Linear Algebra Done Taiwan

Eric Liu

CONTENTS

Chapter 1

What you need to know

1.1 Quick recap

Let V be a vector space over $\mathbb{F}$, and let S be a subset of V. We say

(a) S is **linearly independent** if to write 0 as finite linear combination of S, the coefficients must all be zero.

(b) S **spans** V if all $v \in V$ can be written as finite linear combination of S.

We say S is a **basis** if any of the following equivalent conditions hold true

(a) S is a maximum linearly independent subset of V.

(b) S is linearly independent and spans V.

(c) All elements of V can be uniquely written as some finite linear combination of S.

Suppose V has a basis of cardinality n. Using **Gauss elimination**, one see that all linear independent subset of V must have cardinality not greater than n, and all basis must have cardinality n. Therefore, for vector space with some basis of finite cardinality, it make sense to refer to its **dimension**. Let $\dim(V) = n$. Each linearly independent subset S of V of cardinality n must be a basis, otherwise we may construct a linearly independent set of cardinality greater than n. By an algorithm, one see that

(a) If S spans V and has cardinality greater than n, then there exists a subset of S that is a basis of V.

(b) If B is a basis of V and S is linearly independent, then there exists some subset U of B such that $S \cup U$ is a basis of V.

Let W be another vector space over $\mathbb{F}$, and let $F \in \operatorname{Hom}(V, W)$. It is clear that both image and kernel of F forms a vector space. Therefore, it make sense to define the **rank** of F by

$$\operatorname{rank}(F) \triangleq \dim(\operatorname{Im} F).$$

Clearly, for all basis $B \subseteq V$, we have

$$|X| \leq \dim(\operatorname{Ker} F)$$

where $X \triangleq \{v \in B : F(v) = 0\}$ and $Y \triangleq \{v \in B : F(v) \neq 0\}$. Moreover, because $X \cup Y$ forms a basis, we also have

$$\operatorname{span}\{F(v) \in W : v \in Y\} = \operatorname{Im}(F)$$

which implies

$$\operatorname{rank}(F) \leq |Y|$$

Selecting some $S \subseteq Y$ such that $\{F(v) \in W : v \in S\}$ forms a basis of $\operatorname{Im}(F)$, and replacing $v \in Y - S$ with $v - \sum c_i s_i$ so that $F(v) = 0$, we now see the new X forms a basis for $\operatorname{Ker} F$, and the new Y make $\{F(v) \in W : v \in Y\}$ a basis for W. This implies the **Rank-Nullity Theorem**:

$$\dim(V) = |X| + |Y| = \dim(\operatorname{Ker} F) + \operatorname{rank}(F)$$

We use $V^{\vee}$ to denote the **dual space** of V. Let $F \in \operatorname{Hom}(V, W)$. Its **dual map** $F^{\vee} \in \operatorname{Hom}(W^{\vee}, V^{\vee})$ is defined by

$$F^{\vee}(\varphi) \triangleq \varphi \circ F$$

If V is finite-dimensional with basis $\{v_1, \ldots, v_n\}$, then its **dual basis** $\{\varphi_1, \ldots, \varphi_n\}$ is defined by

$$\varphi_i(v_j) \triangleq \delta_j^i$$

and there exists a natural isomorphism between V and its double dual $(V^{\vee})^{\vee}$ by identifying $v \in (V^{\vee})^{\vee}$ as

$$v(\varphi) \triangleq \varphi(v).$$

Let $\{w_1, \ldots, w_m\}$ be a basis of W with dual basis $\{\xi_i, \ldots, \xi_m\}$. It is clear that under these two basis, the matrices representation of F and $F^{\vee}$ are *transpose* of each other. This together with the observation that $\operatorname{rank}(F) + \dim(\operatorname{ker} F^{\vee}) = \dim(W)$ for finite dimensional V and W, give us a quick proof that the dimensions of column space and row space of a

fixed matrix are always equal. Now, given some n-by-n square matrix A, its **determinant** is by definition

$$\det A \triangleq \sum_{\sigma \in S_n} \left(\operatorname{sgn} \sigma \prod_{k=1}^{n} A_{\sigma(k),k} \right)$$

If one identify $M_n(\mathbb{F})$ with $(\mathbb{F}^n)^n$, determinant can be equivalently defined to be the unique alternating multilinear map from $(\mathbb{F}^n)^n$ to $\mathbb{F}$ such that

$$\det(I) = 1$$

By an algorithm, one see that, for each alternating multilinear map $F : (\mathbb{F}^n)^n \to \mathbb{F}$, we have

$$F(B) = (\det B) \cdot F(I) \text{ for all } B.$$

Therefore, if we observe for each $A \in M_n(\mathbb{F})$ that the map $B \mapsto \det AB$ is indeed alternating multilinear, then we immediately have the celebrated result $\det AB = \det A \det B$ as a corollary. This multiplicative property of determinant allow us to well define determinant for each linear endomorphism F over some finite-dimensional vector space by

$$\det F \triangleq \det([F])$$

where $[F]$ is the matrix representation of F with respect to some basis. Moreover, because the inverse of a linear map is linear if exists, we see that for linear map F,

$$F \text{ is invertible } \iff \det F \neq 0$$

And

$$A \text{ is invertible } \iff \det A \neq 0$$

Given some $F \in \operatorname{End}(V)$, we say $v \in V$ is an **eigenvector** with respect to the **eigenvalue** $\lambda \in \mathbb{F}$ if $v \neq 0$ and $F(v) = \lambda v$. If there exists some basis of V consisting eigenvectors of F, we say F is **diagonalizable**. Let A be some square matrix, if there exists some invertible square matrix P such that PAP^{-1} diagonal, we also say A is **diagonalizable**. Immediately, we see that if V is finite dimensional, then with respect to any basis of V

$$[F] \text{ is diagonalizable } \iff F \text{ is diagonalizable}$$

Let V be finite dimensional. We define the **characteristic polynomial** of F and A by

$$\det(tI - F) \text{ and } \det(tI - A)$$

Obviously, λ is an eigenvalue of F if and only if λ is a root of the characteristic polynomial of F.

