Lecture Notes in Linear Algebra

Kaizhao Liu

May 25, 2025

Contents

0.1 Prologue . . . 5
0.2 Notations . . . 6

1 Basic Concepts: Lost Properties **7**
1.1 Noncommutative-Rings . . . 7
1.2 Polynomials . . . 8
1.2.1 Euclid's algorithm . . . 8
1.2.2 . . . 8
1.2.3 symmetric polynomials . . . 8
1.2.4 . . . 9
1.3 Polynomials Revisited: Structure of Fields . . . 9
1.3.1 Algebraic Field Extensions . . . 9
1.3.2 The Fundamental Theorem of Galois Theory . . . 10
1.4 Matrices . . . 10
1.4.1 Schur Complement . . . 10
1.5 Linear Homomorphism . . . 10
1.5.1 exact sequence . . . 10

2 Structure Theory of a Linear Transformation **13**
2.1 λ-matrix . . . 13
2.1.1 Smith normal form . . . 13
2.1.2 Jordan normal form: alternative treatment . . . 14
2.1.3 The Elementary factors of $f(A)$ . . . 14
2.1.4 Intermezzo: Centers . . . 15
2.2 λ-matrix Revisited: Modules over PID . . . 16
2.2.1 Modules . . . 16
2.2.2 Modules over PID . . . 17
2.3 Decomposition of Linear Operators . . . 17
2.3.1 Cyclic Spaces . . . 18
2.3.2 Irreducible Spaces . . . 18
2.3.3 Centers . . . 18
2.4 Decomposition of Linear Operators Revisited: Modules over PID . . . 19
2.5 ???Linear Homomorphisms . . . 19
2.5.1 ???Semisimple Operators . . . 19

3 Theory of Bilinear Functions **21**
3.1 Bilinear Functions . . . 21
3.2 Dual Vector Spaces . . . 22
3.2.1 Duality . . . 22
3.2.2 Space of Linear Functions . . . 22
3.2.3 Finite Dimensional Vector Spaces . . . 23
3.3 Metric Vector Spaces . . . 23
3.3.1 Orthogonal Direct Sum . . . 24
3.3.2 Isometry . . . 24
3.3.3 Hyperbolic Spaces . . . 24
3.4 Real and Complex Inner Product Spaces . . . 25
3.4.1 Gram-Schmidt Orthogonalization . . . 25
3.4.2 Application: QR Decomposition . . . 26
3.4.3 The Projection Theorem . . . 26

3.4.4 Schur Triangularization . . . 26
3.5 Structure Theory for Normal Operators . . . 26
3.5.1 The Adjoint of a Linear Operator . . . 26
3.5.2 Orthogonal Projections . . . 27
3.5.3 Normal Operators . . . 27
3.5.4 Functional Calculus . . . 27
3.5.5 Application: Positive Operators . . . 27
3.5.6 Application: The Polar Decomposition . . . 27
3.5.7 Normal Operators in Complex Inner Product Spaces: A Summary . . . 28
3.6 Discussion . . . 28

4 Multilinear Algebra **29**
4.1 Tensor Product . . . 29
4.1.1 Tensor Product . . . 29
4.1.2 Kronecker Prodcut: Application to Matrix . . . 29

5 Lie Theory **31**
5.1 BCH Formula . . . 31
5.1.1 Matrix Exponetial . . . 31
5.1.2 Matrix Logarithm . . . 31
5.1.3 BCH Formula for Heisenberg Group . . . 32
5.1.4 General BCH Formula . . . 32
5.2 The Representations of SU(2) and SU(3) . . . 32
5.2.1 The Irreducible Representations of SU(2) . . . 32
5.3 Root Systems . . . 32
5.3.1 Root . . . 32
5.3.2 Example: Rank 2 . . . 34
5.3.3 Example: Rank 3 . . . 34
5.3.4 Additional Properties . . . 34
5.3.5 Application: Classical Lie Algebras . . . 34
5.3.6 Dynkin Diagrams . . . 34
5.3.7 The Root Lattice . . . 34
5.4 Lie Algebra: an Algebraic Viewpoint . . . 34
5.4.1 Basic Definition and Examples . . . 34
5.4.2 Lie Theorem . . . 35
5.4.3 Cartan's Criterion . . . 35
5.4.4 Killing form . . . 35
5.5 Discussion . . . 35

6 Matrix Analysis **37**
6.1 Positive Matrices . . . 37
6.2 Majorization . . . 37
6.3 Matrix Equations . . . 37
6.4 Discussion . . . 37

7 Operators on Hilbert Space **39**
7.1 Hilbert Spaces . . . 39
7.1.1 Examples . . . 39
7.2 Operators . . . 39

8 Miscellaneous Topics **41**
8.1 Gelfand-Tsetlin Integrable System . . . 41
8.2 Unitary Matrices . . . 42
8.3 Matrix Decomposition . . . 42
8.3.1 Positive Definite Matrices . . . 42
8.3.2 . . . 43

0.1 Prologue

In my opinion, **the goal of mathematics** is to develop theories to solve particular questions.

These are the criterion for good theories. The theories should be understandable, that is, if one novice follow the thoery from the beginning to the end, he or she should find all definitions, examples, theorems, proofs, and other mathematical objects as natural as possible. Tricks are not allowed, but integrated in the theory as a natural consequence of the definition.

The questions are from two sources. One is real world, the other is the theory. Some questions are natural and have pratical usage. Other questions can be developed to torture students, can be asked for fun or novelty, being more palyful and having no practical use. For studying purpose, I suggest the secular questions, that is, the first kind of questions. The second kind of questions can be omitted.

Remember the theory is developed for real world questions. In the process of developing the theory, theoritical questions emerge and produce more theories. New thoeries can also be put into real-world use. Therefore, the **relationship between theories and questions** is that they support each other, and can not be divided. When reading this book, you need to keep in mind what are the questions we care about. In this way, you can have a better understanding of the theory. However note that, for a certain problem, there exists other approaches unpresented here. You should compare and contrast them with the approach presented in this book.

Different theories have close relationship. As stated above, one theory can be developed from another theory. One theory can provide method for another theory. All these theories form a mathematician's toolbox.

Please take great care of the process of developing the theory. The ideas and insight can not be stated completely because the fallacy of language. These are left to the readers to feel with hearts. Take special care of the history of the theory.

♠♠♠

WARNING:THIS IS AN UNFINISHED MANUSCRIPT
DO NOT TAKE IT SERIOUSLY

♠♠♠

- **Cayley-Hamilton Theorem**
- **Perron-Frobenius Theorem**
- **Algorithm: division with remainders**
- **Algorithm: Euclid's algorithm**
- **Euclid's ring**
- **prime ideal**
- **principle ideal domain**
- **Unique Factor Domain**
- **Mason-Stothers Theorem**
- **symmetric polynomials**
    1. elementary symmetric polynomials
    2. fundamental theorem of symmetric polynomials: (representation of symmetric polynomials by elementary symmetric polynomials) proof and computation
- **Newton's formula**
    1. derivation
    2. memorization
    3. application
- **Smith normal form**
    1. algorithm *from easy to complex then to easy*
    2. proof

Relationships between similarities of matrix and

0.2 Notations

We use upper case letters for general rings, fields, vector spaces ... $\mathbb{R}$ for real, and $\mathbb{C}$ for complex, $\mathbb{F}$ for either $\mathbb{R}$ or $\mathbb{C}$, $\mathbb{F}_p$ for

Chapter 1

Basic Concepts: Lost Properties

This chapter collects some definitions and theorems out of the scope of linear algebra, but is of its own importance.

1.1 Noncommutative-Rings

This section aims to study the algebra of matrices from the perspective of noncommutative rings.

The goal of next theorem is to express the invert of $x^{-1}+y^{-1}$ by x and y.

Theorem 1.1.1. *If $x, y, (x+y)$ is invertible, then $x^{-1}+y^{-1}$ is invertible with*

$$(x^{-1}+y^{-1})^{-1} = x(x+y)^{-1}y = y(x+y)^{-1}x$$

Proof. To see why this formula is correct, we seek inspiration from commutative rings which we are more familar with. If the ring is commutative, then

$$(x^{-1}+y^{-1})^{-1} = \frac{xy}{x+y}.$$

The formula we seek must reduce to the formual above when the ring is commutative. And the formula we seek must be invariant when we interchange x and y. $xy(x+y)^{-1}$ or something like this can not satisfy the symmetry. So $x(x+y)^{-1}y$ is a good candidate.

We first show that $x(x+y)^{-1}y = y(x+y)^{-1}x$.

$$\begin{aligned} x(x+y)^{-1}y &= (x+y-y)(x+y)^{-1}y \\ &= y - y(x+y)^{-1}y \\ &= y - y(x+y)^{-1}(x+y-x) \\ &= y(x+y)^{-1}x \end{aligned}$$

Now we prove that it's the inverse.

$$\begin{aligned} x^{-1}+y^{-1} &= x^{-1}yy^{-1} + x^{-1}xy^{-1} \\ &= x^{-1}(x+y)y^{-1} \end{aligned}$$

The result follows immediately.

$$\begin{aligned} (x^{-1}+y^{-1})(y(x+y)^{-1}x) &= x^{-1}(x+y)y^{-1}(y(x+y)^{-1}x) \\ &= 1 \end{aligned}$$

□

Remark 1.1.2. Note how we do the tricks. These are typical in ring theory.

Theorem 1.1.3. *If $1-yx$ is invertible, then $1-xy$ is invertible with*

$$(1-xy)^{-1} = 1 + x(1-yx)^{-1}y.$$

Proof. To see why this formula is correct, we seek inspiration from geometric series. Formally,

$$\begin{aligned} (1-xy)^{-1} &= 1 + xy + xyxy + xyxyxy + \cdots \\ &= 1 + x(1 + yx + yxyx + \cdots)y \\ &= 1 + x(1-yx)^{-1}y. \end{aligned}$$

The rest is direct computation. □

1.2 Polynomials

1.2.1 Euclid's algorithm

We list the two fundamental theorems on polynomial ring. These algorithms are the basic tools to portrait the polynomial ring

Theorem 1.2.1 (division with remainder)**.** $f, g \in F[x], \exists q, r$ *with* $\deg r < \deg g$ *s.t.* $f = qg + r$

Theorem 1.2.2 (Bézout's theorem)**.** f *coprime with* $g \in F[x], \exists u, v$ *s.t.* $uf + vg = 1$

1.2.2

Lemma 1.2.3 (Eisenstein's Criterion)**.**

Proof. let $f(x) = a_n x^n + ... + a_0, (a_0, ..., a_n) = 1$

$$p \mid a_0, ..., p \mid a_{s-1}, p \nmid a_s$$

□

Theorem 1.2.4 (Gauss's Lemma)**.** *Let* $f(x) \in \mathbb{Z}[x]$*, and* $f(x)$ *is reducible in* $\mathbb{Q}[x]$*. Then* $f(x)$ *is reducible in* $\mathbb{Z}[x]$*.*

Proof. Let $f(x) = f_1(x)f_2(x), f_i(x) \in \mathbb{Q}[x]$ and $\deg f_i(x) < \deg f(x)$.

$$f(x) = c(f)$$

□

1.2.3 symmetric polynomials

Theorem 1.2.5. *Let* A *be a commutative ring and let* $t_1, ..., t_n$ *be algebraically independent elements over* A*. Let* $f(t) \in A[t_1, ..., t_n]$ *be symmetric of degree* d*. Then there exists a polynomial* $g(X_1, ..., X_n)$ *of weight* $\leq d$ *such that*

$$f(t) = g(s_1, ..., s_n)$$

where each $s_i = s_i(t_1, ..., t_n)$ *is a polynomial in* $t_1, ..., t_n$*.*

Proof. By induction on n. The theorem is obvious if $n = 1$, because $s_1 = t_1$. Assume the theorem is proved for polynomials in n-1 variables.

If we substitute $t_n = 0$ in the expression for $F(X)$, we find

$$(X - t_1) \cdots (X - t_{n-1})X = X^n - (s_1)_0 X^{n-1} + \cdots + (-1)^{n-1}(s_{n-1})_0 X$$

where $(s_i)_0$ is the expression obtained by substituting $t_n = 0$ in s_i. We see that $(s_i)_0$ are precisely the elementary symmetric polynomials in $t_1, ..., t_{n-1}$.

