over F. A non-empty $U \subset V$ is called a vector subspace of V.

if. (1) $u+v \in U$ for any $u, v \in U$.

(2) $au \in U$ for any $u \in U$. $a \in F$.

Pro. If U is a subspace of V then U itself is a vector space under induced operation (i.e. under addition / s.m. defined as in V)

proof: the def. assure the operation on U are correctly defined.

(the axioms are valid in U).

e.g. (1). $\{0\}$ and V are subspace of V.

(2) A line passing through the origin is a subspace of geometric vector plane.

(3) A ray the origin is not a subspace of the g.v.p.

无特殊意义. $f(2)=0$ 也可.

(4). $U = \{ f \in F(t) \mid f(1) = 0 \}$. is a subspace of $F[t]$

(5). $\{(a_1, \dots a_n) \in F^n \mid a_1 + \dots + a_n = 0\}$ is a subspace of F^n

but $\dots = 1$ is not. (addition $= 2$.)

(6) $\{A = (a_1 \dots a_n) \in F^n \mid CA^T = 0\}$ is a subspace of F^n. where $C \in M_{m \times n}(F)$ is a fixed matrix

(7). $\{A \in M_n(\mathbb{R}) \mid A^T = A\}$ is a subspace of $M_n(\mathbb{R})$

proof: If $A, B \in M_n(\mathbb{R})$ and $A^T = A$ $B^T = B$

then $(A+B)^T = A^T + B^T = A + B$.

If $A \in M_n(\mathbb{R})$. $A^T = A$. and $a \in \mathbb{R}$

the $(aA)^T = aA^T = aA$.

Linear combinations and linear span.

Def. Let V be a vector space over F and $v_1 \dots v_n \in V$. A linear combination of the vectors $v_1, \dots v_n$ is a vector in V of the form $a_1v_1 + \dots + a_nv_n$; where $a_1, \dots, a_n \in F$.

The linear span (or span) of vectors $v_1, \dots v_n$ denoted by $\mathrm{Span}(v_1 \dots v_n)$ is the set of all vectors that are linear combinations of $v_1, \dots v_n$. (生成集合)

The vectors $v_1, \dots v_n$ span (生成). V or $v_1 \dots v_n$ is a spanning set of V if $\mathrm{Span}(v_1, \dots v_n) = V$ (注意. 未要求 $v_1, \dots v_n$. 可共线).

e.g. two vector. collinear $\to$ span line

not $\to$ span plane.

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

1. $\text{Span}(v_1 \dots v_n)$ is a subspace

2. $\text{Span}(v_1 \dots v_n)$ is the minimum subspace of V containing $v_1 \dots v_n$

i.e. if U is a subspace of V and $v_1 \dots v_n \in U$ then $\text{Span}(v_1, \dots v_n) \subset U$

proof 1. $u', u \in \text{Span}(v_i)$ then $u = a_1v_1 + \dots + a_nv_n$ $\quad v' = a_1'v_1 + \dots + a_n'v_n$. $\quad a_i, a_i' \in F$.

$u + u' = (a_1 + a_1')v_1 + \dots + (a_n + a_n')v_n \in \text{Span}(v_i)$.

similar. for $a \in F$ ..

proof 2. Any $u \in \text{Span}(v_i)$ can be expressed as $u = a_1v_1 + \dots + a_nv_n$ for some $a_1, \dots a_n \in F$.

If $v_1 \dots v_n \in U$. then $u \in U$ since U is a subspace

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

proof: Suppose that $u, u' \in U_1 \cap U_2$. Then $u, u' \in U_1$. U_1 is a subspace, $u + u' \in U_1$, $u + u' \in U_1 \cap U_2$.

similarly. $a \in F$ ..

Sum and direct sum.

Def. Let $U_1, \dots U_m$ be subspaces of a vector space V. The <u>sum</u> of $U_1 \dots U_m$ is defined as

$$U_1 + \dots + U_m = \{ u_1 + \dots + u_m \mid u_1 \in U_1, \dots u_m \in U_m \}$$ (不是简单相加. 分解后相加. 成子空间).

Pro 2.5. Let. $U_1, U_2, \dots U_m$ be subspace of V. Then

1. $U_1 + \dots + U_m$ is a subspace of V.

2. $U_1 + \dots + U_m$ is the minimum subspace of V containing $U_1 \dots U_m$.

i.e. if U is a subspace of V and $U_1, \dots U_m \subset U$ then $U_1 + U_2 + \dots + U_m \subset U$

(proof $u_i \in U_i \subset U$)

Def. <u>Direct sum</u>

Let $U_1, \dots U_m$ be subspaces of a vector space V.

$V = U_1 \oplus U_2 \oplus \dots \oplus U_m$ if each $v \in V$ can be <u>uniquely</u> represented as $v = u_1 + \dots + u_m$.

where all $u_i \in U_i$

e.g. Let $V = F^n$. $U_1 = \{(*, 0, \dots 0)\}$ $\quad U_2 = \{(0, *, \dots 0)\}$ $\dots$ $U_n = \{(0, 0, \dots, *)\}$

Then $V = U_1 \oplus \dots \oplus U_m$.

Pro 2.6. Let U, W be subspaces of a vector space V. Then $V = U \oplus W$ if and only if

$V = U + W$ and $U \cap W = \{0\}$

proof: If $V = U \oplus W$. then clearly $V = U + W$.

If $v \in U \cap W$. then $0 = v + (-v)$ where $v \in U, -v \in W$. Now uniqueness implies $v = 0$

thus $W \cap U = \{0\}$

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

Then $u - u' = w' - w$ $U \in W$. Thus. $u - u' = w - w' = 0$, $u = u'$, $w = w'$

§ Basis (3).

Def. Linear (in)dependence.

Let V be vector space over F.

Linear dependence : there exist $a_1, a_2, \dots a_n \in F$ not all zero s.t. $a_1 v_1 + \dots + a_n v_n = 0$ (non-trivial linear combination)

Other wise, $v_1, \dots v_n$ are linearly independent.

e.g. Non-zero vector $u, v \in V$ are linear dependent. iff. $v = au$ for some $a \in F$

A set of vectors that contains zero vector is linearly dependent (let the zero vector's cofficient non-0)

3. 3-dimensional vectors are linear dependent iff they are coplanar.

Lemma 3.1. Let V be a vector space, $X \subset Y$ be finited sets of vectors in V. If Y is linearly independent then X is linearly independent, if X is linearly dependent, Y is linearly dependent
(比较定理).

(proof: A non-trivial zero linear combination of elements of X is a non-trivial zero linearly combination of elements of Y.).

Lemma 3.2. (Linear Dependence Lemma)

Let V be vector space over F. Let $v_1, \dots v_n \in V$. be linearly dependent and $v_1 \neq 0$.

Then there exists $2 \le j \le n$. such that $v_j \in \text{Span}(v_1, \dots v_{j-1})$.

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

proof: There are $a_i \in F$. not all zero. $a_1 v_1 + \dots + a_n v_n = 0$.

Let a_j be the nonzero cofficient with maximum index ($a_{j+1} = 0 \dots a_n = 0$).

Then $v_j = -\frac{a_1}{a_j} v_1 - \dots - \frac{a_{j-1}}{a_j} v_{j-1}$.

If $j = 1$. then $a_2 = \dots = a_n = 0$. $\Rightarrow v_1 = 0$. contradiction. Thus $j \geq 2$.

$\text{Span}(v_1 \dots v_{j-1}, v_{j+1} \dots v_n) \subset \text{Span}(v_1 \dots v_n)$. Obviously

To show the inverse inclusion. pick up $u \in \text{Span}(v_1 \dots v_n)$.

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

whence u is a linear combination

Coro 3. Let $v_1 \dots v_n \in V$. be linearly independent and $v \in V$. The vector $v_1, \dots v_n, v$. are linear dependence. iff $v \in \text{Span}(v_1, \dots v_n)$

proof: If $v_1, \dots v_n, v$. are linearly dependent then some of these vector is expressed as a linear combination of preceding one by lemma 3.2 It's can't be one of $v_1 \dots v_n$ due to their linear independence.

Thm 4. Let V be a vector space. If $u_1 \dots u_m \in V$ are l.i. and $v_1 \dots v_n \in V$ are spanning set, $m \le n$.

proof: By coro 3.3. $u_1, v_1, \dots v_n$ are linearly dependent. add. u from l.i. to span.

By Lemma 3.2 there is a vector v_j, $1 \le j \le n$. s.t. $\text{Span}(v_1, \dots v_{j-1}, v_{j+1}, \dots v_n) = V$.
(生成空间的不变性).

now $u_1, u_2, v_{j-1} \dots v_{j+1}, \dots v_n$. are l.d.

exclude v_j' (can't be u_1. by lemma 3.2. also can't be u_2. that is $u_2 \in \text{span}\{u_1\}$.).

One can continue this process: at the ith step one has spanning set $u_1 \dots u_{i-1}, v_j$, of n vectors. Appending u_i result in a l.d. set $u_1 \dots u_i \cdot v_{j_1}, \dots$. By Lemma. 3.2. we can exclude one of the v_i (can't be u_i).

Note that at each step a vector from $v_1 \dots v_n$ is replaced by vector from $u_1, u_2, \dots u_m$. If $m > n$. after nth step. we have $u_1, \dots u_n$. Appending u_{n+1} will result a l.d. set, which contradicts. $u_1 \dots u_m$ are l.i.

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

proof: If $v_1 \dots v_n$ are l.i. we get the l.i set.

If $v_1 \dots v_n$ are l.d.

By Lemma 3.2. a certain vector from this set can be excluded so that the ramaining vectors still span V.

Check the new set. l.i or l.d. l.d. repeat the process. (n is finite. the process finite).

Finite-dimensional vector Space.

Def. A vector space V is finite-dimensional if it has a finite spanning set.

counter-example. Let $F[t]$ is not f.-d..

Assume $f_1 \dots f_n$ span $F[t]$ denote $\max_{1\le i\le n} \deg f_i = m$ Then $t^{m+1} \notin \operatorname{span}(f_1 \dots f_n)$

example. $F[t]_n$ is f.-d. as $F[t]_n = \operatorname{span}(1, t, \dots t^{n})$

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

Let. U a subspace of V.

1° $U=\{0\}$ trivial 2° choose $v_1 \in U$. $\operatorname{span}(v_1) \subset U$. if $\operatorname{span}(v_1) \neq U$. choose $v_2 \in U$ and ~~span(v_1 v_2~~ $v_2 \notin \operatorname{span}(v_1)$.

... i th step. $\operatorname{span}(v_1 \dots v_{i-1}) \subset U$. if the $\operatorname{span}(v_1 \dots v_{i-1}) \neq U$. choose $v_i \in U$. $v_i \notin \operatorname{span}(v_1 \dots v_{i-1})$

Note that $v_1 \dots v_{i-1}, (v_i)$ is l.i (by the process). by coro 3.3)

The process is finite (by Thm 3.4. there is a limit that the number of vectors of l.i. can't ~~be~~ exceed the number of elements in its finte spanning set)

Basis

Def. Let V be a vector space. A basis of V is a finite linear independent spanning set of vectors in V. (验证是 basis. ① l.i. ② spanning).

V has basis $\Leftrightarrow$ V is finite-dimensional.

Thm 3.7 Vectors $u_1 \dots u_n$ in a vector space V over F form basis $\Leftrightarrow$ any vector in V can be uniquely represented a linear combination of $u_1 \dots u_n$.

proof: $\Rightarrow$ Assume two representation. by def. of basis. $u_1 \dots u_m$ is linearly independent $(a_1 - a_1') = (a_2 - a_2') = \dots = 0$.

$\Leftarrow$ If l.d, has two representation. 0 one & non-0 one contradiction.

$u_1 \dots u_n$ are l.i. and can ~~respe~~ span V $\rightarrow$ $u_1 \dots u_n$ is a basis.

e.g. Any non-zero element is a basis of F over F.

$$\underset{\text{vector}}{\alpha} = \underset{\text{scalars}}{(\alpha_0^{-1}\alpha)} \cdot \underset{\text{vector.}}{\alpha_0} \qquad \alpha_0 \in F \setminus 0$$

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

proof: If $(v_1 \dots v_n)$ are spanning. ~.

Otherwise $\exists v_{n+1} \in V \setminus \operatorname{span}(v_1 \dots v_n)$ By coro 3.3. the set $(v_1 \dots v_{n+1})$ are l.i.

repeat. V is finite-dimensional

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

proof: (spanning set + l.i set).

§ Coordinate.(坐标) (4.)

Def. Let V be a vector space with basis $v_1 \dots v_n$. any vector $v \in V$. can be written a unique l.c.
the number $(a_1, \dots, a_n)$ are coordinates of v relative to the basis.
(depend on the ordering of the vectors in the basis).

Dimension. (维度)

Thm 4.1. Any two basis of a vector space V contain the equal of element.
proof: two set. one is l.i. one is span. $n \leq m$. (reverse. $m \leq n$). $\Rightarrow m = n$.

Def. dimension: V be vector space over F (前提). The number of elements in any of its basis is called dimension. ($\dim_F V$ or $\dim V$). $\dim_{\mathbb{C}} \mathbb{C} = 1$. (multiply by complex number)

e.g. $\dim F^n = n$. $\dim M_{m,n}(F) = mn$. $\dim_{\mathbb{R}} \mathbb{C} = 2$. $(\underline{a} + \underline{b}i)$ $\dim F[x]_n = n+1$.
$\dim V \times U = \dim V + \dim U$.

Pro 4.2. Let V be vector space. U be its subspace. Then $\dim U \leq \dim V$. If $\dim U = \dim V$. $U = V$
proof: Let. $u_1 \dots u_n$ a basis of U. $\dim V = m$. $u_1, \dots u_n$ are l.i. in V.
by Thm 3.9. extend the set to a basis contains m element. $n \leq m$.
If $\dim U = \dim V$. then this extended basis must be same size as the set of $u_1 \dots u_n$
Therefore of is a basis of V. $U = \mathrm{span}(u_1 \dots u_n) = V$.

Pro 4.3. Let V be a vector space. $\dim V = n$ and $v_1, \dots v_n \in V$ Then $v_1, \dots v_n$ is a basis of V if they are linearly independent OR span V.
proof: If $v_1 \dots v_n$ are l.i. they can be extended to a basis (Thm 3.9).
Since this basis has n elements
similarly, if $v_1 \dots v_n$ span V then a basis an be extracted from this set
Since this basis will have n elements. $v_1 \dots v_n$ itself form a basis.

Pro 4.4 Let $v_1, \dots v_n$ be a basis of a vector space V over F and $A = (a_{ij}) \in M_n(F)$
Let $u_i = \sum_j a_{ij} v_j$. (系数为行). $\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$. $1 \leq i \leq n$. Then $u_1, \dots u_n$ form a basis. of V iff A is invertible.
proof: Since ~~∀~~ $\dim V = n$. the vector $v_1 \dots v_n$ form a basis of V iff. they are l.i.
$\sum_i c_i u_i = 0$ is trivial $(c_1 = \dots = c_n = 0)$. $\sum_i c_i u_i = \sum_{i,j} c_i a_{ij} v_j = \sum_j (\underline{\sum_i c_i \overset{取系数}{a_{ij}}}) v_j = 0$.
iff $\sum_i (c_i a_{ij})$ for all $1 \leq i \leq n$. which is in turn equivallent to linear independence of the column A. (行向量) ($\Leftrightarrow$ invertibility A).

$$\sum_{i=1}^n c_i a_{i1} = \sum_{i=1}^n c_i a_{i2} = \dots = \sum_{i=1}^n c_i a_{in} = 0. \Rightarrow \left| \begin{array}{l} c_1 (a_{11} \quad a_{12} \quad \dots \quad a_{1n}) + \\ c_2 (a_{21} \quad \dots \quad a_{2n}) + \\ \vdots \\ c_n (a_{n1} \quad \dots \quad a_{nn}) = 0. \end{array} \right.$$

Thm 4.5. Let V be a vector space and U_1 and U_2. be its subspaces. Then

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

proof: Let $u_1 \dots u_m$ be an basis of $U_1 \cap U_2$. Then $u_1 \dots u_m$ are l.i. vectors in U_1.
and thus $\exists v_1 \dots v_l \in U_1$ s.t. $u_1 \dots u_m, v_1 \dots v_l$ form a basis of
Similarly $\exists w_1 \dots w_n \in U_2$. s.t. $u_1 \dots u_m, w_1 \dots w_n$ form a basis of.

Now we need to show $u_1 \dots u_m, v_1 \dots v_l, w_1 \dots w_n$ form a basis of $U_1 + U_2$.
(if so, $\dim(U_1+U_2) = m+l+n$. $\dim U_1 \cap U_2 = m$. $\dim U_1 = m+l$ $\dim U_2 = m+n$.
① any vector $\in U_1+U_2 \to$ written as a sum of a vector from U_1 and a vector from U_2
these two vector can be written as a l.c. this set spans U_1+U_2)
② Suppose that $a_1u_1 + \dots + a_mu_m + b_1v_1 + \dots b_lv_l + c_1w_1 + \dots + c_nw_n = 0$.
which implies. $\underbrace{c_1w_1 + \dots + c_nw_n}_{U_2} = \underbrace{-a_1u_1 - \dots - a_mu_m - b_1v_1 - \dots b_lv_l}_{U_1}$.
Note that the left-hand part lies in U_2 and the right-hand part lies in U_1.
Therefore $c_1w_1 + \dots + c_nw_n \in U_1 \cap U_2$.
Since $u_1 \dots u_m$ is a basis in $U_1 \cap U_2$. one has.
$c_1w_1 + \dots + c_nw_n = d_1u_1 + \dots + d_mu_m$.
Then vectors $u_1, \dots, u_m, w_1, \dots, w_n$ are linearly independent
therefore. $c_i = d_j = 0$. $(1 \le i \le n, 1 \le j \le m)$.
whence. $a_1u_1 + \dots + a_mu_m + b_1v_1 + \dots + b_lv_l = 0$.
The vector ... are also l.i. and therefore ...
the initial linear combination is trivial.

Direct sum revisited.

Pro 4.6 Let $U_1 \dots U_m$ be subspaces of a vector space V. The following conditions are equivalent.
1. $V = U_1 \oplus \dots \oplus U_m \Leftrightarrow$ 2. $V = U_1 + \dots + U_m$ and $U_j \cap (U_1 + \dots + U_{j-1} + U_{j+1} + \dots + U_m) = \{0\}$ for any $i \in [1, m]$

$\Leftrightarrow$ 3. $V = U_1 + \dots + U_m$ and if $u_1 + \dots + u_m = 0$ for $u_i \in U_i$, then $u_i = 0$
4. the union of basis for U_i is a basis of V.
5. $V = U_1 + \dots + U_m$ and $\dim V = \dim U_1 + \dots + \dim U_m$.

proof: $1 \Rightarrow 2$ Let $u_j \in$ (the intersection) $0 + \dots + 0 + u_j + 0 \dots + 0 = u_1 + \dots + u_{j-1} + 0 + u_{j+1} + \dots + u_m$.
by direct sum. $u_j = 0$.
$2 \Rightarrow 3$. Assume. $u_1 + \dots + u_n = 0$. for any $1 \le j \le m$ one has $u_j = -u_1 - \dots - u_m$ $\underset{(u_j=0)}{u_j \in}$ the intersection.
$3 \Rightarrow 4$. spanning: B_i be bases of U_i. $\forall u_i \in U_i \Rightarrow$ combine the l.c. $\Rightarrow$ new l.c.
l.i.: In a zero linear combination of the vectors from $\cup_i B_i$ (group together multiples) of the vectors from $B_1, \dots B_m$. and obtain that the sum is each group equals to 0. It gives zero linear combinations of the vector in each of $B_1 \dots B_m$, thus all the coefficients equals 0.
$4 \Rightarrow 5$. Choose bases B_i in each U_i s.t. $B_i \cap B_j = \phi$ ← if some are corresponding. choose a scalar, be multiple. and denote $B = \cup_i B_i$ since B is a basis of V. then $\dim V = |B| = |B_1| + \dots + |B_m| = \dim U_1 + \dots + \dim U_m$.
Any $v \in V$ is expressed as a linear combination of the vectors from $\cup_i B_i$.
(Group together multiples) of the vectors from $B_1 \dots B_m$ and obtain the expression of v as the sum of vectors from $U_1 \dots U_m$.

(continued).

$5 \Rightarrow 1$. Let $u_1 + \cdots + u_m = u_1' + \cdots + u_m'$ for some $u_i, u_i' \in U_i$ If $u_1 \neq u_1'$ then $u_1 - u_1' = (u_2' - u_2) + \cdots + \cdots \neq 0$.

whence $U_1 \cap (U_2 + \cdots + U_m) \neq \{0\}$

By Thm 4.5. $\dim(U_1 + U_2 + \cdots + U_m) + \dim U_1 \cap (U_2 + \cdots + U_m) = \dim U_1 + \dim(U_2 + \cdots + U_m)$

that is $\dim U_2 + \cdots + \dim U_m < \dim(U_2 + \cdots + U_m)$ which is possible. (use 5. $\dim V = \dim U_1 + \cdots + \dim U_m$)

(Indeed, choose a basis for each subspace $U_2 \dots U_m$, their union spans $U_2 + \cdots + U_m$ but has less element than its dimension.) $n(\text{span}) < n(\text{basis})$.

§5. Linear transformation.

Def. Let V, W be vector space over F. A map $L: V \to W$ is a linear ~~trans~~ transformation if: perserve addition and scalar multiplication

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. $L(0_V) = 0_W$.

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

$$\underset{\mathcal{L}}{0}(\underset{F(\mathbb{R})}{0} \cdot \underset{V}{0}) = \underset{W}{0}$$

2. $id_V \in \mathcal{L}(V)$. (identify $id_V(v) = v$)
3. If $L: F[t] \to F[t]$, $L(f) = f'$. then $L \in \mathcal{L}(F[t])$ polynomial.
4. For $C \in M_{l,n}(F)$. define $L: M_{m,l}(F) \to M_{m n}(F)$, $L(A) = AC$ then $L \in \mathcal{L}(M_{m \times l}(F), M_{m \times n}(F))$
5. translation (平移) $L(\vec{v}) = \vec{v} + \vec{a}$.

Thm 5.1. Let V, W be vector space over F. $v_1 \dots v_n$ be a basis of V, and $w_1 \dots w_n \in W$. There exist a unique $L \in \mathcal{L}(V, W)$ s.t. $L(v_i) = w_i$ for all $i = 1, \dots, n$.

proof: Assume there is $L \in \mathcal{L}(V, W)$ s.t. $L(v_i) = w_i$ for all $i = 1, \dots n$. For any $v \in V$ one has $v = \sum_{i=1}^n a_i v_i$. If $L(v_i) = w_i$ for $i = 1 \dots n$. (uniquely defined).

then $L(v) = L(a_1 v_1 + \cdots + a_n v_n) = a_1 w_1 + \cdots + a_n w_n$. for some $a_1 \dots a_n \in F$. L is <u>unique</u>.

