Linear Algebra A (II) Class Notes

timetraveler314
University of Genshin
timetraveler314@outlook.com

Contents

1. Vector Spaces

1.1. Operations on Subspaces

Theorem 1.1.1

Let $V_1, ..., V_n$ be proper subspaces of a vector space V over an infinite field $\mathbb{F}$. Then the intersection

$$V_1 \cup ... \cup V_n \neq V.$$

Proof: (i) First consider the case where $n = 2$. Choose $v_1 \in V_2 \setminus V_1, v_2 \in V_1 \setminus V_2$. We show that $v = v_1 + v_2 \notin V$.

Suppose not. If $v \in V_1$, then $v_1 = v - v_2$ in V_1, a contradiction. Similarly, if $v \in V_2$, then $v_2 = v - v_1$ in V_2, a contradiction. Thus $v \notin V$.

(ii) (Vandermonde's Determinant) For general $n \in \mathbb{N}$:

Let $v_i \in V \setminus V_i, i = 1, ..., n$. We manage to find a linear combination of them that is not covered by the union of V_i.

Consider $u_a = \sum_{i=1}^{m} a^{i-1} v_i, a \in \mathbb{F}$. Here, $\operatorname{char} \mathbb{F} = 0$. So there exists V_k s.t. $u_a \in V_k$ for infinitely many a.

WLOG, assume $u_{a_1}, \ldots, u_{a_m} \in V_k$, where a_i are distinct. It suffices to show that $v_k \in \operatorname{span}\left\{u_{a_1}, \ldots, u_{a_m}\right\} \subset V_k$, which leads to a contradiction.

To obtain v_k, consider linear combinations of u_{a_i}. We have a Vandermonde matrix:

$$\left(u_{a_1}, \ldots, u_{a_m}\right)\begin{pmatrix} x_1 \\ \cdots \\ x_m \end{pmatrix} = v_k$$

$$\Leftrightarrow (v_1, \ldots, v_m)\underbrace{\begin{pmatrix} 1 & 1 & \ldots & 1 \\ a_1 & a_2 & \ldots & a_m \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{m-1} & a_2^{m-1} & \ldots & a_m^{m-1} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}} = v_k$$

Letting the underbraced part be $\begin{pmatrix} 0 \\ \vdots \\ 1(k\text{−th}) \\ \vdots \\ 0 \end{pmatrix}$, we have a linear combination of u_{a_i} that equals v_k. This is guaranteed by the Vandermonde determinant being nonzero, which gives us a unique solution for $\begin{pmatrix} x_1 \\ \cdots \\ x_m \end{pmatrix}$. □

1.2. Sum of Subspaces

Note 1.2.1

The sum $\underset{i \in \mathcal{I}}{+} V_i$ of subspaces V_i is the smallest subspace containing all V_i. This is the core idea of the sum.

After introducing the operation $+$, we want to check whether the new operation behaves well with respect to the old operations $\cup$ and $\cap$. Here, distributivity does not hold in general, but we have the following result.

Theorem 1.2.1

Let X, Y, Z be subspaces.

(i) $(X \cap Z) + (Y \cap Z) \subset (X + Y) \cap Z$,

(ii) $(X \cap Y) + Z \subset (X + Z) \cap (Y + Z)$.

Proof: (i)

$$X \subset X + Y \Rightarrow (X + Y) \cap Z \subset X \cap Z.$$

Similarly,

$$Y \cap Z \subset (X + Y) \cap Z.$$

(ii)

$$X \cap Y \subset X + Z, Y \cap Z \subset Y + Z.$$

□

2. Linear Mappings

Note 2.1 Linear Mappings and Matrices

Given $T : V \to W$, choose bases $\{e_1, ..., e_n\}$ and $\{\eta_1, ..., \eta_m\}$ for V and W respectively.

We say A is a matrix representation of T if

$$T(e_1, ..., e_n) = (\eta_1, ..., \eta_n)A.$$

Specially, let T be a linear mapping from V to itself. Then A is a matrix representation of T with respect to the same basis.

When basis changes: $\{e_1, ..., e_n\} \to \{f_1, ..., f_n\}$, the transition matrix P is defined by

$$(f) = (e)P.$$

Then the matrix representation of T with respect to the new basis is given by $B = P^{-1}AP$:

$$T(f)P^{-1} = (f)P^{-1}A \Rightarrow T(f) = P^{-1}AP.$$

Generally, $B = Q^{-1}AP$.

2.1. Kernel and Image

Question 2.1.1

Let S, T be subspaces of V. Show that

$$(S + T)/S \simeq T/(S \cap T).$$

Proof:

Note 2.1.1

Upon seeing isomorphism of quotient spaces, we should consider the first isomorphism theorem.

Consider the mapping

$$\varphi : S + T \to T/(S \cap T), s + t \mapsto t + (S \cap T).$$

Its linearity is clear. Now for any $v = s + t \in \ker \varphi$,

$$t + (S \cap T) = 0 + (S \cap T) \Leftrightarrow t \in S \cap T \Leftrightarrow t \in S \Leftrightarrow v \in S.$$

Thus $\ker \varphi = S$. By the first isomorphism theorem,

$$(S + T)/S \simeq T/(S \cap T).$$

□

Corollary 2.1.1

$$\varphi(U) \simeq U/(U \cap \ker \varphi).$$

2.2. Minimal Polynomial and Characteristic Polynomial

Theorem 2.2.1 Primary Decomposition

Consider factorization over $\mathbb{F}$.