5

1.2 Jordan-Chevalley decomposition

Let V be some finite dimensional $\mathbb{F}$-vector space. Given $T \in \mathrm{End}(V)$, we know the kernels of its powers is increasing:

$$0 \subseteq \mathrm{Ker}\, T^1 \subseteq \mathrm{Ker}\, T^2 \subseteq \mathrm{Ker}\, T^3 \subseteq \cdots$$

This sequence grows in good manner. If it stops growing at some points, say, $\mathrm{Ker}\, T^n = \mathrm{Ker}\, T^{n+1}$ for some $n \in \mathbb{Z}_0^+$, then it stops forever:

$$\mathrm{Ker}\, T^n = \mathrm{Ker}\, T^{n+1} = \mathrm{Ker}\, T^{n+2} = \mathrm{Ker}\, T^{n+3} = \cdots$$

In particular, by counting dimensions, we know that this sequence must stops before reaching to the dimension of V in the sense that:

$$\mathrm{Ker}(T^{\dim V}) = \mathrm{Ker}(T^{\dim V + 1}) \tag{1.1}$$

Equation 1.1 allows us to elegantly unify distinct notions within a single framework. For instance, for each eigenvalue λ, we define the **generalized eigenspace** $G(\lambda, T)$ as:

$$G(\lambda, T) \triangleq \mathrm{Ker}(T - \lambda I)^{\dim V}$$

Similarly, we say that T is **nilpotent** if:

$$\mathrm{Ker}\, T^{\dim V} = V$$

Theorem 1.2.1. (Linear independence of generalized eigenspaces) Let V be a finite-dimensional $\mathbb{F}$-vector space and $T \in \mathrm{End}(V)$. If $v_1, \ldots, v_n$ are generalized eigenvectors of T with respect to distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, then they are linearly independent.

Proof. Suppose

$$a_1 v_1 + \cdots + a_n v_n = 0 \tag{1.2}$$

Let k be the largest nonnegative integer such that $(T - \lambda_1 I)^k v_1 \neq 0$. We see:

$$w \triangleq (T - \lambda_1 I)^k v_1 \text{ is an eigenvector with eigenvalue } \lambda_1.$$

Noting that the collection of the endomorphisms:

$$\left\{ (T - \lambda_1 I)^k, (T - \lambda_2 I)^{\dim V}, \ldots, (T - \lambda_n I)^{\dim V} \right\}$$

commutes, we may apply the endomorphism $(T - \lambda_1 I)^k (T - \lambda_2 I)^{\dim V} \cdots (T - \lambda_n I)^{\dim V}$ onto both sides of Equation 1.2 and get

$$\begin{aligned} 0 &= a_1 (T - \lambda_2 I)^{\dim V} \cdots (T - \lambda_n I)^{\dim V} w \\ &= a_1 (\lambda_1 - \lambda_2)^{\dim V} \cdots (\lambda_1 - \lambda_n)^{\dim V} w \end{aligned}$$

which implies $a_1 = 0$. WLOG, we have shown $a_1 = \cdots = a_n = 0$. ∎

Equation 1.1 together with Rank-Nullity Theorem give us a theoretically crucial decomposition of V:

$$V = \operatorname{Ker} T^{\dim V} \oplus \operatorname{Im} T^{\dim V} \tag{1.3}$$

Note that the number $\dim V$ is irrelevant to the true value of decomposition 1.3, in the sense that the same hold true if we replace $\dim V$ with any $k \geq n$, where n is the smallest natural number that makes $\operatorname{Ker}(T^n) = \operatorname{Ker}(T^{n+1})$. However, we in particular consider $\dim V$ for a good reason. As we will later see, decomposition 1.3 plays a central role in decomposition of V into generalized eigenspace of T. See theorem 1.2.3.

We say a polynomial $m \in \mathbb{F}[x]$ is a **minimal polynomial**[1] of T, if $m \in \mathbb{F}[x]$ a nonzero polynomial of smallest degree such that $m(T) = 0 \in \operatorname{End}(V)$. Note that because $\operatorname{End}(V)$ is finite dimensional, such m must exists. We say a polynomial $p \in \mathbb{F}[x]$ **splits over** $\mathbb{F}$ if we may write:

$$p(x) = (x - \lambda_1)^{n_1} \cdots (x - \lambda_m)^{n_m}, \quad \text{where } \lambda_i \in \mathbb{F}$$

Example 1.2.2. The linear transformation

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

from $\mathbb{R}^2$ to $\mathbb{R}^2$ has no eigenvalue in $\mathbb{R}$. Its minimal polynomial is $x^2 + 1$.

Theorem 1.2.3. (Decomposition into generalized eigenspaces) Let V be a finite dimensional $\mathbb{F}$-vector space of dimension ≥ 1, $T \in \operatorname{End}(V)$, and a (hence any) minimal polynomial of T splits over $\mathbb{F}$. If we let $\{\lambda_1, \ldots, \lambda_m\}$ be the set of eigenvalues of T, then:

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$$

Proof. This is proved by induction on the dimension of V. The 1-dimensional base case is trivial. We now prove the inductive case. By equation 1.3, we may decompose V into:

$$V = G(\lambda_1, T) \oplus U, \quad \text{where } U = \operatorname{Im}(T - \lambda_1)^{\dim V}. \tag{1.4}$$

Noting that T and $T - \lambda_1 I$ commutes, we conclude that U is stable under T. Therefore, the restriction $T|_U$ defines an endomorphism $T|_U \in \operatorname{End}(U)$. Let $m \in \mathbb{F}[x]$ be a minimal polynomial of T, and let $m_U \in \mathbb{F}[x]$ be a minimal polynomial of $T|_U$. Because $m(T|_U) = 0$,

---

[1]It is worth mentioning that we didn't prove that the minimal polynomial of T is invariant under field extension. To prove such, one need to show that the minimal polynomial m of $[T]$ over $\overline{\mathbb{F}}$ has only $\mathbb{F}$ coefficients. This follows from noting that if we write $m = m_1 + c_2 m_2 + \cdots + c_k m_k$ where $m_1, \ldots, m_k \in \mathbb{F}[x]$, then $m_2 = \cdots = m_k = 0$, since they have degree strictly smaller than m, as minimal polynomial is defined to be monic.

we know m_U divides m. This then implies m_U also splits over $\mathbb{F}$.