Conversely, for $v \in V$, $v = \sum_{i=1}^n a_i v_i$ put $L(v) = \sum_{i=1}^n a_i w_i$

It defines $L: V \to W$; it remains to prove that it is linear

Let $u, v \in V$. where $v = a_1 v_1 + \cdots + a_n v_n$. $u = b_1 v_1 + \cdots + b_n v_n$.

Then by def. $L(v) = a_1 w_1 + \cdots + a_n w_n$. $L(u) = b_1 w_1 + \cdots + b_n w_n$.

and check $L(u+v)$. $L(av)$...

* to proof maps equal. $f = f' \Leftrightarrow$ 1. $X = X'$ 2. $Y = Y'$ 3. $f(x) = f'(x)$ $\forall x \in X = X'$

Operation

addition. $L + L'$.

composition. $L' \circ L$ (often written as $L'L$ informally the product of L' and L).

Pro 5.2. $\mathcal{L}(V, W)$ w.r.t. the above defined addition and scalar multiplication is <u>a vector space</u>.

In addition the product of linear transformation satisfies the following properties

1. associative. $L_1, L_2, L_3 \in \mathcal{L}(V, W)$. $(L_1 + L_2) + L_3 = L_1 + (L_2 + L_3)$
2. additive identity. zero l.t.
3. inverse of L. is $-L := (-1) \cdot L$ s.t $L + (-L) = 0$
4. distributivity. left/right.

Remark: the composition of the function ~~are~~ is non commutative

Kernel and Image.

Def. The kernel of $L \in \mathcal{L}(V,W)$ is defined by $Ker(L) = \{v \in V \mid L(v)=0\}$

e.g. 1. $Ker(0)$ is V. $(0 \in \mathcal{L}(V,W))$.

2. $Ker(id_V) = \{0\}$.

3. If $L \in \mathcal{L}(F[x])$, $L(f) = f'$ then $Ker(L) = \{f \in F[x] \mid f' = 0\} = F$.

Pro5.3. If $L \in \mathcal{L}(V,W)$ then $Ker(L)$ is a subspace in V. (proof is trivial)

Pro5.4. $L \in \mathcal{L}(V,W)$ is injective iff $Ker(L) = \{0\}$.

proof: Suppose L is injective since $\underline{L(0)=0}$. the condition $L(v)=0$. that $v=0$.
therefore $Ker(L) = \{0\}$. $\hookrightarrow$ from $L(av) = aL(v)$

Conversely. suppose $Ker(L) = \{0\}$ let $v_1, v_2 \in V$. s.t. $L(v_1) = L(v_2)$.
Then $L(v_1 - v_2) = L(v_1) - L(v_2) = 0$. $v_1 - v_2 \in Ker(L)$

Def. The $\underline{\text{image}}$ of $L \in \mathcal{L}(V,W)$ is defined by $Im(L) = \{L(v) \mid v \in V\}$ [the image of vector. v is $L(v)$.]

e.g. 1. $L(f) = f'$, $Im(L) = F[t]_{n-1}$.

pro5.5. If $L \in \mathcal{L}(V,W)$. $Im(L)$ is a subspace of W.

Thm5.6. Let V be a vector space over F. $L \in \mathcal{L}(V,W)$ Then $\dim V = \dim ker(L) + \dim Im(L)$.

proof: Let $u_1 \dots u_m$ be a basis of $ker(L)$. It can be extended $u_1 \dots u_m, v_1 \dots v_n$ a basis of V.

1) remains to proof $\dim Im(L) = n$.

Consider vectors $L(v_1) \dots L(v_n)$ and show that they form a basis $Im(L)$
Obviously. $L(v_1) \dots L(v_n) \in Im(L)$. thus $Span(L(v_1) \dots L(v_n)) \subset Im(L)$.
If $w \in Im(L)$ then $w = L(v)$ for some $v \in V$. and $v = a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_nv_n$.
$\Delta L(a_1u_1 + \cdots a_mu_m) = 0$. $a_1u_1 + \cdots + a_mv_m \in V$.
Then. $w = L(v) = L(\cdots) = b_1L(v_1) + \cdots + b_nL(v_n) \Rightarrow w \in Span(L(v_1) \dots L(v_n))$.

2) show $L(v_1) \dots L(v_n)$ are l.i.

If $c_1L(v_1) + \cdots + c_nL(v_n) = 0$. then $0 = L(c_1v_1 + \cdots + c_nv_n)$
whence $c_1v_1 + \cdots + c_nv_n \in Ker(L)$ and $c_1v_1 + \cdots + c_nv_n = d_1u_1 + \cdots + d_mu_m$.
since $u_1 \dots u_m, v_1 \dots v_n$. are l.i. $c_1 = c_2 = \cdots = c_n = 0$.

Coro5.7. A homogeneous system of linear equations with more variables than equations has nonzero solutions.

proof: $AX = 0$. $A \in M_{n\times m}(F)$ and $m > n$.
The set of its solutions is equal to $ker(L)$. where $L \in \mathcal{L}(M_{m,1} \cdot M_{n,1})$ $L(X) = AX$.
Then $\dim ker(L) = m - \dim Im(L) \geq m - n > 0$.

§6. Matrices.

→ the order matters.

Def. Let V, W be two spaces over F, with bases $\mathcal{B}=(v_1 \dots v_n)$ and $\mathcal{B}'=\{w_1 \dots w_m\}$ respectively. Let $L \in \mathcal{L}(V,W)$. If $L(v_j)=a_{1j}w_1+\dots+a_{mj}w_m$

B中第i个(vi)系数在第i列.

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 matrix of L (relative to the $\mathcal{B}, \mathcal{B}'$ and is denoted by $[L]_{\mathcal{B}\mathcal{B}'}$ or by $[L]_{\mathcal{B}}$ if $V=W$ and $\mathcal{B}=\mathcal{B}'$).

→ matrix

Thm 6.1. V, W two space. $\mathcal{B}, \mathcal{B}'$ two bases. $n=\dim V$, $m=\dim W$.

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 word. the map $\varphi: \mathcal{L}(V,W) \to M_{m,n}(F)$. (denote $\varphi(L)=[L]_{\mathcal{B}\mathcal{B}'}$). is a l.t.

proof: Let $[L]_{\mathcal{B}\mathcal{B}'}=(a_{ij})$. $[L']_{\mathcal{B}\mathcal{B}'}=(b_{ij})$. $\mathcal{B}=\{v_1 \dots v_n\}$ $\mathcal{B}'=\{w_1 \dots w_m\}$.

easy to check. By def. $L(v_j)=\sum_{i=1}^m a_{ij}w_i$ $\quad L'(v_j)=\sum_{i=1}^m b_{ij}w_i$.

$\Rightarrow (L+L')(v_j) = L(v_j)+L'(v_j) = \sum_{i=1}^m (a_{ij}+b_{ij})w_i$ whence. $[L+L']_{\mathcal{B}\mathcal{B}'} = (a_{ij}+b_{ij}) \pm [L]_{\mathcal{B}\mathcal{B}'}+[L']_{\mathcal{B}\mathcal{B}'}$

2 properties. (scalar multiplication check)

Thm 6.2. Let U, V, W be three vector space over F. with $\mathcal{B}=\{u_1 \dots u_l\}$. $\mathcal{B}'=\{v_1 \dots v_m\}$. $\mathcal{B}''=\{w_1 \dots w_n\}$ respectively. $L' \in \mathcal{L}(U,V)$. $L \in \mathcal{L}(V,W)$. Then $[L\circ L']_{\mathcal{B}\mathcal{B}''} = [L]_{\mathcal{B}'\mathcal{B}''}[L']_{\mathcal{B}\mathcal{B}'}$ 不同空间的基.

$$\begin{tikzcd} \underset{\mathcal{B}}{U} \arrow[r, "L'"] \arrow[rr, bend left, "L\circ L'"] & \underset{\mathcal{B}'}{V} \arrow[r, "L"] & \underset{\mathcal{B}''}{W} \end{tikzcd}$$

proof: let. $[L]_{\mathcal{B}'\mathcal{B}''}=(a_{ij}) \in M_{n,m}(F)$ $\quad [L']_{\mathcal{B}\mathcal{B}'}=(b_{ij}) \in M_{m,l}(F)$

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

$$(L\circ L')(u_p) = L(L'(u_p)) = L\Big(\sum_{j=1}^m b_{jp}v_j\Big) = \sum_{j=1}^m b_{jp}L(v_j) = \sum_{j=1}^m\sum_{i=1}^n b_{jp}a_{ij}w_i = \sum_{i=1}^n\Big(\sum_{j=1}^m a_{ij}b_{jp}\Big)w_i$$

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

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

Denote $[v]_{\mathcal{B}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$

Thm 6.3. Let V be a vector space over F with basis $\mathcal{B}=\{v_1 \dots v_n\}$. 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. $V \to M_{n,1}(k)$, $v \mapsto [v]_{\mathcal{B}}$. is a linear transformation.

Thm 6.4. Let V, W be v.s. over F with bases $\mathcal{B}, \mathcal{B}'$. respectively $L \in \mathcal{L}(V,W)$.

(不同空间的向量)

Then. $[L(v)]_{\mathcal{B}'} = [L]_{\mathcal{B}\mathcal{B}'}[v]_{\mathcal{B}}$ for any $v \in V$.

proof: Let $v=\sum_{k=1}^n c_k v_k$. $\quad [L]_{\mathcal{B}\mathcal{B}'}=(a_{ij})$

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

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

$\therefore L\cdot L' \to r$. (两次变回)

$L\cdot r \to r$ (一次换).

△ two l.t. preserve the distance. (in geometric plane): reflection / rotation

orientation "朝向" (与 direction 有区别)

change orientation $\qquad$ preserve the orientation

Change of Basis.

Def. transition matrix. (C). 过渡矩阵.

Let $\mathcal{B}=\{u_1 \dots u_n\}$, $\mathcal{B}'=\{v_1 \dots v_n\}$ be bases of V over F.

Then $v_j=\sum_{i=1}^{n} u_i C_{ij}$ for some C_{ij}. $\underline{C=(C_{ij})_{i,j=1}^{n}}$ from $\mathcal{B}$ to $\mathcal{B}'$. denote by $M_{\mathcal{B}\rightsquigarrow\mathcal{B}'}$

In other word. $M_{\mathcal{B}\rightsquigarrow\mathcal{B}'}=([v_1]_{\mathcal{B}}\ [v_2]_{\mathcal{B}} \dots [v_n]_{\mathcal{B}})$

$(v_1\ v_2 \dots v_n)=(u_1 \dots u_n)\, M_{\mathcal{B}\rightsquigarrow\mathcal{B}'}$ $(*)$. 同空间不同基.

Pro 6.5. Let $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ be bases of V.

1) $M_{\mathcal{B}\rightsquigarrow\mathcal{B}}=E_n$.

2) $M_{\mathcal{B}\rightsquigarrow\mathcal{B}''}=M_{\mathcal{B}\rightsquigarrow\mathcal{B}'}M_{\mathcal{B}'\rightsquigarrow\mathcal{B}''}$ Proof: use equation $(*)$.

3). $M_{\mathcal{B}\rightsquigarrow\mathcal{B}'}$ is invertible. and $M^{-1}_{\mathcal{B}\rightsquigarrow\mathcal{B}'}=M_{\mathcal{B}'\rightsquigarrow\mathcal{B}}$. Proof: let $\mathcal{B}=\mathcal{B}''$ in 2).

Thm 6.6 Let V be a vector space with bases $\mathcal{B}, \mathcal{B}'$, Then for any $v\in V$ 同空间不同向量.

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

let $\mathcal{B}=\{u_i\}_1^n$ $\mathcal{B}'=\{v_i\}_1^n$ $[v]_{\mathcal{B}}=(x_1 \dots x_n)^T$ $[v]_{\mathcal{B}'}=(y_1 \dots y_n)^T$.

$v=x_1u_1+\dots+x_2u_2=y_1v_1+\dots+y_nv_n$. $\Rightarrow (u_1 \dots u_n)(x_1 \dots x_n)^T$

by def. $(v_1 \dots v_n)=(u_1 \dots u_n)M_{\mathcal{B}\rightsquigarrow\mathcal{B}'}$ whence $(u_1 \dots u_n)(x_1 \dots x_n)^T=(u_1 \dots u_n)M_{\mathcal{B}\rightsquigarrow\mathcal{B}'}(y_1 \dots y_n)^T$

$(u_1 \dots u_n)$ are basis. and the l.c. with the given basis is unique.

i.e. $(x_1 \dots x_n)^T=M_{\mathcal{B}\rightsquigarrow\mathcal{B}'}(y_1 \dots y_n)^T$

Thm 6.7. Let $U. V.$ vector spaces. bases $\mathcal{B}.\mathcal{C}$ and $\mathcal{B}'\mathcal{C}'$. For $L\in\mathcal{L}(U.V)$ 不同空间不同基.

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

Proof: Use Thm 6.6 and 6.4. (For $\forall u\in U$)

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

Suppose. u is the ith element of $\mathcal{C}$. then the ith rows of $[L]_{\mathcal{C}.\mathcal{C}'}$ and $M_{\mathcal{C}'\rightsquigarrow\mathcal{B}'}[L]_{\mathcal{B}.\mathcal{B}'}M_{\mathcal{B}\rightsquigarrow\mathcal{C}}$ are equal. u is arbitrary. thus ... □

Lemma 6.8. Let $\mathcal{C}$ be a basis of n-dim vector space V over F and $A\in M_n(F)$ be invertible. Then there exists a basis $\mathcal{B}$ in V s.t. $M_{\mathcal{C}\rightsquigarrow\mathcal{B}}=A$.

proof: Let $A=(a_{ij})$ and $\mathcal{C}=\{v_1 \dots v_n\}$

Define $u_j=\sum_{i=1}^{n}a_{ij}v_i$. since A is invertible. $\mathcal{B}=\{u_1 \dots u_n\}$ is a basis. (Pro 4.4).

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

Thm 6.9. Let vector space $U.V.$ $L\in\mathcal{L}(U,V)$. Then there are a basis $\mathcal{B}$ in U and a basis $\mathcal{B}'$ in V s.t. $[L]_{\mathcal{B},\mathcal{B}'}=\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}$ with $r=\dim \operatorname{Im}(L)$

proof: Let bases $\mathcal{C}$ in U. $\mathcal{C}'$ in V. $[L]_{\mathcal{C}\mathcal{C}'}=\underbrace{P_t \cdots P_1}\, D\, \underbrace{Q_1 \cdots Q_{s-1}Q_s}_{\text{Elementary matrices}}=PDQ$. $D=\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}\begin{cases}\dim U=m\\ \dim V=n\end{cases}$

By Lemma 6.8. $\exists\, \mathcal{B}.\mathcal{B}'$. $M_{\mathcal{C}\rightsquigarrow\mathcal{B}}=Q^{-1}$ $M_{\mathcal{C}'\rightsquigarrow\mathcal{B}'}=P$. $\Rightarrow M_{\mathcal{B}\rightsquigarrow\mathcal{C}}=Q$ $M^{-1}_{\mathcal{B}'\rightsquigarrow\mathcal{C}'}=P$ $\Rightarrow D=[L]_{\mathcal{B}\mathcal{B}'}$ (Thm 6.7)

$\mathcal{B}=\{u_i\}$. $\mathcal{B}'=\{v_n\}$. then $L(u_1)=v_1 \cdots L(u_r)=v_r$. $L(u_{r+1})=0 \dots L(u_m)=0$.

$\dim \operatorname{Im}(L)=\dim \operatorname{Span}(L(u_1) \dots L(u_m))=\dim \operatorname{span}(v_1 \dots v_r)=r$.

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

proof: Choose bases. $B.B'$. $\dim Im(L) = rank\,[L]_{B.B'}$ (by thm 6.9).

$[L]_{CC'} = M^{-1}_{B' \to C'} [L]_{B,B'} M_{B \to C}$ and $M^{-1}_{B' \to C'} \cdot M_{B \to C}$ are invertible.

(previous. thm. matrices multiply invertible matrices. rank unchange)

证同构: ① 是 l.t.
② { bijective / invertible
bijective [in $\varphi \in Ker(L)$ / sur. $\forall w \in W$.

§7. Isomorphism (同构).

Pro 7.1. If $L \in \mathcal{L}(V,W)$ is invertible. then the inverse map is also a l.t.

Def. isomorphism. an invertible l.t.

Vector spaces V and W over F are isomorphic if there is an isomorphism from V to W

Pro 7.2. 1. any vector space is isomorphic to itself (reflexity)

2. if V is isomorphic to W then W is isomorphic to V. (symmetry)

3. V isomorphic to W and W isomorphic to U then V isomorphic to U. (transitivity)

$\Rightarrow$ isomorphic is a equivalent relation. (denote. by $V \cong W$).

* notation: $\bar{x} = \{ y \in X \mid x \sim y \}$ the equivalence class of x.

Thm 7.3. Two vector spaces over F are isomorphic $\Leftrightarrow$ their dimensions. are equal

proof: "$\Rightarrow$" L is biject. injectivity $\Rightarrow$ $Ker(L) = 0$ (Pro 5.4).
surjectivity $\Rightarrow$ $Im(L) = W$ $\Rightarrow$ $\dim W = \dim V$ (Thm 5.6).

"$\Leftarrow$" suppose. $\dim V = \dim W > 0$,

Choose a basis. $\{v_1 \dots v_n\}$ $\{w_1 \dots w_n\}$. $\exists! L \in \mathcal{L}(V,W)$. $L(v_i) = w_i$ (by Thm 5.1)

injective: any $v \in Ker(L)$ $v = \sum_{i=1}^{n} a_i v_i$. $0 = L(v) = \sum_{i=1}^{n} a_i L(v_i) = a_1 w_1 + \dots + a_n w_n$.

$\{w_1 \dots w_n\}$ are l.i. $\Rightarrow a_i = 0$. $\Rightarrow v = 0 \Rightarrow$ injective

surjective. $w \in W$. $w = b_1 w_1 + \dots + b_n w_n = b_1 L(v_1) + \dots + b_n(v_n) = L(b_1 v_1 + \dots + b_n v_n)$

Coro 7.4. Any vector space V with $\dim V = n$. is isomorphic to the coordinate space F^n and. the column space $M_{n.1}(F)$. If B is a basis of V. then $\varphi: V \to M_{n,1}(F)$. $\varphi(v) = [v]_B$ is an isomorphism.

proof: "the former part. use the Thm 7.3.
Let $B = \{v_i\}$. φ is l.t. (Thm 6.3).
$\boxed{in}$ $v \in Ker(\varphi)$ $v = 0 \cdot v_1 + \dots + 0 \cdot v_n = 0$. $Ker(\varphi) = \{0\}$
$\boxed{sur}$ for $\forall w \in (a_1 \dots a_n)^T \in M_{n,1}(F)$ $\exists v = a_1 v_1 + \dots a_n v_n \in V$. $\Rightarrow [v]_B = w$.

Thm 7.5. Let $V.W$ be vector space over F. Then $\mathcal{L}(V,W)$ is isomorphic $M_{m,n}(F)$ where $m = \dim W$, $n = \dim V$. Moreover. if $B. B'$ are bases. of $V.W$.

$\varphi: \mathcal{L}(V,W) \to M_{m,n}(F)$. $(\varphi(L) = [L]_{B.B'})$ is an isomorphism.

proof: φ is l.t. (by Thm 6.1). Let $B = \{v_i\}$. $B' = \{w_i\}$.

$\boxed{in}$ If $L \in Ker(\varphi)$. and $v \in V$. Coordinates are 0. $L(v) = 0$. $Ker(\varphi) = \{0_n\}$

$\boxed{sur}$ let $A = (a_{ij}) \in M_{m,n}(F)$. Thm 5.1. implies $\exists L \in \mathcal{L}(V,W)$ s.t. $L(v_j) = \sum_{i=1}^{} a_{ij} w_i$ $1 \le j \le n$.

Then $[L]_{BB'} = A$. $\Rightarrow Im(\varphi) = M_{m,n}(F)$.

Coro 7.6. If $V.W$ are vector space over F. $\dim \mathcal{L}(V,W) = \dim V \cdot \dim W$

Def. $L: V \to V$ (l.t.) is called a linear operator on V. (the set denote by $\mathcal{L}(V)$)

Pro 7.7 Let V be a vector space and $L \in \mathcal{L}(V)$. The following statement are equivalent.

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

proof: (calculate the dimension by previous. thm).

§8. Eigenvector (特征向量) and Eigenvalue (特征值).

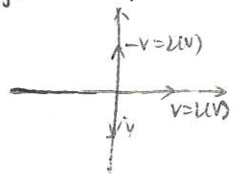
V vectors space $L \in \mathcal{L}(V)$.

Def. eigenvalue: $\underline{\lambda \in F}$. s.t $L(u) = \lambda u$ for some $u \in V$. (u is nonzero)

eigenvector. $u \in F$. s.t $L(u) = \lambda u$. for some $\lambda \in F$ (λ can be zero).

$\lambda \xLeftrightarrow{\text{be associated to}} u$. (线性变换的特征值/向量?)

e.g. 1. L reflection across x-axis



two eigenvalue: $1, -1$

eigenvector: lying on x-axis or y-axis (w.r.t. 1 and -1).

e.g. 2. $L(A) = A^T$. eigenvalue: 1. ; -1

eigenvector: symmetric matrix. ; $\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$

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

proof the additivity. / multiplication scalar.

Def. Eigenspace: the subspace $V_\lambda(L)$ is the eigenspace of L associated to λ.

note $V_\lambda(L) \neq \{0\}$ iff λ is an eigenvalue of L.

Def. Let $A \in M_n(F)$. $\underline{\text{An eigenvalue of } A}$ is $\lambda \in F$ s.t. $AX = \lambda X$. for a nonzero column vector $X \in M_{n,1}(F)$. $\underline{\text{A column eigenvector}}$ of A is $\underline{X} \in M_{n,1}(F)$ s.t. $AX = \lambda X$. for some $\lambda \in F$. (矩阵的特征值/列向量).

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

1). If $\lambda \in F$. is e.value of $L \in \mathcal{L}(V)$. then λ is an e.value of $[L]_B$ for any basis of V.

(let X is $[u]_B$. the coefficient matrix). $[L]_B [u]_B = [L(u)]_B = [\lambda u]_B = \lambda [u]_B$.

2). If $\lambda \in F$. is e.value of $[L]_B$ for a basis B of V. then λ is an eigenvalue of L.

proof. $\exists X \in M_{n,1}(F)$. $X \neq 0$. s.t. $[L]_B X = \lambda X$. 计 $X = [u]_B$ for $u \in V$.

then $[L(u)]_B = [L]_B [u]_B = [L]_B X = \lambda X = \lambda [u]_B = [\lambda u]_B$.

that is $L(u) = \lambda u$. note that $u \neq 0$. since $[u]_B = X \neq 0$.