Let $m_\varphi = p_1^{r_1} ... p_k^{r_k}$ be the factorization of the minimal polynomial of φ. Then

$$V = \bigoplus \ker p_i^{r_i}.$$

Let $f_\varphi = p_1^{s_1} ... p_k^{s_k}$ be the factorization of the characteristic polynomial of φ. Then

$$V = \bigoplus \ker p_i^{s_i}.$$

They are both the primary decomposition of V.

Denote $K(f)$ by the kernel of $f(\varphi)$.

Theorem 2.2.2 Projection Map as Polynomials

Let $\varphi : V \to V$ be a linear mapping. Then the projection map $\pi_i : V \to \ker p_i^{r_i}$ is a polynomial of φ.

Proof: The case when $k = 1$ is trivial. We prove the general case where $k > 2$.

Let $g_j = p_1^{r_1} ... \hat{p}_j^{r_j} ... p_k^{r_k}$. Then

$$K(g_j) = \ker g_j(\varphi) = \bigcup_{t \neq j} \ker p_t^{r_t}.$$

We see that all g_j are coprime. By Bezout's identity, there exist h_j s.t.

$$h_1 g_1 + ... + h_k g_k = 1.$$

Let $x = x_1 + ... + x_k, x_j \in \ker p_j^{r_j}$. Then

$$\begin{aligned} \sum_j x_j = x &= \sum_j h_j(\varphi) g_j(\varphi)(x) \\ &= \sum_{j,t} h_j(\varphi) g_j(\varphi)(x_t) \\ &= \sum_j h_j(\varphi) g_j(\varphi) x_j \left(\text{Because } K(g_j) = \ker g_j(\varphi) = \bigcup_{t \neq j} \ker p_t^{r_t} \right). \end{aligned}$$

Notice the primary decomposition is a direct sum. Then

$$x_j = (h_j \cdot g_j)(\varphi) x_j, j = 1, ..., k$$

This shows that $\pi_i = h_i(\varphi) g_i(\varphi)$. $\square$

Example 2.2.1

Let $\varphi \in \mathcal{L}(V)$ be diagonalizable. If W is a φ-invariant subspace, then $\varphi|_W$ is also diagonalizable.

Proof (1):
1. Expand a basis $w_1, ..., w_m$ of W to a basis $w_1, ..., w_n$ of V.

Then

$$\varphi(w_1, ..., w_m) = (w_1, ..., w_n)\begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix},$$

where A_1 is the matrix representation of $\varphi|_W$.

Note that

$$\begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix}^k = \begin{pmatrix} A_1^k & * \\ 0 & A_3^k \end{pmatrix},$$

This shows that for any polynomial f, $f(A) = \begin{pmatrix} f(A_1) & * \\ 0 & f(A_3) \end{pmatrix}$.

- φ is diagonalizable, so m_A is a product of distinct linear factors.

$m_A(A) = 0 \Rightarrow m_A(A_1) = 0$, so $m_{A_1} \mid m_A$, which shows that m_{A_1} is a product of distinct linear factors $\Rightarrow$ diagonalizable. □

Proof (2): By factorizing the space.

$$V = \bigoplus V_{\lambda_i}(\varphi)$$

In fact, $W = \bigoplus\left(W \cap V_{\lambda_{i(\varphi)}}\right)$. For all $w \in W$, $w = w_1 + ... + w_i, w_k \in V_{\lambda_k}$. We prove that $w_i \in W$.

Consider the projection $p_i : V \to V_{\lambda_i}$, p_i **is a polynomial of** φ.

$$w_i = p_{i(w)} = g_{i(\varphi)}(w) \in W (W \text{ is } \varphi\text{-invariant}).$$

The decomposition $W = \bigoplus\left(W \cap V_{\lambda_{i(\varphi)}}\right)$ gives □

Example 2.2.2

Let A, B be diagonalizable, $AB = BA$. Show that A, B is simultaneously diagonalizable.

Proof: Let V_{λ_i} be the eigenspace of A corresponding to λ_i. Then V_{λ_i} is B-invariant.

B is diagonalizable on V_{λ_i}, so V_{λ_i} is the direct sum of eigenspaces of B.

$V = \bigoplus V_{\lambda_i}$ is the direct sum of eigenspaces of B. This shows that A, B is simultaneously diagonalizable. □

Actually, the primary decomposition is not strong enough. When we expand the field to $\mathbb{C}$, we may get more information.