Clearly, every eigenvalues of $T|_U$ lies in $\{\lambda_2, \ldots, \lambda_m\}$. To see that $T|_U$ actually has spectrum $\{\lambda_2, \ldots, \lambda_m\}$, let $v_2 \in V$ be an eigenvector of T with eigenvalue λ_2. Decomposition 1.4 allows us to write:

$$v_2 = v_1 + (T - \lambda_I)^{\dim V} w, \quad \text{where } v_1 \in G(\lambda_1, T) \text{ and } w \in V$$

Because $(T - \lambda_2)v_2 = 0$, we know

$$0 = (T - \lambda_2)v_1 + (T - \lambda_1 I)^{\dim V}(T - \lambda_2 I)w$$

Clearly, $(T - \lambda_2)v_1 \in G(\lambda_1, T)$. This together with direct sum 1.4 implies $(T - \lambda_2)v_1 = 0$, which by linear independence of generalized eigenvectors implies $v_1 = 0$, as desired.

Finally, we may now use inductive hypothesis to decompose U into:

$$U = G(\lambda_2, T|_U) \oplus \cdots \oplus G(\lambda_m, T|_U)$$

WLOG, it remains to show:

$$G(\lambda_2, T|_U) = G(\lambda_2, T)$$

Arbitrary select $v \in G(\lambda_2, T)$. We may decompose $v = a_1 v_1 + \cdots + a_m v_m$, where $v_1 \in G(\lambda_1, T)$ and $v_n \in G(\lambda_n, T|_U)$ for $2 \leq n \leq m$. Because $(T - \lambda_2 I)^{\dim V} v_2 = 0$, we know

$$(T - \lambda_2 I)^{\dim V}(a_1 v_1 + a_3 v_3 + \cdots + a_m v_m) = 0$$

This implies $a_1 v_1 + a_3 v_3 + \cdots + a_m v_m \in G(\lambda_2, T)$. It now follows from linear independence of generalized eigenvectors that $a_1 = a_3 = \cdots = a_m = 0$. We have shown $v \in G(\lambda_2, T|_U)$. ∎

Given some finite dimensional $\mathbb{F}$-vector space V, and some $T \in \text{End}(V)$, we are particularly concerned with the existence and uniqueness of the **Jordan-Chevalley decomposition** of T, i.e, some diagonalizable $S \in \text{End}(V)$ such that

(a) $N \triangleq T - S$ is nilpotent.

(b) S and N commute.

If minimal polynomial of T splits over $\mathbb{F}$, then Theorem 1.2.3 assert the existence of such decomposition by requiring S to map $v \in G(\lambda_i, T)$ to $\lambda_i v$. To see such decomposition is unique, let $S \in \text{End}(V)$ be some Jordan-Chevalley decomposition of T and decompose V into

$$V = E(\lambda_1, S) \oplus \cdots \oplus E(\lambda_k, S)$$

where $E(\lambda_i, S)$ are the eigenspaces of S corresponding to distinct eigenvalues λ_i. Because T, S commute and $T - S$ is nilpotent, we may conclude that λ_i are indeed the eigenvalues of T and

$$E(\lambda_i, S) \subseteq G(\lambda_i, T) \text{ for all } i$$

This together with the decomposition we obtained from Theorem 1.2.3:

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$$

shows the uniqueness of Jordan-Chevalley decomposition. Interestingly, we may also take a completely algebraic approach to prove:

(a) the existence of Jordan-Chevalley decomposition.

(b) the uniqueness of Jordan-Chevalley decomposition.

(c) that $S = p(T)$ for some polynomial $p \in \overline{\mathbb{F}}[x]$.

without ever invoking linear independence of generalized eigenvectors, Theorem 1.2.3, nor any result that depends on them. The argument proceeds as follows: Be premise, the minimal polynomial $m \in \mathbb{F}[x]$ of T splits over $\mathbb{F}$:

$$m(x) = (x - \xi_1)^{r_1} \cdots (x - \xi_k)^{r_k}, \quad \text{where } \xi_i \in \mathbb{F}$$

Consider the polynomials $f_i \in \mathbb{F}[x]$ defined by

$$f_i(x) \triangleq \prod_{j=1; j \neq i}^{k} (x - \xi_j)^{r_j}.$$

Since $\gcd(f_1, \ldots, f_k) = \text{const.}$, Bézout's identity says that there exist polynomials $q_i \in \mathbb{C}[x]$ such that

$$1 = \sum_{i=1}^{k} q_i f_i.$$

Define $\pi_i \in \text{End}(V)$ by $\pi_i \triangleq (q_i f_i)(T)$. Because m divides $f_i f_j$ for $i \neq j$, we conclude that

$$\pi_i \pi_j = 0 \quad \text{for } i \neq j.$$

This, together with $\sum \pi_j = 1$, shows that π_i are **projections**:

$$\pi_i^2 = \pi_i \quad \text{for all } i.$$

9

It is clear that V is a direct sum of the images of these projections. Define

$$S \triangleq \sum_{i=1}^{k} \xi_i \pi_i.$$

Since S is a polynomial in T, we conclude that S and N commute. The fact that N is nilpotent follows from the definitions of f_i. We now prove the uniqueness. Let $T = S' + N'$ be another Jordan-Chevalley decomposition. We are required to prove $S - S' = 0$. Clearly $\{S, N, S', N'\}$ all commute with each other. This implies $S - S' = N - N'$ is nilpotent. Note that $S - S'$ is diagonalizable[2] . This now implies $S - S' = 0$, since no units are zero-divisor.