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

4). If $X \in M_{n,1}(F)$ is a column eigenvector of $[L]_B$ for a basis B of V then $X = [u]_B$ where $u \in V$ is a eigenvector of V.

0 is an eigenvalue iff $\ker(L) \neq \{0\}$.

Thm 8.3 Let $L \in \mathcal{L}(V)$ and $v_1 \dots v_n \in V$ be nonzero eigenvectors associated to distinct eigenvalue $\lambda_1 \dots \lambda_n \in F$. Then $v_1 \dots v_n$ are l.i.

(Assume the converse. apply Lemma 3.2.)

Pro 8.4 Let $L \in \mathcal{L}(V)$ and $\lambda_1 \dots \lambda_m$ be distinct eigenvalues of L. Then the sum $V_{\lambda_1}(L) + \cdots + V_{\lambda_m}(L)$ is direct. More. $\dim V_{\lambda_1}(L) + \dots + \dim V_{\lambda_m}(L) \le \dim V$

Proof: Applying Pro 4.6. Assume $u_1 + \dots + u_m = 0$, where $u_j \in V_{\lambda_j}(L)$. by thm 8.3. $u_1 = \dots = u_m$.
$\Downarrow$
that is. the sum is direct.

The purpose of learn eigenvalue

In general case. $[L]^2_{B,B'} \ne [L^2]_{B,B'} = [L]_{B,B} [L]_{B,B'}$. ($B = B'$ we have the equality).

$\Rightarrow$ we try to find nice forms: $[L]_B = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$. $B = \{u_1 \dots u_m\}$. $\Rightarrow L(u_i) = \lambda_i u_i$ (t.b.c)

$B = \{e_1, \dots e_n\}$ be a basis of V. $\forall k$. $\mathrm{span}(e_1 \dots e_k)$ is L-invariant $\Leftrightarrow [L]_B$ is upper triangle.

Invariant Subspace (不变子空间)

Def. Let $L \in \mathcal{L}(V)$. A subspace U of V is invariant under L. or L-invariant if $L(u) \in U$ for any $u \in U$

e.g. rotation about the rotation. by $\phi = 0, \pi$ (only). (子空间是直线).

Pro 8.5 $V_\lambda(L)$ is L-invariant. (the converse does not hold. counter-e.g: $\mathbb{R}[t]_n$. invariant, not eigen—

proof: If $v \in V_\lambda(L)$. then $L(v) = \lambda v \in V_\lambda(L)$ (scalar multiplication in subspace)

Def. Restriction (限制)

Let $L \in \mathcal{L}(V)$. and U be an L-invariant subspace of V. The map $L|_U : U \to U$. given by $(L|_U)(u) = L(u)$. := the restriction of L to U. (对值域/定义域的缩小).

Pro 8.6 The restriction is correctly defined and belongs to $\mathcal{L}(U)$.

proof: $u \in U$. $(L|_U)(u) = L(u) \in U$. thus $L|_U : U \to U$. $L|_U$ is l.t. (trivially).

Pro 8.7 V. $L \in \mathcal{L}(V)$. Let $V_1 \dots V_m$ be L-invariant subspace of a vector space V with bases $B_1 \dots B_m$ respectively, and $V = V_1 \oplus \cdots \oplus V_m$. Then $[L]_B = \mathrm{diag}([L|_{V_1}]_{B_1}, \dots [L|_{V_m}]_{B_m})$ → 可对多个维度进行一个划分.
where $B = B_1 \cup \dots \cup B_m$ is a basis of V.

proof: let $v_i \in V_i$. V_i is L-invariant. $L(v_i) \in V_i$. All the cooridinates of $L(v_i)$ are zero relative to B except the B_i. These coordinates are exactly the coordinates of $L|_{V_i}(v_i) = L(v_i)$ relative to B_i

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

§ 9. Characteristic polynomial. (特征多项式)

Def. $A \in M_n(F)$. $\chi_A(t) = |tE_n - A|$

Def. trace (迹) of $A = (a_{ij}) \in M_n(F)$.

$$=: \mathrm{Tr}\,A = \sum_{i=1}^{n} a_{ii} \quad (\text{det of diag matrix})$$

Pro 9.1. If $A \in M_n(F)$, then $\chi_A(t) = t^n - \mathrm{Tr}A \cdot t^{n-1} + \cdots + (-1)^n |A|$ (此处省略不是省略类似项，是其余项不易表达).

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

1) $\deg \chi_A = n$. coefficient of t^n is 1.

2). only the diag product can produce ct^{n-1}. the coefficient of t^{n-1} coincides with the coefficient of t^{n-1} in $\prod_{1 \le i \le n} (t - a_{ii})$. i.e. $-\mathrm{Tr}A$

3). the constant term $= \chi_A(0) = |-A| = (-1)^n |A|$

Lemma 9.2. If $A, B \in M_n(F)$ are such that $A = UBU^{-1}$ for an invertible matrix $U \in M_n(F)$

proof: $\chi_A(t) = |tE_n - A| = |tE_n - UBU^{-1}| \overset{*}{=} |U(tE_n - B)U^{-1}| = |U| \cdot |tE_n - B| \cdot |U|^{-1} = |tE_n - B|$

$* \; |tE_n - UBU^{-1}| = |UtE_nU^{-1} - UBU^{-1}| \Leftarrow UtE_nU^{-1} = tUE_nU^{-1} = t(U \cdot U^{-1})(U \cdot U)^{-1} = tE_n$

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

proof: let. $U = M_{B \to B'}$ use Lemma 9.2.

Def. Let $L \in \mathcal{L}(V)$. $\chi_{[L]_B}(t)$ is called characteristic polynomial of L

(since $\chi_{[L]_B}(t)$ does not depend on the choice of B).

$\mathrm{Tr}[L]_B$ $|[L]_B|$ are trace of L. and determinant of L.

(since $\mathrm{Tr}[L]_B$ and $|[L]_B|$ do not depend on the choice of B.)

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

proof: let $A \in M_n(F)$. If $\lambda \in F$ is an eigenvalue of A. then the system of linear equations $\lambda X = AX$, nonzero X. $(\lambda E_n - A)X = 0$ has a non-zero solution. whence $\lambda E_n - A$ is singular. 系数行列式 $= 0$. $\chi_A(\lambda) = 0$. λ is root of $\chi_A(t)$.

Therefore. $\chi_A(\lambda) = |\lambda E_n - A| = 0$. (all implications are reversible)

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

Coro 9.5. ~~Let F be algebraically closed.~~ V be a nonzero vector space over $\mathbb{C}$. $L \in \mathcal{L}(V)$

Then L has an eigenvalue. (在复数域上多项式. deg ≥ 1. 必有根).

$\boxed{\text{Algorithm}}$ find eigenvalue and its associated eigenvectors u. 两种语言(运算关系)的转换.
(based on Pro 8.2. and Pro 9.4) 矩阵(行.列向量) $\Leftrightarrow$ 线性变换.

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

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

III. Find the roots of $\chi_A(t)$; (roots are eigenvalue of L. as well as $[L]_{\mathcal{B}}$.)

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

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

Def. Let $f \in F[X]$. $f = \sum_{i=0}^{n} a_i x^i$. $A \in M_m(F)$. $L \in \mathcal{L}(V)$.
Then $f(A) \in M_m(F)$. $f(L) \in \mathcal{L}(V)$ are defined.

$$f(A) = \sum_{i=0}^{n} a_i A^i \qquad (A^0 = E_n)$$

$$f(L) = \sum_{i=0}^{n} a_i L^i \qquad (A^0 = id_V. \quad L^2 = L \circ L \ \cdots).$$

Pro 9.7. $(f+g)(A) = f(A) + g(A)$ $\qquad (fg)(A) = f(A)\,g(A)$.
$(af)(A) = a\ f(A)$ $\qquad (f \circ g)(A) = f(g(A))$.

$f(L)$ has these four properties as well.

proof: (follow the def.).

$$fg = \sum_{\ell} \left(\sum_{i+j=\ell} a_i b_j\right) x^{\ell}$$

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

Coro 9.8 If. $f, g \in F[X]$. $A \in M_m(F)$. $L \in \mathcal{L}(V)$ then
$f(A)\,g(A) = g(A)\,f(A)$. $\quad f(L)\,g(L) = g(L)\,f(L)$ (矩阵可交换的罕见情况)

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

Lemma 9.10 (8.8). Let $L \in \mathcal{L}(V)$ and $p \in F[X]$. Then $Ker(p(L))$ and $Im(p(L))$ are L-invariant.

proof: If $v \in Ker(p(L))$ then $p(L)(v) = 0$.
show $p(L)(L(v)) = 0$.

$$\Rightarrow p(L)(L(v)) = (p(L)\,L)(v) \overset{\text{Coro.9.8}}{=} (L\,p(L))(v) = L(p(L)(v)) = L(0) = 0.$$

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

show. $L(v) = p(L)(L(u))$.

$$L(v) = L(p(L)(u)) = (L\,p(L))(u) = (p(L) \cdot L)(u) = p(L)(L(u))$$

Thm. (Hamilton - Cayley).

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

△ Warning! we can't simply substitute. t by A. in $\chi_A(t) = |tE_n - A|$.
if so we obtain. matrix = number. it's faulty.

$A \in M_n(F)$
Remind: $AB = BA = \det A \cdot E_n$
(adjoint matrix)

proof: 1) Let $\chi_A(t) = c_0 + c_1 t + \cdots + c_n t^n$ $c_n = 1$.

Let $B = \mathrm{cof}(tE_n - A)$. be the cofactor matrix of $tE_n - A$.

i.e. $(tE_n - A)B = \chi_A(t) E_n$.

Since the entrices of B are cofactors $B \in M_n(F[t]_{n-1})$.

Therefore. let $B = B_0 + tB_1 + \cdots + t^{n-1}B_{n-1}$. $B_i \in M_n(F)$.

then substitute B in previous equality.

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

$$\Rightarrow \begin{cases} -AB_0 = c_0 E_n & (1) \\ B_0 - AB_1 = c_1 E_n & (2) \\ \vdots \\ B_{n-2} - AB_{n-1} = c_{n-1} E_n & (n) \\ B_{n-1} = c_n E_n & (n+1) \end{cases}$$

$(1) + A(2) + \cdots + A^{n-1}(n) + A^n(n+1)$.

$\Rightarrow 0 = c_0 E_n + c_1 A + \cdots + c_n A^n$

i.e. $\chi_A(A) = 0$ (0 is a matrix!).

2) For $L \in \mathcal{L}(V)$. choose a basis $\mathcal{B}$ in V.

Then $\chi_L = \chi_{[L]_\mathcal{B}}$ and $[\chi_L(L)]_\mathcal{B} = [\chi_{[L]_\mathcal{B}}(L)]_\mathcal{B} = \chi_{[L]_\mathcal{B}}([L]_\mathcal{B}) = 0$. by. Lemma 9.9. $f = \chi_{[L]_\mathcal{B}}$

the matrix of the operator is $0 \Rightarrow$ the operator itself is 0.

§ 10 Special types of operators and matrices.

Def. An operator $L \in \mathcal{L}(V)$ is called diagonalizable if its matrix relative to some basis of V is diagonal

Thm 10.1. Let $L \in \mathcal{L}(V)$. $\lambda_1 \ldots \lambda_m \in F$. be the distinct eigenvalues of. Four equipvalent.

1. L is diagonalizable
2. V has a basis consisting of the eigenvectors of L.
3. $V = V_{\lambda_1}(L) \oplus \cdots V_{\lambda_m}(L)$
4. $\dim V = \dim V_{\lambda_1}(L) + \cdots + \dim V_{\lambda_m}(L)$

$1 \Leftrightarrow 2$. The. $[L]_\mathcal{B}$. w.r.t. $v_1 \ldots v_n$ has the form $\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_m \end{pmatrix}$ iff $L(v_j) = \lambda_j v_j$. for all $j = 1, \ldots, n$.

$2 \Rightarrow 3$. by. 2. any vector of V is a l.c. of these eigenvectors. thus. $V = V_{\lambda_1}(L) + \cdots + V_{\lambda_m}(L)$.
Pro 8.4. the sum is direct

$3 \Rightarrow 4$. Pro 4.6.

不一定basis.要证.

$4 \Rightarrow 2$: Let $\dim V = n$. let. $\mathcal{B}_j$ for each subspace $V_{\lambda_j}(L)$. join the together $\{v_1 \ldots v_n\}$ consisting of ei.ve.
show l.i. group. the vector from same $V_{\lambda_j}(L)$. $u_1 + \cdots + u_m = 0$. Thm 8.3. implies $u_1 = u_2 = \cdots = u_m = 0$.
each u_j is a l.c. of basis vector of $V_{\lambda_j}(L)$. thus all $a_i = 0$. l.i. set of n-vectors in n-dim space $\Rightarrow$ basis

Coro 10.2. Let $\dim V = n$ and $L \in \mathcal{L}(V)$ have n distinct eigenvalues.

Then L is diagonalizable.

proof: Let $v_1 \dots v_n$ be ei. ve. associate with $\lambda_1 \dots \lambda_n$. by thm 8.3 theyare l.i. and they form a basis.

Def. matrix diagonalizable.

$A \in M_n(F)$. if. $A = UDU^{-1}$ for a diagonal matrix $D \in M_n(F)$. and an invertible matrix $U \in M_n(F)$

Pro 10.3. 1. Let B be a basis in V. and $L \in \mathcal{L}(V)$. L is diagonalizable $\Leftrightarrow$ $[L]_B$ diagonalizable.

2. Let $A \in M_n(F)$. Then $L \in \mathcal{L}(M_{n,1}(F))$. $L(X) = AX$ is diag.. $\Leftrightarrow$ A is diag..

Proof. 1. "$\Rightarrow$" L is diag.. C is a basis such that $[L]_C$ is diagonal. let $U = M_{B \to C}$

$[L]_B = U [L]_C U^{-1}$ (by thm 6.7).

"$\Leftarrow$" let $[L]_B = UDU^{-1}$ for a diagonal matrix $D \in M_n(F)$. and U is invertible.

Lemma 6.8 asserts that $\exists\, C$ of V s.t. $M_{B \to C} = U$.

Then $[L]_C = U^{-1} [L]_B U = D$ is diagonal.

2. Let F be standard basis of $M_{n,1}(F)$.

Then $[L]_F = A$, one can use 1.

理论1.2.无本质区别.
应用时代表不同的已知条件.

$$\text{已知 } L \; \underset{2}{\overset{1}{\rightleftarrows}} \; \text{已知矩阵}$$

Nilpotent operators and matrices.

Def. An operator is nilpotent if its matrix relative to some basis of V is strictly upper triangular. i.e. only the elements $\boxed{\text{above}}$ the diagonal may be nonzero (线上也是0).

Pro 10.4. Four equivalent properties. ($N \in \mathcal{L}(V)$. $\dim V = n$)

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

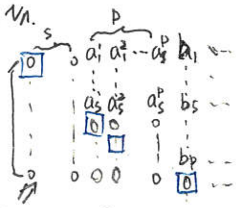
obv. obv.

proof: $2 \Rightarrow 3$. $\chi_N(N) = N^n = 0$. (可能是n次为0. 也可能是从第j次到n次均为0).

$4 \Rightarrow 3$. Let $v_1 \dots v_n$ be a basis of V. Then. $\exists\, j_1 \dots j_n$ s.t. $N^{j_k}(v_k) = 0$ for $1 \le k \le n$.

If $j = \max(j_1 \dots j_n)$. then $N^j(v_k) = 0$ for $1 \le k \le n$ (低次为0 高次必为0).

$\Rightarrow N^j(v) = 0$ for any $v \in V$ since v can be represented as l.c. of $v_1 \dots v_n$.



$3 \Rightarrow 1$. $\ker(N) \subset \ker(N^2) \subset \dots \subset \ker(N^j) = N$.

Choose a basis of $\ker(N)$. extend it to a basis of $\ker(N^2) \dots \ker(N^j) = V$.

The matrix N relative to this basis. is strictly upper triangular.

let $\{e_1 \dots e_s\}$ form a basis of $\ker(N)$. extend to $\ker(N^2)$. we let $\{e_1, \dots, e_s, e_1' \dots e_p'\}$ form a basis of $\ker(N^2)$.

$N(e_1) = \dots N(e_s) = 0$. $\quad N^2(e_1') = N(N(e_1')) = 0$. $\quad N(e_1') \in \ker(N)$. $\quad N(e_1') = \sum_{i=1}^{s} a_i e_i \quad N(e_p')$ same.

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

Def. A matrix $A \in M_n(F)$ is called nilpotent if $A^j = 0$ for some $j \in \mathbb{N}$.

N is nilpotent $\Leftrightarrow$ $[N]_B$ is nilpotent.

Pro 10.6 A matrix $A \in M_n(F)$ is nilpotent if and only if $A = UQU^{-1}$ for a strictly upper triangle matrix $Q \in M_n(F)$ and invertible matrix $U \in M_n(F)$

Proof: Consider $L \in \mathcal{L}(M_{n,1}(F))$ $L(X) = AX$. Remind $[L]_F = A$. where F is standard basis.

Clearly L is nilpotent if and only if A is nilpotent. (转化为向量空间问题).

$\Rightarrow$ If L is nilpotent and C is a basis of in $M_{n,1}(F)$ s.t $Q = [L]_C$ is strictly triangular matrix

Denote by U the matrix which has C as its column. then $M_{F \to C} = U$. (提供一种找基C使得 $[L]_C$ strictly triangle 的方法☆).

$\Leftarrow$ $\boxed{Q^n = 0}$ $A^n = (UQU^{-1})^n = UQ^nU^{-1} = 0$.

A conclusion

	Diagonalizable	Nilpotent.
Matrix.	$B = UDU^{-1}$	$A = UQU^{-1}$, $\exists j, A^j = 0$
Operator	$\exists B$ s.t. $[L]_B$ is diagonal	$\exists B$ s.t $[L]_B$ is nilpotent.
	$\forall C$ $[L]_C$ is diagonalizable	$\forall C$. $[L]$ is nilpotent

§11 Generalized Eigenvectors.

Def. generalized eigenvector (associated to λ).

A nonzero vector $(L - \lambda \mathrm{id}_V)^j(v) = 0$. for some $j \in \mathbb{N}$. ($(L - \lambda \mathrm{id}_V)(v) = 0$ for ei. vector)

generalized eigenspace. (denote by $V(\lambda, L) = \{v \in V \mid (L - \lambda \cdot \mathrm{id}_V)^j(v) = 0 \text{ for some } j > 0\}$

The set of all generalized eigenvectors of L associated to λ. together with 0.

Pro 11.1 Let $L \in \mathcal{L}(V)$. and $\lambda \in F$ (proposition extend from eigen v/s.)

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

proof. "1 $\Rightarrow$" there $\exists m \geq 0$. $(L - \lambda \mathrm{id}_V)^m(v) \neq 0$. $(L - \lambda \mathrm{id}_V)^{m+1}(v) = 0$.

Then $L(u) = \lambda u$. $u \neq 0$. for $u = (L - \lambda \mathrm{id}_V)^m(v)$ $\Rightarrow$ $(L - \lambda \mathrm{id}_V)(u) = 0$. $\Rightarrow$ $L(u) = \lambda u$.

"1 $\Leftarrow$" $V_\lambda(L) \subset V(\lambda, L)$

Thm 11.3. V be v.s over $\mathbb{C}$.

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. $(L - \lambda_j \mathrm{id}_V)|_{V(\lambda_j, L)}$ is nilpotent.

$L(v) = (L - \lambda \mathrm{id}_V)(v) + \lambda v$

proof 3&2. $\exists j$. s.t. $(L - \lambda \mathrm{id}_V)^j(v) = 0 \Rightarrow (L - \lambda \mathrm{id}_V)^{j-1}(L - \lambda \mathrm{id}_V)(v) = 0$. $(L - \lambda \mathrm{id}_V)(v) \in V(\lambda, L)$.

that is. $V(\lambda, L)$ is $L - \lambda \mathrm{id}_V$-invariant. (which implies. $V(\lambda, L)$ is L-invariant).

Then one can consider $(L - \lambda \mathrm{id}_V)|_{V(\lambda, L)}$ which is nilpotent by Pro 10.4.

proof 1: factorization $\chi_L(t) = (t-\lambda_1)^{k_1} \cdots (t-\lambda_m)^{k_m}$

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

Then the greatest common divisor of $\chi_1, \chi_2, \dots \chi_m$ is 1.

$\exists f_1 \dots f_m \in \mathbb{C}[t]$. $f_1\chi_1 + f_2\chi_2 + \cdots + f_m\chi_m = 1$. (1)

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

补了 kj 项. L 大于 k 次均为 0.

Since $(L-\lambda_j id_V)^{m_j} f_j(L)\chi_j(L) = f_j(L)\chi_L(L) = 0$. by H-C thm. $W_k \subset V(\lambda_k, L)$. for any $1 \le k \le m$

(1) implies. $f_1(L)\chi_1(L) + \dots + f_m(L)\chi_m(L) = id_V$.

whence $W_1 + \cdots + W_m = V$. since $W_j \subset V(\lambda_j, L)$. $V(\lambda_1, L) + \cdots + V(\lambda_m, L) = V$ (then check direct)

Assume $v_j \in V(\lambda_j, L) \cap U_j$. where $U_j = V(\lambda_1, L) + \cdots + V(\lambda_{j-1}, L) + V(\lambda_{j+1}, L) + \cdots + V(\lambda_m, L)$

$\Rightarrow$ want to use Pro 4.b.(2).

thus we can write $v_j = v_1 + \cdots + v_{j-1} + v_{j+1} + \cdots v_m$. $v_i \in V(\lambda_i, L)$ (need to show v_j are zero).

Let $n = \dim V$ Then $(L-\lambda_i id_V)^{k_i'}(v_i) = 0$. where $k_i' = \dim V(\lambda_i, L) \le n$.

whence, $(L-\lambda_i id_V)^n(v_i) = 0$. for all $1 \le i \le m$.

(得到一个使 $v_j \mapsto 0$ 的线性变换).

Therefore $(L-\lambda_j id_V)^n(v_j) = 0$ and $(L-\lambda_1 id_V)^n \cdots (L-\lambda_{j-1} id_V)^n (L-\lambda_{j+1} id_V)^n \cdots (L-\lambda_m id_V)^m (v_j) = 0$ *

the polynomial $(t-\lambda_j)^n$ and $p_j(t) = (t-\lambda_1)^n \cdots (t-\lambda_{j-1})^n (t-\lambda_{j+1})^n \cdots (t-\lambda_m)^n$ are co-prime.

whence $h_j(t)(t-\lambda_j)^n + g_j(t)p_j(t) = 1$. for some $h_j, g_j \in \mathbb{C}[t]$ (建立所得线性变换与 id_V 的关系 即有 $v_j \mapsto v_j = 0$)

Then $h_j(L)(L-\lambda_j id_V)^n + g_j(L)p_j(L) = id_V$.

whence $v_j = h_j(L)(L-\lambda_j id_V)^n(v_j) + g_j(L)p_j(L)(v_j) = 0$.