Note 2.2.1 Space Decompositions

1. Primary Decomposition $V = \bigoplus V_{\lambda_i}(\varphi)$
2. Cyclic Subspace Decomposition
3. Irreducible φ-invariant Decomposition

2.3. Cyclic Spaces

Consider the Frobenius matrix

$$A = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}$$

Then $A\boldsymbol{e}_1 = \boldsymbol{e}_2, A\boldsymbol{e}_2 = \boldsymbol{e}_3, \dots, A\boldsymbol{e}_{n-1} = -a_{n-1}\boldsymbol{e}_n, A\boldsymbol{e}_n = -a_0\boldsymbol{e}_1 - a_1\boldsymbol{e}_2 - \dots - a_{n-1}\boldsymbol{e}_{n-1}$.

We see that a basis can be $\boldsymbol{e}_1, A\boldsymbol{e}_1, A^2\boldsymbol{e}_1, \dots, A^{n-1}\boldsymbol{e}_1$. This is such a good property that we want to generalize it.

Definition 2.3.1 Cyclic Spaces

Let $\varphi : V \to V, \xi \in V$. Let

$$W = \langle \xi, \varphi(\xi), \varphi^2(\xi), \dots, \varphi^{n-1}(\xi) \rangle.$$

W is a φ-invariant subspace of V generated by ξ. We call W a φ-cyclic subspace of V generated by ξ.

What is the minimal polynomial of $\varphi|_W$? How do we construct a basis for W?

While trying to construct the maximum number of linearly independent vectors, we may encounter the following problem:

$$a_1\xi + a_2\varphi(\xi) + \dots + a_{n-1}\varphi^{n-2}(\xi) = a_n\varphi^{n-1}(\xi).$$

This is equivalent to finding $f(x) \in \mathbb{F}[x], f(\varphi) = 0$. This reduces to finding an annihilating polynomial of ξ s.t. $f(\varphi)\xi = 0$.

Of all the polynomials, the monic polynomial of the smallest degree is unique. Thus we can define the minimal annihilating polynomial of ξ.

Lemma 2.3.1

Let $\xi \neq 0$. If the minimal annihilating polynomial of ξ w.r.t. φ is $f(x)$, Then

$$\xi, \varphi(\xi), \dots, \varphi^{\deg f - 1}(\xi)$$

is a basis of W.

A corollary: $\dim W = \deg m_\xi(x)$.

Proof: Assume there exists some $b_i \neq 0$ s.t.

$$b_0\xi + b_1\varphi(\xi) + \ldots + b_{\deg f-1}\varphi^{\deg f-1}(\xi) = 0.$$

Then $b_0 + b_1 x + \ldots + b_{\deg f-1}x^{\deg f-1} \neq 0$ is an annihilating polynomial of ξ with degree less than $\deg f$, a contradiction. □

We call m_ξ, m_φ both "annihilating". What is the relationship between them?

Theorem 2.3.1 Annihilating Polynomials

$$m_\xi \mid m_\varphi.$$

Can they be the same?

Lemma 2.3.2 Steps towards Cyclic Decomposition (Hard)

(1) $\exists \xi \in V$ s.t. $m_\xi = m_\varphi$. (An unique ξ that cyclically generates V)

(2) $\exists U$ s.t. U is φ −invariant, and

$$V = \langle \xi, \ldots, \varphi^{s-1}(\xi)\rangle \oplus U.$$

Note 2.3.1

(2) tells us that the complement of the cyclic subspace is still φ −invariant. Now go on with U till $U = \{0\}$, then we get the cyclic decomposition.

After this, we get a basis of V with nicer properties compared to the primary decomposition.

By virtue of dual spaces, we may prove the result in a more general setting.

Proof: (1) Let $m_\varphi = p_1^{r_1} \ldots p_s^{r_s}$.

i. $s = 1$. Then $m_\varphi = p^s, p$ is irreducible. By the minimality of s, $\exists \xi, p^{s-1}(\varphi)(\xi) \neq 0$.

By $m_\xi \mid m_\varphi$, $m_\xi = p^k, k \leq s$. If $k < s$, then $p^k(\varphi)(\xi) = 0$, a contradiction. Thus $k = s$.

ii. $s \geq 2$. By primary decomposition,

$$V = \ker p_1^{r_1}(\varphi) \oplus \ker p_2^{r_2}(\varphi) \oplus \ldots \oplus \ker p_s^{r_s}(\varphi).$$

Consider the restriction map $\varphi|V_i$**.** The minimal polynomial of $\varphi|V_i$ is $p_i^{r_i}$. By i.,

$$\exists \xi_i \in V_i, m_{\xi_i} = p_I^{r_i}.$$

Construct $\xi = \sum \xi_i$**.** We show that $m_\xi = m_\varphi$.

By definition, $m_\xi(\varphi)(\sum \xi_i) = \sum m_\xi(\varphi)(\xi_i) = 0$. Each term $m_\xi(\varphi)(\xi_i) \in V_i$. Utilize the **direct sum** property, $m_\xi(\varphi)(\xi_i) = 0$. Thus $m_{\xi_i} \mid m_\xi$. Because p_i are coprime, $m_\varphi \mid m_\xi \Rightarrow m_\varphi = m_\xi$.

(2) (GTM 23, last chapter). □

Apply the lemma repeatedly, we get the cyclic decomposition.

How to prove that a space is cyclic?