---

[2]Two commuting diagonalizable linear operator S, S' over finite dimensional vector space are always simultaneously diagonalizable. Prove this by first proving that $E(\lambda, S)$ is stable under S' and then consider $V = E(\lambda_1, S) \oplus \cdots \oplus E(\lambda_k, S)$

1.3 Jordan forms

Theorem 1.3.1. (Jordan forms of nilpotent operators) Let V be a finite-dimensional $\mathbb{F}$-vector space of dimension ≥ 1, and let $N \in \operatorname{End}(V)$ be nilpotent. There exists an unique element $(n_1, \ldots, n_m)$ in the set:

$$\{(n_1, \ldots, n_k) \mid \quad k \in \mathbb{N}, \quad \text{and } n_1 \geq \cdots \geq n_k \geq 0, \quad \text{and } n_i \in \mathbb{N} \cup \{0\}\}$$

such that there exists a tuple $v_1, \ldots, v_m \in V$ that makes:

(a) the tuple $N^{n_1}v_1, \ldots, Nv_1, v_1, \ldots, N^{n_m}v_m, \ldots, Nv_m, v_m$ forms a basis for V.

(b) $N^{n_1+1}v_1 = \cdots = N^{n_m+1}v_m = 0$.

Proof. The uniqueness follows from the observation that if $(n_1, \ldots, n_m)$ is such an element, then for all $i \in \mathbb{N}$, we have:

$$\dim\left(\operatorname{Ker} N^i\right) = \sum_{j=1}^{i} j d_j, \quad \text{where } d_j \text{ is the amount of } l \text{ that makes } n_l + 1 \geq j$$

We now prove the existence via induction.[3] The 1-dimensional base case is trivial. We now prove the inductive case. Because N is nilpotent, we know $\dim(\operatorname{Im} N) < \dim V$. Noting that $\operatorname{Im} N$ is stable under N, we see $N|_{\operatorname{Im} N} \in \operatorname{End}(\operatorname{Im} N)$. From inductive hypothesis, we have $v_1, \ldots, v_n \in \operatorname{Im} N$ and nonnegative integer $k_1, \ldots, k_n$ such that

$$\left\{ N^{k_1}v_1, \ldots, Nv_1, v_1, \ldots, N^{k_n}v_n, \ldots, Nv_n, v_n \right\} \text{ form a basis for } \operatorname{Im} N.$$

and

$$N^{k_1+1}v_1 = \cdots = N^{k_n+1}v_n = 0.$$

Because $v_1, \ldots, v_n \in \operatorname{Im} N$, we may let $u_1, \ldots, u_n \in V$ satisfy $v_j = Nu_j$ for all j. To see

$$\left\{ N^{k_1+1}u_1, \ldots, Nu_1, u_1, \ldots, N^{k_n}u_n, \ldots, Nu_n, u_n \right\} \text{ is linearly independent,} \qquad (1.5)$$

suppose some finite linear combination equals to 0. By applying N to this finite linear combination, we see the only possible nonzero coefficients are those of $N^{k_j+1}u_j$, otherwise

$$\left\{ N^{k_1}v_1, \ldots, Nv_1, v_1, \ldots, N^{k_n}v_n, \ldots, Nv_n, v_n \right\} \text{ would not be linearly independent.}$$

Knowing that the only possible nonzero coefficients are those of $N^{k_j+1}u_j = N^{k_j}v_j$, we may conclude that even these coefficients are zero, since $\left\{ N^{k_1}v_1, \ldots, N^{k_n}v_n \right\}$ is linearly independent in the first place. We may now expand Set 1.5 to a basis

$$\left\{ N^{k_1+1}u_1, \ldots, Nu_1, u_1, \ldots, N^{k_n+1}u_n, \ldots, Nu_n, u_n, w_1, \ldots, w_p \right\} \qquad (1.6)$$

---

[3]Note that the proof for the existence of decomposition of finite-dimensional vector space over algebraically closed field into generalized eigenspaces also relies on induction.

for V. Because $\left\{N^{k_1+1}u_1, \ldots, N^2u_1, Nu_1, \ldots, N^{k_n+1}u_n, \ldots, N^2u_n, Nu_n\right\}$ is a basis for $\operatorname{Im} N$, we may subtract each w_i with some element of

$$\operatorname{span}\left\{N^{k_1}u_1, \ldots, Nu_1, u_1, \ldots, N^{k_n+1}u_n, \ldots, Nu_n, u_n\right\}$$

so that Set 1.6 form a desired basis for V. ∎

Let V be some n-dimensional $\mathbb{F}$-vector space and $T \in \operatorname{End}(V)$ has minimal polynomial that splits over $\mathbb{F}$. Finally it follows at once from theorem 1.2.3 and theorem 1.3.1 that T admits a **Jordan basis**. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T, and let r_i be the dimension of the largest Jordan block with eigenvalue λ_i. It is clear that the minimal polynomial $m \in \mathbb{F}[x]$ of T when required to be monic, take the form:

$$m(x) = (x - \lambda_1)^{r_1} \cdots (x - \lambda_m)^{r_m}$$

This give a proof of the **Cayley-Hamilton Theorem**, that is, $p(F) = 0$, where p is the characteristic polynomial of F.