* $(L-\lambda_1 id_V)^n \cdots (L-\lambda_m id_V)^n (v_1 + \cdots + v_{\lambda m}) = \underbrace{(L-\lambda_1 id_V)^n \cdots (L-\lambda_m id_V)^n}_{\text{commutate}}(v_1) + \cdots (v_2) + \cdots (v_m)$.

more clear, we let $T_i = (L-\lambda_i id_V)^n$. $\Rightarrow T_i(v_i) = 0$ $=0$.

$T_1 \cdots T_{j-1} T_{j+1} \cdots T_m (v_1 + \cdots v_{j-1} + v_{j+1} + \cdots + v_m) = T_2 \cdots T_m \underbrace{T_1(v_1)}_{=0} + T_1 T_3 \cdots T_m \underbrace{T_2(v_2)}_{=0} + \cdots + \cdots = 0$.

Coro 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

Use the sum of dim to proof it. trivial.

Lemma 11.5. Let $N \in \mathcal{L}(V)$. and $\operatorname{Ker}(N^j) = \operatorname{Ker}(N^{j+1})$ for $j > 0$. then $\operatorname{Ker}(N^j) = \operatorname{Ker}(N^{j+m})$ for any $m \in \mathbb{N}$. In other words if two consecutive (连续的) terms in the chain

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

are equal then all the subsequent terms are equal.

proof: It suffices to prove $\operatorname{Ker}(N^{j+k}) = \operatorname{Ker}(N^{j+k+1})$ for any $k \in \mathbb{N}$.

Let $v \in \operatorname{Ker}(N^{j+k+1})$ and $N^{j+k+1}(v) = N^{j+1}(N^k(v)) = 0$.

Then $N^k(v) \in \operatorname{Ker}(N^{j+1}) = \operatorname{Ker}(N^j)$ whence $N^{j+k}(v) = N^j(N^k(v)) = 0$.

Jordan basis for a nilpotent operator

Thm 11.6. Let V be a vector space over F, $N \in \mathcal{L}(V)$ be a nilpotent operator. Then there exist $v_1 \dots v_s \in V$ and $m_1, \dots m_s \in \mathbb{N}$. s.t.

$$N^{m_1}(v_1), \dots, N(v_1), v_1$$
$$N^{m_2}(v_2), \dots N(v_2), v_2.$$
$$\vdots$$
$$N^{m_s}(v_s), \dots N(v_s), v_s$$

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

(In Fact. let the forming basis be $\mathcal{B}$.

$$[N]_{\mathcal{B}} = \begin{bmatrix} \underbrace{\begin{smallmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{smallmatrix}}_{m_1} & & & \\ & \underbrace{\begin{smallmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{smallmatrix}}_{m_2} & & \\ & & \ddots & \\ & & & \underbrace{\begin{smallmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{smallmatrix}}_{m_s} \end{bmatrix}$$

proof: Induction on $\dim V$. ($\dim V = 1$. a nilpotent operator must be 0).

Suppose $\dim V > 1$, then $\dim Im(N) < \dim V$ (otherwise. N is invertible. $N^j = 0$. $N^{j-1} = N^j \cdot N^{-1} = 0$ $\Rightarrow N = 0 \Rightarrow N$ not invertible. contradicts)

(which implies $N = 0$).

if $Im N = 0$, we can take an arbitrary basis $v_1 \dots v_s$ in V and put $m_1 = \dots = m_s = 0$.

Thus $Im(N)$ is N-invariant (by Lemma 9.10)

(by introduction. we hypothesize the theorem holds when the dim $< \dim V$.

we find $Im(N)$ satisfy the condition. (i.e. applies the theorem to $N|_{Im(N)}$).

gives $v_1 \dots v_s \in Im(N)$. (suppose $v_i = N(u_i)$ for some $u_i \in V$).

$$\begin{matrix} N^{m_1+1}(u_1) & \cdots & N(u_1) \\ \vdots & & \vdots \\ N^{m_s+1}(u_s) & \cdots & N(u_s) \end{matrix}$$

forms a basis of $Im(N)$. (let it be B') 有3基. 扩大. 证 l.i./span.

and $N^{m_1+2}(u_1) = \dots N^{m_s+2}(u_s) = 0$.

consider l.c. $a^1_{m_1+1} N^{m_1+1}(u_1) + \dots + a^1_1 N(u_1) + \dots + a^s_{m_s+1} N^{m_s+1}(u_s) + \dots + a^s_1 (u_s) + a^1_0 u_1 + \dots + a^s_0 u_s = 0$.

Let $N(LHS) = N(RHS)$. we now that coefficient all equals 0.

the set $B'' = \{ B' \cup (u_1 \dots u_s) \}$ are l.i.

Extend B'' with $w_1 \dots w_t \in V$ to a basis of V.

$N(w_j) \in Im(N)$. $N(w_j)$ is a l.c of B'.

$N(w_j) = N(x_j)$ x_j is a l.c of $\{ B' \setminus \{N^{m_1+1}(u_1) \dots N^{m_s+1}(u_s)\} \cup (u_1, \dots, u_s) \}$.

put $u_{s+j} = w_j - x_j$. $1 \le j \le t$. the vectors $\{ B'' \cup (u_{s+1}, \dots u_{s+t}) \}$ forms a basis of V.

where $m_{s+1} = \dots = m_{s+t} = 0$. since $N(u_{s+j}) = N(w_j) - N(x_j) = 0$.

w_j are basis vectors. x_j are l.c. of other basis vectors.

Some equivalent thm

$\chi_L(t) = (t-\lambda_1)^{r_1} (t-\lambda_2)^{r_2} \cdots (t-\lambda_s)^{r_s}$. $V = V_1 \oplus V_2 \oplus \dots \oplus V_s$.

$V_i = \{ v \in V \mid (L - \lambda_i id_V)^{r_i}(v) = 0 \}$

§12 Jordan normal form

Def. Jordan block. $J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \cdots \lambda & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}$

Jordan matrix (block diagonal matrix each block is Jordan matrix).

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

($\lambda_1, \lambda_2, \ldots \lambda_s$ may not be all distinct). 相等的λ实质是重根拆了用.

Jordan basis for $L \in \mathcal{L}(V)$, a basis whose matrix of L relative to this basis is a Jordan matrix.

The matrix is called Jordan normal form of L.

Thm 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. (a Jordan normal form of the operator is unique up to the order of the Jordan blocks) 除去Jordan块的排列顺序外是唯一的.

排列顺序等价于 $\chi_L(t)$ 中 $(t-\lambda_i)^{r_i}$ 的排列顺序

Proof: By thm 11.4 $\dim V$ = the sum of dim of eigenspace (direct sum).
Moreover, if $\lambda_i \in F$ is eigenvalue, the restriction of $L - \lambda_i \mathrm{id}_V$ on $V(\lambda_i, L)$ is nilpotent.
Then Thm 11.6 implies that the matrix of $(L - \lambda_i \mathrm{id}_V)|V(\lambda_i, L)$ relative to a certain basis has a form $\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$

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

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

⇒ Jordan chain

$L(u_1) = \lambda u_1$
$L(u_2) = u_1 + \lambda u_2$
$\vdots$
$L(u_m) = u_m + \lambda u_{m-1}$

A Jordan chain starts from an eigenvector. If $\mathcal{J}$ is a Jordan basis $C = [L]_{\mathcal{J}} - \lambda E_n$, it satisfies $CX_1 = 0$, $CX_2 = X_1, \ldots CX_m = X_{m-1}$ where $X_j = [u_j]_{\mathcal{J}}$, $1 \le j \le m$

Coro 12.2. Let V be a vector space over $\mathbb{C}$ and $L \in \mathcal{L}(V)$. [1] The sum of the orders of the Jordan blocks associated with λ is equal to its multiplicity as a root of the characteristic polynomial. [2] The number of the Jordan blocks associated with λ is equal to $\dim \operatorname{Ker}(L - \mathrm{id}_V)$ ⇒ i.e. dim of eigenspace associate with λ. 用的时候反过来.

特征多项式重根数(λ) → λ-Jordan块的阶数.

proof: [1] J be a Jordan normal form of L and $\dim V = n$. Then $tE_n - J$ has determinant $\prod_{i=1}^{n}(t - c_i)$ (where $c_1, \ldots c_n$ are diagonal entries of J. The multiplicity of λ as a root of $\chi_L(t) = \det(tE_n - J)$ equals the number of occurences of λ among $c_1 \ldots c_n$ which in turns equals the sum of the orders of the Jordan blocks associated with λ.

proof[12]: Let s be the number of the Jordan blocks associated with λ and $u_1, \dots u_s$ be the starting vectors for the Jordan chains.
The form of the Jordan matrix implies that $v \in V$ is an eigenvector of L associated with λ if and only if its coordinates relative to the Jordan basis <u>other than</u> (除3) $u_1, \dots u_s$ are all zero. Therefore $u_1, \dots u_s$ form a basis of $\ker(L - \lambda \mathrm{id}_V)$.

<u>Thm 12.4</u> Let $A \in M_n(\mathbb{C})$. Then there exists an invertible matrix $U \in M_n(\mathbb{C})$ and a Jordan matrix $J \in M_n(\mathbb{C})$ such that $A = UJU^{-1}$. Moreover, such the Jordan matrix J is unique up to the order of the Jordan blocks

Proof. Consider $L \in \mathcal{L}(M_{n,1}(\mathbb{C}))$, $L(X) = AX$. Remind that $[L]_{\mathcal{F}} = A$, where $\mathcal{F}$ is the standard basis of $M_{n,1}(\mathbb{C})$. By Thm 12.1. there exists a basis $\mathcal{J}$ such that $[L]_{\mathcal{J}} = J$ is a Jordan matrix.
Let $U \in M_n(\mathbb{C})$ be the transition matrix from $\mathcal{F}$ to $\mathcal{J}$.
Then $A = [L]_{\mathcal{F}} = U[L]_{\mathcal{J}}U^{-1} = UJU^{-1}$ as required.

Jordan basis

Uniqueness: Let $A = U'J'U'^{-1}$. By Lemma 6.8 $\exists\ \mathcal{J}'$ of $M_{n,1}(\mathbb{C})$ s.t. $M_{\mathcal{F} \sim \mathcal{J}} = U'$
Then $[L]_{\mathcal{J}'} = U'^{-1}[L]_{\mathcal{F}}U' = U'^{-1}AU' = J'$ i.e. $\mathcal{J}'$ is a Jordan basis.
Thm 12.1 implies J and J' are equal up to the order of the blocks.

<u>Def.</u> The Jordan matrix J from the above thm. is called Jordan normal form of A.

<u>Algorithm:</u> For any $A \in M_n(\mathbb{R})$. $A = UJU^{-1}$. find J. and U.

1. Compute eigenvalues of $L: x \mapsto Ax$.
→ we find orthogonal complent of $\operatorname{Im} C$
2. denote $C = A - \lambda E$. compute. $\boxed{C^T z = 0}$ the basis vector of the solution $z_1, z_2, \dots z_s$.
solve. $CX = 0$. (solution is basis vector of Jordan chain with length≥ 1).
→ find vectors in $\ker C$.
extend. $\hat{C} = \begin{bmatrix} C \\ z_1^T \\ \vdots \\ z_s^T \end{bmatrix}$.
solve. $\hat{C}x = 0$. (solution is basis vector of Jordan chain with length ≥ 1)
$\Rightarrow X_1^1.\ X_2^1 \dots X_t^1$

X_1^s							
X_1^{s-1}	X_2^t						
$\dots$		$\dots$					
X_1^2	X_2^2	$\dots$					
X_1^1	X_2^1	X_3^1	$\dots$	X_t	X_{t+1}	$\dots$	X_{t+k}

solve. $\hat{C}X' = X_j^1$ → consisting Jordan chain length >2.
→ inconsistent length $=2$.
solve. $CX = X_j^1. \Rightarrow X_j^2, \{X_j^1, X_j^2\}$

$$U = (X_1^{(1)} | X_1^{(2)} | \cdots | X_1^{(s)} | X_2^{(1)} | \cdots | X_{t+1} | X_{t+2} | \cdots | X_{t+k})$$

$J = \operatorname{diag}(J(\lambda_1), \dots J(\lambda_q))$. Check: $\underline{\underline{AU}} = UJ$

An alternative form of Jordan blocks. (has 1 on the subdiagonal).

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

$J' \Rightarrow$ each diag block equal $J_n'(\lambda)$

Application of Jordan normal form.

Thm 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 s.t. $[L]_{\mathcal{J}'} = J'$. Moreover, J' is unique up to the order of the Jordan blocks $J_n'(\lambda)$

Pf: match the indices of Jordan basis w.r.t. Jordan chain.

the matrix relative the basis is J'.

Lemma 12.6.

$$J_m(\lambda)^N = \begin{bmatrix} \lambda^N & \binom{N}{1}\lambda^{N-1} & \binom{N}{2}\lambda^{N-2} & \cdots & \binom{N}{n-2}\lambda^{N-n+2} & \binom{N}{n-1}\lambda^{N-n+1} \\ 0 & \lambda^N & \binom{N}{1}\lambda^{N-1} & \cdots & \binom{N}{n-3}\lambda^{N-n+3} & \binom{N}{n-2}\lambda^{N-n+2} \\ & & \lambda^N & & & \\ & & & \ddots & & \\ & & & & \lambda^N & \binom{N}{1}\lambda^{N-1} \\ & 0 & & & & \lambda^N \end{bmatrix}$$

it is assumed that $\binom{N}{m} = 0$ for $m > N$.

Pf: by induction. $\binom{N+1}{m} = \binom{N}{m-1} + \binom{N}{m}$

Remark: In this way we can compute. $A^N = (UJU^{-1})^N = UJ^NU^{-1}$

Lemma 12.8. Let $Q \in M_n(\mathbb{C})$ be nilpotent. Then there exist $S \in M_n(\mathbb{C})$. s.t. $S^2 = E_n + Q$.

Pf: First shows. for any $m \geq 0$. $\exists f_m \in \mathbb{R}[t]_m$. s.t. $f_m(0) = 1$ and $f_m^2(t) = 1 + t + g_m(t) \cdot t^{m+1}$ for some. $g_m \in \mathbb{R}[t]$

Then put $S = f_{n-1}(Q)$. $S^2 = f_{n-1}^2(Q) = E_n + Q + g_{n-1}(Q) \cdot Q^n = E_n + Q$.

Remark. $f(t) = 1 + \sum_{j=1}^{n-1} \dfrac{\binom{\frac{1}{2}}{j}}{j!} t^j$ (n th Taylor polynomial for $\sqrt{1+t}$).

whence $S = E_n + \sum_{j=1}^{n-1} \dfrac{\binom{\frac{1}{2}}{j}}{j!} Q^j$.

Pro 12.9. Let $A \in M_n(\mathbb{C})$. be invertible. Then there exists. an invertible $B \in M_n(\mathbb{C})$. s.t. $A = B^2$

Pf: Thm 12.4 implies. $\exists$ Jordan J. invertible U. $A = UJU^{-1}$ ($J^{(1)} \dots J^{(s)}$ its blocks.)

For any $1 \leq k \leq s$. $J^{(k)} = J_m(\lambda) = \lambda E_m + Q$. Q is strictly upper tri: matrix

A is invertible. $\lambda \neq 0$. and then $\lambda^{-1}J^{(k)} = E_m + \lambda^{-1}Q$.

by Lemma 12.8. $\exists$. S_k s.t. $S_k^2 = \lambda^{-1}J^{(k)}$ $(\sqrt{\lambda}S_k)^2 = J^{(k)}$.

$\Rightarrow B = USU^{-1}$. $S = \begin{bmatrix} \sqrt{\lambda}S_1 & & \\ & \ddots & \\ & & \sqrt{\lambda}S_k \end{bmatrix}$

§ 13. Quadratic and Bilinear forms.

Def. Quadratic form. f is variable $X_1, X_2, \dots X_n$. a <u>homogeneous polynomial</u> of second degree. (each monomial has total degree 2)

$$f(X_1, X_2, \cdots X_n) = \sum_{i,j} a_{ij} X_i X_j$$

$A_f = (a_{ij})$ matrix of a quadratic form. symmetric

$$f(X_1, X_2, \dots, X_n) = (X_1, X_2, \dots X_n) \cdot A_f \, (X_1, X_2, X_3, \dots X_n)^T$$

Also, we can define quadratic form $\sum_{i \le j} b_{ij} X_i X_j$ (to avoid repetition $b_{ij} = a_{ij} + a_{ji}$ $= 2a_{ij}$)

Change variable. $X_1'. X_2' \dots X_n'$

$$\begin{cases} X_1 = C_{11} X_1' + C_{12} X_2' + \cdots + C_{1n} X_n' \\ X_2 = C_{21} X_1' + \cdots + C_{2n} X_n' \\ \vdots \\ X_n = C_{n1} X_1' + \cdots + C_{nn} X_n' \end{cases}$$

(where transition matrix $C = (C_{ij})_{i,j=1}^n$ is invertible)

new quadratic form $\tilde{f}$ $\quad \tilde{f}(X_1', X_2', \dots X_n') = f(X_1, X_2, \dots X_n)$ is equivalent to f

$$(X_1', X_2', \dots X_n') A_{\tilde{f}} (X_1', \dots X_n')^T = (X_1, \dots X_n) A_f (X_1, \dots X_n)^T$$

$$= ((X_1', \dots X_n') \cdot C^T) \cdot A_f \cdot (C \cdot (X_1', \dots X_n')^T)$$

$$\Rightarrow A_{\tilde{f}} = C^T A_f C \quad (\text{or } A' = C^T A C \text{ simply}).$$

<u>Thm 13.2.</u> Lagrange diagonalization theorem.

- For any quadratic form in variables $X_1, X_2, \dots X_n$. over a field F there exist an invertible linear change of variables $(X_1 \dots X_n)^T = C \cdot (X_1' \dots X_n')^T$. such that $f((X_1, \dots X_n) \cdot C^T) = d_1 (X_1')^2 + d_2 (X_2')^2 + \cdots + d_n (X_n')^2$ for some $d_1, \dots d_n \in F$.

- For any symmetric matrix A (over F) $\exists$ invertible matrix C. such that $C^T A C$ is a diagonal matrix.

Algorithm

$$\begin{cases} X_1' = C_{11} X_1 + \cdots + C_{1n} X_n \\ X_2' = C_{21} X_1 + C_{22} X_2 + \cdots + C_{2n} X_n \\ \vdots \\ X_n' = C_{n1} X_1 + \cdots + C_{nn} X_n \end{cases} \quad \begin{pmatrix} X_1' \\ \vdots \\ X_n' \end{pmatrix} = C \cdot (X_1, \dots X_n)^T$$

$\Rightarrow$ this can be diagonal.

to obtain this. ① we eliminate. $X_i X_i$ and obtain $d_1 X_1^2 + g(X_2, \dots X_n)$ repeat ...

② if coefficient of X_i^2 all 0. but coefficient of $X_1 X_2 \ne 0$. $X_1 = (X_1' + X_2')$. $X_2 = (X_1' - X_2')$

$$f(X_1 \dots X_n) = (X_1 \dots X_n) A \cdot (X_1 \dots X_n)^T \sim f(X_1', \dots X_n') = (X_1' \dots X_n') A' (X_1' \dots X_n')^T$$

$A = C^T A' \cdot C \quad A'$ can be diagonal

A modern def. of quadratic form

Def. Let V be a vector space over F. A $\boxed{\text{function}}$ $q: V \to F$ is a quadratic form on V if there exists a basis (can be any basis) $B = \{v_1 \dots v_n\}$ and homogenous quadratic polynomial f in variables $X_1, \dots, X_n$ such that for any $x_1, x_2, \dots x_n \in F$ one has

$$q(x_1 v_1 + \cdots + x_n v_n) = f(x_1, x_2, \dots x_n)$$

Thus. to compute the value of f. on arbitrary $v \in V$.

take $v = \sum_{i=1}^{n} x_i v_i \qquad (x_1, \dots, x_n)^T = [v]_B$

then compute the corresponding $f(x_1, x_2, \dots x_n)$

Remark

"$\exists$ a basis" $\to$ "any basis." we have $(X_1', X_2', \dots X_n') C^T = (X_1, X_2, \dots, X_n)$.

homogeneous quadratic polynomial f for quadratic form $q: V \to F$.

$\Longleftrightarrow$ (same role as). $[L]_B$ for linear operator $L: V \to V$

(Since then, we write $f = [q]_B$. to write $f_{q,B}$ stressing the choice of basis)

Def. Gram matrix.

Given a quadratic form q and a basis $B = \{v_1, \dots v_n\}$. the symmetric matrix A. such that. $q(v) = (x_1, \dots x_n) A (x_1 \dots x_n)^T$ for any vector $v = x_1 v_1 + \dots + x_n v_n$.

A is Gram matrix. of the quadratic form q relative to the basis B

Def. bilinear form.

Let V be a vector space over a field F. Let $h: V \times V \to F$ be a function of two vector arguments. h is called bilinear form on a vector space V if it satisfies:

$h(au_1 + u_2, v) = a h(u_1, v) + h(u_2, v)$ for all $u, v_1, v_2 \in V$ and $a \in F$.

$h(u, av_1 + v_2) = a h(u, v_1) + h(u + v_2)$ for all $u, v_1, v_2 \in V$ and $a \in F$.

(if we fixed one of the argument, we obtain a linear function).

Given basis. $B = \{v_1 \dots v_n\}$. $(u, v) \in V \times V$. ($u = \sum x_i v_i$, $v = \sum y_i v_i$).

$$h(u,v) = \sum_{i=1}^{n} x_i\, h(v_i, y_1 v_1 + \dots + y_n v_n) = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} y_j h(v_i, v_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j h(v_i, v_j).$$

Def. Gram matrix of bilinear form. (线性空间中表示基的转换关系). $L: V \to F^n$

$x \mapsto [h(x, v_1) \dots h(x, v_n)]^T$.

Let $G_{h,B} = (g_{ij})$ be a matrix s.t. $g_{ij} = h(v_i, v_j)$. Then $G_{h,B}$ is called a Gram matrix of a bilinear form h relative to the basis B.

$\Downarrow$

$[L]_{B,\varepsilon} = G_{h,B}$

$$\sum_{i,j=1}^{n} x_i y_j h(v_i, v_j) = (x_1 \dots x_n)\, G_{h,B}\, (y_1 \dots y_n)^T.$$

$h(u,v) = x^T G_{h,B}\, y \qquad x = [u]_B \quad y = [v]_B$ (coordinates column)

Def. $h: V \times V \to F$. is symmetric. if $h(u,v) = h(v,u)$ for any $u, v \in V$. (Gram matrix symmetric).