Theorem 2.3.2

W is cyclic iff $\dim W = \deg m_{\varphi|_W}$.

2.4. Irreducibility

Definition 2.4.1 Irrreducible

V is called φ-irreducible if $V \neq V_1 \oplus V_2$, where V_1, V_2 are φ-invariant subspaces of V.

Theorem 2.4.1

V is φ-irreducible $\Leftrightarrow$ V is φ-cyclic, $m_\varphi = f^k, k \geq 1$ (f is irreducible).

Proof: $\Rightarrow$: Let $m_\varphi = f_1^{k_1} ... f_s^{k_s}$ be the irreducible factorization of m_φ.

Then $V = \bigoplus \ker f_i^{k_i}$. Because $\ker f_i^{k_i}$ is φ-invariant, V is not φ-irreducible if $s \geq 2$. Thus $s = 1$.

By the cyclic lemma, $\exists \xi \in V$ s.t. $m_\xi = m_\varphi = f^k$. Then $W = \langle \xi, \varphi(\xi), ..., \varphi^{k-1}(\xi) \rangle$ is φ-cyclic.

Moreover, by irreducibility, $W = V$.

$\Leftarrow$: Proof by contradiction. Let $V = V_1 \oplus V_2$, V_1, V_2 are φ-invariant. Then $m_\varphi = \operatorname{lcm}\left(m_{\varphi|_{V_1}}, m_{\varphi|_{V_2}}\right)$. WLOG, we have $m_{\varphi|_{V_1}} = f^k = m_\varphi$. By analyzing the dimension, we get a contradiction. □

Hence we are able to determine whether a space is irreducible while decomposing it. Then any V must have a decomposition into irreducible cyclic subspaces.

Diving into the irreducible subspaces, we can give a concrete characterization.

For each irreducible space, let e be the cyclic element,

$$m_\varphi = f^k$$
$$f(\varphi) = \varphi^p + a_{p-1}\varphi^{p-1} + ... + a_1\varphi + a_0 = 0.$$

Utilizing the relation, we can construct a basis of V.

$$\begin{array}{rcr} e, \varphi_e, & ..., & \varphi^{p-1}e, \\ f(\varphi)_e, f(\varphi)\varphi(e), & ..., & f(\varphi)\varphi^{p-1}(e), \\ f^2(\varphi)_e, f^2(\varphi)\varphi(e), & ..., & f^2(\varphi)\varphi^{p-1}(e), \\ & ... & \\ f^{k-1}(\varphi)_e, f^{k-1}(\varphi)\varphi(e), & ..., & f^{k-1}(\varphi)\varphi^{p-1}(e). \end{array}$$

The kp elements together form a basis. Let's prove it.

Proof: M denotes the space spanned by the above kp elements.

It suffices to show $\varphi(M) \subseteq M$ because of irreducibility. (If there is more than one invariant subspace, M won't be irreducible)

By direct calclation, ... □

The matrix of φ under this basis is a block matrix

$$\begin{pmatrix} F & & \\ A & F & \\ & A & F \end{pmatrix}, F = \begin{pmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ 0 & 0 & \ldots & 1 & -a_{p-1} \end{pmatrix}, A = \begin{pmatrix} 0 & 0 & \ldots & 1 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}$$

Example 2.4.1 Jordan Normal Form

When $\mathbb{F} = \mathbb{C}$, every polynomial can be completely factored, giving fine enough matrix as Jordan blocks. $m_\varphi = (x-a)^k \Rightarrow \begin{pmatrix} a & & & & \\ 1 & a & & & \\ & 1 & a & & \\ & & \ddots & \ddots & \\ & & & 1 & a \end{pmatrix}$

Example 2.4.2

When $\mathbb{F} = \mathbb{R}$, $m_\varphi(x) = (x^2 + ax + b)^k$, $F = \begin{pmatrix} 0 & -b \\ 1 & -a \end{pmatrix}$.

Example 2.4.3

$\varphi : V \to V, m_\varphi(x) = f^k, f(x)$ is irreducible.

Solution: (1) $m_\varphi = f$. $V = \bigoplus V_i$, $m_{\varphi|_{V_i}} = f$.

Then V can be decomposed into $\frac{\dim V}{\deg f}$ irreducible subspaces.

(2) □

2.5. Appendix : Linear Maps on Matrix Spaces

Example 2.5.1

Let $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{p \times q}$, consider

$$T : X \in \mathbb{F}^{n \times p} \mapsto AXB \in \mathbb{F}^{m \times q}.$$

It's easy to check that T is linear. What is the matrix representation of T?

Consider the vectorization of matrices.

Definition 2.5.1 Vectorization

Let $A \in \mathbb{F}^{m \times n}$. The vectorization of A is defined as

$$X = (x_1, \ldots, x_p) \in \mathbb{F}^{n \times p} \mapsto \operatorname{vec}(X) = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathbb{F}^{np \times 1}.$$

Then our desired matrix representation is given by

$$\underbrace{\mathrm{vec}(AXB)}_{\substack{\text{coordinate of } T(X) \\ mq}} = \underbrace{(B^T \otimes A)}_{\substack{\text{matrix representation of } T \\ mq \times np}} \underbrace{\mathrm{vec}(X)}_{\substack{\text{coordinate of } X \\ np}},$$

where $\otimes$ is the Kroenecker product.