Let V be a finite dimensional $\mathbb{F}$-vector space and $T \in \operatorname{End}(V)$ has minimal polynomial of degree d. Let $v \in V$. Clearly,

$$\left\{v, Tv, T^2v, \ldots, \right\} \subseteq \operatorname{span}\left\{v, \ldots, T^{d-1}v\right\}$$

This implies that if V is T-cyclic, then $d = \dim V$. Conversely, if $d = \dim V$,

1.4 Analytic functions on matrices

Let $N \in \mathrm{M}_n(\mathbb{C})$ be nilpotent and $\lambda \in \mathbb{C}$. Because λI and N commute, we may apply binomial theorem to have:

$$(\lambda I + N)^m = \sum_{k=0}^{m} \binom{m}{k} \lambda^{m-k} N^k$$

In fact, given $p_d(z) \triangleq a_d(z - z_0)^d + \cdots + a_0 \in \mathbb{C}[z]$, we have

$$
\begin{aligned}
p_d(\lambda I + N) &= \sum_{m=0}^{d} a_m ((\lambda - z_0)I + N)^m \\
&= \sum_{m=0}^{d} a_m \sum_{k=0}^{m} \binom{m}{k} (\lambda - z_0)^{m-k} N^k \\
&= \sum_{k=0}^{n-1} \left(\sum_{m=k}^{d} a_m \binom{m}{k} (\lambda - z_0)^{m-k} \right) N^k = \sum_{k=0}^{n-1} \frac{p_d^{(k)}(\lambda)}{k!} N^k
\end{aligned}
$$

Consider the power series $\sum a_k (z - z_0)^k$ whose convergence disk $B_R(z_0) \subseteq \mathbb{C}$ contains λ. By basic complex analysis, we know the function $g : B_R(z_0) \to \mathbb{C}$ defined by:

$$g(z) \triangleq \sum_{k=0}^{n} a_k (z - z_0)^k$$

is holomorphic and allow term-by-term differentiation. Therefore, we may write:

$$\sum_{k=0}^{\infty} a_k ((\lambda - z_0)I + N) = \lim_{d \to \infty} p_d(\lambda I + N) = \lim_{d \to \infty} \sum_{k=0}^{n-1} \frac{p_d^{(k)}(\lambda)}{k!} N^k = \sum_{k=0}^{n-1} \frac{g^{(k)}(\lambda)}{k!} N^k \quad (1.7)$$

Now, let $U \subseteq \mathbb{C}$ be open and $f : U \to \mathbb{C}$ holomorphic with $\lambda \in U$. Because of equation 1.7, we see that, independent of choice of $z_0 \in U$ that has λ lying in the convergence disk of f at z_0, the series:

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(z_0)}{k!} ((\lambda - z_0)I + N)^k = \sum_{k=0}^{n-1} \frac{f^{(k)}(\lambda)}{k!} N^k$$

always converges to the same matrix. Because of such, we may define $f(\lambda I + N) \in M_n(\mathbb{C})$ with the assurance that every possible computation agree. Moreover, not only does this definition make philosophical sense, letting $h : U \to \mathbb{C}$ be another holomorphic function,

13

because

$$\frac{(fh)^{(k)}}{k!} = \sum_{r=0}^{k} \binom{k}{r} f^{(r)} h^{(k-r)}, \quad \text{for all } k \in \mathbb{N} \cup \{0\}$$

we also see the good behavior:

$$f(\lambda I + N)h(\lambda I + N) = (fh)(\lambda I + N)$$

Setting f, h to both be some branches of the multi-valued function $z \mapsto z^{(\frac{1}{r})}$, immediately we see that for all $r \in \mathbb{N}$, there always exists some $A \in M_n(\mathbb{C})$ such that $A^r = \lambda I + N$. (Note that by chain rule, $\sqrt{z}$ always have derivative $\frac{1}{\sqrt{z}}$ where the two $\sqrt{z}$ agree.)

Note that one should be careful in attempt to define value for holomorphic function f on matrix $A \in M_n(\mathbb{C})$ not of the form $\lambda I + N$, since even if the definition can be made precises by considering Jordan form, there may not be a single power series $\sum a_n (z - z_0)^n$ that makes

$$f(A) = \sum a_n (A - z_0 I)^n$$

However, we shall make the remark that **entire functions** are free of this ill situation. Therefore, for every $A \in M_n(\mathbb{C})$, considering its Jordan's form $A = PJP^{-1}$, we see block by block that

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(z_0)}{k!}(A - z_0 I) = P \cdot \left(f(J(\lambda_1)) \oplus \cdots \oplus f(J(\lambda_n)) \right) \cdot P^{-1}$$

where

$$f(J(\lambda)) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\lambda)}{k!} N^k, \quad \text{with } J(\lambda) = \lambda I + N$$

The fact that we may generalize entire function $\mathbb{C} \to \mathbb{C}$ to $M_n(\mathbb{C}) \to M_n(\mathbb{C})$ is very important with the following application.

Let $A, B \in M_n(\mathbb{F})$ with $\mathbb{F} \in \{\mathbb{C}, \mathbb{R}\}$. The **Frobenius norm** $\|\cdot\|_F$ is defined by:

$$\|A\|_F \triangleq \sum_{i,j} |a_{i,j}|^2 = \text{tr}(A^*A)$$

By Cauchy-Schwarz inequality, we see that the Frobenius norm:

$$\|AB\|^2 = \sum_{i,j} \left(\sum_k |A_{i,k} B_{k,j}| \right)^2$$

$$\leq \sum_{i,j} \left(\sum_k |A_{i,k}|^2 \right) \left(\sum_k |B_{k,j}|^2 \right)$$

$$= \sum_{i,j,k,l} |A_{i,k}|^2 |B_{j,l}|^2 = \sum_{i,k} |A_{i,k}|^2 \sum_{j,l} |B_{j,l}|^2 = \|A\|^2 \|B\|^2$$

is submultiplicative. Therefore, the **matrix exponential**, which we know take the form

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

make e^{At} converges uniformly in $t \in \mathbb{R}$ on any bounded interval. Because of such, we may perform term-by-term differentiation and get

$$\frac{de^{At}}{dt} = Ae^{At}$$

This implies that for any first order linear system of ODE:

$$\begin{cases} y'(t) = Ay(t) \textbf{ (DE)} \\ y(0) = y_0 \textbf{ (IV)} \end{cases}$$

by ODE uniqueness theorem, the solution must have the closed form:

$$y(t) \triangleq e^{At} y_0$$

Moreover, because every general solution to $y' = Ay$ has an initial value at 0, we now see that the general solution space is just the column space of e^{At}. We close this section by showing how to compute the matrix exponential using Jordan decomposition.