Thm13.3 (Same notion in quadratic and symmetric bilinear form)

Let $h: V\times V \to F$. be a symmetric bilinear form on a vector space V.
Then $q_h(v) = h(v,v)$ is a quardratic form.

- A polarization of the quadratic form q. (itself is a bilinear form).

Let $q: V\to F$. be quadratic form a vector space V. Then $h_q(u,v) = \frac{1}{2}(q(u+v) - q(u) - q(v))$.
the Construtions are inverse to each other.

Pf: 1. h_p is linear on both vector arguments.
it is sufficient to proof. in a certion basis $\mathcal{B}$ it is computed by $h(u,v) = [u]_B^T G_{h,B} [v]_B$

consider $h_A(u,v) = x^T A y$ (A is the matrix of quadratic form q).

consider a bilinear form $h_A(u,v) = x^T A y$ where $x = [u]_B$ and $y = [v]_B$.

In order to prove $h_A = h_q$. we need to check.

$$x^T A y = \frac{1}{2}((x+y)^T A (x+y) - x^T A x - y^T A y)$$

(first construct quadratic form q. then construct a bilinear form h_q. check $h_q(v,v) = q(v)$.

In another direction. first construct bilinear form h, then consider $q = q_h$.

$h_q(u,v) = \frac{1}{2}(h(u+v, u+v) - h(u,u) - h(v,v)) = h(u,v)$.

Def. Axiomatic def of quadratic form.

A function $q: V \to F$ which is the function $h_q: V\times V \to F$. of two vector given by $h_q = \frac{1}{2}(q(u+v) - q(u) - q(v))$ is bilinear

Proposition 13.4. Let V be quadratic space $h: V\times V \to F$. be a symmetric bilinear form and $q: V\to F$ is corresponding quadratic form. For an invertible $L \in \mathcal{L}(V)$.
the following statement are equivalent.

- L preserves the bilinear form h. : $h(L(u), L(v)) = h(u,v)$. for $\forall u,v \in V$
- L preserves the quadratic form q: $q(L(v)) = q(v)$ for $\forall v \in V$

by Thm 13.3. $h_q(L(u), L(v)) = \frac{1}{2}(q(L(u)+L(v)) - q(L(u)) - q(L(v)))$

$$= \frac{1}{2}(q(u+v) - q(u) - q(v)) = h_q(u,v)$$

Def. Orthogonal Operator

Invertible operator $L: V\to V$ is said to be orthogonal w.r.t. a quadratic (symmetric bilinear) form if this operator preserve the given form

$$x^T G y = ([L]_B x)^T G ([L]_B y) = x^T([L]_B^T G [L]_B) y$$

Pro 13.5. n-dim vs V. over F. basis $\mathcal{B}$. h bilinear form. Let G be a Gram matrix, $A \in M_n(F)$ is the matrix of an operator on V w.r.t h relative to $\mathcal{B}$. $\Leftrightarrow$ $G = A^T G A$.

(by uniqueness of Gram matrix).

(if $G = E$ (standbasis.). $A^T A = E$. A is orthogonal

Def. Vectors orthogonal.

$u, v \in V$ are orthogonal w.r.t h. if $h(u,v)=0$. (we write $u \perp_h v$)

Def. Orthogonal Complement.

For a subset (not nessary.) $U \subset V$ we define its ~~orthogonal~~ orthogonal complent w.r.t h by

$U^{\perp h} = \{v \in V \mid h(v,u)=0 \text{ for all } u \in U\}$.

(if qua or bilinear form is given by context. omit the index h)

Pro 13.6 Properties of Orthogonal Complement (正交补集).

1. $U^{\perp}$ is a subspace in V
2. $U \subset (U^{\perp})^{\perp}$
3. $U_1 \subset U_2 \Rightarrow U_1^{\perp} \supset U_2^{\perp}$
4. $(U_1+U_2)^{\perp} = U_1^{\perp} \cap U_2^{\perp}$
5. $(U_1+\cdots+U_k)^{\perp} = U_1^{\perp} \cap U_2^{\perp} \cap \cdots \cap U_k^{\perp}$
6. $U = \mathrm{Span}(u_1, u_2, \ldots u_k)$ then $U^{\perp} = u_1^{\perp} \cap \cdots \cap u_k^{\perp}$
7. $\dim U^{\perp} \geq n - \dim U$

proof: 2) $\forall u \in U$, $\forall v \in U^{\perp}$ s.t. $h(v,u)=0$. i.e. $v \perp u$.

$u \in \{v \in V \mid h(u,v)=0. \text{ for all } x \in U^{\perp}\}$. i.e. $U \subset (U^{\perp})^{\perp}$.

3). $\forall x \in U_2^{\perp}$. $\forall u_2 \in U_2 \Rightarrow h(u_2, x)=0$

thus for any $u_1 \in U_1 \subset U_2$, $h(u_1; x)=0$ holds.

thus $x \in \{v \in V \mid h(x, u_1)=0 \text{ for all } u_1 \in U_1\} = U_1^{\perp}$.

4) $(U_1+U_2)^{\perp} \subset U_i^{\perp}$ from 3).

"$\supset$". $\forall x \in U_1^{\perp} \cap U_2^{\perp}$. $h(x,u_1)=0$, $h(x,u_2)=0$ for all $u_1 \in U_1$. $u_2 \in U_2$.

$h(x, u_1+u_2) = h(x,u_1)+h(x,u_2)=0$. for all $u_1+u_2 \in U_1+U_2$.

5) from 4). use induction

6). particular case in 5) $U_i = \mathrm{span}(u_i)$. $u_i^{\perp} = (\mathrm{span}(u_i))^{\perp}$.

7) Let $U = \mathrm{Span}(u_1, \ldots u_k)$ where $k = \dim U$

Consider a map $L: V \mapsto F^k$ assign v the column $(h(v,u_1), \ldots h(v,u_k))^T$.

Check L is linear

By def. $\mathrm{Ker}(L) = \bigcap_{i=1}^{k} u_i^{\perp} = U^{\perp}$ by 6)

$\dim U^{\perp} = \dim \mathrm{Ker}\, L = \dim V - \dim \mathrm{Im}(L)$ $\quad \mathrm{Im}(L) \subset F^k$.

Def. non-degenerate symmetric bilinear

Symmetric bilinear form $h: V \times V \to F$ is non-degenerate if $V^{\perp h} = 0$

i.e. $\forall u \neq 0$. $\exists v \in V$ s.t. $h(u,v) \neq 0$.

conversely. $\exists u \neq 0$. $\langle u \rangle^{\perp h} = V$ (→ notation. all vectors orthogonal to u). the bilinear form h and corresponding quadratic form q are called degenerate.

Pro 13.7. The following conditions are equivalent.

1. h is non-degenerate

2. For any basis $\mathcal{B}$. the $G_{h,\mathcal{B}}$ has non-zero determinant.

3. there exists a basis $\mathcal{B}$ such that Gram matrix has non-zero determinant.

proof $1° \Rightarrow 2°$ Take $\mathcal{B} = \{v_1 \dots v_n\}$. Consider a linear map $L: V \to F^n$ $x \mapsto (h(x,v_1), h(x,v_2), \dots h(x,v_n))^T$

$[L]_{\mathcal{B},\mathcal{E}} = G_{h,\mathcal{B}}$. ($\mathcal{E}$ is the standard basis F^n)

h is non-degenerate. $\forall x \in V$ (nonzero) the column $L(x)$. has nonzero component.

(otherwise $x \in \bigcap_{i=1}^{n} u_i = V^\perp = 0$.)

Therefore $\mathrm{Ker}(L) = 0$. and homogeneous l.s. $[L]_{\mathcal{B},\mathcal{E}} \cdot \boldsymbol{x} = 0$ has only trivial solution

$2° \Rightarrow 3°$ Obvious.

$3° \Rightarrow 1°$ similarly. construction. $\mathcal{B}$. L. $[L]_{\mathcal{B},\mathcal{E}} = G_{h,\mathcal{B}}$

Hint: in standard basis. $h\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) = \sum_{i=1}^{n} x_i y_i$

(to find the subspace $\langle u \rangle^\perp$).

Def. Restriction.

For any subspace $U \sqsubset V$. and h, q on V can consider a restriction $h|_U$ or $q|_U$.

It is usual restriction of a domain for arbitrary map.

Trivially, the map $h|_U : U \times U \to F$. remains to be bilinear

Moreover, the quadratic form U corresponding to $h|_U$ would be equal the restriction $q|_U$.

For Gram matrix. let $\mathcal{B}_U = \{v_1, v_2, \dots v_k\}$. extend. $\mathcal{B} = \{v_1, \dots v_k, v_{k+1}, \dots v_n\}$.

$$G_{h,\mathcal{B}} = \begin{bmatrix} G_{h|U,\mathcal{B}_U} & \cdots \\ \cdots & \cdots \end{bmatrix}$$ (upper-left corner of size $k \times k$)

Thm 13.9. (Orthogonal Completion for non-degenerate quadratic form).

Let V be n-dim v.s. and. h be a non-degenerate symmetric bilinear form on V. then:

1. For any subspace $U \sqsubset V$ one has $\dim U^\perp = n - \dim U$
2. $(U^\perp)^\perp = U$; If $U_1^\perp = U_2^\perp$ then $U_1 = U_2$ for any subspaces $U_1, U_2 \sqsubset V$.
3. For any subspace $U_1, U_2 \sqsubset V$ one has $(U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp$

proof: By prove of Pro 13.6.7. $\dim U^\perp = \dim \ker L = \dim V - \dim \mathrm{Im}(L)$.

We need to show L is a surjective. ($L : V \to F^k$). (we need the lemma to proof).

Lemma 13.10. Let $W \sqsubset U$ is a proper subspace. Then there exists non-zero linear functional $\varphi: U \to F$. s.t. $\varphi|_W \equiv 0$.

proof: Consider $\forall$ non-degenerate $\tilde{h}$ on U. $W^{\perp \tilde{h}} \neq 0$ ($\dim W^{\perp \tilde{h}} \geq \dim U - \dim W$).

take. nonzero $u \in W^{\perp \tilde{h}}$ one obtains. $x \mapsto \tilde{h}(x,u) = 0$ being restriction on W.

but $\tilde{h}$ is non-degenerate. contradict. (all non-degenerate is impossible)

Resuming the proof. Assume the converse. $\exists \varphi: F^k \to F$. $\varphi|_{\mathrm{Im}(L)} = 0$. $\varphi \circ L|_{\mathrm{Im}L} = 0$.

φ is linear. $\varphi((x_1, \dots x_n)^T) = a_1 x_1 + \dots + a_n x_n$. ($\varphi$ is nonzero. $\exists a_i \neq 0$)

$(\varphi \circ L)(v) = a_1 h(v, u_1) + \dots + a_k h(v, u_k) = h(v, \sum^k a_i u_i) \Rightarrow h$ non-degenerate. $\Rightarrow \sum_{i=1}^{k} a_i u_i = 0 \Rightarrow \{u_1 \dots u_n\}$ l.i.

proof2. by Pro 13.6.2. $U \subset (U^\perp)^\perp$. then check $\dim U = \dim (U^\perp)^\perp$.

Applying 1). $\dim U^\perp + \dim (U^\perp)^\perp = n$ (subspace $U^\perp$).

$\dim U + \dim U^\perp = n$ (subspace U).

proof3. sufficient to check. $((U_1 \cap U_2)^\perp)^\perp = (U_1^\perp + U_2^\perp)^\perp$ ("$\perp$" can be strip off)

by pro 13.6.4. $(U_1^\perp + U_2^\perp)^\perp = (U_1^\perp)^\perp \cap (U_2^\perp)^\perp = U_1 \cap U_2$.

(give us a way to compute the intersection. $U_1 \cap U_2 = (U_1^\perp + U_2^\perp)^\perp$.

Thm 13.12. (Riesz representation theorem (baby version)).

Let V be n-dim v.s. and h is a non-degenerate bilinear form.

Then for any bilinear functional $\varphi \in \mathcal{L}(V, F)$. there exists a unique vector v_φ such that $\varphi(x) = h(x, v_\varphi)$ for any vector $x \in V$.

e.g. h is the dot product.

§ 14. Orthogonal decomposition (正交分解).

Def. Gram–Schmidt. orthogalization.

Let V be a vector space with a given bilinear form h and q is corresponding qua. form.

Orthogonalization is a process of find <u>orthogonal basis</u>.

i.e. B_0 $\{v_1, \dots v_n\}$. s.t. $G_{h,B}$ is diagonal. ($h(v_i, v_i) = 0$ for all $i = j$.).

Pro 14.1. In any vector space V with a symmetric bilinear form h there exists an orthogonal basis.

(restatement of Lagrange diagonalization theorem).

Proof. If $h(u,v) = 0$ for all $u, v \in V$. Any basis be orthogonal.

If $\exists$ u, v $h(u,v) \neq 0$. we can choose $v_1 \in V$.: $q(v_1) \neq 0$. denote $V' = u_1^{\perp}$.

We can prove that. $V = \text{span}(v_1) \oplus V'$. ($\dim V' = n-1$).

Then proof by induction. hypothesis: the pro holds for V' and bilinear form $h_{V'}$.

So we can let $B_{V'} = \{v_2, v_3, \dots, v_n\}$. then by $V = \text{span}(v_1) \oplus V'$ and Pro 4.6.

$\{v_1, \dots v_n\}$ is a basis of V. $h(v_1, v_i) = 0$. ($i \geq 2$. $v_i \in V' = v_1^{\perp}$).

and $h(v_i, v_j) = 0$ ($i \neq j \geq 2$) by hypothesis.

thus we have. $h(v_i, v_j) = 0$ ($i \neq j$, $1 \leq i, j \leq n$).

Deduce. Lagrange diagonalization from Pro 14. 1.

For given symmetric matrix A. consider h on F^n which has A as its Gram matrix.

relative to basis $\mathcal{E} = \{e_1, e_2, \dots, e_n\}$. (standard basis).

that is to say. define h by $h(x, y) = x^T A y$

By pro 14. 1. $\exists$ $\mathcal{E}'$. $C = M_{\mathcal{E} \to \mathcal{E}'}$. such that Gram matrix A' has diagonal form.

$\Rightarrow A' = C^T A C$

Gram–Schmidt orthogonalization process.

<u>Thm 14. 2.</u> Let V be v.s over F. with h (s. b. f.). Given a basis $\mathcal{B} = \{v_1, v_2, \dots v_n\}$. suppose.

for any $k = 1 \dots n$ the restriction $h_{\text{span}(v_1 \dots v_k)}$ is non-degenerate.

Then there exist an orthogonal basis $\mathcal{E} = \{e_1, e_2, \dots, e_n\}$. s.t. $e_k \in \text{span}(v_1, \dots v_k)$. for any k.

△ [The $M_{B_v \to B_e}$ is lower triangular

△ the upper-left $k \times k$ minor in the Gram matrix $G_{h,B}$ is non-zero. (the condition) equivalent.

△ span ($e_1, \dots e_k$) = span ($v_1, \dots v_k$).

proof by induction. $n = 1$ is trivial.

Apply $V' = \text{span}(v_1, \dots v_{n-1})$. have already found. $B_0 = \{e_1, e_2, \dots e_{n-1}\}$. dim $V' = n$ be basis.

since $v_n \notin V' = \text{span}(e_1 \dots e_{n-1})$. we obtains a linearly independent system $\{e_1, \dots e_{n-1}, v_n\}$.

let us. modify the vector v_n in $\{e_1 \dots e_{n-1}\ v_n\}$ to $\{e_1 \dots e_{n-1},\ v_n'\}$. v_n' is orthogonal.

let. $v_n' = v_n - a_1 e_1 - a_2 e_2 - \dots - a_{n-1} e_{n-1}$

$$h(v_n', e_k) = h(v_n, e_k) - \sum_{i=1}^{n-1} a_i h(e_i, e_k) = h(v_n, e_k) - a_k h(e_k, e_k)$$

since $h(v_n', e_k) = 0$. thus. $a_k = \dfrac{h(v_n, e_k)}{h(e_k, e_k)}$

($h(e_k, e_k) \neq 0$. Since. $G_{h|V', \mathcal{E}'}$. is diagonal and $\det G_{h|V', \mathcal{E}'} \neq 0$.)

So $\quad e_n = v_n' = v_n - \sum_{i=1}^{n-1} \dfrac{h(v_n, e_k)}{h(e_k, e_k)} e_k$. and obtain the basis □.

Def. orthonormal basis

v.s V. s.b.f. h The basis $B = \{v_1, \dots v_n\}$. is said to be orthonormal. if the Gram matrix $G_{h,B}$ is equal to Identity matrix.

i.e. B is orthogonal and $h(e_k, e_k) = 1$ for $k = 1, 2, \dots n$.

Remark: Not alway the orthonormal basis exists.

In the case $F = \mathbb{R}$. and $h(v,v) > 0$ for all $v \in V$ ($v \neq 0$) (orthonormal basis exists 的充分条件).
we can make further step in Thm 14.2. denoting $e_k' = \dfrac{e_k}{\sqrt{h(e_k, e_k)}}$ → "length"
obtain orthonormal basis $\{e_1', e_2', \dots, e_n'\}$.

Remark: Let $V = \mathbb{R}^n$. h is dot product. $G_{h,B} = E_n$. means B is standard basis.
$h(\vec{x}, \vec{x}) = x_1^2 + \dots + x_n^2 > 0$ for ($\vec{x} \neq 0$).

Pro 14.3. Consider $V = \mathbb{R}^n$ with bilinear form given by standard dot product

For a matrix $A \in M_n(\mathbb{R})$ the following conditions are equivalent:

1. A linear operator $x \mapsto A \cdot x$ is a orthogonal operator on $\mathbb{R}$.
2. A is orthogonal matrix ($A^T A = E$) ⇒ $\boxed{\text{def}}$
3. Columns of matrix A constitute (形成). an orthonormal basis in $\mathbb{R}^n$.
4. $AA^T = E$
5. If $x_1^T, x_2^T, \dots x_n^T$ are rows of A then $\{x_1 \dots x_n\}$ is an orthonormal basis in $\mathbb{R}$.

proof: 1) ⇔ 2) Pro 13.5

2) ⇔ 3) $h(x, y) = x^T y$ Let $A = [y_1, y_2, \dots y_n]$

$h(y_i, y_j) = y_i^T y_j$ (ith row of A^T. j th column of A) $A^T A = (h(y_i, y_j))$

4) ⇔ 5). (Follow the 2) ⇔ 3))

2) ⇔ 4). First check A is invertible

$n = \operatorname{rank}(A^T A) \le \operatorname{rank} A^T, \operatorname{rank}(A)$. $\quad \operatorname{rank}(A) = n$.

$A^T A = E \iff A^T A A^{-1} = E A^{-1} \quad A^T = A^{-1}$.

$AA^T = AA^{-1} = E$.

Coro 14.5. Let A be a symmetric matrix s.t. all its upper left minors are non-zero. Then there exists a low triangular matrix L s.t. LAL^T is diagonal.

<u>Pf</u>: Let $A = G_{h,B}$.

By thm 14.2. $\exists\ \mathcal{E}_0 = \{e_1 \dots e_n\}$ is orthogonal.

$M_{B\sim\mathcal{E}}$ is upper triangular. Thus. $G_{h,\mathcal{E}} = M_{B\sim\mathcal{E}}^{\ T} A\, M_{B\sim\mathcal{E}}$ (put $L = M_{B\sim\mathcal{E}}^{T}$).

$*\ B = -B^T$ skew-symmetric matrix

Def. decomposition into orthogonal direct sum and orthogonal decomposition.

Let V (v.s.) h (s.b.f.). $V = U_1 \oplus U_2 \oplus \cdots \oplus U_k$. if it is a direct sum decomposition s.t. for any $x \in U_i$ and $y \in U_j$ one has $h(x,y) = 0$. provided $i \neq j$.

We use the notation $V = U_1 \perp U_2 \perp \cdots \perp U_k$.

Pro 14.6. Let $U \sqsubset V$ be a subspace s.t. $h|_U$ is non-degenerate. Then there is an orthogonal direct sum decomposition $V = U \perp U^\perp$. Moreover, if there is another orthogonal decomposition $V = U \perp W$ then $W = U^\perp$.

proof: 1. First we check $U \cap U^\perp = 0$.

Let $u \in U \cap U^\perp$ for any $v \in U$, $h(u,v) = 0$. (since $u \in U^\perp$).

thus we have $u = 0$. since $h|_U$ is non-degenerate.

Second. By Pro 13.6.7. $\dim U^\perp \geqslant \dim V - \dim U$.

By thm 4.5. $\dim(U^\perp + U) = \dim U + \dim U^\perp \geqslant \dim V$. $\Rightarrow V = U^\perp + U$.

by pro 4.6. $U^\perp \oplus U = V$.

2. $W \sqsubset U^\perp$. ($\forall w \in W$. s.t. $h(w,u) = 0$ for any $u \in U$. thus $w \in U^\perp$).

$\dim W = \dim V - \dim U = \dim U^\perp$.

Def: Orthogonal Projector. (正交投影算子). 注意.正交投影算子的前提.

Let. <u>$h|_U$ non-degenerate</u>. Let <u>$V = U \perp W$</u> is the decomposition into orthogonal direct sum. Define a linear operator $P: V \to V$ in the following way. Take $v \in V$ and consider its unique representation $v = u + w$ where $u \in U$ and $w \in W$. Put $P(v) = u$. Then the operator P is called orthogonal projector onto subspace U. $P = \mathrm{Proj}_U$.

Pro 14.8. Suppose that $\{u_1 \dots u_k\}$ is <u>an orthogonal basis</u> of the subspace $U \sqsubset V$ and $h(u_i,u_i) \neq 0$ $(i = 1 \dots k)$.

Then for any vector v. $\mathrm{Proj}_U(v) = \sum_{i=1}^{k} \frac{h(v,u_i)}{h(u_i,u_i)} u_i$ ($\mathrm{Proj}_{U^\perp}(v) = v - \sum_{i=1}^{k} \frac{h(v,u_i)}{h(u_i,u_i)} u_i$)

Proof: $v = \sum_{i=1}^{k} \frac{h(v,u_i)}{h(u_i,u_i)} u_i + \left(v - \sum_{i=1}^{k} \frac{h(v,u_i)}{h(u_i,u_i)} u_i\right)$ (first summand $\in U$)

We need to check. second summand $\in U^\perp$

$$h\left(v - \sum_{i=1}^{k} \frac{h(v,u_i)}{h(u_i,u_i)} u_i,\ u_k\right) = h(v,u_k) - \sum_{i=1}^{k} \frac{h(v,u_i)}{h(u_i,u_i)} h(u_i, u_k) = h(v,u_k) - \frac{h(v,u_k)}{h(u_k,u_k)} h(u_k,u_k) = 0.$$

In special case $U = V$. and $\{u_1 \dots u_n\}$ is a <u>orthonormal</u> basis in V. $\Rightarrow v = \sum_{i=1}^{n} h(v,u_i) u_i \Rightarrow$ <u>Fourier decomposition</u> of v. w.r.t. basis $\{v_1 \dots v_n\}$.