Definition 2.5.2 Kroenecker Product

$$C \otimes D = \begin{pmatrix} c_{11}D & c_{12}D & ... & c_1 nD \\ c_{21}D & c_{22}D & ... & c_2 nD \\ \cdots & \cdots & \cdots & \cdots \\ c_{m1}D & c_{m2}D & ... & c_{mn}D \end{pmatrix}.$$

$\dim \ker T$?

Theorem 2.5.1

Let $A, B \in \mathbb{F}^{n \times n}$, f, g be coprime polynomials, $f(A) = g(B) = 0$.

Prove that $T : X \in \mathbb{F}^{n \times n} \mapsto AX - XB \in \mathbb{F}^{n \times n}$ is an isomorphism.

Proof: It suffices to show that $\ker T = 0$.

If $AX = XB$, then $f(A)X = Xf(B) = 0$. By the coprimality of f, g,

$$uf + vg = 1(\text{Bezout}).$$

$$\begin{aligned} X = IX &= (uf + vg)(A)X \\ &= 0 + v(A)Xg(B) = 0. \end{aligned}$$

Thus $\ker T = 0$. □

The form $AB - BA$ is quite common in Lie algebra. We may consider the following.

Theorem 2.5.2

Define communitator $[A, B] = AB - BA$. Then

If $A, B \in \mathrm{Hom}(V, V)$ s.t. $[A, B] = AB - BA = I$ Then

(1) there exists no annihilating polynomial of A.

(2) V must be infinite-dimensional.

Proof: (1) By easy induction,

$$A^k B - BA^k = kA^{k-1}.$$

Then for any polynomial f, $f(A)B - Bf(A) = f'(A)$.

(The core idea is utilizing the fact that derivative reduces the degree of the polynomial, thus giving an infinite descent.)

Suppose there exists an annihilating polynomial f of A. Then $f(A)B - Bf(A) = f'(A) = 0$.

Then $f^{(k)}$ is an annihilating polynomial of A for all k. This is a contradiction since for some k, $f^{(k)} = \text{const}$, and the constant is nonzero.

(2) Suppose V is finite-dimensional. Then

$$I, A, A^2, ..., A^{(\dim V)^2}$$

are linearly dependent. Thus there exists a nontrivial linear combination of them that equals 0. This gives an annihilating polynomial of A, a contradiction. □

Question 2.5.1 $AB - BA = C$

Let A, B, C be n-dimensional matrices. If $AB - BA = C$ and $AC = CA$, then C is non-invertible. Furthermore, C is nilpotent.

Proof: We can use a clever identity $A^k B - BA^k = kCA^{\{k-1\}}$.

The proof is done by induction.

Note 2.5.1 Why did we think of this?

To prove that C is nilpotent, we should increase the power of C in the given conditions. This can be easily achieved by repeatedly left-multiplying C and simplifying the expression.

Consider the linear map $\Phi : X \mapsto XB - BX$. If C is not nilpotent, i.e., $C^k \neq 0$ for all k, then each k is an eigenvalue of Φ, which contradicts the finite-dimensional case.

→ StackExchange, https://math.stackexchange.com/questions/811160/ab-ba-a-implies-a-is-singular-and-a-is-nilpotentb □

3. Linear Spaces with Metric

Definition 3.1 Orthogonal Complement

Let $f : V \times V \to \mathbb{F}$ be a (anti-)symmetric non- bilinear form. The orthogonal complement of U is defined as

$$U^{\perp} = \{v \in V \mid \forall u \in U, f(u, v) = 0\}.$$

$$V = U \oplus W \vee U \subset W$$

3.1. Multilinear Functions

Definition 3.1.1

A function $f : V \times ... \times V \to \mathbb{F}$ is called a k-linear function if it is linear in each of its arguments.

Now, to see how two vectors relate to each other, we consider bilinear functions, of which the most important are symmetric and antisymmetric bilinear functions.

Definition 3.1.2

A bilinear function $f : V \times V \to \mathbb{F}$ is called symmetric if $f(x, y) = f(y, x)$ for all $x, y \in V$.

A bilinear function $f : V \times V \to \mathbb{F}$ is called antisymmetric if $f(x, y) = -f(y, x)$ for all $x, y \in V$.

Actually, bilinearty can be deduced from linearity in the first argument, and use (anti)symmetry to deduce linearity in the second argument.

3.1.1. Gram Matrix : Representing Bilinear Functions

Given a basis $B = \{v_1, ..., v_n\}$ of V, we can represent a bilinear function $f : V \times V \to \mathbb{F}$ by a matrix M such that $f(x, y) = x^T M y$ for all $x, y \in V$. $M_{ij} = f(v_i, v_j)$.

Definition 3.1.1.1 Gram Matrix

The above matrix M is called the Gram matrix of f with respect to the basis B. It is denoted by $[f]_B$.

Now it suffice to study the properties of the matrix M to understand the bilinear function f. For example, if M is symmetric, then f is symmetric.