We now carry out induction on d. If $d = 0$, our assertion is trivial. Assume $d > 0$, and assume our assertion proved for polynomials of degree $< d$. Let $f(t_1, ..., t_n)$ have degree d. There exists a polynomial $g_1(X_1, ..., X_{n-1}$ of weight $\leq d$ such that

$$f(t_1, ..., t_{n-1}, 0) = g_1((s_1)_0, ..., (s_{n-1})_0)$$

We note that $g_1(s_1, ..., s_{n-1})$ has degree $\leq d$ in $t_1, ..., t_n$. The polynomial

$$f_1(t_1, ..., t_n) = f(t_1, ..., t_n) - g_1(s_1, ..., s_{n-1})$$

has degree $\leq d$ in $t_1, ..., t_n$ and is symmetric. We have

$$f_1(t_1, ..., t_{n-1}, 0) = 0$$

Hence f_1 has a root t_n, and by symmetry,

$$f_1 = s_n f_2(t_1, ..., t_n)$$

f_2 has degree $\leq d - n < d$. By induction, there exists a polynomial g_2 in n variables and weight $\leq d - n$ such that

$$f_2(t_1, ..., t_n) = g_2(s_1, ..., s_n)$$

We obtain

$$f(t) = g_1(s_1, ..., s_{n-1}) + s_n g_2(s_1, ..., s_n)$$

and each term on the right has weight $\leq d$. This completes the proof. □

Corollary 1.2.6.

1.2.4

Theorem 1.2.7 (Mason-Stothers Theorem)**.** *Let $a(t), b(t), c(t)$ be relatively prime polynomials over $\mathbb{C}$ such that $a+b=c$. Then*

$$\max \deg \{a,b,c\} \leq \deg \operatorname{rad}(abc) - 1$$

Proof. Dividing by c and let $f = \frac{a}{c}, g = \frac{b}{c}$.Then $f+g=1$, where f, g are rational functions. Differentiating, we get $f'+g'=0$, which we rewrite as

$$\frac{f'}{f}f + \frac{g'}{g}g = 0$$

So that

$$\frac{b}{a} = \frac{g}{g} = -\frac{\frac{f'}{f}}{\frac{g'}{g}}$$

□

Theorem 1.2.8 (Newton's Formula)**.**

Proof. Let $h(x) = \prod_{i=1}^{n}(x-x_i)$. Then

$$h'(x) = \sum_{i=1}^{n} \frac{h(x)}{x-x_i}$$

$$x^{k+1}h'(x) = \sum_{i=1}^{n} \frac{x^{k+1} - x_i^{k+1} + x_i^{k+1}}{x-x_i} h(x)$$

$$\text{if } k \mathcal{E} n, 0 = s_k - \sigma_1 s_{k-1} + \cdots + (-1)^n \sigma_n s_{k-n}$$

$$\text{if } k < n, (-1)^k (n-k)\sigma_k = s_k - \sigma_1 s_{k-1} + \cdots + (-1)^k \sigma_k s_0$$

For example,

$$D(x_1,x_2,x_3) = \begin{vmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix}$$

$$s_0 = 3, s_1 = \sigma_1, s_2 = \sigma_1^2 - 2\sigma_2, s_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3, s_4 = \sigma_1^4 - 4\sigma_1^2\sigma_2 + 4\sigma_1\sigma_3 + 2\sigma_2^2$$

□

1.3 Polynomials Revisited: Structure of Fields

1.3.1 Algebraic Field Extensions

Definition 1.3.1 (extension field)**.** A field F is said to be an extension field of K provided that K is a subfield of F.

If F is an extension field of K, then F is a vector space over K. The dimension of the K-vector space F will be denoted by $[F:K]$. F is said to be a finite dimensional extension or infinite dimensional extension of K according as $[F:K]$ is finite or infinite.

Theorem 1.3.2. *Let F be an extension field of E and E an extension field of K. Then [F:K]=[F:E][E:K]. Futhermore [F:K] is finite if and only if [F:E] and [E:K] are finite.*

In the situation $K \in E \in F$ of the above theorem, E is said to be an intermediate field of K and F.

1.3.2 The Fundamental Theorem of Galois Theory

Definition 1.3.3. Let E and F be extension fields of a field K. A nonzero map $\sigma : E \longrightarrow F$ which is both a field and a K-module homomorphism is called a K-homomorphism. Similarly if a field automorphism $\sigma \in \operatorname{Aut} F$ is a K-homomorphism, then σ is called a K-automorphism of F. The group of all K-automorphism of F is called the Galios group if F over K and is denoted $\operatorname{Aut}_K F$.

Theorem 1.3.4. *Let F be an extension field of K and $f \in K[x]$. If $u \in F$ is a root of f and $\sigma \in \operatorname{Aut}_K F$, then $\sigma(u) \in F$ is also a root of f.*

Example 1.3.5. If $F = \mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2})(\sqrt{3})$, then since $x^2 - 3$ is irreducible over $\mathbb{Q}(\sqrt{2})$

Theorem 1.3.6. *Let F be an extension field of K, E an intermediate field and H a sbugroup of $Aut_K F$. Then*

$$(i) H' = \{v \in F | \sigma(v) = v \forall \sigma \in H\} \text{ is an intermediate field of the extension}$$
$$(ii) E' = \{\sigma \in Aut_K F | \sigma(u) = u \forall u \in E\} = Aut_E F \text{ is a subgroup of } Aut_K F$$

The field H' is called the fix field of H in F.

Definition 1.3.7. Let F be an extension field of K s.t. the fixed field of the galios group $Aut_K F$ is K itself. Then F is said to be a Galios extension of K.

Theorem 1.3.8 (Fundamental Theorem of Galios Theory)**.** *If F is a finite dimensional Galios extension of K, then there is a one to one correspond between the set of all intermediate fields of the extension and the set of all subgroups of the Galios group $Aut_K F$ such that:*

1.4 Matrices

1.4.1 Schur Complement

Definition 1.4.1. Suppose $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$. If A_{11} is invertible, define the Schur complement of A to be $A \setminus A_{11} = A_{22} - A_{21}A_{11}^{-1}A_{12}$.

Theorem 1.4.2. $\begin{pmatrix} I & O \\ -A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ O & A \setminus A_{11} \end{pmatrix}$

$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ O & I \end{pmatrix} = \begin{pmatrix} A_{11} & O \\ A_{21} & A \setminus A_{11} \end{pmatrix}$

$\begin{pmatrix} I & O \\ -A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ O & I \end{pmatrix} = \begin{pmatrix} A_{11} & O \\ O & A \setminus A_{11} \end{pmatrix}$

If A is Hermitian

Theorem 1.4.3 (Haynsworth)**.**

1.5 Linear Homomorphism

1.5.1 exact sequence

Exact sequence provide a sophisticated method for describing elementary properties of linear mappings.

Definition 1.5.1. Suppose that

$$0 \longrightarrow F \xrightarrow{\varphi} E \xrightarrow{\psi} G \longrightarrow 0$$

is a short exact sequence, and assume that $\chi : E \longleftarrow G$ is a linear mapping such that

$$\psi \circ \chi = \iota$$

Then χ is said to split the sequence and the sequence

$$0 \longrightarrow F \xrightarrow{\varphi} E \underset{\chi}{\overset{\psi}{\rightleftarrows}} G \longrightarrow 0$$

is called a split short exact sequence.

Theorem 1.5.2. *Consider the short exact sequence*

$$0 \longrightarrow E_1 \xrightarrow{i} E \xrightarrow{\pi} E/E_1 \longrightarrow 0$$

Then the relation $\chi \underset{}{\overset{}{\rightleftarrows}} \operatorname{Im}\chi$ *defines a bijection between linear mappings* $\chi : E \longleftarrow E/E_1$ *which split the sequence, and complementary subspaces of* E_1 *in* E.

Proof. □

Theorem 1.5.3. *A short exact sequence is split if and only if there exist a linear mapping* $\omega : F \longleftarrow E$ *such that* $\omega \circ \varphi = \iota$

Theorem 1.5.4. *Given an exact sequence*

$$E \xrightarrow{\varphi} F \xrightarrow{\psi} G \xrightarrow{\chi} H$$

Then

$$\varphi \text{ is surjective} \Longleftrightarrow \chi \text{ is injective}$$

Proof.

$$\begin{aligned}
\chi \text{ is injective} &\Longleftrightarrow \ker\chi = \{0\} \\
&\Longleftrightarrow \operatorname{Im}\psi = \{0\} \\
&\Longleftrightarrow \ker\psi = F \\
&\Longleftrightarrow \operatorname{Im}\varphi = F \\
&\Longleftrightarrow \varphi \text{ is surjective}
\end{aligned}$$

□

Theorem 1.5.5 (5-lemma)**.** *Assume a commutative diagram of linear maps where both horizontal sequence are exact.*

$$\begin{array}{ccccccccc}
E_1 & \xrightarrow{\alpha_1} & E_2 & \xrightarrow{\alpha_2} & E_3 & \xrightarrow{\alpha_3} & E_4 & \xrightarrow{\alpha_4} & E_5 \\
\downarrow{\scriptstyle\varphi_1} & & \downarrow{\scriptstyle\varphi_2} & & \downarrow{\scriptstyle\varphi_3} & & \downarrow{\scriptstyle\varphi_4} & & \downarrow{\scriptstyle\varphi_5} \\
F_1 & \xrightarrow{\beta_1} & F_2 & \xrightarrow{\beta_2} & F_3 & \xrightarrow{\beta_3} & F_4 & \xrightarrow{\beta_4} & F_5
\end{array}$$

Then :

i) If φ_4 *is injective and* φ_1 *is surjective, then*

$$\ker\varphi_3 = \alpha_2(\ker\varphi_2)$$

ii) If φ_5 *is injective and* φ_2 *is surjective, then*

$$\operatorname{Im}\varphi_3 = \beta_3^{-1}(\operatorname{Im}\varphi_4)$$

iii) If maps $\varphi_1, \varphi_2, \varphi_4, \varphi_5$ *are linear isomorphisms, then so is* φ_3.

Proof. □

Chapter 2

Structure Theory of a Linear Transformation

This chapter aims to find the canonical form of linear transformations. We reduce a linear space to its indecomposable subspaces, then show that these indecomposable subspaces are in fact cyclic.

To fully understand this procedure, we provide several extra viewpoints, including λ-matrix theory and module theory. We will show that module theory interprets all these results most adequately.

2.1 λ-matrix

When dealing with the question about if a matrix can be diagonalized or reduced to some other simple form, we often use the theory of λ-matrix, the concept of invariant factors, determinant factors, elementary factors, minimal polynomials, characteristic polynomials, and so on. Remember, the theoretical method is very powerful in solving relevant questions. So do not compute complicated and unsturctural calculation! Use your structural observation! So normally, an insightful observation can simplified the proof to a few sentences.

2.1.1 Smith normal form

Every λ-matrix can be reduced to a λ-matrix with only diagonal elements by elementary operations on λ-matrix.

The elementary operations of λ-matrix is essentially the same as that of ordinary matrix, but note that the elements of the matrix are polynomials, which forms a integral domain$F[\lambda]$, so the invertible elements of $F[\lambda]$ consist of $F \setminus \{0\}$. Therefore, to make the elementary λ-matrix with respect to an elementary operation invertible, we must restrict the second kind of elementary operations on $F \setminus \{0\}$ instead of $F[\lambda]$.