If h is dot product, $h|_{U}$ ($U \sqsubset V$) is non-degenerate. $h > 0$ ($|\vec{a}| > 0$, $a \neq 0$)

Pro 14.10. Let $U \subset \mathbb{R}^n$. h is the dot product on $\mathbb{R}^n$ and $\mathcal{Y} = \{y_1, \dots y_k\}$ is the basis in U.
Denote by Y the matrix with the columns $y_1, \dots, y_k$. Then

1. $k \times k$ matrix $Y^T Y$ is Gram matrix $G_{h|U, \mathcal{Y}}$ $(h(y_i, y_j) = y_i^T y_j)$.
2. If $\mathcal{Y}$ is an orthonormal basis in U then $n \times n$ matrix YY^T is the matrix of orthogonal projector $Proj_U$ relative to the standard basis in $\mathbb{R}^n$.
3. In general case, $Y(Y^T Y)^{-1} Y^T$ is the matrix of projector $Proj_U$ relative to the standard basis in $\mathbb{R}^n$.

proof: 2). Let $\{y_{k+1}, \dots y_n\}$ be a basis of $U^\perp$. Then $\{y_1 \dots y_n\}$ is the basis in $\mathbb{R}^n$.
By thm 5.1. it is sufficient to check $Proj_U(y_i) = (YY^T)(y_i)$. (for all $i = 1 \dots n$).
We have $(YY^T)Y = Y(Y^T Y) = Y \cdot E_k = Y$. thus we check $i = 1, \dots k$.
For those $j > k$ one has $(YY^T) y_j = Y(Y^T y_j) = 0$. $\longrightarrow$ the ith component $= y_i^T y_j$ where $y_i \in U$. $y_j \in U^\perp$.
3) similarly.

Pro 14.11. Let $\mathbb{R}^n$ be equipped with a standard dot product. The matrix $P \in M_n(A)$ defines an orthogonal projector on same subspace. (usual way $x \mapsto P \cdot x$), if and only if $P^2 = P$ and $P^T = P$

Pf "$\Rightarrow$" trivial
"$\Leftarrow$" Denote subspace $\operatorname{Im} P$ by U. (if $x \in \operatorname{Im} P$. $\exists z$. $x = Pz$. $Px = PPz = P^2 z = Pz = x$.)
Then for any $x \in U$, one has $P\vec{x} = \vec{x}$
Further, given $y \in U^\perp$. for any $x \in U$. $h(x, P(y)) = x^T P y = x^T P^T y = h(P(x), y) = 0$.
thus. $P(y) \in U^\perp$. other hand, $P(y) \in \operatorname{Im} P = U$. $\Rightarrow P(y) \in U^\perp \cap U = 0$.
$\forall v \in \mathbb{R}^n$ $v = x + y$ where $x \in U$. $y \in U^\perp$.
$P(v) = P(x) + P(y) = P(x) = x$. thus P is an orthogonal projector onto U.

Euclidean Space.

Def. metric space - one can measure a distance between any two points.
normed space - a vector space equipped with a length function

Def. Euclidean Space (real inner space) (V, h).
Let V be a real vector space equipped with a symmetric bilinear form h such that $h(v,v) \geq 0$ for any $v \in V$ and $h(v,v) = 0$ if and only if $v = 0$.

The corresponding quadratic form is called positive definite quadratic form.
Its Gram matrix is also called a positive definite matrix

e.g. the column space $\mathbb{R}^n$ equipped with dot product. (main example)
Given an arbitrary (V, h). chose $B = \{v_1, \dots v_n\}$ orthonormal, define a linear map. $\theta_B : V \to \mathbb{R}^n$ $v \to \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ (coordinates)
θ is isomorphism and Euclidean; given bilinear form is preserved by θ. if $x = \theta(v)$ and $y = \theta(w)$.
then $h(v, w) = x^T y$ ($\Rightarrow$ dot product)

Def. An isomorphism of the Euclidean space.

Let (V,h). (V,h') be Euclidean space. The linear map $\theta: V \to V$ is called an isomorphism of these Euclidean spaces if. $\forall u,v \in V$. $h(u,v) = h'(\theta(u), \theta(v))$

Remark. h be a bijection automatically.

since $\mathrm{Ker}\,\theta = 0$ as $\theta(v) = 0 \Leftrightarrow h'(\theta(v), \theta(v)) = 0 \Leftrightarrow h(v,v) = 0 \Leftrightarrow v = 0$

Difference between (V,h) and $(\mathbb{R}^n, \cdot)$. (only difference)

In the space $\mathbb{R}^n$ the orthonormal basis is supposed to be already chose (standard basis) but in V there are no preferable choice of the basis. (all orthonormal basis are equal naturally).

Def. length of vector v.

Let (V,h) be a Euclidean space. The quantity $\sqrt{h(v,v)}$ is called the length of vector $v \in V$ and denote by $\|v\|$.

(θ preserves the length of any vector $v \in V$).

Pro 14.12. (Cauchy-Буняковский-Schwarz inequality).

For any two vector $u,v \in V$. in Euclidean space one has.

1. $|h(u,v)| \leq \|u\| \cdot \|v\|$
2. $\|u+v\| \leq \|u\| + \|v\|$. (triangle equality).

Pf: 1. Consider $h(tu+v, tu+v) = h(u,u) \cdot t^2 + 2h(u,v) \cdot t + h(v,v) \in \mathbb{R}[t]$

non-negative for any $t \in \mathbb{R}$. $4h(u,v)^2 - 4h(u,u)h(v,v) \leq 0$. $\square$

2. by 1. $\|u+v\|^2 = h(u+v, u+v) = h(u,u) + 2h(u,v) + h(v,v) \leq \|u\|^2 + 2\|u\| \cdot \|v\| + \|v\|^2 = (\|u\| + \|v\|)^2$

Least Squares.

Def. (Distance). In Euclidean space V one can measure the distance between to point for $u,v \in V$. one defines $\mathrm{dist}(u,v) = \|u-v\|$

Problem. For given $v \in V$. find $u_0 \in U$ s.t. $\mathrm{dist}(v, u_0) = \min\limits_{u \in U} \mathrm{dis}(v,u)$

u_0 should be a base of the altitude drawn from v to U $u_0 = \mathrm{Proj}_U(v)$.

(check $\forall u. \|v-u_0\| \leq \|v-u\|$. $v-u = \underset{U^\perp}{(v-u_0)} + \underset{U}{(u_0-u)}$).

Def. least square solution.

Let $A: \mathbb{R}^n \to \mathbb{R}^m$ is a linear map given by the matrix $A_{m \times n}$. Consider (possibly inconsistent l.s.).

$A \cdot \mathbf{x} = \mathbf{b}$ (where $b \in \mathbb{R}^m$). A column x_0 is called a least square solution to the system:

if the norm of the difference $\| Ax_0 - b \|$ is minimal as possible. $\|Ax_0 - b\| \leq \|Ax - b\|$ for any $x \in \mathbb{R}^m$

[if $\mathbb{R}^m$ and $\mathbb{R}^n$ are equipped with dot product "the norm of the vector $y = (y_1 \dots y_m)^T$" is equal to $\|y\| = \sqrt{y_1^2 + \dots + y_k^2}$ "the norm of": vector $\to$ length.]

Pro 14.15. 1. The set of all least square solutions of a given linear system coincide with the set of ~~all set of~~ all exact solutions of the system $A\cdot x = b_0$ where $b_0 = \mathrm{Proj}_{\mathrm{Im}A}(b)$

2. The set of all least square solutions of a given linear system coincide with the set of all exact solutions of the (so called normal linear system) $A^TAx = A^Tb$

3. When $\mathrm{rank}(A) = n$, then the least square solution of l.s. is unique, and given by

$$x_0 = (A^TA)^{-1}A^Tb$$

proof 1. If x_0 is the least square solution, the point $Ax_0 \in \mathrm{Im}A \subset \mathbb{R}^m$ is the closest point to $\vec{b}$ in subspace $\mathrm{Im}A$. by Problem, denote $b_0 = \mathrm{Proj}_{\mathrm{Im}A}(b)$. the closest point is unique. $Ax_0 = b_0$.

Proof 2. We need the Lemma first.

Lemma 14.16. Under above notation. Let $U = \mathrm{Im}A \subset \mathbb{R}^m$. For any $y, y' \in \mathbb{R}^m$ one has

$$\mathrm{Proj}_U(y) = \mathrm{Proj}_U(y') \Leftrightarrow A^Ty = A^Ty'.$$

Pf: $U^\perp = \ker(A^T)$. (A^Ty row vector of $A^T \in U$).

Let $y = u + z$ and $y' = u' + z'$ where $u, u' \in U$. $z, z' \in U^\perp$ (orthogonal decomposition)

Then $A^Tu = A^Tu + A^Tz = A^Ty$. similarly. $A^Tu' = A^Ty'$

Thus. $A^Ty = A^Ty' \Leftrightarrow A^T(u-u') = 0 \Leftrightarrow u-u' \in \ker(A^T) = U^\perp \Leftrightarrow u-u' \in U\cap U^\perp = 0 \Leftrightarrow u = u'$ □

Continue: x is ~~the~~ a least square solution of $Ax = b \Leftrightarrow Ax = \mathrm{Proj}_{\mathrm{Im}A}(b)$

and We have $\mathrm{Proj}_{\mathrm{Im}A}(Ax) = \mathrm{Proj}(b) \Leftrightarrow A^TAx = A^Tb$

Proof 3. $\mathrm{rank}(A)$ = the numbers of column. column vectors are l.i. their form a basis of their span $U = \mathrm{Im}(A)$. A^TA is a Gram matrix of the restriction of dot product on U w.r.t. the basis. (A^TA is invertible due to the positive definiteness).

Lagrange Interpolation.

Def.

x	c_1	c_2	c_3	$\cdots$	c_{n-1}	c_n
$f(x)$	b_1	b_2	b_3	$\cdots$	b_{n-1}	b_n

$c_i \in F$. $b_i \in F$.
the point c_i interpolation nodes. b_i value at the node c_i.

Pro 14.17. There exists a unique $p \in F[t]_{n-1}$ such that $p(c_i) = b_i$. That is, the interpolation problem has an exact solutions in the class of polynomial of degree less or equal to $n-1$.

Pf: Consider map $\theta: F[X]_{n-1} \to F^n$. $p(X) \mapsto (p(c_1) \ldots p(c_n))^T$.

Check the linearity.

Check the bijectivity. (it's sufficient to check $\ker(\theta) = 0$.

$p \in \ker\theta \Leftrightarrow p(c_i) = 0$. $\deg(p) \leq n-1$. can't have n different roots $\Leftrightarrow p = 0$

Remark: For the θ. $\theta(X^k) = (c_1^k, c_2^k, \ldots, c_n^k)^T$.
so if we let the standard basis of F^n. $B = \{1, X, X^2, \cdots X^{n-1}\}$ of $F[X]_{n-1}$
we have $[\theta]_{B,\mathcal{E}}$ be the Vandermonde matrix $= \begin{bmatrix} 1 & c_1 & \cdots & c_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & c_n & \cdots & c_n^{n-1} \end{bmatrix}$

Lagrange interpolation formula.

Consider a polynomial which correspond to the vectors of standard basis in F^n under the θ.
(i.e. $\theta(L_i(X)) = e_i$) $\theta: p \to \begin{pmatrix} p(c_1) \\ \vdots \\ p(c_n) \end{pmatrix}$.

For any $i = 1, \dots, n$. there is a unique polynomial. $L_i \in F[X]_{n-1}$.

which have

x	c_1	c_2	$\cdots$	c_i	$\cdots$	c_n
$f(x)$	0	0	$\cdots$	1	$\cdots$	0

$$L_i(X) = \frac{(X-c_1)(X-c_2)\cdots(X-c_{i-1})(X-c_{i+1})\cdots(X-c_n)}{(c_i-c_1)(c_i-c_2)\cdots(c_i-c_{i-1})(c_i-c_{i+1})\cdots(c_i-c_n)}$$

let the original polynomial be $f(X)$.

$\theta(f(X)) = b_1e_1 + \cdots + b_ne_n = \theta(b_1L_1(X) \cdots + b_nL_n(X))$ $f(X) = \sum_{i=1}^{n} b_iL_i(X)$ (where $f(c_i) = b_i$).

Pro 14.18. $f(X) = \sum_{i=1}^{n} f(c_i) \prod_{k \neq i} \frac{X-c_k}{c_i-c_k}$

which gives the unique of the interpolation problem. (find the polynomial)

§15 Adjoint Operator

Def Hermitian form: (Since it's impossible to define positivity for complex bilinear form).
An Hermitian form on a complex vector space V is a function $f: V \times V \to \mathbb{C}$ s.t.
the following axioms hold:

1. $f(au+v, w) = af(u,w) + f(v,w)$. for any $u, v, w \in V$, $a \in \mathbb{C}$.
2. $f(v,u) = \overline{f(u,v)}$. for any $u, v \in V$.

For second axiom $f(w, au+v) = \overline{f(au+v, w)} = \overline{af(u,w) + f(v,w)} = \bar{a} f(w,u) + f(w,v)$.

If we fix the first argument. $f(w,x) = \varphi(x)$. $\varphi(x)$ is antilinear.

$\varphi(x+y) = \varphi(x) + \varphi(y)$. $\varphi(\lambda x) = \bar{\lambda}\varphi(x)$

Similar as bilinear form. For a given basis $B = \{v_1, \dots, v_n\}$. $G_{f,B} = (f(v_i, v_j))_{i,j=1}^{n}$

If. $[u]_B = x$. $[v]_B = y$. $f(u,v) = x^T G_{f,B} \bar{y}$

$f(u,v) = \overline{f(v,u)} \Leftrightarrow x^T G_{f,B} \bar{y} = \overline{x^T G_{f,B} \bar{y}} = \bar{x} \overline{G_{f,B}} y = y^T \overline{G_{f,B}^T} \bar{x}$

Def Complex matrix $A \in M_n(\mathbb{C})$ is called hermitian symmetric if $\overline{A^T} = A$.

The Gram matrix of hermitian form is symmetric.

If $C = M_{B \to B'}$ $G_{f,B'} = C^T G_{f,B} \bar{C}$

Pro 15.1. (Lagrange diagonalization form). For any hermitian symmetric matrix $A \in M_n(\mathbb{C})$ there exist an invertible complex matrix C s.t. $C^T A \bar{C}$ is real diagonal.

pf: $f(v,v) = \overline{f(v,v)}$ so $f(v,v) \in \mathbb{R}$.

If $A = \overline{A^T}$ then $A^T = \bar{A}$. $\det A = \det(A^T) = \det(\bar{A}) = \overline{\det A}$, $\det A$ is a real number.

Def. (positive definite).

An hermitian form $f: V\times V \to \mathbb{C}$ is called positive definite if $f(v,v)>0$. for any non-zero vector $v\in V$.

A is called positive definite matrix if $x^T A \bar{x} > 0$. for any non-zero column $x \in \mathbb{C}^n$.

(The Sylvester criterion for positive definite matrice holds also)

Def. A complex inner space (or unitary space) is a complex vector space equipped with a positive definite hermitian form.

Postive definite hermitian form is frequently called. unitary scalar product. (or even simply scalar product).

Def. The operator on a unitary space which preserve unitary scalar product is called unitary operator. The matrix defining unitary operator on a standard. column space $\mathbb{C}^n$ is called a unitary matrix. (保持向量模长不变，夹角不变).

Remark. the term. "unitary" is a complex counterpart of the term "orthogonal" in real case

For the matrix of the unitary operator. $G_{f,B} = A^T G_{f,B} \bar{A}$ (where $A=[L]_B$

For unitary matrix U. s.t. $U^T \bar{U} = E$ (且 $|\det U| = 1$)

Def. The matrix $\overline{A^T}$ is usually called an hermitian conjugate or hermitian transpose to A.

Remark: If (V,f) is a complex inner space (or real inner space). then $f(u_1,u_2)$. will be denoted be $(u_1,u_2)_V$.

Adjoint operator

Pro 15.2. For any linear map $L: V \to V'$ there exists a unique linear map $L^*: V' \to V$ s.t. for any pair of vectors $v\in V$, $v' \in V'$ one has. $f(L(v), v')_{V'} = f(v, L^*(v'))_V$

Def: L^* is called an adjoint (operator) of a linear map L.

Pf: choose orthonormal basis in V and V' be B and B'. $\Rightarrow [L(v)]_{B'}^T G \overline{[v']_{B'}} = [v]_B^T G \overline{[L^*(v')]_B}$

For $x=[v]_B$. $y=[v']_{B'}$. [we need to check $x^T A^T \bar{y} = x^T \overline{A' y}$]

Since $x^T A^T \bar{y} = x^T \cdot (\overline{\bar{A}^T y})$ we let $A^* = \overline{A^T}$, which defined the adjoint operator

Uniqueness: for given $v' \in V'$ consider a linear map $x \xrightarrow{\theta} (L(x), v')_{V'}$ from V to $\mathbb{C}$ or $\mathbb{R}$.

By the Riesz representation thm. (前提是 non-degenerate) θ can be given by a scalar product by a unique vector $u\in V$: $(L(x), v')_{V'} = \theta(x) = (x,u)_V$. $L^*(v') = u$ nessarily.

Coro 15.3. The matrix of adjoint linear map w.r.t. orthonormal bases is the hermitian conjugate to the matrix of a given linear map relative to the same bases. ($\overline{A^T} = A^*$).

(if the basis the arbitrary. $A^T G' = G \overline{A^*}$

Def. self-adjoint. the matrix of operator $\bar{A} = A^T$.

A linear operator $L: V \to V$ on a inner space V is called self-adjoint if $L^* = L$.

Pro15.4. Characteristic polynomial of the self-adjoint operator is always a polynomial with real coefficients, all its complex roots are real, so the characteristic polynomial is completely factorizable over the $\mathbb{R}$.

Pf: Choose an orthonormal basis. The matrix A of the operator s.t. $A^T = \bar{A}$
the characteristic polynomials of the matrices A and A^T coincide.
$\bar{A}$'s polynomial can be obtained by applying complex conjugation to all coefficient of A.

take $\lambda \in \mathbb{C}$. $\lambda \bar{x} = A\bar{x}$. $\lambda x^T \bar{x} = (\lambda x^T) \cdot \bar{x} = x^T A^T \bar{x} = x^T \bar{A} \bar{x} = x^T \bar{\lambda} \bar{x} = \bar{\lambda} x^T \bar{x} \Rightarrow \bar{\lambda} = \lambda$.

Pro15.5. For a given self-adjoint operator on an inner space its eigenvectors belonging to the different eigenvalues are orthogonal to each other

Pf: Let a standard complex inner space $\mathbb{C}^n$ equipped with a hermition dot product with orthogonal basis. and the self-adjoint operator is given by A (s.t. $A^T = \bar{A}$).

We suppose $Ax = \lambda x$ and $Ay = \mu y$.

$$\mu x^T \bar{y} = \bar{\mu} x^T \bar{y} = x^T \bar{\mu} \bar{y} = x^T \bar{A} \bar{y} = x^T A^T \bar{y} = \lambda x^T \bar{y}$$

since $\mu \neq \lambda$. $x^T \bar{y} = 0$.

self-adjoint operator. can find orthonormal basis consist of eigenvector. (different eigenspace $\perp$. in an eigenspace: orthogonalization

Pro15.6. (Properties of adjoint map).

Let $A_1, A_2: V \to U$ and $B: W \to V$ be a linear maps between (complex) inner spaces. Then:

- $(A^*)^* = A$.
- $(A_1 + A_2)^* = A_1^* + A_2^*$, $(aA)^* = \bar{a} A^*$
- $(AB)^* = B^* A^*$

Pf: Just turn the problem of operator into the problem of matrix.
or directly proof: $(AB(w), u)_U = (A(B(w)), u)_U = (B(w), A^*(u))_V = (w, B^*A^*(u))_W$.

Pro 15.7. (Kernel and Image of adjoint map).

Let $L: V \to V'$ be a linear map between inner space and $L^*: V \to V$ be an adjoint map of L.

Then $Ker(L) = (Im(L^*))^\perp$, $Im(L) = (Ker(L^*))^\perp$

Pf: Two equality can transformation mutually. since $(L^*)^* = L$. $(U^\perp)^\perp = U$.

1°. "$\subseteq$". $v \in Ker(L)$ $\forall v' \in V'$, $0 = (L(v), v')_{V'} = (v, L^*(v'))_V$ v' is orthogonal to any vector in $Im L^*$.

2° "$\supseteq$". $v \in (Im(L^*))^\perp$, $\forall v' \in V'$, $L^*(v') \in Im(L^*)$ $(v, L^*(v'))_V = 0 = (L(v), v')_{V'}$.
$L(v) \in (V')^\perp$ since the hermition form is positive definite thus $(V')^\perp = \{0\}$.
$L(v) = 0$. $v \in Ker(L)$

Pro 15.8. Let $L: V \to V$ be a linear operator on a unitary (or Euclidean space V). Then for any subspace $U \subset V$ if U is invariant under L. then $U^{\perp}$ is invariant under L^*.

Pf: Let us take $\forall w \in U^{\perp}$ and check that $L^*(w) \in U^{\perp}$.

for any $u \in U$. $(u, L^*(w))_V = (\underset{U}{L(u)}, \underset{U^{\perp}}{w})_V = 0$.

Canonical form of self-adjoint operator.

$(Q \cdot Q^T = E)$

we need to prove that ① for any real symmetric matrix A. $\exists$ an orthogonal matrix Q. s.t. $Q^{-1}AQ$ is diagonal. ② for any hermitian symmetric matrix $B \in M_n(\mathbb{C})$. $\exists$ unitary matrix. U s.t. $U^{-1}BU$ is diagonal.

Algorithm for ①: 1) Consider linear operator $A: \mathbb{R}^n \to \mathbb{R}^n$ $x \mapsto Ax$.

2). we have $A = [A]_F$. $D = [A]_B = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$

$\Rightarrow D = Q^{-1}AQ$.. Q is the transition matrix $F \to B$

3). compute eigenvalue. eigenvector.
find orthogonal basis. vectors $\to$ $\tilde{Q}$ having those vectors in the column.

4) $\tilde{Q}$ divide column by length. obtain orthogonal Q.

Thm 15.12. Consider a self-adjoint operator $L: V \to V$ on an inner space V. Then there exists an orthonormal basis s.t. the matrix of the given operator relative to this basis is diagonal.