3.2. Euclidean Spaces and Symplectic Spaces

3.2.1. Euclidean Spaces and Inner Products

When M is symmetric and positive definite, we have an inner product similar to the dot product in $\mathbb{R}^n$. This is called an inner product space.

Example 3.2.1.1 Inner Product of Matrices

When $V = M_{n \times n}(\mathbb{R}) \simeq \mathbb{R}^{n^2}$, we define

$$f(A, B) = \mathrm{tr}(AB^T).$$

The transpose here is necessary to make the function positive definite, because we have $f(A, A) = \mathrm{tr}(AA^T) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 \geq 0$.

Given basis $B = \{E_{ij} \mid 1 \leq i, j \leq n\}$, the Gram matrix is $[f]_B = I_{n^2}$.

Example 3.2.1.2 Inner Product of Functions

When $V = C^0[0, 1]$, we define

$$\varphi(f, g) = \int_0^1 f(t) g(t) \, \mathrm{d}t.$$

Given basis $B = \{1, t, t^2, ...\}$, the Gram matrix is $[\varphi]_B = \left(\frac{1}{i+j+1}\right)_{i,j \geq 0}$.

$[\varphi]_B = \left(\frac{1}{i+j+1}\right)$ is not easy to deal with. However, by the general property of Gram matrices, we can deduce that φ is symmetric and positive definite.

In Euclidean spaces, inner products give rise to geometric intuition.

Exercise 3.2.1.1

Let $\dim V = n$, (V, f) is an inner product space. Show that if

$$B = \left(f(v_i, v_j)\right)_{1 \le i, j \le n},$$

then

$$\det B \ne 0 \Leftrightarrow v_1, \dots, v_n \text{ is a basis of } V.$$

Proof: If $\det B \ne 0$, then B is invertible. Let $x \in V$ be such that $f(x, v_i) = 0$ for all i. Then $x^T B = 0$, so $x = 0$. Assume $v_1, \dots, v_n$ are not linearly independent. Then there exists $x \in V$ such that $x = \sum_{i=1}^{n} a_i v_i$ and $a_i \ne 0$ for some i. Then $f(x, v_i) = a_i f(v_i, v_i) = 0$, so $x = 0$, a contradiction. □

3.3. Orthogonal Complements

Definition 3.3.1

Given a subset S of inner product space V, the orthogonal complement of S is the set

$$S^\perp = \{v \in V \mid \langle v, s \rangle = 0, \forall s \in S\}.$$

The orthogonal complement of S is well-defined i.e. S exists and is unique.

Theorem 3.3.1

$$V = \operatorname{span} S \oplus S^\perp.$$

Proof Sketch: WLOG, we assume S is a subspace by $S^\perp = \operatorname{span} S^\perp$.

If $v \in S \cap S^\perp$, then

$$\langle v, v \rangle = 0 \Rightarrow v = 0.\text{(by positive definiteness)}$$

Besides, $\dim S^\perp = \dim V - \dim S$, so that $V = \operatorname{span} S \oplus S^\perp$. □

Theorem 3.3.2 Pythagorean Theorem

If $V = W_1 \oplus W_2$ and $W_1 \perp W_2$. Then $\|v\|^2 = \|w_1\|^2 + \|w_2\|^2$ for all $v = w_1 + w_2$, $w_1 \in W_1$, $w_2 \in W_2$.

3.4. Exercises

Question 3.4.1

Prove that there are most $n + 1$ vectors in $V(\dim V = n)$ such that the angle between any two of them is obtuse.

Proof: Assume $v_1, ..., v_{n+1}, v_{n+2}$ are such vectors. Then there exists $\lambda_1, ..., \lambda_{n+1}$ which are not all zero such that $\sum_{i=1}^{n+1} \lambda_i v_i = 0$. Then

$$\lambda_1 \langle v_1, v_{n+2} \rangle + ... + \lambda_{n+1} \langle v_{n+1}, v_{n+2} \rangle = 0.$$

Obviously λ_i are not all of the same sign, so let

$$u = \sum_{i:\lambda_i > 0} \lambda_i v_i = - \sum_{i:\lambda_i < 0} \lambda_i v_i,$$

we have

$$\begin{aligned} 0 &\leq \langle u, u \rangle \\ &= \sum_{i:\lambda_i > 0, j:\lambda_j < 0} \lambda_i (-\lambda_j) \langle v_i, v_j \rangle < 0, \end{aligned}$$

a contradiction. □

Proof (Alternative, Gram-Schmidt): Note a interesting property of Gram-Schmidt process:

$$w_k = v_k - \sum_{i<k} \frac{\langle v_k, w_i \rangle}{\langle w_i, w_i \rangle} w_i.$$

By induction, we see

$$\begin{gathered} w_k \text{ is orthogonal to } w_i \text{ for all } i < k, \\ w_k \text{ and } v_{k+1}, ..., v_n \text{ are obtuse to each other.} \end{gathered}$$

Because

$$\langle w_k, v_{k+l} \rangle = \underbrace{\langle v_k, v_{k+l} \rangle}_{<0} - \underbrace{\sum_{i<k} \langle v_k, w_i \rangle \frac{\langle w_i, v_{k+l} \rangle}{\langle w_i, w_i \rangle}}_{>0 \text{ by IH.}}$$

so □

3.5. Spectral Theorem

3.5.1. The Complex Spectral Theorem

Now in the perspective of invariant subspaces, we may prove the spectral theorem in a way different from what we did on matrices (i.e. Schur → diagonalization).