Example 2.1.1. Let $f(x)$ be relatively prime to $g(x)$ and $h(x)$, we have already know that

$$\begin{pmatrix} f(x)g(x) & \\ & h(x) \end{pmatrix} \sim \begin{pmatrix} g(x) & \\ & f(x)h(x) \end{pmatrix}$$

Now, as an exercise, lets find out $U(x)$ and $V(x)$ that make

$$U(x) \begin{pmatrix} f(x)g(x) & \\ & h(x) \end{pmatrix} V(x) = \begin{pmatrix} g(x) & \\ & f(x)h(x) \end{pmatrix}$$

That is, finding expicit operations that transform the first matrix to the second. Because f is relatively prime to h and g, $\exists u, v$ s.t. $fu + hv = 1$,$\exists p, q$ s.t. $fp + gq = 1$.

$$\begin{aligned} \begin{pmatrix} f(x)g(x) & \\ & h(x) \end{pmatrix} &\sim \begin{pmatrix} fg & hgv \\ & h \end{pmatrix} \\ &\sim \begin{pmatrix} fg & g \\ & h \end{pmatrix} \\ &\sim \begin{pmatrix} g & fg \\ h & \end{pmatrix} \end{aligned}$$

$$\sim \begin{pmatrix} g & \\ h & -fh \end{pmatrix}$$
$$\sim \begin{pmatrix} g & \\ qgh & -fh \end{pmatrix}$$
$$\sim \begin{pmatrix} g & \\ & fh \end{pmatrix}$$

Remark 2.1.2. This result can be used to compute the Smith normal form of a diagonal λ-matrix.

Theorem 2.1.3 (Calculation of Minimal Polynomials)**.** *The last invariant factor of $\lambda I_{n\times n} - A$ is the minimal polynomial of A, $A \in M_n(F)$, where F is a field .*

2.1.2 Jordan normal form: alternative treatment

Except for the λ-matrix treatment above, there is another observation that leads to a new way to compute the Jordan normal form, and we exhibit the treatment below.

The key observation: rank $J_n(0)^i = \max(n-i, 0)$

Given $A \in M_{n\times n}(F)$,We first calculate the eigenvalues of λ of A. Then, for every λ, compute the rank of $(\lambda I - A)^i$ until it becomes a constant. Denote the rank correspond to i by r_i, and let $r_0 = n$. Suppose $d_i = r_{i-1} - r_i$, then d_i means that we have at least d_i $J_i(\lambda)$ in the Jordan normal form, so we have $c_i = d_i - d_{i+1}$ $J_i(\lambda)$ Jordan blocks in the Jordan normal form.

Example 2.1.4. Define

$$D : \mathbb{C}[x]_n \longrightarrow \mathbb{C}[x]_n$$
$$f \longmapsto f'$$

We choose a basis for $\mathbb{C}[x]_n$: $(1, x, \cdots, x^n)$. Then the matrix of D is

$$\begin{pmatrix} 0 & 1 & & & \\ & & 2 & & \\ & & & \ddots & \\ & & & & n \\ & & & & 0 \end{pmatrix}$$

$$\det(\lambda I - A) = 0 \Longrightarrow \lambda = 0$$

Due to $\operatorname{rank} D = n$, there is only one Jordan block, therefore the Jordan normal form of D is immediately obtained.

Now let's find the transition matrix. The key observation is: there is only one cyclic subspace, so we only need to find a generator for this space, and take the set of elements generated by the generator to be the basis.

$$P = \begin{pmatrix} \frac{n!}{0!} & & & & \\ & \frac{n!}{1!} & & & \\ & & \frac{n!}{2!} & & \\ & & & \ddots & \\ & & & & \frac{n!}{n!} \end{pmatrix}$$

From the example we can see, for certain structures, we prefer to use the treatment above from the perspective of rank instead of computing with λ-matrix. Readers should choose the best treatment according to different situations.

2.1.3 The Elementary factors of $f(A)$

We assume $A \in M_{n\times n}(\mathbb{C})$.

Theorem 2.1.5. *The elementary factors of $f(A)$ can be obtained by the following procedure:*
Given an elementary factor of A

$$(\lambda - \lambda_0)^p$$

it corresponds to elementary factors of $f(A)$ through the following way:

$$\begin{array}{cc} & (\lambda - f(\lambda_0))^p, \text{when } p\mathcal{E}1 \text{ and } f'(\lambda_0) \neq 0 \\ \underbrace{(\lambda - f(\lambda_0))^{q+1}}_{h} & \underbrace{(\lambda - f(\lambda_0))^q}_{k-h}, \text{when } p > 1 \text{ and } f'(\lambda_0) = \cdots = f^{(k-1)}(\lambda_0) = 0, f^{(k)}(\lambda_0) \neq 0 (k < p) \end{array}$$

$$\text{where } p = qk + h(q\mathcal{E}0, k > h\mathcal{E}0)$$

$$\underbrace{\lambda - f(\lambda_0)}_{p}, \text{when } p > 1 \text{ and } f'(\lambda_0) = \cdots = f^{(p-1)}(\lambda_0) = 0$$

Proof. Suppose

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \cdots, (\lambda - \lambda_u)^{p_u}$$

are the elementary factors of A. Then $A = TJT^{-1}$, where J is a Jordan matrix. Therefore, $f(A) = Tf(J)T^{-1}$ □

2.1.4 Intermezzo: Centers

In this section, we are interested in the study of the centers of a matrix. We begin by a well-known result.

Theorem 2.1.6.

Proof. □

In the following discussion, we base on a simple but useful theorem which eliminate a lot of possible situations.

Theorem 2.1.7. *Let A and B be $n \times n$ and $m \times m$ matrix respectively. Suppose the minimal polynomials of A and B are relatively prime, then the euqation $XA = BX$ only have zero solution.*

Proof.

$$Xm_A(A) = m_A(B)X = 0 \text{ and } m_B(B)X = 0$$

$$\exists u, v \text{ s.t. } um_A + vm_B = 1 \Longrightarrow 0 = u(B)m_A(B)X + v(B)m_B(B)X = X$$

□

Remark 2.1.8. Actually, this is a special case of the general discussion about matrix equation. We may take a deeper look at it in 6.3.

And we need another theorem to reduce a general matrix to some simple cases

Theorem 2.1.9. *If two matrices are similar, than there centers are isomorphic.*

Proof. Let A, B be two matrices.Suppose P is an invertible matrix s.t. $P^{-1}AP = B$. If $X \in C(A)$, then $P^{-1}XP \in B$, this gives a homomorphism. By symmetry this is a isomorphism. □

Now let's discuss the center of a companion matrix.

Theorem 2.1.10. *Let $A \in M_{(n\times n)}(F)$ be a companion matrix of a polynomial $p(x)$. Then $C(A) = F[A]$, $\dim C[A] = n$.*

Proof. Let $B \in C(A)$, then $BA = AB$. Let $p(x) = x^n + \sum_{i=0}^{n-1} a_i x^i$. Let $\operatorname{vec}\alpha_1, \ldots, \operatorname{vec}\alpha_n$ be the basis, then

$$\begin{aligned} A\operatorname{vec}\alpha_1 &= \operatorname{vec}\alpha_2 \\ &\vdots \\ A\operatorname{vec}\alpha_{n-1} &= \operatorname{vec}\alpha_n \\ A\operatorname{vec}\alpha_n &= -\sum_{0}^{n-1} a_i \operatorname{vec}\alpha_{i+1} \end{aligned}$$

Denote $B \operatorname{vec} \alpha_i$ by $\operatorname{vec} \beta_i$. If $AB = BA$, then $AB \operatorname{vec} \alpha_i = BA \operatorname{vec} \alpha_i$ for all i. Therefore,

$$\begin{aligned} A \operatorname{vec} \beta_1 &= \operatorname{vec} \beta_2 \\ &\vdots \\ A \operatorname{vec} \beta_{n-1} &= \operatorname{vec} \beta_n \\ A \operatorname{vec} \beta_n &= -\sum_0^{n-1} a_i \operatorname{vec} \beta_{i+1} \end{aligned}$$

So given β_1, we can calculate other β_i's and the last equality is trivially satisfied by the property of A. We get $\dim C(A) = n$. On the other hand, $\dim F[A] = n$ and $F[A] \subset C(A)$, so $C(A) = F(A)$. □

Remark 2.1.11. To deal with companion matrix, or something with a similar cyclic structure, it is insightful to consider its action on the basis.

Now let's discuss the center of a Jordan block.

Theorem 2.1.12. *Given $a \in F$, let $J_n(a)$ denote the $n \times n$ Jordan block with diagonal elements a. Then $C(J_n(a)) = F[J_n(0)]$ and $\dim C(J_n(a)) = n$.*

Proof. Let $B \in C(J_n(a))$. Note that $J_n(a) = aI + J_n(0)$, so $J_n(a)B = BJ_n(a) \iff J_n(0)B = BJ_n(0)$. Similarly, let $\operatorname{vec} \alpha_1, \ldots, \operatorname{vec} \alpha_n$ be the basis, then

$$\begin{aligned} A \operatorname{vec} \alpha_1 &= \operatorname{vec} \alpha_2 \\ &\vdots \\ A \operatorname{vec} \alpha_{n-1} &= \operatorname{vec} \alpha_n \\ A \operatorname{vec} \alpha_n &= 0 \end{aligned}$$

for the rest of the proof, we can argue exactly as in the proof of 2.1.10 □

We can use 2.1.7 to generalize 2.1.12 to a Jordan canonical form.

Theorem 2.1.13. *Let $A \in M_{n\times n}(F)$ be a Jordan canonical form,i.e.*

$$A = diag\{J_{n_!}(\lambda_1), J_{n_2}(\lambda_2), \cdots, J_{n_s}(\lambda_s)\}$$

with the restriction that none of the λ_i's are equal. Then $C(A) = \{B = diag\{B_1, B_2, \cdots, B_n\} \,|B_i \in F[J_{n_i}(0)]\}$, $\dim C(A) = n, C(A) = F[A]$.

Actually, there is a general theorem stated as below:

2.2 λ-matrix Revisited: Modules over PID

2.2.1 Modules

Definition 2.2.1. Let M be an R-module. A submodule of the form

$$\langle v \rangle = Rv = \{rv | r \in R\}$$

for $v \in M$ is called a cyclic submodule generated by v.

Definition 2.2.2. Let M be an R-module. The annihilator of an element $v \in M$ is

$$\operatorname{ann}(v) = \{r \in R | rv = 0\}$$

and the annihilator of a submodule N of M is

$$\operatorname{ann}(N) = \{r \in R | rN = \{0\}\}$$

It is easy to see that $\operatorname{ann}(v)$ and $\operatorname{ann}(N)$ are ideals of R. For $x \in M$ we define

$$\zeta_x : R \longrightarrow Rx, a \mapsto ax$$

then this is a module homomorphism. $\ker(\zeta_x)$=$\operatorname{ann}(x)$, and $Rx \simeq R/\operatorname{ann}(x)$.

Definition 2.2.3. Let M be an r-module. A subset of M is a basis if it is linearly independent and spans M. An R-module M is said to be free if M has a basis.

Theorem 2.2.4.

2.2.2 Modules over PID

In this section, we consider modules over principle ideal domains.

Lemma 2.2.5. *Let R be a PID, M be a free module with rank n. Then every submodule of M is a free R-module with rank $\leq n$.*

Let M' be a finitely generated module over a PID R, $x_1, \cdots, x_m$ be a set of generators. Construct a free R-module M of rank m, $e_1, \cdots, e_m$ be a set of basis, then the homomorphism from M to M'

$$\eta : \sum_{i=1}^{m} a_i e_i \twoheadrightarrow \sum_{i=1}^{m} a_i x_i$$

is surjective. Denote $N = \ker(\eta)$, then $M/N \simeq M'$, where N can be interpreted as the collection of the relationship between M''s generators $x_1, \cdots, x_m$. Conversely, given a submodule N of M, we can construct a finitely generated $R-$module M'.

Let $f_1, \cdots, f_n$ be a set of generators that generate N. Let

$$f_i = \sum_{j=1}^{m} a_{ji} e_j, i = 1, \cdots, n$$

i.e.

$$(f_1, \cdots, f_n) = (e_1, \cdots, e_m)A$$

where $A = (a_{ij})$ is a $m \times n$ matrix. Make a basis transformation, let $e'_1, \cdots, e'_m$ be another basis of M. Let

$$(e_1, \cdots, e_m) = (e'_1, \cdots, e'_m)P$$

where $P = (p_{ij}) \in M_{m\times m}(R)$ is a invertible matrix. Conversely, if P is invertible, then $e'_1, \cdots, e'_m$ is a set of basis of M. Similarly, utilize an invertible matrix $Q \in M_{n\times n}(R)$ to be a transformation of N's generators

$$(f_1, \cdots, f_n) = (f'_1, \cdots, f'_n)Q$$

where the restriction that Q is invertible guarantees that $f'_1, \cdots, f'_n$ still generates N. Therefore

$$(f_1, \cdots, f_n) = (e'_1, \cdots, e'_m)PAQ^{-1}$$

If we are given a free module M of rank m and its basis $e_1, \cdots, e_m$, the matrix A portraits the submodule N, as well as the quotient module M/N. The discussion above explains the relationship between the transformation of basis as well as generators and the transformation of matrix. This leads to our following definition.