Given square matrix $A \in M_{n \times n}(\mathbb{F})$ with Jordan decomposition:

$$A = QJQ^{-1} = Q(S + N)Q^{-1} \text{ where diagonal } S \text{ and nilpotent } N \text{ commutes.}$$

Because the series $\sum \frac{X^n}{n!}$ always absolutely converge, and because S and N commute, we may copy the argument of Merten's Theorem for Cauchy product to see that

$$e^A = Qe^{S+N}Q^{-1} = Q(e^S e^N)Q^{-1}$$

15

1.5 Later

Example 1.5.1.

$$\exp \begin{bmatrix} t\lambda & t & 0 \\ 0 & t\lambda & t \\ 0 & 0 & t\lambda \end{bmatrix} = \begin{pmatrix} e^{t\lambda} & te^{t\lambda} & \frac{t^2 e^{t\lambda}}{2} \\ 0 & e^{t\lambda} & te^{t\lambda} \\ 0 & 0 & e^{t\lambda} \end{pmatrix}$$

Because of such

$$e^{\lambda I + N} = \sum_{k=0}^{n-1} \frac{e^\lambda}{k!} N^k$$

$$\sin(\lambda I + N) = \sum_{k=0}^{n-1} \frac{}{k!}$$

1.6 Powers of Jordan blocks

In this section, we denote Jordan block by $J(\lambda, n)$. For example, by $J(\sqrt{5}, 2)$ we mean:

$$\begin{pmatrix} \sqrt{5} & 1 \\ 0 & \sqrt{5} \end{pmatrix}$$

This section only prove one result:

Theorem 1.6.1. (Powers of Jordan blocks) Let $\mathbb{F}$ has zero characteristic, $\lambda \in \mathbb{F}$, and $n, t \in \mathbb{N}$. Then:

$$(J(\lambda, n))^t \text{ has the Jordan form: } \begin{cases} J(\lambda^t, n) & \text{if } \lambda \neq 0 \\ \begin{pmatrix} J(0, n-t+1) & 0 \\ 0 & 0 \end{pmatrix} & \text{if } \lambda = 0 \text{ and } t < n \\ 0 & \text{if } \lambda = 0 \text{ and } t \geq n \end{cases}$$

Proof. (iii) is clear from computation. (ii) is also clear from computation. Indeed, if $t < n$, then:

$$J(0, n)^t = P \begin{pmatrix} J(0, n-t+1) & 0 \\ 0 & 0 \end{pmatrix} P^{-1}$$

where

$$P = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_{t+1} & \mathbf{e}_{t+2} & \cdots & \mathbf{e}_n & \mathbf{e}_2 & \mathbf{e}_3 & \cdots \mathbf{e}_t \end{bmatrix}$$

We now prove (i). For simplicity, from now on we write J in place of $J(\lambda, n) = I + \lambda N$. Clearly, J^t is always an upper triangular matrix whose diagonal entries are all λ^t, and its eigenvalues in $\overline{\mathbb{F}}$ are all λ^t. We are required to prove that J^t has only one Jordan block, i.e., $(J^t - \lambda^t I)^{n-1} \neq 0$. Using binomial theorem, we may compute:

$$J^t = (N + \lambda I)^t = \sum_{k=0}^{t} \binom{t}{k} \lambda^{t-k} N^k$$

This implies:

$$J^t - \lambda^t I = a_{n-1} N^{n-1} + \cdots + a_1 N^1, \quad \text{where } a_i \triangleq \binom{t}{i} \lambda^{t-i}$$

which implies:

$$(J^t - \lambda^t I)^{n-1} = a_1^{n-1} N^{n-1} \neq 0$$

∎

Clearly, Theorem 1.6.1 give us the **Spectral Theorem for Polynomials**, whatever version it is for you.

17

1.7 Spectral theorem

In this section, $\mathbb{F}$ is always a subfield of $\mathbb{C}$. Let V be a $\mathbb{F}$-vector space. By a **norm** on V, we mean some **positive-definite** functional $\| \cdot \| : V \to \mathbb{R}$ that satisfies **absolute homogeneity** and **triangle inequality**. In this context, for $\| \cdot \|$ to be positive-definite, it must satisfy

$$\|v\| = 0 \implies v = 0$$

Observing that

$$\|0\| = \|0 + v\| \leq \|0\| + \|v\|, \quad \text{for all } v \in V$$

we see norm must also be nonnegative. By an **inner product** on V, we mean some **positive-definite** map $\langle \cdot, \cdot \rangle : V^2 \to \mathbb{F}$ that is linear in the first argument and satisfies **conjugate symmetry**:

$$\langle v, w \rangle = \overline{\langle w, v \rangle} \quad \text{for all } v, w \in V$$

In this context, for $\langle \cdot, \cdot \rangle : V^2 \to \mathbb{F}$ to be positive-definite, it has to satisfy

$$\langle v, v \rangle > 0 \quad \text{for all } v \neq 0$$

Let $M \in \mathrm{M}_n(\mathbb{F})$. We say M is **Hermitian** if $M = M^*$. If M is Hermitian, clearly $\mathbf{x}^* M \mathbf{x} \in \mathbb{C}$ must be real for all $\mathbf{x} \in \mathbb{F}^n$. We say M is **positive-definite** if M is Hermitian and $\mathbf{x}^* M \mathbf{x} > 0$ for all nonzero $\mathbf{x} \in \mathbb{F}^n$.