Proof: Given an normal operator A s.t. $AA^* = A^* A$, it suffices to show

$$\mathbb{C}^n = \bigoplus_{\lambda} \ker(\lambda I - A),$$

and $\ker(\lambda I - A)$ are orthogonal to each other. It is clear that

$$\mathbb{C}^n = \ker(\lambda I - A) \oplus \ker{(\lambda I - A)}^{\perp}.$$

Now we need to show that $\ker{(\lambda I - A)}^{\perp}$ is A-invariant (then we can apply induction on $A|_{\ker{(\lambda I - A)}^{\perp}}$).

For all $v \in \ker{(\lambda I - A)}^{\perp}$,

$$
\begin{aligned}
&Av \in \ker{(\lambda I - A)}^{\perp} \\
\Leftarrow\ & Av \perp w, \forall w \in \ker(\lambda I - A), \\
\Leftarrow\ & 0 = \langle Av, w \rangle = \langle v, A^* w \rangle, \\
\Leftarrow\ & A^* w \in \ker \ker(\lambda I - A), \forall w \in \ker(\lambda I - A), \\
\Leftarrow\ & \ker(\lambda I - A) \text{ is } A - \text{invariant (by commutativity)}.
\end{aligned}
$$

□

Example 3.5.1.1

A is Hermitian $\Rightarrow$ A^2 is Hermitian.

Question 3.5.1.1 Canonical Form of Real Normal Matrix

Let A be a normal real matrix, then A is (orthogonally) similar to

$$\operatorname{diag}\left(a_1, ..., a_k, \begin{pmatrix} b_1 & c_1 \\ -c_1 & b_1 \end{pmatrix}, ..., \begin{pmatrix} b_l & c_l \\ -c_l & b_l \end{pmatrix}\right).$$

Actually, the building blocks

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

Proof: Merge the complex eigenspaces into

$$W = \ker(A^2 - 2bA + (b^2 + c^2)I).$$

Try to find $v, w \in W$ s.t. $A(v, w) = (v, w)\begin{pmatrix} b & c \\ -c & b \end{pmatrix}$.

$$
\begin{aligned}
Av = bv - cw &\Rightarrow \boldsymbol{w} = -c^{-1}(A - bI)\boldsymbol{v} \\
&\Rightarrow Aw = cv + bw. (\text{Note: } A^2 v = 2bAv - (b^2 + c^2)v)
\end{aligned}
$$

By the above restriction, find

$$
\begin{gathered}
1. w \perp v \\
2. \operatorname{span}{(v, w)}^{\perp} \text{ is } A \text{ -invariant}.
\end{gathered}
$$

Then the result follows from simple induction. □

4. Review

4.1. Multilinear Functions

Example 4.1.1

Let $f_1, f_2 \in V^*$, if $\forall v \in V, f_1(v) f_2(v) = 0$, prove that $f_1 = 0$ or $f_2 = 0$.

Proof: (Matrix)

$$[f_1(v_1), ..., f_n(v_n)] = (a_1, ..., a_n),$$
$$[f_2(v_1), ..., f_n(v_n)] = (b_1, ..., b_n).$$

$$(a_1, ..., a_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}(b_1, ..., b_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0.$$

Choosing $(x_i) = e_i$, $a_i b_i = 0, \forall i$.

Choosing $(x_i) = e_i + e_j$, $a_i b_i + a_j b_j = 0, \forall i, j$.

If $f_1 \neq 0$, then WLOG there exists i s.t. $a_1 \neq 0$, then $b_1 = 0$. Then $a_1 b_j + a_j b_1 = a_1 b_j = 0, \forall j$, so $b_j = 0, \forall j$. □

Question 4.1.1

If $g : V \times V \to \mathbb{F}$ is a bilinear function s.t.

$$g(v_1, v_2) = 0 \Rightarrow g(v_2, v_1) = 0,$$

prove that g must be symmetric or antisymmetric.

Example 4.1.2

Let g be a nonzero alternating bilinear function on V, then g cannot be decomposed into the product of two linear functions.

$$\nexists f_1, f_2, g(\alpha, \beta) = f_1(\alpha) f_2(\beta)$$

Proof: Proof by contradiction. Assume $g(\alpha, \beta) = f_1(\alpha) f_2(\beta)$.

$$g(\alpha, \beta) = f_1(\alpha) f_2(\beta) = -f_1(\beta) f_2(\alpha).$$

Let $\beta = \alpha$, By Example 4.1.1, we have $f_1 = 0$ or $f_2 = 0$, a contradiction since g is not identically zero. □

Theorem 4.1.1 Algebraic Problem, Geometric Approach

A real symmetric matrix A of order n is orthogonally similar to a matrix whose diagonal entries are zero

$$\Leftrightarrow \operatorname{tr} A = 0.$$

Note 4.1.1

$$g(v_i + v_j, v_i + v_j) = g(v_i, v_i) + g(v_j, v_j) + 2g(v_i, v_j).$$

So studying the properties of the induced quadratic form $q(v) = g(v, v)$ gives us information about the bilinear function g.