Definition 2.2.6. Let $A, B \in M_{m\times n}(R)$. If $\exists\ P \in M^{\times}_{m\times m}(R)$ and $Q \in M^{\times}_{n\times n}(R)$ s.t.

$$B = PAQ$$

then we call A is equivalent to B on R.

Theorem 2.2.7. *Let M be a free module on R with rank m, N be its submodule. The there exists a basis $e_1, \cdots, e_n$ of M s.t. $d_1e_1, \cdots, d_re_r$ be a basis of N, satisfying*

$$d_i | d_{i+1}, i = 1, \cdots, r-1$$

$d_1, \cdots, d_r$ are uniquely determined by N except for units, and they are called the invariant of N or the invarinat factors of M/N. r is the rank of N, $m - r$ is the rank of M/N.

Actually, by using the matrix representation, this theorem is essentially the same to the theorem of Smith normal form of λ-matrix. The only difference is that $F[\lambda]$ is not only a PID, but also a Euclid's ring. Therefore, we need to generalize the Smith normal form to PID.

Theorem 2.2.8.

2.3 Decomposition of Linear Operators

In this section, we focus on how to decompose a vector space(module) with respect to a linear operator into several simple subspaces(submodules).

First, we need to introduce some important concepts that will be used extensively

2.3.1 Cyclic Spaces

2.3.2 Irreducible Spaces

2.3.3 Centers

In this part, we consider the center of a linear transformation φ.

Let f be any polynomial. Then $\ker(f)$ is invariant under every $\phi \in C(\varphi)$. In fact if $v \in \ker(f)$ is any vector, then

$$f(\varphi)\phi v = \phi f(\varphi)v = 0$$

Next consider the decomposition of V into generalized eigenspaces of φ and of ϕ,

$$V = V_1 \oplus \cdots V_r \quad (\text{for } \varphi)$$

and

$$V = W_1 \oplus \cdots W_s \quad (\text{for } \phi)$$

and the corresponding projection operators in V, π_i and ρ_j. Since the mappings π_i and ρ_j are respectively polynomials in φ and ϕ, it follows that

$$\pi_i \circ \rho_j = \rho_j \circ \pi_i \quad \forall i, j$$

Now define linear transformations τ_{ij} in V by

$$\tau_{ij} = \pi_i \circ \rho_j$$

Then we obtain that

$$\tau_{ij}^2 = \pi_i \circ \rho_j \circ \pi_i \circ \rho_j = \pi_i^2 \circ \rho_j^2 = \pi_i \circ \rho_j = \tau_{ij}$$

and hence the τ_{ij} are again projection operators in V.

Since

$$\operatorname{Im} \tau_{ij} \subset V_i \cap W_j$$

and

$$\sum_{i,j} \tau_{ij} = (\sum_i \pi_i) \circ (\sum_j \rho_j) = Id_V$$

it follows that

$$\operatorname{Im} \tau_{ij} = V_i \cap W_j$$

and

$$V = \sum_{i,j} V_i \cap W_j$$

Theorem 2.3.1. *Let $V = V_1 \oplus \cdots V_s$ be any decomposition of V as a direct sum of subspaces. Then the subspaces V_j are invariant under φ if and only if the projection operators σ_j are contained in $C(\varphi)$.*

Proof. If $\sigma_j \in C(\varphi)$, then V_j are invariant under φ. Conversely, if the V_j are invariant under φ, we have for each $v \in V_j$ that $\varphi v \in V_j$, and hence $\sigma_j \varphi v = \varphi v = \varphi \sigma_j v$ while $\sigma_l \varphi v = 0 = \varphi \sigma_l v$ for $l \neq j$. Thus σ_l commute with φ. □

Theorem 2.3.2 (Cecioni-Frobenius Theorem)**.**

Theorem 2.3.3 (bicommutant)**.**

$$C^2(\varphi) = F[\varphi]$$

Proof. Clearly $C^2(\varphi) \supset F[\varphi]$. Conversely, suppose $\phi \in C^2(\varphi)$ is any linear transformation and let

$$V = V_1 \oplus \cdots \oplus V_s$$

be a decomposition of V into cyclic subspaces w.r.t. φ. Let a_i be any fixed generator of the space V_i.

Denote by φ_i the linear transformation in V_i induced by φ and let μ_i be the minimum polynomial of φ_i. Then $\mu_i | \mu$ so we can write $\mu = \mu_i v_i$. We may assume that $\mu_1 = \mu$.

Now the V_i are invariant under φ, so the projection operators commute with φ. Hence they commute with ϕ as well, and V_i are invariant under ϕ. In particular $\phi a_i \in V_i$, therefore $\phi a_i = g_i(\varphi)a_i$. Thus if $h(\varphi)a_i \in V_i$ is an arbitrary vector in V_i we obtain

$$\phi h(\varphi)a_i = h(\varphi)\phi a_i = h(\varphi)g_i(\varphi)a_i = g_i(\varphi)h(\varphi)a_i$$

so $\phi_i = g_i(\varphi)$ where ϕ_i denotes the restriction of ϕ to V_i.

Our goal is to show $\phi = g_1(\varphi)$.

Consider now linear transformations $\chi_i (i \neq 1)$ in V defined by

$$\chi_i x = x \quad x \in F_j \ j \neq i$$
$$\chi_i f(\varphi)a_i = v_i(\varphi)f(\varphi)a_1$$

To show that χ_i is well-defined it is sufficient to prove that

$$f(\varphi)a_i = 0 \Longrightarrow v_i(\varphi)f(\varphi)a_1 = 0$$

This is because $f(\varphi)a_i = 0$, then $\mu_i | f$ and so $\mu | v_i f$.

χ_i commutes with φ, and hence with ϕ. On the other hand,

$$\chi_i \phi a_i = \chi_i g_i(\varphi)a_i = v_i(\varphi)g_i(\varphi)a_1$$
$$\phi\chi_i a_i = \phi v_i(\varphi)a_1 = v_i(\varphi)\phi a_1 = v_i(\varphi)g_1(\varphi)a_1$$

whence

$$v_i(\varphi)[g_i(\varphi) - g_1(\varphi)]a_1 = 0$$

This relation implies $\mu | v_i(g_i - g_1)$, so that $\mu_i | g_i - g_1$. This last relation yields that for any vector $x \in V_i$,

$$\phi x = g_i(\varphi)x = g_1(\varphi)x$$

that is $\phi = g_1(\varphi)$. □

2.4 Decomposition of Linear Operators Revisited: Modules over PID

Theorem 2.4.1 (primary decomposition)**.** *Let V be finite-dimensional and let $\tau \in \mathcal{L}(V)$ have minimal polynomial*

$$m_\tau(x) = p_1^{e_1}(x) \cdots p_n^{e_n}(x)$$

where the polynomials $p_i(x)$ are distinct monic primes. Then the $F[x]$-module V_τ is the direct sum

$$V_\tau = \oplus_{i=1}^n V_{p_i}$$

where

$$V_{p_i} = \frac{m_\tau(x)}{p_i^{e_i}(x)} V = \{v \in V | p_i^{e_i}(\tau)(v) = 0\}$$

is a primary submodule of V_τ of order $p_i^{e_i}(x)$. In vector space terms, V_{p_i} is a τ-invariant subspace of V and the minimal polynomial of $\tau|_{V_{p_i}}$ is $p_i^{e_i}(x)$

2.5 ???Linear Homomorphisms

To study abstract mathematics well, the key is to transfer freely between the abstract world and the computational world. For example, semisimple homomorphism can correspond to block matrix. One aspect is that we can apply abstract structures in real world problem, another aspect is that we can subtract abstract structures from concrete examples.

2.5.1 ???Semisimple Operators

Definition 2.5.1 (simple,irreducible)**.** A linear operator T on a vector space V is simple if V has no nontrivial T-invariant subspace.

Definition 2.5.2 (semisimple,completely reducible)**.** A linear operator T on a vector space is semisimple if every T-invariant subspace has a complementary T-invariant subspace.

Remark 2.5.3. This definition origins from semisimple representation in representation theory. A semisimple representation is a linear representation of a group or an algebra that is a direct sum of simple representations(a nonzero representation that has no proper nontrivial subrepresentation).

Chapter 3

Theory of Bilinear Functions

Now we study the theory of bilinear functions. This chapter begins with a general definition of bilinear functions, then studies a more special case where the bilinear function is nondegenerate and induces a kind of duality. If we impose some other restrictions, then we would obtain metric vector spaces, and the more familiar inner product spaces.

The duality provides us a powerful language to deal with the canonical form of normal operators, which is very difficult in the language of matrices. The bridge between these two viewpoints is that the transposition of a matrix is the adjoint of the operator corresponding to that matrix in the standard inner product space.

We also know that the transposition of a matrix corresponds to the dual mapping of that matrix. And it is exactly the duality identify the dual mapping from the space of linear functions to the original space. These two different approaches lead to the same end.

3.1 Bilinear Functions

Definition 3.1.1 (bilinear functions)**.** Let V and W be vector spaces and F be a field. Then a mapping $B : V \times W \to F$ satisfying

$$\begin{aligned} B(\lambda v_1 + \mu v_2, w) &= \lambda B(v_1, w) + \mu B(v_2, w) \quad \forall v_1, v_2 \in V \ w \in W \\ B(v, \lambda w_1 + \mu w_2) &= \lambda B(v, w_1) + \mu B(v, w_2) \quad \forall v \in V \ w_1, w_2 \in W \end{aligned}$$

Remark 3.1.2. The space of all bilinear functions is denoted by $\mathrm{Bil}(V, W; F)$.

Definition 3.1.3 (radicals)**.** A bilinear function B in $V \times W$ determines two subspaces $N_V \subset V$ and $N_W \subset W$ defined by

$$\begin{aligned} N_V &= \{v | B(v, W) = 0\} \\ N_W &= \{w | B(V, w) = 0\} \end{aligned}$$

They are called the left radical and the right radical respectively.

Definition 3.1.4 (non-degenerate)**.** If $N_V = N_W = \{0\}$, then B is called non-degenerate.

Definition 3.1.5 (orthogonality)**.** Two vectors $v \in V$ and $w \in W$ are called orthogonal if $B(v, w) = 0$, written $v \perp w$. Orthogonality for subspaces is defined similarly.

Definition 3.1.6 (orthogonal complement)**.** Let V_1 be a subspace of V, then the vectors of W which are orthogonal to V_1 forms a subspace $V_1^\perp$ of W. $V_1^\perp$ is called the orthogonal complement of V_1. In the same way $W_1 \subset W$ determines an orthogonal complement $W_1^\perp \subset V$.

Theorem 3.1.7. *$Hom(W, V^\vee) \simeq Bil(V, W; F) \simeq Hom(V, W^\vee)$*

Remark 3.1.8. Notice that for finite dimensional spaces V, W, a necessary condition for non-degenerate bilinear functions to exist is that $\dim V = \dim W$.

Corollary 3.1.9. *Let V, W be finite dimensional vector spaces, then*

$$\dim V - \dim N_V = \dim W - \dim N_W$$

Theorem 3.1.10. *Let V, W be finite dimensional vector spaces with $\dim V = \dim W$, then $\forall B \in Bil(V, W; F)$, TFAE:*
(1) B is non-degenerate.
(2) $N_V = \{0\}$
(3) $N_W = \{0\}$

3.2 Dual Vector Spaces

In this section, we focus on non-degenerate bilinear functions.

3.2.1 Duality

Definition 3.2.1 (dual spaces)**.** Suppose V, W is a pair of vector spaces, and assume that a fixed non-degenerate bilinear function, $\langle , \rangle$, in $V \times W$ is defined. Then V and W will be called **dual** with respect to the billinear function $\langle , \rangle$.