Let W be an inner product $\mathbb{F}$-vector space. Let $\{w_1, \ldots, w_n\}$ be a basis. The matrix $M \in \mathrm{M}_n(\mathbb{F})$ defined by $G_{i,j} \triangleq \langle w_i, w_j \rangle$ is called the **Gram matrix**. G is clearly Hermitian. To see it is moreover positive-definite, simply compute

$$\mathbf{x}^* G \mathbf{x} = \left\langle \sum_i \overline{x_i} w_i, \sum_i \overline{x_i} w_i \right\rangle > 0, \quad \text{for all } \mathbf{x} \in \mathbb{F}^n$$

Note that clearly we have

$$\left\langle \sum \overline{c_i} w_i, \sum \overline{d_i} w_i \right\rangle = \mathbf{c}^* G \mathbf{d}$$

Note that if $M \in \mathrm{M}_n(\mathbb{F})$ is positive definite, then M define an inner product on $\mathbb{F}^n$ by

$$\langle \mathbf{x}, \mathbf{y} \rangle_M = \mathbf{y}^* M \mathbf{x}$$

With an inner product equipped on V, we may discuss the **orthogonality** of vectors. Given a countable set $\{v_1, v_2, v_3, \ldots\}$ of non-zero vectors, we use the term **Gram-Schmidt process** to refer to the process of defining

$$e_n \triangleq v_n - \frac{\langle v_n, e_{n-1}\rangle}{\|e_{n-1}\|^2}e_{n-1} - \cdots - \frac{\langle v_n, e_1\rangle}{\|e_1\|^2}e_1$$

so that $\{e_1, e_2, e_3, \ldots\}$ become orthogonal while

$$\mathrm{span}(v_1, \ldots, v_n) = \mathrm{span}(e_1, \ldots, e_n) \quad \text{for all } n.$$

Although implicit, Gram-Schmidt process may be considered one of the most important notions in inner product space. For example, the Gram-Schmidt process, together with **Pythagorean Theorem**, can be used to prove the **Cauchy-Schwarz Inequality**, which in term shows that the functional $f : V \to \mathbb{R}$ induced by

$$f(v) \triangleq \sqrt{\langle v, v\rangle}$$

indeed satisfies triangle inequality, thus forming a norm. In addition, given an endomorphism T of some finite-dimensional complex vector space V, by applying Gram-Schmidt process to a Jordan basis, we get an orthogonal basis under which T becomes an upper triangular matrix, called its **Schur's form**.

It should be noted that putting an inner product on a vector space only allow better knowledge. For example, a set of nonzero orthogonal vectors must be linearly independent. This is critical when required to prove an inner product space V has some orthogonal decomposition

$$V = W_1 \oplus \cdots \oplus W_n$$

since the direct sum part would be implied by that they are orthogonal.

Let V, W be two inner product space over $\mathbb{F}$, and let $F : V \to W$ be some linear map. If some linear map $F^\dagger : W \to V$ satisfies

$$\langle Fv, w\rangle = \langle v, F^\dagger w\rangle \quad \text{for all } v \in V, w \in W,$$

we say $F^\dagger$ is an **adjoint** of F. If both V and W are finite-dimensional, then $F^\dagger$ exists uniquely: Given orthogonal bases $\{e_1, \ldots, e_n\} \subseteq V, \{f_1, \ldots, f_m\} \subseteq W$, the adjoint $F^\dagger$ is uniquely defined by:

$$\overline{B_{j,i}} = \frac{\langle e_i, e_i\rangle}{\langle f_j, f_j\rangle}A_{i,j}, \quad \text{where } B = [F^\dagger]_{f_1,\ldots,f_m}^{e_1,\ldots,e_n} \text{ and } A = [F]_{e_1,\ldots,e_n}^{f_1,\ldots,f_m}$$

19

Therefore, with respect to orthonormal basis, the matrix representation of $F^\dagger$ is simply the complex conjugation of that of F. It is then obvious that the double adjoint of a linear map between two finite-dimensional inner product spaces over $\mathbb{F}$ is itself.

Given finite dimensional $\mathbb{C}$-vector space V and $T \in \text{End}(V)$, if T is **normal**, i.e. T commutes with its adjoint, then by direct computation, we see that its Schur's form must be diagonal. This result is called **Spectral Theorem for complex finite-dimensional vector space**. If V is over $\mathbb{R}$ instead of $\mathbb{C}$, then T being normal is not enough for T to be orthogonally diagonalized, since we do not have Schur's decomposition in the first place.[4]

Theorem 1.7.1. (Spectral Theorem for real finite-dimensional vector space) Let V be a finite-dimensional real vector space, and let $T \in \text{End}(V)$. If T is **self-adjoint**, then there exists an orthogonal eigenbasis for V with respect to T.

Proof. The is proved by induction on the dimension of V. The two base cases correspond to the 0-dimensional and 1-dimensional cases, both of which are trivial. We now prove the inductive case.

Let $v^\perp$ denote the space of vectors orthogonal to v. Because T is self-adjoint, if v is an eigenvector of T, then $v^\perp$ is stable under T. This reduce the problem into proving the existence of an eigenvector.