Pf: by introduction on the dimension. let $\dim V = n$. $\dim = 1$ is trivial.
Assume $\dim' = n-1$ has been proved.

By Pro 15.4. we can chosen $\lambda_1 \in \mathbb{R}$. and corresponding $u_1 \in V$. i.e. $L(u_1) = \lambda u_1$.
we assume $\|u_1\| = 1$ (if not, divide by its length).
the subspace. $\text{span}(u_1)$ is L-invariant.
Therefore the $W = \text{span}(u_1)^{\perp}$ is also L-invariant (by Pro 15.8 and $L = L^*$)., $\dim W = n-1$.
We can consider $L|_W$ and the restriction of the given scalar product on W.
Applying the hypothesis to $L_W: W \to W$: we choose an orthonormal basis in W consisting of eigenvectors of $L|_W$. (i.e. $\exists u_2, \dots u_n \in W$ s.t. $L(u_i) = \lambda_i u_i$. $2 \le i \le n$)
since $V = \text{span}(u_1) \perp W$. we have orthonormal basis $B = \{u_1 \dots u_n\}$ satisfied the requirements.
and the matrix is $[L]_B = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$.

§16. Canonical form of normal operator

Normal operator on unitary space

Def. (normal operator)

An operator $L: V \to V$ on an inner space ($\mathbb{R}$ or $\mathbb{C}$) is called normal if $LL^* = L^*L$

i.e. the product with an adjoint is commutative.

e.g. 1) self-adjoint operator ($L = L^*$).

2) skew-symmetric operator ($L^* = -L$).

3) orthogonal or unitary operator ($L^*L = Id$).

Lemma 16.1. Operator $L: V \to V$ on an inner space preserves scalar product $\Leftrightarrow$ $L^{-1} = L^*$

Pf: "$\Rightarrow$". $(u,v) = (L(u), L(v))$. for any $u, v \in V$.

$\Leftrightarrow (L(u), L(v)) = (u, L^*L(v)) \Leftrightarrow (u, L^*L(v) - v) = (u, L^*L(v)) - (u, v) = 0$.

$u \neq 0$. the product non-degenerate. thus $L^*L(v) = v = Id_V$. $\Rightarrow L^{-1} = L^*$.

Def. (isometry 等距变换 operator). the common word for orthogonal and unitary operator.

Thm 16.2. Let L be a normal operator on an unitary space. Then there exists an orthonormal basis consisting of eigenvectors of L.

Coro 16.3. Let $A \in M_n(\mathbb{C})$. s.t. $A\bar{A}^T = \bar{A}^T A$. Then there exists a unitary matrix U. s.t. $U^T A \bar{U}$ is diagonal.

Consider $x \mapsto Ax$. is normal. Use Thm 15.13 and 16.2. Denote $U^T = Q^{-1}$. Q is $M_{\mathcal{F} \to \mathcal{B}}$ orthonormal

($U^T\bar{U} = E_n$ by unitary)

Lemma 16.4. Let $T: V \to V$ be a linear operator s.t. $L \circ T = T \circ L$. Then the eigenspace $U = V_\lambda(L)$ is invariant under T.

Pf. We need to prove that for any eigenvector u of L the vector $T(u)$ would be also an eigenvector of L corresponding to the same eigenvalue λ. (i.e. $L(T(u)) = \lambda T(u)$).

$L \circ T(u) = T \circ L(u) = T(\lambda u) = \lambda T(u)$.

Pf of 16.2. Let λ be an eigenvalue of L. Denote $U = V_\lambda(L)$. By Lemma 16.4. U is L^*-invariant.

By Pro 15.8. $U^\perp$ is L-invariant ($(L^*)^* = L$).

Therefore we proceed by induction on $\dim V$.

$V = U \perp U^\perp$, both of them are L and L^* invariant.

$(L|_W(x), y)_W = (L(x), y)_V = (x, L^*(y))_V = (x, L^*|_W(y))_W$ for any x, y.

Then consider $L|_{U^\perp}$. $(L|_{U^\perp})^* = L^*|_{U^\perp}$ and the restriction $L|_{U^\perp}$ is also normal.

thus we can apply inductive hypothesis. find an orthonormal basis in $U^\perp$ consisting of eigenvectors of L).

(用的是第二归纳法: $V = U \perp U^\perp$ 证 $L|_U$, $L|_{U^\perp}$ normal. 两个子空间分别运用归纳假设)

Canonical Form of Normal Operator (on Euclidean vector space).

normal operator may have imaginary eigenvalues. we can not find corresponding eigenvector in real vector space for this imaginary eigenvalues.

Lemma 16.5 Let $L: V \to V$ be a normal operator on Euclidean space $\lambda = a+bi$ where $b \neq 0$. Consider a polynomial $p(t) = (t-\lambda)(t-\bar{\lambda}) = t^2 - 2at + (a^2+b^2)$. Let $T: V \to V$. $T = p(L) = L^2 - 2aL + (a^2+b^2)Id$. Then the L-invariant subspace $U = \ker(T)$ has an orthonormal basis such that the matrix of $L|_U$ has block-diagonal form with 2×2 blocks of the form $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

Pf: 1° $\ker T$ is an eigenspace of T corresponding to the zero eigenvalue. ($\ker T = V_0(T)$).

$\ker T$ is L^*-invariant (by Lemma 16.4). L-invariant (by Pro 8.5).

($\ker T \neq 0$ since the matrix T w.r.t. arbitrary basis has zero determinant,

$[T]_B [V]_B = 0$ has non-zero solution). $[T] = ([L]-\lambda E)([L]-\bar{\lambda}E)$,

2° Denote by $S: U \to U$ an operator $L|_U - aId_U$. $S^* = L^*|_U - \bar{a}Id_U$, $S^*S = SS^*$

we want to check $S + S^* = 0$. ① $L^2 - 2aL + a^2 Id = -b^2 Id \Leftrightarrow S^2 = -b^2 Id_U$

② $(S^*)^2 = (S^2)^* = -b^2 Id_U = S^2$.

Consider $(S - S^*)(S + S^*) = S^2 - (S^*)^2 - (SS^* + S^*S) = 0$.

Assume the converse. take $x \in U$. $y = (S+S^*)(x) \neq 0$ we have $(S-S^*)(y) = 0$. $S(y) = S^*(y) \neq 0$.

$0 < (S(y), S^*(y)) = (S^2(y), y) = (-b^2 y, y) = -b^2(y,y) < 0$. contradicts.

3. $\forall u \in U$. $(u \neq 0)$. assume $\|u\| = 1$. $S(u) \neq ku$. (otherwise. $S^2(u) = k^2 u \neq -b^2 u$).

Denote $v = b^{-1} S(u) \Rightarrow S(v) = -bu$. $S(u) = bv$ ' $S^*(v) = bu$. $S^*(u) = -bv$.

thus, span (u,v) is invariant under S, S^*, thus invariant under L, L^*. (denote by U_0)

since $b(v,u) = (S(u), u) = (u, S^*(u)) = (u, -S(u)) = -(S(u), u)$

thus $(v,u) = 0$. $v \perp u$, $\|v\| = \|u\| = 1$. (Check)

$[S]_{B_0} = \begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix}$. $B_0 = \{u, v\}$ is an orthonormal basis.

$[L|_{U_0}]_{B_0} = [S]_{B_0} + aE = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

$U_0^\perp$ is L, L^*-invariant by Pro 15.8.

4. Prove for any L-invariant subspace U s.t. $p(L)|_U = 0$. $\exists B_{U_0}$ have the properties.

Proceed by induction on $\dim U$. if $\dim U = \dim U_0$ done.

find orthonormal basis in U_0. s.t. $L|_{U_0^\perp}$ has a block-diag form. (Induction hypothesis).

Union of basis in $U_0^\perp$ and B_0 we obtain the desired orthogonal basis.

Thm 16.6. Let L be a normal operator on the Euclidean space V. Then there exists an orthonormal basis $\mathcal{B}$ s.t. the matrix $[L]_{\mathcal{B}}$ has a block-diagonal form. with the blocks of size 1×1 and 2×2. where 1-dimensional blocks corresponds to real eigenvalue and 2×2 blocks of the form $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. correspond to pairs of complex eigenvalue $a\pm bi$
(the canonical form)

Pf: Proceed by introduction on $\dim V$.

Let $\lambda \in \mathbb{R}$. root. make induction step as in Thm 16.2.

$\lambda \notin \mathbb{R}$. $\lambda = a+bi$. consider U as in Lemma 16.5. $U, U^{\perp}$ are L, L^*-invariant. apply induction hypothesis to $U^{\perp}$ and apply Lemma 16.5 to U.

Find the canonical form. of matrix A. $A' = \begin{bmatrix} \lambda & & \\ & \begin{smallmatrix} a & -b \\ b & a \end{smallmatrix} & \\ & & \ddots \end{bmatrix} = Q^{-1}AQ$. Q is orthogonal ($\bar{Q}^T Q = E$)

1) compute the eigenvalue / vector. real roots is simple.
2) find orthonormal basis in imaginary-eigenvalue subspace.
3) denote Q be columns of orthonormal basis.

Canonical form of isometries.

Lemma 16.8. Let λ be an eigenvalue of unitary (or orthogonal) matrix A. Then $\lambda \cdot \bar{\lambda} = 1$.
(λ belongs to the unit circle in complex plane $\cos\varphi + i\sin\varphi$).

Pf: We know that $A^T \cdot \bar{A} = E$. Then $Ax = \lambda x$ implies that $x^T A^T = \lambda x^T$ and $\bar{A}\bar{x}^T = \bar{\lambda}\bar{x}$.

Therefore $x^T\bar{x} = x^T A^T \bar{A}\bar{x} = \lambda x^T \bar{\lambda}\bar{x} = \lambda\bar{\lambda} x^T \bar{x}$.

Since $x^T\bar{x} > 0$. then $\lambda\bar{\lambda} = 1$.

Coro 16.9. Let $L: V \to V$ be a unitary operator on a complex inner space V. Then there exists an orthonormal basis $\mathcal{B}$ in V such that the matrix $[L]_{\mathcal{B}}$ is diagonal where all the components belong to unit circle.

Pf: By Thm 16.2. the orthonormal basis can be made of eigenvectors.
$L(u) = \lambda u$. By Lemma 16.8. $\lambda \in \cos\varphi + i\sin\varphi$.

in 3-dim. $\begin{pmatrix} 1 & & \\ & \cos\varphi & -\sin\varphi \\ & \sin\varphi & \cos\varphi \end{pmatrix}$ $\begin{pmatrix} -1 & & \\ & \cos\varphi & \sin\varphi \\ & \sin\varphi & \cos\varphi \end{pmatrix}$.
only two possibilities.
compute the trace. (preserve)

Thm 16.10. Let $L: V \to V$ be an orthogonal operator on a Euclidean space. Then there exists an orthonormal basis such that the matrix of L relative to this basis is block diagonal with blocks of size 1×1 and 2×2 where $[\pm1]$. $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$.
Moreover, assuming that φ can be equal to π one can regard that there is at most one 1×1 block. In the case $\det(L) = 1$. this 1×1 block necessarily contain $+1$.

Pf: L is normal ($L^* = L^{-1}$). If $\lambda = a+bi$. (non-real). apply lemma 16.8. $a = \cos\varphi$. $b = \sin\varphi$.

If λ is real $|\lambda| = 1$. $\Rightarrow \lambda = \pm1$.

2×2 blocks of the canonical form in Thm 16.6. are just rotation 2×2 matrix.

When $\det(L) = 1$. there are even numbers of -1. organize a pair of two -1. into $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$.

$\det L = 1 \Rightarrow$ rotation

$\det L = -1 \Rightarrow$ reflection $L' = \begin{pmatrix} -1 & & \\ & 1 & \\ & & 1 \end{pmatrix}$. $L'' = \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix} \to$ reflection in 0 space. (central symmetric)

Coro 16.11. (Euler rotation thm).
Every isometry in 3-dim Euclidean space preserving origin and orientation. (i.e. having postive determinant) is a rotation around around a certain axis.

scalar product = inner product (positive definite)
standard ~~pr~~ scalar product on $\mathbb{C}^n$ ($\mathbb{R}^n$) = dot product.

Reflection (example of the isometry operators on inner space).

Def. (Reflection) Let $V = U \perp W$ be orthogonal decomposition for any $v \in V$ consider a unique representation $v = u + w$ where $u \in U$, $w \in W$.
The linear map $s_W : V \to V$ which assigns $u - w$ to the vector $v = u + w$ is a reflection in the subspace U. The subspace U is called a mirror of the reflection S_W.
(Closely related with orthogonal projector).

Pro 16.13. Reflection is an isometry. (i.e. it preserves scalar product).
(particular. $S_U = \mathrm{span}(v_0)$; S_U is usually denoted by s_{v_0} and called a reflection along v_0.)

Pro 16.14. For any $x \in V$. one has $s_{v_0} = x - 2\dfrac{(x, v_0)}{(v_0, v_0)} \cdot v_0$.

Pf: Let $x = kv_0 + u$. where $u \perp v_0$. Then $(x, v_0) = k(v_0, v_0) + (u, v_0) = k(v_0, v_0)$.

$$x - 2\frac{(x, v_0)}{(v_0, v_0)} \cdot v_0 = kv_0 + u - 2kv_0 = u - kv_0 = s_{v_0}(x).$$

Pro 16.15. Let $P = \mathrm{Proj}_U$ where $V = U \perp W$. Then $Id_V - 2P$ is the reflection S_U in W

Pf: For $v = u + w$ where $u \in U$, and $w \in W$ one has $P(v) = u$.

$$(Id - 2P)(v) = v - 2u = w - u. = S_U(v).$$

Pro 16.16. Reflection L is a selfadjoint operator such that $L^2 = Id$.

Pf: Let $L = S_U$. where $V = U \perp W$. $(Id - 2P)^* = Id - 2P^*$.
Check $P = \mathrm{Proj}_U$ is self-adjoint. $(P(v), v') = (u, v') = (u, u' + w') = (u, u') = (v, P(v'))$.

$$L^2(v) = L(L(v)) = L(w - u) = w + u = v.$$

Remark: $LL^* = Id$. (orthogonal) we have $L^2 = Id \iff L = L^*$ only selfadjoint orthogonal operator satisfies. $L^2 = Id$.

§17. Polar Decomposition.

Positive operators.

In real inner space. we find correspondence between selfadjoint operators and quadratic forms, then we can reformulate them abstractly to avoid matrix notation.

Pro 17.1. Let V be a real inner space. Then given a self-adjoint operator $L: V \to V$ the formula $h_L(u,v) = (L(u), v)$ define a symmetric bilinear form. From thes other hand, given a symmetric bilinear form h one can uniquely define an operator L_h s.t. $h(u,v) = (u, L_h(v))$.

Pf: "$\Rightarrow$" $(u,v) \mapsto (L(u), v)$ is linear. (L is linear. scalar product is linear).
h_L is symmetric by symmetry of scalar product. and $L^* = L$
$h_L(u,v) = (L(u), v) = (u, L(v)) = (L(v), u) = h_L(v,u)$.
"$\Leftarrow$" Consider a fixed v. $x \mapsto h(x,v)$ By Riesz representation thm. any linear functional on inner space can be given by a scalar product with a suitable vector (i.e. $\exists!\, w \in V$, s.t. $h(x,w) = (x,w)$ for $x \in V$)
Denote w by $L(v)$. $L(v+v') = L(v) + L(v')$ $((x, L(v)+L(v')) = (x, L(v+v')))$. L is linear.
$\Rightarrow h(x,v) = (x, L(v))$, check L is self-adjoint $(x, L(v)) = (L(x), v)$

Pro 17.2. Let V be a complex inner space. For a given hermitian form $f: V \times V \to \mathbb{C}$. there exists a unique linear operator $L_f: V \to V$ s.t. $(x, L_f(v)) = f(x,v)$ for any $x, v \in V$. Moreover, L_f is selfadjoint.

And in another direction, given a selfadjoint operator L on an unitary space one can define another hermitian form $f_L(u,v) = (u, L(v))$.

Remark: if $B = \{e_1, \dots, e_n\}$ is orthonormal basis in an inner space V and $L: V \to V$ is a self-adjoint operator then the matrix $[L]_B = G^T_{h_L, B}$

$L(e_j) = \sum_{i=1}^{n} (L(e_j), e_i) e_i \xrightarrow{\text{Fourier decomposition}} [L]_{ij} = (L(e_j), e_i) = h_L(e_j, e_i)$

Def. (positive operator)
A self-adjoint operator $L: V \to V$. on an inner space V. is called strictly positive if the corresponding bilinear (hermitian) form is positive definite. i.e. $(L(u), u) > 0$. for any $u \neq 0$

Pro 17.3. Let $L: V \to V$ be an invertible operator on an inner space. Then the operator LL^* is strictly positive (selfadjoint) operator.

Pf: $(LL^*(u), u) = (L^*(u), L^*(u)) > 0$. since $L^*(u) \neq 0$ for $u \neq 0$. (invertible)

Pro 17.4. For a given self-adjoint operator L. following statement are equivalent:

- L is positive
- All the roots of characteristic polynomial $\chi_L(t)$ are positive real numbers.
- All the upper-left minor of the matrix L relative to an orthonormal basis are positive.

Pf: 1) $\Rightarrow$ 2). consider $\lambda \in \mathbb{R}$. corresponding eigenvector $v \in V$. Then $0 < (L(v), v) = \lambda (v,v) \Rightarrow \lambda > 0$.
2) $\Rightarrow$ 1) consider an orthonormal basis $\{v_1, \dots, v_n\}$ consisting of eigenvectors of L. Then $L(v_i) = \lambda_i v_i$ where $\lambda_i > 0$ and $(v_i, v_j) = 0$ for $i \neq j$. Then for any $v = \sum_{i=1}^{n} x_i v_i$.
one has $(L(v), v) = \left(\sum_{i=1}^{n} \lambda_i x_i v_i, \sum_{i=1}^{n} x_i v_i \right) = \sum_{i=1}^{n} \lambda_i |x_i|^2 > 0$

1) ⇒ 3). it's sufficient to prove that determinant of the Gram matrix of positive definite form is positive. definite. (the restriction of the positive definite form is p.d.).
determinant is just the product of all eigen values of the L, which is positive.

3) ⇒ 1) Consider hermitian (bilinear) form f_L. G be Gram matrix.
Given orthonormal basis in the inner space need not be orthogonal w.r.t f_L.
We can apply the Gram-Shmidt orthogonalization process and obtain a diagonal matrix $G' = C^T G \bar{C}$ where C is upper triangular with units on the diag.
$G'_k = C_k G_k \bar{C}_k$. $|G'_k| = |G_k|$. (G' has the same upper left minors as G).
⇒ G' is diag with positive component ⇒ define a p.d.b. (hermitian).

Pro 17.5. For positive self-adjoint operator L on an inner space V there exists a unique positive operator $T: V \to V$ such that $T^2 = L$.

Pf: First to derive existence of canonical form of the self-adjoint operator
There exists an orthonormal basis such that the matrix $[L]$ w.r.t. this basis is diagonal with positive elements since they are equal to eigenvalues.
then consider $D = \begin{bmatrix} \sqrt{\lambda_1} & & & 0 \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ 0 & & & \sqrt{\lambda_n} \end{bmatrix}$ (T. w.r.t orthonormal basis).

For uniqueness. Let $\lambda_1 \dots \lambda_k$ be all different eigenvalue.
by Thm 10.1 $V = V_{\lambda_1}(L) \perp \cdots \perp V_{\lambda_k}(L)$.
We denote $T|_{V_{\lambda_i}(L)}$. for any $v = \sum_{i=1}^{k} v_i$. $v_i \in V_{\lambda_i}(L)$.
$T(v) = \sum_{i=1}^{k} \sqrt{\lambda_i}\, v_k$.

Take any μ which is eigenvalue of T. $L = T^2$
Then $L|_{V_\mu(T)}(v) = \mu^2 v$ $\quad V_\mu(T) \subset V_{\mu^2}(L) \Rightarrow \mu = \sqrt{\lambda_i}$ for a certain λ_i.
any in this, belong to this

different eigenspaces of T include into different eigenspaces of L.
renumbering. $\mu_i = \sqrt{\lambda_i}$. $V_{\mu_i}(T) \subset V_{\lambda_i}(L)$
we have two orthogonal. direct sum decomposition, and the summand coincides.

Polar decomposition.

consider scaling transformation $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$.

the unit circle is given by. $\dfrac{x'^2}{a^2} + \dfrac{y'^2}{b^2} = 1$. (an ellipse). [unit sphere ⇒ ellipsoid].

on unit circle (sphere). $\|x\| = 1$. $x' = Ax$ belongs to the unit~. iff $\|A^{-1}x'\| = 1$.
$(A^{-1}x')^T (A^{-1}x') = x'^T A^{-1T} A^{-1} x' = 1$. $(A^{-1T}A^{-1})$ is positive symmetric.
then there exists an orthogonal transformation C. s.t. $C^T(A^{-1T}A^{-1})C$ is diag. (positive)
Consider another c.s. s.t. $\vec{x} = C\vec{y}$. the set $\{x \mid x^T A^{-1T} A^{-1} x = 1\} = \{y \mid y^T (C^T A^{-1T} A^{-1} C) y = 1$
$= \{y \mid \dfrac{y_1^2}{\lambda_1^2} + \cdots + \dfrac{y_n^2}{\lambda_n^2} = 1\}$ .where $\lambda_1^2 \dots \lambda_n^2$ are all eigenvalues of AA^T
($C^T A^{-1T} A^{-1} C$ is diag. matrix with λ_i^{-1} on the diagonal). $(AA^T)^{-1}$ has eigenvalue λ_i^{2-1}.

Thm 17.6. (Polar decomposition in inner space).

For any invertible operator $L: V \to V$ on an inner space there exists a positive selfadjoint operator $S: V \to V$ and isometry $Q: V \to V$ s.t. $L = SQ$.

Remark: Geometric meaning L: unit sphere $\to$ ellipsoid.

S. a scaling operator. obtain ellipsoid from unit sphere by some scaling which is given suitable orthonormal basis by diagonal matrix.

$S^{-1}L = E_n$ (to sphere itself). hence orthogonal.

Pf: By Pro 17.3 and 17.5. there exist. a positive selfadjoint operator S s.t. $S^2 = LL^*$. Consider $Q = S^{-1}L$ and prove Q is isometry.

$$QQ^* = S^{-1}L(S^{-1}L)^* = S^{-1}LL^*(S^{-1})^* = S^{-1}S^2 \cdot S^{-1} = Id_V.$$

Remark: we have right/Left polar decomposition $L = SQ$. $L = QS_1$.

(in general. ~~$Q \neq Q_1$.~~ $S \neq S_1$. the equality holds. iff L is normal).

In fact. $Q = Q_1$ in two kinds of polar decomposition.