Theorem 3.2.2. *Let V, W be a pair of vector spaces which are dual with respect to a scalar product $\langle , \rangle$. Then an injective linear map $\Phi : V \longrightarrow W^\vee$ is defined by*

$$\Phi(v)(w) = \langle v, w \rangle$$

Remark 3.2.3. If E has finite dimension, then Φ is also surjective, hence an isomorphism.

Definition 3.2.4 (dual mapping)**.** Suppose that V, V' and W, W' are two pairs of dual spaces and $\varphi : V \to W$ and $\varphi^* : W' \to V'$ are linear mappings. The mappings φ and φ^* are called dual if

$$\langle \varphi v, w' \rangle = \langle v, \varphi^* w' \rangle \quad \forall v \in V, w' \in W$$

Uniqueness comes immediately.

Theorem 3.2.5. *There exists at most one dual mapping to a given linear mapping $\varphi : V \to W$.*

Let us discuss the operations of dual mappings.

Theorem 3.2.6. *Let $\sigma : V \to W$ and $\tau : V \to W$ and $r \in F$*
(1) $(\sigma + \tau)^* = \sigma^* + \tau^*$
(2) $(r\tau)^* = r\tau^*$

Now we look at the kernel and image spaces.

Theorem 3.2.7. *Let $\sigma : V \to W$,*
(1) $\ker \varphi^* = (\operatorname{Im} \varphi)^\perp$
(2) $\ker \varphi = (\operatorname{Im} \varphi^*)^\perp$

Theorem 3.2.8. *Consider two subapsces V_1 and V_2 of V, then*

$$(V_1 + V_2)^\perp = V_1^\perp \cap V_2^\perp$$

3.2.2 Space of Linear Functions

Let V be a vector space and $V^\vee$ be the space of linear functions in V, then they are dual w.r.t. the natural evaluation. The space of linear functions have three important results, which are not valid for arbitrary pairs of dual spaces.

The first result is the existence of dual mapping.

Theorem 3.2.9. *Let W, W' be arbitrary dual spaces and $\varphi : V \to W$ be a linear mapping. Then a dual mapping $\varphi^* : W' \to V^\vee$ exists and is given by*

$$(\varphi^* w')(v) = \langle \varphi x, w' \rangle \quad \forall w' \in W', v \in V$$

Theorem 3.2.10. *Suppose $\varphi : V \to W$ is a linear mapping, and consider the dual mapping $\varphi^* : W^\vee \to V^\vee$. Then*

$$\operatorname{Im} \varphi^* = (\ker \varphi)^\perp$$

Theorem 3.2.11. *If $W \subset V$ is any subspace, then*

$$W^{\perp\perp} = W$$

Let us consider direct decompositions.

Theorem 3.2.12. *Suppose $V = V_1 \oplus V_2$. Then $V^\vee = V_1^\perp \oplus V_2^\perp$ and the pairs $V_1^\perp, V_2$ and $V_2^\perp, V_1$ are dual w.r.t. the induced evaluation.*

Remark 3.2.13. Thus the induced injections $V_1^\perp \to V_2^\vee$ and $V_2^\perp \to V_1^\vee$ are surjective, and hence $V^\vee = V_1^\vee \oplus V_2^\vee$. Finally $(V_1^\perp)^{\perp\perp} = V_1^\perp$ and $(V_2^\perp)^{\perp\perp} = V_2^\perp$.

3.2.3 Finite Dimensional Vector Spaces

For finite dimensional vector spaces, the dual spaces have some nice property, and the most important is remark3.2.2, which has established the relationship between a general dual space and the specific dual space of linear functions. Thus, we can transport the properties from the space of linear functions to any general dual space.

3.3 Metric Vector Spaces

In this section, all vector spaces are assumed finite-dimensional. We restrict our attention to two special kinds of bilinear form, namely the symmetic ones and the alternate ones. Note that we do not assume non-degeneration, because actually that is what we are going to deal with.

Definition 3.3.1. A bilinear form $V \times V \to F$ is
(1) **symmetric** if $\langle x, y\rangle = \langle y, x\rangle \quad \forall x, y \in V$
(2) **skew-symmetric** if $\langle x, y\rangle = -\langle y, x\rangle \quad \forall x, y \in V$
(3) **alternate** if $\langle x, y\rangle = 0 \quad \forall x \in V$

Definition 3.3.2. The pair $(V, \langle , \rangle)$ is called a **metric vetor space** if $\langle , \rangle$ is symmetric, skew-symmetric, or alternate.

Remark 3.3.3. A metric vector space with a symmetric/alternate form is called an orthogonal/symplectic geometry.

We exhibit the relationship between alternate and skew symmetric bilinear forms.

Theorem 3.3.4. *Let V be a vector space over a field F.*
(1) If $char(F) = 2$, then

$$alternate \implies symmetric \iff skew\text{-}symmetric$$

(2) If $char(F) \neq 2$, then

$$alternate \iff skew\text{-}symmetric$$

The reason we focus on symmetric or alternate bilinear forms is that only these kinds of bilinear forms guarantee the symmetric of orthogonality.

Theorem 3.3.5. *Let V be a vector space with a bilinear form. TFAE:*
(1) Orthogonality is a symmetric relation, i.e. $x \perp y \implies y \perp x$
(2) V is a metric vector space, i.e. the form is symmetric or alternate.

Now we study two types of degenerate behaviors that a vector may possess. The first case is that a vector may be orthogonal to itself, and the more severe case is that it may be orthogonal to every vector in V.

Definition 3.3.6 (isotropic)**.** Let V be a metric vector space. A nonzero $v \in V$ is isotropic if $\langle v, v\rangle$=0. V is isotropic if it contains at least one isotropic vector.

Definition 3.3.7 (degenerate)**.** A vector $v \in V$ is degenerate if $v \perp V$.

Remark 3.3.8. In a metric vetor space V, the left radical is equal to the right radical, called the radical of V and denoted by $\mathrm{rad}(V)$. It is exactly the set of all degenerate vectors.

Definition 3.3.9. If S is a subspace of a metric vector space V, then $\mathrm{rad}(S)$ denotes the set of vectors in S that are degenerate in S.

Remark 3.3.10. $\mathrm{rad}(S) = S \cap S^\perp$

3.3.1 Orthogonal Direct Sum

Definition 3.3.11. A metric vector space V is the orthogonal direct sum of the subspace S and T if $V = S \oplus T$ and $S \perp T$, denoted by $V = S \odot T$.

Theorem 3.3.12. *$V = \mathrm{rad}(V) \odot S$, where S is non-degenerate.*

Proof. S is any vector space complement of $\mathrm{rad}(V)$. □

Lemma 3.3.13. *Let S be a subspace of a metric vector space V. If either V or S is non-degenerate, the linear map $\tau : V \to S^\vee$ defined by*

$$\tau x = \langle \cdot, x \rangle |_S$$

is surjective and has kernel $S^\perp$.

Proof. Use the Riesz representation theorem. □

Theorem 3.3.14. *Let S be a subspace of V. If S is non-degenerate, then*

$$\dim(S) + \dim(S^\perp) = \dim(V)$$

Hence TFAE:
(1) $V = S + S^\perp$
(2) S is non-degenerate
(3) $V = S \odot S^\perp$

Proof. Application of the above lemma. □

Theorem 3.3.15. *Let S be a subspace of V. If V is non-degenerate, then*

$$\dim(S) + \dim(S^\perp) = \dim(V)$$

Hence:
(1) $S^{\perp\perp} = S$
(2) $\mathrm{rad}(S) = \mathrm{rad}(S^\perp)$
(3) S is non-degenerate if and only if $S^\perp$ is non-degenerate

3.3.2 Isometry

We now turn to a discussion of structure-preserving maps on metric vector spaces.

Definition 3.3.16. Let V and W be metric vector spaces. A linear isomorphism is called an isometry if it preserves the metric. If an isometry exists from V to W, we write $V \approx W$.

Remark 3.3.17. If V is a nondegenerate orthogonal/symplectic geometry, an isometry of V is called an orthogonal/symplectic transformation.

3.3.3 Hyperbolic Spaces

A special type of two-dimensional metric vector space plays an important role in the structure theory of metric vector spaces, so we single out this construction in this subsection.

Definition 3.3.18. Let V be a metric vector space. A **hyperbolic pair** is a pair of vectors $u, v \in V$ for which

$$\langle u, u \rangle = \langle v, v \rangle = 0, \langle u, v \rangle = 1$$

The subspace $H = \mathrm{span}(u, v)$ is called a **hyperbolic plane**, and any space which is an orthogonal direct sum of several hyperbolic planes is called a **hyperbolic space**.

3.4 Real and Complex Inner Product Spaces

For complex inner product spaces, we are no long considering bilinear forms, but sesquilinear forms. Here we impose positive-definiteness, but the vector space need not to be finite dimensional.

Definition 3.4.1 (inner product)**.** Let V be a vector space over $F = \mathbb{R}$ or $F = \mathbb{C}$. An inner product on V is a positive-definite sesquilinear function $\langle , \rangle : V \times V \longrightarrow F$.

Example 3.4.2 (standard inner product on $\mathbb{R}^n$)**.**

$$\langle (r_1, \cdots, r_n), (s_1, \cdots, s_n) \rangle = r_1 s_1 + \cdots + r_n s_n$$

Example 3.4.3 (standard inner product on $\mathbb{R}^n$)**.**

$$\langle (r_1, \cdots, r_n), (s_1, \cdots, s_n) \rangle = r_1 \bar{s}_1 + \cdots + r_n \bar{s}_n$$

Example 3.4.4. The vectpr space $C[a, b]$ of all continous complex valued functions on the closed interval $[a, b]$ is a complex inner product space under the inner product

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}\mathrm{d}x$$

The following lemma is a simple conseqence of the positive definiteness.

Lemma 3.4.5. *If $\langle u, x \rangle = \langle v, x \rangle$ for every $x \in V$, then $u = v$.*

The next result points out one of the main differneces between real and complex inner product spaces and will play a key role in later work.

Theorem 3.4.6. *Let $\tau \in \mathcal{L}(V)$.*
1)

$$\langle \tau v, w \rangle = 0 \ \forall v, w \in V \Longrightarrow \tau = 0$$

2)If V is a complex inner product space, then

$$\langle \tau v, v \rangle = 0 \ \forall v \in V \Longrightarrow \tau = 0$$

but this does not hold for real inner porduct spaces.

Proof. Part 1) follow directly from the above lemma. □

Definition 3.4.7 (norm)**.** If V is an inner product space, the **norm** of $v \in V$ is defined by

$$\|v\| = \sqrt{\langle v, v \rangle}$$

Recalling that an isometry is an isomorphism which preserve the inner product, we have the following equivalent definition.

Theorem 3.4.8. *An isomorphism $\tau \in Hom(V, W)$ is an isometry if and only if it preserves the norm.*

Recalling the definition of an orthogonal direct sum, in an inner product space, we have the following properties.

By positive definiteness, we have a non-degenerate property.

Theorem 3.4.9. *For a subspace $S \subset V$,* $\mathrm{rad}(S) = \{0\}$.

We also have the uniqueness of an orthogonal direct sum decomposition.

3.4.1 Gram-Schmidt Orthogonalization

The Gram-Schmidt Orthogonalization process guarantee that for a finite dimensional inner product space,

Theorem 3.4.10. *If $V = S \odot T$, then $T = S^{\perp}$.*

3.4.2 Application: QR Decomposition

Theorem 3.4.11 (QR decomposition)**.**

Theorem 3.4.12 (Cholesky decomposition)**.**

3.4.3 The Projection Theorem

Theorem 3.4.13. *If S is a finite-dimensional subspace of an inner product space V, then $V = S \odot S^{\perp}$.*

We can prove it by the general theory of metric vector space, but for inner product spaces, we can approach the problem by Gram-Schmidt Orthogonalization.

Theorem 3.4.14.

3.4.4 Schur Triangularization

3.5 Structure Theory for Normal Operators

Throughout this section, all vector spaces are assumed to be finite-dimensional unless noted. Also the field F is either $\mathbb{R}$ or $\mathbb{C}$.

3.5.1 The Adjoint of a Linear Operator

We begin by limiting our attention within inner product spaces. We want to study a special type of linear operator, namely the normal operators. First we recall a definition.