Proof Sketch: $\Rightarrow$ is trivial. Now assume $\operatorname{tr} A = 0$.

Discuss the problem on $V/\mathbb{R}$. g is a symmetric bilinear function, so it induces a quadratic form $q(v) = g(v, v)$.

Choose a basis, $q(v) = f(X) = X^T A X$ where $v = (\boldsymbol{v}_1, ..., \boldsymbol{v}_n)^{-1} X$.

The change of variable $X = PY$ gives $f(X) = Y^T P^T A P Y$.

Induction on the dimension of the space.

If □

Theorem 4.1.2

> **Definition 4.1.1** Zero Cone of a Quadratic Form
>
> The zero cone of a quadratic form q is the set $S = \{v \in V \mid q(v) = 0\}$.

S is a subspace of $V \Leftrightarrow q$ is p.s.d. or n.s.d.

Proof Sketch:

$$\begin{aligned} S &= \{v \in V \mid q(v) = 0\} \\ &= \left\{X \in \mathbb{R}^n \mid X^T A X = 0\right\} \\ &\overset{X=PY}{=} \left\{Y \in \mathbb{R}^n \mid y_1^2 + ... + y_p^2 - y_{p+1}^2 - ... - y_r^2 = 0\right\} \end{aligned}$$

If q is not definite, then there exists n linearly independent vectors in S, so $S = V$, a contradiction.

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

Every row is a vector in S.

If q is p.s.d., then there exists a basis η s.t $S = \langle \eta_{p+1}, ..., \eta_n \rangle$. If q is n.s.d., the same argument applies. □

Question 4.1.2

$A^H = A$ is p.d. then $\det A \le a_{11} a_{22} ... a_{nn}$.

Proof: By induction.

$$A = \begin{pmatrix} A_1 & \alpha \\ \alpha^T & a_{nn} \end{pmatrix} \to \begin{pmatrix} A_1 & 0 \\ 0 & a_{nn} - 2\alpha^T A_1^{-1} \alpha \end{pmatrix}.$$

$$\det A = \det(A_1) \underbrace{\left(a_{nn} - 2\alpha^T A_1^{-1} \alpha\right)}_{>0} \leq \det A_1 a_{nn}.$$

"$=$" holds iff $\alpha = 0$ (by p.d. of A_1^{-1}), i.e. A is diagonal. □

5. Miscellaneous

Proof:

考虑n阶分块对称矩阵

$$\begin{pmatrix} \mathbf{0} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{B} \end{pmatrix}, \boldsymbol{B}^T = \boldsymbol{B}.$$

先设 $\boldsymbol{B}$ 可逆. 做分块矩阵的初等行变换，不改变行列式是否为0的性质.

$$\begin{pmatrix} \mathbf{0} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{B} \end{pmatrix} \to \begin{pmatrix} -\boldsymbol{A}^T \boldsymbol{B}^{-1} \boldsymbol{A} & \mathbf{0} \\ \boldsymbol{A}^T & \boldsymbol{B} \end{pmatrix}.$$

则行列式

$$\det \begin{pmatrix} -\boldsymbol{A}^T \boldsymbol{B}^{-1} \boldsymbol{A} & \mathbf{0} \\ \boldsymbol{A}^T & \boldsymbol{B} \end{pmatrix} = |\boldsymbol{B}| \cdot \left|\boldsymbol{A}^T \boldsymbol{B}^{-1} \boldsymbol{A}\right|.$$

下面讨论 $\boldsymbol{B}$ 是 $k\left(k < \frac{n}{2}\right)$ 阶矩阵的情况. 此时 $\boldsymbol{B}^{-1}\boldsymbol{A}$ 是 $k \times (n-k)$ 阶矩阵，且$k < \frac{n}{2} \Rightarrow k < n - k$. 因此由 Cauchy-Binet 公式的直接推论，$\left|\boldsymbol{A}^T \boldsymbol{B}^{-1} \boldsymbol{A}\right| = 0$.

故在 $k < \frac{n}{2}$ 时我们可以断定 $\det\begin{pmatrix} \mathbf{0} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{B} \end{pmatrix} = 0$.

若 $\boldsymbol{B}$ 并不可逆，考虑存在 $\varepsilon > 0$，微扰 $\forall 0 < t < \varepsilon, \boldsymbol{B}' = \boldsymbol{B} + tI$ 可逆. 由于 $\det\begin{pmatrix} \mathbf{0} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{B}+tI \end{pmatrix}$ 是 t 的多项式，故 $\det\begin{pmatrix} \mathbf{0} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{B} \end{pmatrix} = \lim_{t \to 0^+} \det\begin{pmatrix} \mathbf{0} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{B}+tI \end{pmatrix} = 0$. □