Coro 17.7 & 8. For any invertible $A \in M_n(\mathbb{R})$ (rep. $\mathbb{C}$) ; $\exists$ symmetric positive definite $S \in M_n(\mathbb{R})$ (rep. hermitian symmetric matrix $\underset{S = \overline{S}^T}{S} \in M_n(\mathbb{C})$). and orthogonal matrix Q (rep. unitary $U \in M_n(\mathbb{C})$) such that. $A = SQ$ ($A = SU$)

Coro 17.9 & 10. For any invertible matrix $A \in M_n(\mathbb{R})$. $\exists$ diag. $D \in M_n(\mathbb{R})$. with positive elements. and orthogonal matrice $Q_1, Q_2 \in M_n(\mathbb{R})$. s.t. $A = Q_1 D Q_2$.

(if $A \in M_n(\mathbb{C})$. Q_1, Q_2 be unitary).

Pf: For invertible A. we take polar decomposition $A = SQ$.

$\exists$ $S = CDC^{-1}$. C is orthogonal. D is diag.

then $A = CDC^{-1}Q$. denote $Q_1 = C$. $Q_2 = C^{-1}Q$. are orthogonal.

Arbitrary operator on an inner space.

Thm 17.11. For any operator $L: V \to V$ on a complex inner space there exists an orthonormal basis B s.t. the matrix $[L]_B$ is uppertriangular.

Pf: by induction on $\dim V$.

Take eigen.v. $v_1 \in V$. s.t. $\|v_1\| = 1$. correspond. $\lambda_1 = 1$. Denote $U = \mathrm{span}(v_1)^\perp$ $P = \mathrm{Proj}\, U$. orthogonal complement of eigenspace.

take orthonormal $\{u_2, \dots u_n\}$ basis in U. $B = \{v_1, u_2, \dots, u_n\}$. $[L]_B \begin{bmatrix} \underset{1\times1}{\lambda_1} & 0 \\ & \boxed{A}_{n-1} \end{bmatrix}$ (arbitrary).

$\exists A = PLP|_U$. ($\forall u \in U$. $PLP(u) = PL(u) = P(a_1 v_1 + a_2 u_2 + \cdots a_n u_n) = a_2 u_2 + \cdots + a_n u_n$

For u_i. $a_{i2}, \dots a_{in}$ define the ith column)

Apply induction hypothesis to U and $PLP|_U$. There exist orthonormal $B' = \{v_2', \dots v_n'\}$ s.t. $[L']_{B'}$ upper.tri

consider $B = \{v_1, v_2', \dots v_n'\}$, $\Rightarrow [L]_B = \begin{bmatrix} \lambda_1 & \\ & [PLP]_{B'} \end{bmatrix}$

Coro 17.12. For any matrix $A \in M_n(\mathbb{C})$. $\exists$ a unitary matrix U s.t. $U^{-1}AU$ is upper-tri.
(in real inner space. exists $U^{-1}AU$ be block triangular form with maxi-size 2.)
(U is transition matrix. from standard basis to orthogonal basis. i.e. columns with orthogonal basis.
we can simply check this by computing every elements)

§18. Singular value decomposition.

Consider arbitrary operator $L: V \to U$. between different inner space.

$L^*L: V \to V$. $LL^*: U \to U$ both non-negative and self-adjoint.

[$(LL^*(u), u) = (L^*(u), L^*(u)) \geq 0$ (since $L^*(u) = 0 \not\Rightarrow u = 0$).
$(L^*L(v), v) = (L(v), (L^*)^*(v)) \geq 0$.]

<u>Proposition 18.1</u>. Under above circumstances strictly positive eigenvalues of L^*L and LL^* are the same.

impossible to have negative ei.v.
$A \cdot \bar{A}^T$

<u>Pf</u>: Take λ of L^*L and the corresponding $v \in V$.

$L^*L(v) = \lambda v \neq 0$. as $\lambda > 0$. and $v \neq 0$.

$LL^*(L(v)) = L(\lambda v) = \lambda L(v)$. λ is a eigenvalue of LL^*

<u>Fact</u>: any two triangular matrice A. B. (size $m \times n$, $n \times m$).
the characteristic polynomial of AB and BA related.
i.e. $\exists$ monic $p(t)$. $\deg p(t) = k$.
s.t. $\chi_{AB} = (-1)^m t^{m-k} p(t)$. $\chi_{BA} = (-1)^n t^{n-k} p(t)$.
$\Rightarrow$ non-zero eigenvalues coincide and multiplicities are the same.

Def. (singular value). Positive square roots of non-zero eigenvalues of LL^* are called singular values of the linear map L.

<u>Thm 18.2</u> (Singular value decomposition). Consider a linear operator $L: V \to U$. Then there are exists orthonormal bases B_V in V and B_U in U s.t. the matrix $[L]_{B,B'}$ has the block form $\begin{pmatrix} D_r & 0 \\ 0 & 0 \end{pmatrix}$ where D_r is a diagonal $r \times r$ matrix with singular values of L on its diagonal.
(the block is very similar to Thm 6.9).
Denote. $r = \dim \operatorname{Im}(L)$. Then $\dim \operatorname{Ker}(L) = n - r$, $\dim \operatorname{Ker}(L)^{\perp} = r$
Choose an orthonormal basis in $\operatorname{Ker}(L)^{\perp} \to \operatorname{Im} L^*$ and denote it by $\{v_1', \dots, v_r'\}$ and orthonormal basis in $\operatorname{Ker}(L)$ $\{v_{r+1}, \dots, v_n\}$. orthonormal basis in $\operatorname{Im}(L)$ and $\operatorname{Im}(L)^{\perp}$: $\{u_1', \dots u_r'\}$, $\{u_{r+1}', \dots, u_n'\}$

Then the matrix of L. $[L]_{V\to V'} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$. rank $A = r$. size $A = r$ invertible.

Analogously, $[L^*L]_{V'\to V'} = \begin{pmatrix} \bar{A}^T A & 0 \\ 0 & 0 \end{pmatrix}$

Find the canonical form of self-adjoint operator L^*L. find orthonormal basis $\{v_1 \ldots v_r\}$ in $\mathrm{Ker}(L)^\perp$ $(\mathrm{Im}\, L^*)$

s.t. the matrix of $[L^*L]_{V\to V} = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$ (square value of singular value of A on diag.).

Denote $L^*L(v_i) = \sigma_i^2 v_i$. $(L(v_i), L(v_i)) = (v_i, L^*L(v_i)) = [v_i]^T \cdot \sigma_i^2 [v_i] = \sigma_i^2$

denote $u_i = \frac{L(v_i)}{\sigma_i}$. $\{u_i\}$ forms a orthonormal basis of $\mathrm{Im}(L)$.

denote. $B_V = \{v_1 \ldots v_r, v_{r+1} \ldots v_m\}$. $B_U = \{u_1 \ldots u_r, u_{r+1} \ldots u_m\}$.

$[L]_{B_V, B_U} = \begin{bmatrix} D_r & 0 \\ 0 & 0 \end{bmatrix}$ $\left(L(v_i) = \begin{cases} \sigma_i u_i & 1 \le i \le r. \\ 0 & , i \ge r+1 \end{cases} \right).$

Coro 18.3. Let $A \in M_{m,n}(\mathbb{R})$ an arbitrary matrix of rank r. Then there exists. orthogonal matrices $Q_1 \in M_m(\mathbb{R})$, $Q_2 \in M_n(\mathbb{R})$ s.t. $Q_1 A Q_2 = \begin{pmatrix} D_r & 0 \\ 0 & 0 \end{pmatrix}$

(D_r has singular values on its diag.).

we just need to check the transition matrix of standard to orthonormal basis is orthog. (let it be Q_1^{-1})

For invertible operator. $A = Q_1 D Q_2 = \underbrace{(Q_1 D_n Q_1^{-1})}_{S} \underbrace{(Q_1 Q_2)}_{Q}$. we can deduce the polar decomposition.

Operator norm and relative errors

operator norm (define for a linear map between two inner spaces $L: V \to U$.)

measures the "size" of certain linear operators.

Def. For $L: V \to U$. its norm $\|L\|_2$ is equal to $\sup_{\|x\|=1} \|L(x)\|$.

By the def. for all $v \in V$ one has $\|L(v)\| \le \|L\|_2 \cdot \|v\|$.

Pro. 18.4. Let $L, L_1: V \to U$ and $L_2: W \to V$ be linear maps between inner spaces.

Then 1. $\|L + L_1\|_2 \le \|L\| + \|L_1\|$ and $\|\alpha L\|_2 \le |\alpha| \cdot \|L\|_2$.

2. $\|L_1 L_2\|_2 \le \|L_1\|_2 \cdot \|L_2\|_2$

Pf: $\|L(x)\| \le \|L\|_2 \cdot \|x\|$

$\|(L+L_1)(x)\| = \|L(x) + L_1(x)\| \le (\|L\|_2 + \|L_1\|) \|x\|$

Def. (absolute error). two vectors $x, \tilde{x} \in V$ where $\tilde{x}$ is assumed to be an approximation to x. the difference $\|x - \tilde{x}\|$ is called an absolut error.

(relative error). $\frac{\|x - \tilde{x}\|}{\|x\|}$

the estimated $\|L(x) - L(\tilde{x})\| \le \|L\|_2 \cdot \|x - \tilde{x}\|$

Remark: For relative error: applicable. only for invertible map.

(since in the case $L(x) = 0$ the relative error is undefined.)

$$\frac{\|L(x) - \tilde{L}(x)\|}{\|L(x)\|} \le \|L\|_2 \frac{\|x - \tilde{x}\|}{\|x\|} \cdot \frac{\|x\|}{\|L(x)\|} \le \|L\|_2 \cdot \|L^{-1}\|_2 \cdot \frac{\|x - \tilde{x}\|}{\|x\|}.$$

(estimate. rel. errors as rel. error in measurement)

Def. Condition number. $K(L) = \|L\|_2 \cdot \|L^{-1}\|_2$ (for L)

When the condition number is big, small relative errors in data could cause big relative error in result. (behaviour similar to non-invertible map).

Pro 18.5 For any operator $L: V \to U$ between inner space the operator norm $\|L\|_2$, is equal to the greatest singular number of L, it is denoted usually by $\sigma_1(L)$

If L is an invertible linear map between the inner spaces of dimension n then the condition number is equal to $K(L) = \frac{\sigma_1(L)}{\sigma_n(L)}$ where $\sigma_1(L) \geq \sigma_2(L) \geq \cdots \geq \sigma_n(L)$ are all singular values of A in decreasing order

Remark: $L^{-1}(L^{-1})^* = (L^*L)^{-1}$. the greastest singular eigenvalue of L^{-1} is σ_n^{-1} (since postive eiv. of LL^* and L^*L coincides.) $\|L\|_2 \cdot \|L^{-1}\|_2 = \sigma_1(L) \cdot \frac{1}{\sigma_n(L)}$

Pf: Apply singular value decomposition, take orthonormal basis $\{v_1 \dots v_r, v_{r+1}, \dots v_n\}$ s.t. $\|L(v_i)\| = \sigma_i$ ($L(v_i) = \sigma_i v_i$).

Then for any $x \in V$ take decomposition $x = \sum_1^n c_i v_i$ $\|x\|^2 = \sum_1^n |c_i|^2$

Let us estimate $\|L(x)\|^2 = \|c_1 L(v_1) + \dots + c_n L(v_n)\| = |c_1|^2 \sigma_1^2 + \dots + |c_n|^2 \sigma_n^2$

Therefore $\|L(x)\|^2 \leq \sigma_1^2 \|x\|^2$

Approximation by low rank linear map.

Thm 18.6. Let $L: V \to U$ be a linear map between inner space and $\sigma_1 \geq \sigma_2 \geq \dots$ are singular values of L. Then for any $k < \operatorname{rank}(L)$:

- given an arbitrary linear map $T: V \to U$ s.t. $\operatorname{rank}(T) = k$, one has $\|L - T\|_2 \geq \sigma_{k+1}$
- there exists a linear map $L_k: V \to U$ such that $\operatorname{rank}(L_k) = k$ and $\|L - L_k\|_2 = \sigma_{k+1}$.

§19. Angles in Euclidean space

Recall: angle $\alpha \in [0,\pi]$ between two non-zero vectors u, v, in the real inner space. $\mathcal{V}$ is defined by $\cos\alpha = \frac{(u,v)_{\mathcal{V}}}{\|u\| \cdot \|v\|}$

Define angle between $v \in V$ (non-zero) and some subspace $\mathcal{U} \subset \mathcal{V}$ (in Euclidean space): (geometric / algebric, two equivalent definitions).

Def. The angle between v and $\mathcal{U}$ is equal to $\frac{\pi}{2}$. if $v \in \mathcal{U}^{\perp}$ and is equal to the angle between v and orthogonal projection $\mathrm{Proj}_{\mathcal{U}}(v)$ when $v \notin \mathcal{U}^{\perp}$.

(vector & space).

(if v is unit. $\cos\alpha = \|\mathrm{Proj}_{\mathcal{U}}(v)\|$.)

Pro 19.1. The angle between v and $\mathcal{U}$ is equal to the minimum angle between v and $u \in \mathcal{U}$. when $u \neq 0$ ranges the subspace $\mathcal{U}$.

Pf: the case. $u \in \mathcal{U}^{\perp}$ is trivial.

take $u_0 = \mathrm{Proj}_{\mathcal{U}}(v)$. for any $u \in \mathcal{U}$. orthogonal decomposition $u = ku_0 + w$. $w \in \mathrm{span}(u_0)^{\perp} \cap \mathcal{U}$

$w \perp v$ $[(w,v) = (w, u_0 + (v-u_0)) = (\underset{\mathrm{span}(u_0)^{\perp}}{w}, u_0) + (\underset{\mathcal{U}}{w}, \underset{\mathcal{U}^{\perp}}{v-u_0}).]$

$(v, u_0)_{\mathcal{V}} = (u_0 + (v-u_0), u_0) = (u_0, u_0) > 0$. the angle is acute (锐角).

I/ $k=0$. $u \perp v$. angle u, v is not minimal

II/ $k<0$. $(u,v) = (ku_0 + w, v) = k(u_0, v) < 0$. obtuse angle. not minimal $[(w,v) = (w, u_0) + (\underset{\mathcal{U}}{w}, \underset{\mathcal{U}^{\perp}}{w_0}) = 0]$.

III/ $k>0$ divide u be k. assume $u = u_0 + w'$.

$(u,v) = (u_0 + w', v) = (u_0, v)$.

Compare: $(u_0, v)_{\mathcal{V}} = (u, v)_{\mathcal{V}}$. $\|u\|^2 = \|u_0\|^2 + \|w\|^2 \geq \|u_0\|$.

$$\Rightarrow \frac{(u_0, v)_{\mathcal{V}}}{\|u_0\| \cdot \|v\|} \geq \frac{(u,v)_{\mathcal{V}}}{\|u\| \cdot \|v\|}$$

In the case of two arbitrary non-zero subspaces $\mathcal{U}_1, \mathcal{U}_2 \subset V$. in the Euclidean space we can define minimal angle between $\mathcal{U}_1$ and $\mathcal{U}_2$.

Def. The minimal angle θ_{min} between $\mathcal{U}_1$ and $\mathcal{U}_2$ is the minimal angle between 2 non-zero vectors $u_1 \in \mathcal{U}_1$ and $u_2 \in \mathcal{U}_2$ (u_1 ranges $\mathcal{U}_1$, u_2 ranges $\mathcal{U}_2$).

(space & space).

$$\cos\theta_{min} = \max_{u_1 \in \mathcal{U}_1,\, u_2 \in \mathcal{U}_2} \frac{(u_1, u_2)_{\mathcal{V}}}{\|u_1\| \cdot \|u_2\|} = \max_{v \in \mathcal{U}_1,\, \|v\|=1} \|\mathrm{Proj}_{\mathcal{U}_2}(v)\|.$$

(equivalent: minimal value of angle between $v \in \mathcal{U}_1$ and the subspace $\mathcal{U}_2$. when $v \neq 0$ ranges $\mathcal{U}_1$.)

Pro 19.2. Let $P_1 = \mathrm{Proj}_{\mathcal{U}_1}$. $P_2 = \mathrm{Proj}_{\mathcal{U}_2}$. $\mathcal{U}_1, \mathcal{U}_2 \subset$ Euclidean space V. Then $\cos\theta_{min} = \|P_1 P_2\|_2$.

Remark: orthogonal projector is self-adjoint. For any linear map. $\|L\|_2 = \|L^*\|_2$.

$\Rightarrow \|P_1P_2\|_2 = \|P_2P_1\|_2$. (the norm equals the greatest singular value).

"$\geq$"

Pf: $\forall$ unit $u_1 \in U_1$. $\cos\theta_{min} \geq \|P_2(u_1)\|$ ($\|P_2(u_1)\|$ angle between u_1 and U_2.)

(by def. of norm) $\cos\theta_{min} \geq \|P_2|_{U_1}\|_2$. where $P_2|_{U_1} : U_1 \to V$.

Since. $\forall v \in V$. $\|P_2 P_1(v)\| \leq \|P_2|_{U_1}\|_2 \cdot \|P_1(v)\| \leq \|P_2|_{U_1}\|_2 \cdot \|v\| \Rightarrow \|P_2 P_1\|_2 \leq \|P_2|_{U_1}\|_2$

$\forall u \in U_1$. $\|P_2|_{U_1}(u)\| = \|P_2(P_1(u))\| \leq \|P_1 P_2\|_2 \cdot \|u\| \Rightarrow \|P_2|_{U_1}\|_2 \leq \|P_2 P_1\|_2$.

$\Rightarrow \|P_2 P_1\|_2 = \|P_2|_{U_1}\|_2$. $\Rightarrow \cos\theta_{min} \geq \|P_2 P_1\|_2$.

"$\leq$".

$\cos\theta_{min} = \max\limits_{v \in U_1, \|v\|=1} \|P_2(v)\|$, $\exists \underbrace{v_0}_{unit}$ $\cos\theta_{min} = \|P_2(v_0)\|$.

For any arbitrary unit $v \in U_1$. $\|P_2(v)\| \leq \|P_2|_{U_1}\|_2$.

i.e. $\cos\theta_{min} = \|P_2(v_0)\| \leq \|P_2|_{U_1}\|_2 = \|P_2 P_1\|_2$.

Remark: When two subspaces have non-zero intersection then the minimal angle between this subspace is equal to 0.
e.g. minimal angle between any two planes in 3-dim space is always zero.

Def. (Principal angle). $\theta_1 \leq \theta_2 \leq \cdots \leq \theta_k$ are principal angles between U_1 and U_2 if

$\cos\theta_1 \geq \cos\theta_2 \geq \cdots \geq \cos\theta_k$ is a decreasing sequence of singular values of $P_1 P_2$.

By Singular value decomposition. it means that $\exists$ orthonormal basis $v_1, v_2, \ldots v_k$ in U_1 s.t. vectors $P_2(v_1), \ldots P_2(v_k)$. are orthogonal to each other. and $\|P_2(v_i)\| = \cos\theta_i$. for $i=1\ldots k$.

(Consider $P_1 P_2 : V \to V$. $[P_1 P_2]_B = \begin{bmatrix} D_r & 0 \\ 0 & 0 \end{bmatrix}$. $P_1 P_2(v_i) = P_2(v_i)$ since $v_i \in U_1$.)

Case 1. two plane in $\mathbb{R}^3$. two singular value: 1 and $\cos\alpha$. (α is the usual angle between these plane)

case 2. two plane in $U_1. U_2 \subset \mathbb{R}^4$. for $L = P_2 P_1$. $L^* L = P_1^* P_2^* P_2 P_1 = P_1 P_2^2 P_1 = P_1 P_2 P_1$.

$L L^* = P_2 P_1 P_2$. have at least zero eigenvalues and two other common eigenvalues $\sigma_1^2 \geq \sigma_2^2$. (can be 0)

Pro 19.3. There exists. an orthonormal basis $\{v_1, v_2\}$ in U_1. and an orthonormal basis $\{u_1, u_2\}$. in U_2 s.t.

$P_2(v_1) = \sigma_1 u_1$. $P_2(v_2) = \sigma_2 u_2$

$P_1(u_1) = \sigma_1 v_1$ $P_1(u_2) = \sigma_2 u_2$.

Pf: Since $U_1^\perp \subset V_0(L^* L)$. we can choose orthonormal basis v_1, v_2, v_3, v_4. consisting of eigenvectors of $L^* L$ s.t. $v_3. v_4 \in U_1^\perp$ $v_1. v_2 \in U_1$. (the self-adjoint operator. orthonormal basis consisting of eigenvector). self-adjoint. (factorizable).

$\Rightarrow P_1 P_2 P_1(v_1) = \sigma_1^2 v_1$ $P_1 P_2 P_1(v_2) = \sigma_2^2 v_2$. $u_i = \frac{L(v_i)}{\sigma_i}$. u_i forms an orthonormal basis of $Im(L)$

In the basis $\{u_1, u_2, u_3, u_4\}$ of U_2 corresponding to singular value decomposition We can take $u_1, u_2 \in U_2$

Thus we find the orthonormal bases $\{u_1, u_2\}$. $\{v_1. v_2\}$. in $U_2 . U_1$. (just show the orthonormal basis consist of eigenvector).

s.t. $P_2(v_1) = L(v_1) = \sigma_1 u_1$. $P_2(v_2) = L(v_2) = \sigma_2 u_2$.

if $\sigma_1 \neq 0$. $P_1(u_1) = \frac{1}{\sigma_1} P_1(\sigma_1 u_1) = \frac{1}{\sigma_1} P_1 P_2(v_1) = \frac{1}{\sigma_1} P_1 P_2 P_1(v_1) = \frac{1}{\sigma_1} \cdot \sigma_1^2 \cdot v_1 = \sigma_1 v_1$.

$\sigma_1 = 0$. $u_1 \in Im(L)^\perp \cap U_2$. $u_4 \in Ker(L^*) = Im(L)^\perp$ $Im(L) = Im(P_2 P_1) \subset U_2$.

$0 \cdot v_1 = 0 = L^*(u_1) = P_1 P_2(u_4) = P_1(u_4)$.

1° $Im(P_2 P_1) = U_2$. $\{u_1, u_2\}$

2° $Im(P_2 P_1) \subsetneq U_2$.

Def. $\theta_{min}. \theta_{max} \in [0, \frac{\pi}{2}]$. s.t. $\cos\theta_{min} = \sigma_1$ .and $\cos\theta_2 = \sigma_2$.
are called minimal and maximum angle between two planes in $\mathbb{R}^4$.

$U_2 = Im(L) \perp (Im(L)^\perp \cap U_2)$.

u_1 u_2 (系数可调,满足σ_2).

(in 3-dim. $\theta_{min} = 0$. $v_1 = u_1$.
If $\theta_{min} = \theta_{max} = \frac{\pi}{2}$. the plane orthogonal).