Definition 3.5.1. Let V and W be finite-dimensional inner product spaces over F and let $\tau \in \mathrm{Hom}(V, W)$. Then there is a unique function $\tau^* : W \to V$ defined by

$$\langle \tau v, w \rangle = \langle v, \tau^* w \rangle$$

Remark 3.5.2. This is a special case of the more general construction of adjoint operators, where we restrict the nondegenerate bilinear form to the inner product.

Here are some basic properties of the adjoint operator. These properties essentially rely on the properties of inner product spaces.

Theorem 3.5.3. *Let V and W be finite-dimensional inner product spaces over F and let $\sigma, \tau \in Hom(V, W)$ and $r \in F$,*

(1) $(\sigma + \tau)^* = \sigma^* + \tau^*$

(2) $(r\tau)^* = \bar{r}\tau^*$

(3) $\tau^{**} = \tau$*, and so* $\langle \tau^* v, w \rangle = \langle v, \tau w \rangle$

(4) If $V = W$*, then* $(\sigma\tau)^* = \tau^*\sigma^*$

(5) If τ *is invertible, then* $(\tau^{-1})^* = (\tau^*)^{-1}$

(6) If $V = W$ *and* $p(x) \in \mathbb{R}[x]$*, then* $p(\tau)^* = p(\tau^*)$

Theorem 3.5.4. *If $\tau \in End(V)$ and S is a subspace of V, then*

(1) S is τ-invariant if and only if $S^{\perp}$ is τ^*-invariant

(2) $(S, S^{\perp})$ reduces τ if and only if S is both τ-invariant and τ^*-invariant

Theorem 3.5.5. *Let V and W be finite-dimensional inner product spaces over F and let $\tau \in Hom(V, W)$.*

(1) $\ker(\tau^*) = \mathrm{Im}(\tau)^{\perp}$ *and* $\mathrm{Im}(\tau^*) = \ker(\tau)^{\perp}$

(2) $\ker(\tau^*\tau) = \ker(\tau)$ *and* $\ker(\tau\tau^*) = \ker(\tau^*)$

(3) $\mathrm{Im}(\tau^*\tau) = \mathrm{Im}(\tau)$ *and* $\mathrm{Im}(\tau\tau^*) = \mathrm{Im}(\tau^*)$

(4) $(\rho_{S,T})^* = \rho_{T^{\perp}, S^{\perp}}$

3.5.2 Orthogonal Projections

An operator ρ is a projection operator if and only if it is idempotent. In an inner product space, we can single out some special projection operators.

Definition 3.5.6. A projection of the form $\rho_{S,S^\perp}$ is said to be orthogonal.

Remark 3.5.7. Some care must be taken to avoid confusion between orthogonal projections and two projections that are orthogonal to each other, that is $\rho\sigma = \sigma\rho = 0$.

Here is a characterization of orthogonal projections.

Theorem 3.5.8. *Let V be a finite-dimensional inner product space. TFAE:*
(1) ρ is an orthogonal projection
(2) ρ is idempotent and self-adjoint
(3) ρ is idempotent and does not expand lengths

3.5.3 Normal Operators

Definition 3.5.9. A linear operator τ on an inner product space V is normal if it commutes with its adjoint.

Theorem 3.5.10. *Let $\tau \in End(V)$ be normal*
(1) The following are also normal:
a. $\tau|_S$ if τ reudces $(S, S^\perp)$
b. τ^*
c. τ^{-1} if τ is invertible
d. $p(\tau)$, for any polynomial $p(x) \in F[x]$
(2) For any $v, w \in V$, $\langle \tau v, \tau w\rangle = \langle \tau^* v, \tau^* w\rangle$, and in particular $\|\tau v\| = \|\tau^* w\|$, and so

$$\ker(\tau^*) = \ker(\tau)$$

(3) For any integer $k\mathcal{E}1$,

$$\ker(\tau^k) = \ker(\tau)$$

(4) The minimal polynomial $m_\tau(x)$ is a product of distinct prime monic polynomials
(5) $\tau v = \lambda v \iff \tau^* v = \bar{\lambda} v$
(6) If S and T are submodules of V_τ with relatively prime orders, then $S \perp T$, and in particular, if λ and ν are distinct eigenvalues of τ, then $\mathcal{E}_\lambda \perp \mathcal{E}_\nu$

Using the general structure above, we can obtain the spectral theorem for normal operators.

Theorem 3.5.11 (spectral theorem for normal operators: complex case)**.** *Let V be finite-dimensional complex inner product spaces and let $\tau \in End(V)$. TFAE:*
(1) τ is normal
(2) τ is unitary diagonalizable
(3) τ has an orthogonal spectral resolution

$$\tau = \lambda_1\rho_1 + \cdots + \lambda_k\rho_k$$

Theorem 3.5.12 (spectral theorem for normal operators: real case)**.**

3.5.4 Functional Calculus

3.5.5 Application: Positive Operators

3.5.6 Application: The Polar Decomposition

It is well-known that any nonzero complex number z can be written in the polar form $z = \rho e^{i\theta}$. We can do the same for any nonzero linear operator τ on a finite-dimensional complex inner product space.

Theorem 3.5.13. *Let τ be a nonzero linear operator on a finite-dimensional complex inner product space V.*
(1) There exist a positive operator ρ and a unitary operator ν for which $\tau = \nu\rho$. Moreover, if τ is invertible, then the decomposition is unique.
(2) There exist a positive operator σ and a unitary operator μ for which $\tau = \sigma\mu$. Moreover, if τ is invertible, then the decomposition is unique.

Any unitary operator μ has the form $\mu = e^{i\sigma}$, where σ is a self adjoint operator. This gives the following corollary.

Corollary 3.5.14. *Let τ be a nonzero linear operator on a finite-dimensional complex inner product space V. Then there is a positive operator ρ and a self-adjoint operator σ for which τ has the polar decomposition*

$$\tau = \rho e^{i\sigma}$$

Moreover, if τ is invertible, then the decomposition is unique.

Normal operators can be characterized using the polar decomposition.

Theorem 3.5.15. *Let $\tau = \rho e^{i\sigma}$ be a polar decomposition of a nonzero operator τ. Then τ is normal if and only if $\rho\sigma = \sigma\rho$.*

3.5.7 Normal Operators in Complex Inner Product Spaces: A Summary

Theorem 3.5.16. *Let $(V, \langle , \rangle)$ be a finite-dimensional inner product space, $\tau \in End(V)$. TFAE:*
(1) τ is normal
(2) $\|\tau\alpha\| = \|\tau^*\alpha\|, \forall\alpha \in V$
(3) $\tau = \tau_1 + i\tau_2$, where T_1, T_2 are self-adjoint and commutative
(4) If $\alpha \in v, c \in \mathbb{C}$, $\tau\alpha = c\alpha$, then $\tau^*\alpha = \bar{c}\alpha$
(5) τ is diagonalizable
(6) $\exists g \in \mathbb{C}[x]$ s.t. $T^* = g(T)$
(7) If a subspace $W \subset V$ is τ-invariant, then W is T^*-invariant
(8) $\tau = \rho\mu$, ρ is positive and μ is unitary with $\rho\mu = \mu\rho$
(9) τ has an orthogonal spectral resolution
(10) $\mathrm{tr}(T^*T) = \sum_{i=1}^{n} |\lambda_i|^2$, where λ_i are the eigenvalues of T

Theorem 3.5.17. *If S, T, ST are normal, then TS is normal.*

Proof. First, notice that the eigenvalues of ST are the same as the eigenvalues of TS.

Next, $\mathrm{tr}((TS)^*(TS)) = \mathrm{tr}(S^*T^*TS) = \mathrm{tr}(T^*TSS^*) = \mathrm{tr}(TT^*S^*S) = \mathrm{tr}(T^*S^*ST) = \mathrm{tr}((ST)^*ST) = \sum_{i=1}^{n} |\lambda_i|^2$. So TS is normal. □

3.6 Discussion

Section ??? follows Roman [2013]

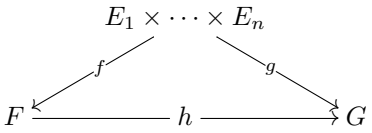
Chapter 4

Multilinear Algebra

4.1 Tensor Product

4.1.1 Tensor Product

The tensor product is a universal object in the category, whose objects are multilinear maps of a fixed set of modules $E_1, \cdots, E_n$. The morphism from f to g is defined by h which makes the following diagram commutative:



Theorem 4.1.1. *A tensor product exists and is uniquely determined up to a unique isomorphism.*

Proof. By abstract nonsense, we know of course a tensorproduct is uniquely determined.

Let M be the free module generated by the set of all n-tuples $(x_1, \cdots .x_n)$,$(x_i \in E_i)$, i.e. generated by the set $E_1 \times \cdots \times E_n$. Let N be the submodule generated by all the elements of the following type:

$$(x_1, \cdots, x_i + x_i', \cdots, x_n) - (x_1, \cdots, x_i, \cdots, x_n) - (x_1, \cdots, x_i', \cdots, x_n)$$

$$(x_1, a \cdots, x_i, \cdots, x_n) - a(x_1, \cdots, x_i, \cdots, x_n)$$

for all $x_i \in E_i, x_i' \in E_i, a \in R$. We have the canonical injection

$$E_1 \times \cdots \times E_n \longrightarrow M$$

of our set into the free module generated by it. We compose this map with the canonical map $M \longrightarrow M/N$ on the factor module, to get a map

$$\varphi : E_1 \times \cdots \times E_n \longrightarrow M/N$$

We contend that φ is multilinear and is a tensor product.

It is obvious that φ is multilinear. Our definition was adjusted to this purpose. Let

$$f : E_1 \times \cdots \times E_n \longrightarrow G$$

be a multilinear map. □

The module M/N will be denoted by

$$E_1 \otimes \cdots \otimes E_n$$

4.1.2 Kronecker Prodcut: Application to Matrix

Definition 4.1.2.

Definition 4.1.3 (vectorization)**.** $\text{vec} : \mathbb{R}^m \otimes \mathbb{R}^n \cong \mathbb{R}^{mn}$ by converting a matrix to a vector is an isomorphism.

Theorem 4.1.4. *Suppose $A : k \times l, X : l \times m, B : m \times n$. Then*

$$\text{vec}(AXB) = B^t \otimes A \,\text{vec}(X)$$

Proof. Write $X = (x_1, \cdots, x_l)$, where x_i is a $m \times 1$ vector.

$$RHS = \begin{pmatrix} b_{11}A & \cdots & b_{m1}A \\ \vdots & & \vdots \\ b_{1n}A & \cdots & b_{mn}A \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_l \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m b_{i1}Av_i \\ \vdots \\ \sum_{i=1}^m b_{in}Av_i \end{pmatrix}$$

$$LHS = \operatorname{vec}((Av_1, \cdots, Av_l)B) = \operatorname{vec}(\sum_{i=1}^m b_{i1}Av_i, \cdots, \sum_{i=1}^m b_{in}Av_i)$$

□

Lemma 4.1.5. $\operatorname{tr}(BCD) = (\operatorname{vec}(B^T))^T (I \otimes C) \operatorname{vec}(D)$

Proof. $\operatorname{tr}(BCD) = (\operatorname{vec}(B^T))^T \operatorname{vec}(CD) = (\operatorname{vec}(B^T))^T (I \otimes C) \operatorname{vec}(D)$ □

Theorem 4.1.6. *Let $A \in M_{n\times n}(\mathbb{F})$, $B \in M_{q\times q}(\mathbb{F})$. Then the eigenvalues of $A \otimes I_q + I_n \otimes B$ are $\lambda_i(A) + \lambda_j(B)$, $1 \le i \le n, 1 \le j \le q$.*

Proof. Take the Jordan canonical form. □

Theorem 4.1.7.

Chapter 5

Lie Theory

In this chapter, we exhibit the very beautiful thoery of Sophus Lie, emphasising on the application of linear algebra.

5.1 BCH Formula

5.1.1 Matrix Exponetial

The exponential is the mechanism for passing information from the Lie algebra to the Lie group.

Let X be an $n \times n$ real or complex matrix. We wish to define the exponential of X, denoted e^X, by the usual power series

$$e^X = \sum_{m=0}^{\infty} \frac{X^m}{m!}$$

Actually, regardless of the norm, this series converge. For simplicity, we use the Frobenius norm. This norm satisfies the inequalities

$$\|X + Y\| \leq \|X\| + \|Y\|$$

$$\|XY\| \leq \|X\| \, \|Y\|$$

Proof. The first of these is the trianglular inequality, and the second follows from the Schwarz inequality. □

5.1.2 Matrix Logarithm

Lemma 5.1.1. *The function $f(z) = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(z-1)^m}{m}$ is defined and analytic in a circle of radius 1 about $z = 1$.*

For all z with $|z - 1| < 1$,

$$e^{\log z} = z$$

For all u with $|u| < \log 2$, and $|e^u - 1| < 1$,

$$\log e^u = u$$

Definition 5.1.2. For any $n \times n$ matrix A, define $\log A$ by

$$\log A = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A-1)^m}{m}$$

5.1.3 BCH Formula for Heisenberg Group

5.1.4 General BCH Formula

5.2 The Representations of SU(2) and SU(3)

5.2.1 The Irreducible Representations of SU(2)

We will use the following basis for $\mathfrak{sl}(2;\mathbb{C})$:

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}; Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Lemma 5.2.1. *Let u be an eigenvector of $\pi(H)$ with eigenvalue $\alpha \in \mathbb{C}$. Then*

$$\pi(H)\pi(X)u = (\alpha + 2)\pi(X)u$$

$$\pi(H)\pi(Y)u = (\alpha - 2)\pi(Y)u$$

Proof. □

5.3 Root Systems

5.3.1 Root

Definition 5.3.1 (Root System)**.** A root system is a finite-dimensional real vector space V with an inner product $\langle \cdot, \cdot \rangle$, together with a finite collection R of nonzero vectors in V satisfying the following properties:
(i) The vectors in R span V.
(ii) $\alpha \in R \Longrightarrow -\alpha \in R$.
(iii) If $\alpha \in R$, then the only multiples of α in R are α and $-\alpha$.
(iv) $\alpha, \beta \in R \Longrightarrow w_\alpha \cdot \beta \in R$, where

$$w_\alpha \cdot \beta = \beta - 2\frac{\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle}\alpha$$

(v) $\forall \alpha, \beta \in R$, $2\frac{\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle}$ is an integer.
$\dim V$ is called the rank of the root system and the elements of R is called roots.

There are some redundancy in this definition since $w_\alpha(\alpha) = -\alpha$, but our definition is more intuitive.

Theorem 5.3.2 (Weyl Group)**.** *Let (V, R) be a root system, then the Weyl group W of R is the subgroup of the orthogonal group of V, generated by w_α where $\alpha \in R$.*

Theorem 5.3.3 (Direct Sum)**.** *Suppose (V, R) and (W, S) are root systems. Consider the vector space $V \oplus W$, with the natural inner product. Then $R \cup S$ is a root system in $V \oplus W$, called the direct sum of V and W. Here we assume the natural identification.*

Definition 5.3.4 (Irreduciblility)**.** A root system (V, R) is called reducible if there exists an orthogonal decomposition $V = V_1 \oplus V_2$ with $\dim V_1 > 0$ and $\dim V_2 > 0$ s.t. every element of R is either in V_1 or in V_2. If no such decomposition exists, (V, R) is called irreducible.

Definition 5.3.5 (Equivalence)**.** Two root systems (V_1, R_1) and (V_2, R_2) are said to be equivalent if $\exists T : V_1 \to V_2$ which is an invertible linear transformation s.t. T maps R_1 onto R_2 and s.t. $\forall \alpha \in R_1, \beta \in V_1$, we have

$$T(w_\alpha \cdot \beta) = w_{T\alpha} \cdot T\beta.$$

A map T with this property is called an equivalence.

What are the root systems?

Theorem 5.3.6. *Suppose α, β are noncollinear roots and $\langle \alpha \rangle \geq \langle \beta \rangle$. Then one of the following holds:*
(i) $\langle \alpha, \beta \rangle = 0$
(ii) $\langle \alpha, \alpha \rangle = \langle \beta, \beta \rangle$ and the angle between α and β is $\frac{\pi}{3}$ or $\frac{2\pi}{3}$
(iii) $\langle \alpha, \alpha \rangle = 2\langle \beta, \beta \rangle$ and the angle between α and β is $\frac{\pi}{4}$ or $\frac{3\pi}{4}$
(iv) $\langle \alpha, \alpha \rangle = 3\langle \beta, \beta \rangle$ and the angle between α and β is $\frac{\pi}{6}$ or $\frac{5\pi}{6}$

Proof. Let θ be the angle between α and β.

$$4\frac{\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle}\frac{\langle\beta,\alpha\rangle}{\langle\beta,\beta\rangle} = 4\cos^2\theta \in \mathbb{Z}$$

□

Corollary 5.3.7. *Suppose α, β are roots.*
(i) If the angle between α and β is strictly obtuse, then $\alpha+\beta$ is a root.
(ii) If the angle between α and β is strictly acute, then $\alpha-\beta$ is a root.

Proof. Discuss case by case. □

Now we discuss the dual of a root system.

Definition 5.3.8 (Co-Root)**.** If (V, R) is a root system, then for each root $\alpha \in R$, define the co-root H_α by

$$H_\alpha = 2\frac{\alpha}{\langle\alpha,\alpha\rangle}$$

The set of all co-roots is denoted $R^\vee$ and is called the dual root system to R.

This is actually an inversion with respect to the ball centered at the origin with radius $\sqrt{2}$.

Theorem 5.3.9 (Duality)**.** $(R^\vee)^\vee = R$

Proof. The dual system can be regarded as an inversion with respect to the ball centered at the origin with radius $\sqrt{2}$, and inversion is a kind of involution. □

Theorem 5.3.10 (Dual Root System)**.** *$R^\vee$ is a root system and its Weyl group is the same as the that of R.*

Next we construct the base if a root system.

Definition 5.3.11 (Base)**.** A subset Δ of R is called a base for R if the following conditions hold:
(i) Δ is a basis for V as a vector space.
(ii) Each root can be expressed as a linear combination of elements of Δ with integer coefficients and in such a way that the coefficients are either all non-negative or all nonpositive.

The roots for which the coefficients are non-negative are called positive roots and the others are called negative roots. The set of positive roots relative to a fixed base Δ is denoted R^+. The elements of Δ are called simple positive roots.

Lemma 5.3.12. *If α, β are distinct elements of a base, then $\langle\alpha,\beta\rangle \le 0$.*

Proof. Use Corollary 5.3.7. □

Lemma 5.3.13. *If V is a finite-dimensional real vector space and R is a finite subset of V not containing 0, then there exists a hyperplane M that does not contain any element of R.*

Proof. The union of finite collection of hyperplanes can not be all of V. □

Definition 5.3.14 (Indecomposability)**.**

Theorem 5.3.15. *Suppose (R, V) is a root system, W is a hyperplane not containing any element of R, and R^+ is the set of roots lying on the fixed side of V. Then the set of indecomposable elements of R^+ is a base for R.*

Theorem 5.3.16. *Given any base Δ for R, there exists a hyperplane V s.t. Δ arises as in Theorem 5.3.15.*

Theorem 5.3.17 (Base for Dual Root System)**.** *If Δ is a base for R, then the set of all co-roots H_α, $\alpha \in \Delta$, is a base for the dual root system $R^\vee$.*

Finally we explore the integral elements.

Definition 5.3.18 (Integral Element)**.** An element $v \in V$ is called an integral element if for all $\alpha \in R$, the quantity

$$2\frac{\langle\mu,\alpha\rangle}{\langle\alpha,\alpha\rangle}$$

is an integer.

Definition 5.3.19. If Δ is a base for R, then an integral element μ is called dominant integral if

$$2\frac{\langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle} \geq 0, \quad \forall \alpha \in \Delta$$

It is called strictly dominant if the inequality is strict.

Remark 5.3.20. An integral element is dominant if and only if it is contained in closed fundamental Weyl chamber, and strictly dominant if and only if contained in the open fundamental Weyl chamber.

Lemma 5.3.21. *If $v \in V$ has the property that*

$$2\frac{\langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle}$$

is an integer for all $\alpha \in \Delta$, then v us an integral element.

Proof. Use Theorem 5.3.17. □

Definition 5.3.22 (Fundamental Weights)**.**

Definition 5.3.23 (Higher and Lower)**.**

5.3.2 Example: Rank 2

5.3.3 Example: Rank 3

5.3.4 Additional Properties

5.3.5 Application: Classical Lie Algebras

5.3.6 Dynkin Diagrams

The classification of root systems is given in terms of an object called the Dynkin diagram.

Suppose $\Delta = \{\alpha_1, \cdots, \alpha_r\}$ is a base for a root system R. Then the Dynkin diagram for R is a graph having vertices $v_1, \cdots, v_r$. Between any two vertices, we place either no edge, one edge, two edge, or three edge corresponding to the four cases in Theorem 5.3.6.

Theorem 5.3.24. *Every irreducible root system is isomorphic to precisely one root system from the following list:*
(i) A_n, $n \geq 1$
(ii) B_n, $n \geq 2$
(iii) C_n, $n \geq 3$
(iv) D_n, $n \geq 4$
(v) G_2, F_4, E_6, E_7, E_8

5.3.7 The Root Lattice

5.4 Lie Algebra: an Algebraic Viewpoint

5.4.1 Basic Definition and Examples

Definition 5.4.1. Let $\mathfrak{g}$ be a vector space on $\mathbb{F}$, call $\mathfrak{g}$ a Lie algebra on $\mathbb{F}$ if there is a bineary operation [,] satisfying:

$$(1)[x, y] = -[y, x]$$
$$(2)[k_1 x_1 + k_2 x_2, y] = k_2[x_1, y] + k_2[x_2, y]$$
$$(3)[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$$

Remark 5.4.2. If $[x, y] = 0 \forall x, y \in \mathfrak{g}$, then we call $\mathfrak{g}$ a trivial Lie algebra

Example 5.4.3. If $\dim \mathfrak{g} = 1$, then $\mathfrak{g}$ is trivial.

5.4.2 Lie Theorem

5.4.3 Cartan's Criterion

Lemma 5.4.4.

(i) If s is diagonalizable, then ad_s is diagonalizable

(ii) If n is nilpotent, the ad_n is nilpotent

(iii) If $z = s + n$ is the Jordan-Chevalley decomposition of z, then $ad_z = ad_s + ad_n$ is the Jordan-Chevalley decomposition of ad_z

Proof. (i)(iii) can be proved by direct computation. For (ii), ad_n acts on x by adding n on either side of x. Acting enough times can make it vanish. □

5.4.4 Killing form

We begin with a lemma.

Lemma 5.4.5. *Let $B(X, Y)$ be a symmetric nondegenerate bilinear form on $\mathbb{F}^{n\times n}$ with the associative property*

$$B(XY, Z) = B(X, YZ)$$

Then $\exists c \neq 0 \in \mathbb{F}$ s.t. $B = c\,\mathrm{tr}(XY)$

This lemma can be easily proved by direct computation, but here we make use of the general theory of bilinear form to exihibit an elegant proof.

Proof. Note that $\mathrm{tr}(XY)$ is another nondegenerate bilinear form on $\mathbb{F}^{n\times n}$, so there exists an isomorphism φ of $\mathbb{F}^{n\times n}$, such that

$$B(X, Y) = \mathrm{tr}(\varphi(X)Y)$$

By the property of $B(X, Y)$,

$$\mathrm{tr}(\varphi(X)Y) = \mathrm{tr}(\varphi(XY)) = \mathrm{tr}(X\varphi(Y))$$

Thus $\varphi(X)Y = \varphi(XY) = X\varphi(Y)$, so φ is multiplication by a constant. □

Remark 5.4.6. An ideal in $\mathbb{F}^{n\times n}$ is either $\{0\}$ or $\mathbb{F}^{n\times n}$ itself, that is, $\mathbb{F}^{n\times n}$ is a simple algebra.

Now we can define the killing form on a finite-dimensional Lie algebra

5.5 Discussion

Section root system follows Hall [2004].

Chapter 6

Matrix Analysis

In this chapter, we examine some analysis properties of matrix. We consider matrices on field $\mathbb{C}$.

6.1 Positive Matrices

Theorem 6.1.1. *There are several characterization of positive matrices.*

6.2 Majorization

Definition 6.2.1. Let $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$. Define $x^{\uparrow}$ and $x^{\downarrow}$ be the vectors obtained by rearranging the coordinates of x in the increasing and the decreasing orders respectively.

Remark 6.2.2. $x_j^{\uparrow} = x_{n+1-j}^{\downarrow}$, $1 \leq j \leq n$

Definition 6.2.3. We say that x is majorized by y, in symbols $x \prec y$, if the following conditions are satisfied:

(i) $\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow}$, $\forall 1 \leq k \leq n$

(ii) $\sum_{j=1}^{n} x_j^{\downarrow} = \sum_{j=1}^{n} y_j^{\downarrow}$

Remark 6.2.4. The two conditions (i) and (ii) can be repalced by

(i′) $\sum_{j=1}^{k} x_j^{\uparrow} \geq \sum_{j=1}^{k} y_j^{\uparrow}$, $\forall 1 \leq k \leq n$

(ii′) $\sum_{j=1}^{n} x_j^{\uparrow} = \sum_{j=1}^{n} y_j^{\uparrow}$

Definition 6.2.5.
We say that x is weakly submajorized by y, in symbols $x \prec_w y$ if condition (i) is fulfilled.
We say that x is weakly supermajorized by y, in symbols $x \prec^w y$ if condition (i′) is fulfilled.

6.3 Matrix Equations

6.4 Discussion

Bhatia [1997]

Chapter 7

Operators on Hilbert Space

Often, topics of Hilbert spaces is not contained in a book of linear algebra. But Hilbert Spaces is so much like the finite-dimensional Euclidean spaces, it is natural to begin our journey of functional analysis after the study of inner product spaces. Many of the definitions and theorems are modification of previously established results.

7.1 Hilbert Spaces

7.1.1 Examples

Example 7.1.1. Let I be any set and $l^2(I)$ denotethe set of all functions

Example 7.1.2. Let (X, Ω, μ) be a measure space consisting of a set X, a σ-algebra Ω of subsets of X, and a countably additive measure μ defined on Ω with values in the non-negative extended real numbers

Recall that an absolutely continuous function on the unit interval $[0, 1]$ has a derivative a.e. on $[0, 1]$.

Example 7.1.3. Let $\mathcal{H}$ = the collection of all absolutely continous functions $f : [0, 1] \to \mathbb{F}$ s.t. $f(0) = 0$ and $f' \in L^2(0, 1)$. If $\langle f, g\rangle = \int_0^1 f'(t)g'(t)\mathrm{d}t$ for f and g in $\mathcal{H}$, then $\mathcal{H}$ is a Hilbert space.

7.2 Operators

Chapter 8

Miscellaneous Topics

In this chapter, we discuss several advanced topics making use of linear algebra.

8.1 Gelfand-Tsetlin Integrable System

We begin by listing some facts which are useful to the following discussion.

Theorem 8.1.1 (Cauchy's Interlace Theroem)**.** *Let A be a Hermitian matrix of order n, and let B be a principal submatrix of A of order $n-1$. If $\lambda_n \leq \lambda_{n-1} \leq \cdots \leq \lambda_2 \leq \lambda_1$ lists the eigenvalues of A and $\mu_n \leq \mu_{n-1} \leq \ \leq \mu_3 \leq \mu_2$ the eigenvalues of B, then $\lambda_n \leq \mu_n \leq \lambda_{n-1} \leq \mu_{n-1} \leq \cdots \leq \lambda_2 \leq \mu_2 \leq \lambda_1$*

Proof 1. This is an immediate consequence of the Courant-Fischer minimax theorem. □

Proof 2. Compute the characteristic polynomial of the matrix in two different ways, one diagonal, the other submatrix diagonal and compare the coefficients.

$$A \sim \begin{pmatrix} \mu_2 & & & & a_2 \\ & \mu_3 & & & a_3 \\ & & \ddots & & \vdots \\ & & & \mu_n & a_n \\ \bar{a}_2 & \bar{a}_3 & \cdots & \bar{a}_n & d \end{pmatrix} \sim \begin{pmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \ddots & & \\ & & & \lambda_{n-1} & \\ & & & & \lambda_n \end{pmatrix}$$

$$\prod_{i=1}^{n}(\lambda - \lambda_i) = (\lambda - d - \sum_{i=2}^{n} \frac{|a_i|^2}{\lambda - \mu_i}) \prod_{i=2}^{n}(\lambda - \mu_i)$$

$$|a_k|^2 = \frac{\prod_{i=1}^{n}(\mu_k - \lambda_i)}{\prod_{i\neq k, i=2}^{n}(\mu_k - \mu_i)}, k = 2, \cdots, n$$

□

Definition 8.1.2 (Gelfand-Tsetlin Integrable System)**.** The phase space $M = \text{Hermite}(n \times n)$. Let $A = (a_{ij})$ where a_{ij} be the coordinate function of (i, j). Define a Poisson bracket on M by $\{a_{ij}, a_{kl}\} = \delta_{jk}a_{il} - \delta_{li}a_{kj}$. This is Gelfand-Tsetlin integrable system.

Lemma 8.1.3. *$Tr(A^k)$ is a Casimir function.*

Proof. We want to show: $\left\{\text{Tr}(A^k), a_{ij}\right\} = \mathcal{L}_{X_{a_{ij}}} \text{Tr}(A^k) = 0$. Noticing that $\text{Tr}(A^k)$ is a symmetric function of eigenvalue functions, we want to show $X_{a_{ij}}$ is a tangent vector of a similarity transformation.

$$X_{a_{ij}}(a_{kl}) = \{a_{ij}, a_{kl}\} = \delta_{jk}a_{il} - \delta_{li}a_{kj}$$

$$X_{a_{ij}} = (\delta_{jk}a_{il} - \delta_{li}a_{kj})\frac{\partial}{\partial a_{kl}}$$

$$\mathcal{L}_{X_{a_{ij}}}(A) = [E_{ij}, A]$$

$$g^t_{X_{a_{ij}}}(A) = e^{tE_{ij}}Ae^{-tE_{ij}}$$

□

Theorem 8.1.4. *Denote $A^{(k)}$ the left-top $k\times k$ submatrix of $A \in$ Hermite$(n\times n)$. Then $\left\{Tr(A^{(k)^j}, Tr(A^{(s)^t})\right\}$=0.*

Proof. We do Thimm's trick.

$$M_{1\times 1} \subset M_{2\times 2} \subset \cdots \subset M_{n\times n}$$
$$C^\infty(M_{n\times n}) \longrightarrow C^\infty(M_{(n-1)\times(n-1)}) \longrightarrow \cdots \longrightarrow C^\infty(M_{1\times 1}) \text{ by restriction}$$

□

8.2 Unitary Matrices

Theorem 8.2.1. *Suppose $U \in U(n)$ is unitary and symmetric, then $U = P^{-1}\Lambda P$, where $P \in O(n)$ and Λ is unitary diagonal.*

Proof. The key is to separate the real and imaginary part of U, i.e. $U = A + iB$ with $A, B \in \mathbb{R}^{n\times n}$.

Notice that A and B are symmetric, $U^*U = \bar{U}U = (A - iB)(A + iB)$, so that $A^2 + B^2 = I$ and $AB = BA$. Therefore, $\exists P \in O(n)$, s.t. $A = P^{-1}\Lambda_1 P, B = P^{-1}\Lambda_2 P$ and Λ_1, Λ_2 are real diagonal. Thus $U = P^{-1}(\Lambda_1 + i\Lambda_2)P := P^{-1}\Lambda P$, where Λ is diagonal. Λ is unitary because U is unitary. □

Corollary 8.2.2. *Suppose $U \in U(n)$ is unitary and symmetric, then $U = B^tB$, where $B \in U(n)$.*

Theorem 8.2.3. *Any unitary matrix $U \in U(n)$ can be written as $U = P_1\Lambda P_2$, where $P_1, P_2 \in O(n)$ and Λ is a diagonal matrix.*

Proof. Notice that U^tU is a symmetric unitary matrix, therefore $U^tU = P^tDP$ where $P \in O(n)$ and D is diagonal. Let Λ be a square root of D, that is, $\Lambda^2 = D$, then Λ is also a unitary diagonal matrix, and $U^tU = P^t\Lambda^t\Lambda P$. Let $P_1 = UP^{-1}\Lambda^{-1}, P_2 = P$, then $P_1^tP_1 = I$. As U, D, Λ are unitary, $P_1^*P_1 = I$, so $P_1^t = P_1^*$. Therefore $P_1 \in O(n)$. So $U = P_1\Lambda P_2$ satisfies the required conditions. □

Corollary 8.2.4 (QS Decomposition)**.** *$\forall U \in U(n)$, $\exists Q \in O(n), S \in U(n)$ s.t. $U = QS$ and $S = f(U^tU)$, for some $f \in \mathbb{C}[x]$*

Proof. By the previous theorem, $U = P_1\Lambda P_2 = P_1P_2P_2^{-1}\Lambda P_2$. We know $\Lambda = f(D)$, so $P_2^{-1}\Lambda P_2 = f(P_2^{-1}DP_2) = f(U^tU)$. □

8.3 Matrix Decomposition

In this section, we collect some results of matrix decomposition.

8.3.1 Positive Definite Matrices

Throughout this subsection, $F = \mathbb{R}$ or $F = \mathbb{C}$.

Lemma 8.3.1. *If $P, Q \in \mathbb{C}^{n\times m}$, then*

$$P^*P = Q^*Q \iff \exists U \in U(n), \text{ s.t. } Q = UP$$

Proof. Denote $A = P^*P = Q^*Q$, then $\ker A = \ker P = \ker Q$. Consider an inner product on $\mathbb{C}^m/\ker A$ defined by $\langle X, Y\rangle_A = Y^*AX$, and an inner product on $\mathbb{C}^n$ defined by $\langle X, Y\rangle_0 = Y^*X$. Then $P : \mathbb{C}^m/\ker A \to \operatorname{Im} P$, $Q : \mathbb{C}^m/\ker A \to \operatorname{Im} Q$ are both isometries. So there exists a isometry $U : \operatorname{Im} P \to \operatorname{Im} Q$. Extending U to an isometry of $\mathbb{C}^n$ gives the result. □

Theorem 8.3.2 (Cholesky Decomposition)**.** *If A is positive definite, then there exist an invertible matrix P s.t. $A = P^*P$. Futhermore, if we impose the following restrictions:*
(1) P is upper diagonal
(2) the diagonal elements of P are positive real numbers
then P is unique.

Example 8.3.3. For $A \in \mathbb{R}^n$, TFAE:
(1) A is a product of two positive definite matrices
(2) A is diagonalizable and all the eigenvalues of A are positive

Proof. If $A = S_1S_2$, by Cholesky decomposition, $S_2 = P^tP$, so $PAP^{-1} = PS_1P^t$ is still positive definite.

The other direction is evident. □

Remark 8.3.4. For semi-positive definite matrices, a similar result holds, but the matrix P may not be invertible, and may not be unique.

8.3.2

Theorem 8.3.5 (LU Decomposition)**.** *For $A \in GL_n(F)$, TFAE: (1) the ordered principal minors of A are nonzero (2) $\exists L, U \in GL_n(F)$, L is lower trianglular, U is upper trianglular, and $A = LU$*
Futhermore, if all the diagonal elements of L(or U) are 1*, then the decomposition is unique.*

Theorem 8.3.6 (Bruhat Decomposition)**.**

Bibliography

Rajendra Bhatia. *Matrix analysis*. Graduate texts in mathematics. Springer, New York, 1997. ISBN 0387948465. 37

Brian C Hall. *Lie groups, Lie algebras, and representations: an elementary introduction*, volume 222 of *Graduate Texts in Mathematics*. Springer, New York, NY, 2004. ISBN 0387401229. 35

Steven Roman. *Advanced linear algebra*, volume 135 of *Graduate texts in mathematics*. Springer, 3rd ed. edition, 2013. ISBN 0387978372. 28