Let f be some real polynomial that sends T to zero. By Fundamental Theorem of Algebra, we may write

$$0 = f(T) = (T^2 + b_1 T + c_1 I) \cdots (T^2 + b_n T + c_n I)(T - \lambda_1 I) \cdots (T - \lambda_m I)$$

for some real b_i, c_i, λ_i such that $b_i^2 < 4c_i$ for all i. Because T is self-adjoint, for each i, by Cauchy-Schwarz inequality, we may compute for all $v \neq 0$ that

$$\begin{aligned}
\langle (T^2 + b_i T + c_i I)v, v \rangle &= \langle T^2 v, v \rangle + b_i \langle Tv, v \rangle + c_i \langle v, v \rangle \\
&= \langle Tv, Tv \rangle + b_i \langle Tv, v \rangle + c_i \|v\|^2 \\
&\geq \|Tv\|^2 - b_i \|Tv\| \cdot \|v\| + c_i \|v\|^2 \\
&= \left(\|Tv\| - \frac{b_i \|v\|}{2} \right)^2 + \left(c_i - \frac{b_i^2}{4} \right) \|v\|^2 > 0
\end{aligned}$$

This implies $T^2 + b_i T + c_i I$ are invertible, which implies for some j, the map $T - \lambda_j I$ is not invertible, i.e., for some j, the real number λ_j is an eigenvalue. ∎

---

[4]Actually, we do have Schur's decomposition in the sense that we may first decompose a real matrix into a complex Jordan matrix, and then applying the Gram-Schmidt process. Yet, because the matrix representation of adjoint of T when the underlying field is $\mathbb{R}$ is just the transpose of that of T instead of the conjugate transpose, direct computation yield nothing.

Recall that earlier we proved that for all finite dimensional $\mathbb{F}$-vector space V and $T \in \text{End}(V)$ we have:

$$V = \text{Ker}(T^k) \oplus \text{Im}(T^k), \quad \text{for all } k \geq n$$

where n is the smallest natural number that makes $\text{Ker}(T^n) = \text{Ker}(T^{n+1})$. The spectral property in fact greatly improve this result. If T satisfies the spectral property, then we know $n = 1$, as spectral property guarantee diagonalizablity. Moreover, the decomposition would clearly always be orthogonal for any k.

Theorem 1.7.2. (Classification of quadratic form via spectral theorem) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. If $F \in \mathbb{F}[x_1, \ldots, x_n]$ is a homogeneous polynomial of degree 2, then there exists some $\varphi \in \text{End}(\mathbb{F}^n)$ such that

$$F = \left(\sum_{i=1}^{n} c_i x_i^2 \right) \circ \varphi, \quad \text{for some } c_i \in \mathbb{F}$$

Proof. Let $A \in M_n(\mathbb{F})$ be a square matrix such that:

$$F = \sum_{i,j} a_{i,j} x_i x_j$$

Clearly, we have:

$$F(x_1, \ldots, x_n) = \mathbf{x}^t A \mathbf{x}$$

By spectral theorem, there exists some $P \in M_n(\mathbb{F})$ and diagonal $D \in M_n(\mathbb{F})$ such that $A = P^t D P$. The proof then follows from letting diag $D = \{c_1, \ldots, c_n\}$ and $\varphi(\mathbf{x}) \triangleq P\mathbf{x}$. $\blacksquare$

1.8 SVD

Let V, W be two finite-dimensional inner product space over $\mathbb{F}$, where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Given $F \in \mathrm{Hom}(V, W)$, clearly $F^\dagger F \in \mathrm{End}(V)$ is self-adjoint. Therefore, by Spectral Theorem, there exists some orthonormal eigenbasis α of V with respect to $F^\dagger F$. Because for every pair $e_1, e_2 \in \alpha$, whether $F(v_1), F(v_2)$ are both nonzero or not, we have:

$$\langle F(e_1), F(e_2) \rangle = \langle e_1, F^\dagger F(e_2) \rangle = 0$$

By deleting the zeros and negation of some elements, we may extend $\{F(e) \in W : e \in \alpha\}$ to an orthonormal basis β for W such that every non-zero entry of the $\dim(V)$-by-$\dim(W)$ $\mathbb{C}$-matrix $[F]_\alpha^\beta$ is a positive real number and lies in the diagonal. This process is called the **singular value decomposition of F**.

Polar decomposition. Let $A \in \mathrm{M}_n(\mathbb{F})$ with $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. We are concerned whether there exists semi positive definite $S \in \mathrm{M}_n(\mathbb{F})$ and orthogonal (unitary) $P \in \mathrm{M}_n(\mathbb{F})$ such that $A = PS$ or $A = SP$, for the sake of giving the decomposition $z = xe^{i\theta}$ a matrix analog with $S = x$ and $P = e^{i\theta}$. SVD shows that such decomposition always exists.

1.9 Centralizer

Let $\mathbb{F}$ be a field and $A \in M_n(\mathbb{F})$. Clearly, its **centralizer**

$$C(A) \triangleq \{B \in M_n(\mathbb{F}) : AB = BA\}$$

forms a $\mathbb{F}$-vector subspace of $M_n(\mathbb{F})$. If A admits an Jordan form in $\mathbb{F}$, we may give a formula for the dimension of $C(A)$ depending on it:

Theorem 1.9.1. (Dimension of centralizer of matrix that has Jordan form) If $A \in M_n(\mathbb{F})$ has the Jordan form:

$$A = J(\lambda_1, d_{1,1} \oplus \cdots \oplus d_{1,k_1}) \oplus \cdots \oplus J(\lambda_t, d_{t,1} \oplus \cdots \oplus d_{t,k_t}), \quad \lambda_i \in \mathbb{F}$$

where $d_{i,1} \geq \cdots \geq d_{i,k_i}$ for all i. Then:

$$\dim(C(A)) = \sum_{l=1}^{t} \sum_{i=1}^{k_l} \sum_{j=1}^{k_l} \min\{d_{l,i}, d_{l,j}\}$$

$$\dim(C(A)) = \sum_{i=1}^{t} \sum_{j=1}^{k_i} (2j-1)(d_{i,j} - 1)$$

Proof. Denote $V \triangleq \mathbb{F}^n$. Give V a $\mathbb{F}[T]$-module structure, where $\mathbb{F}[T]$ is a formal polynomial ring, by letting $T \cdot v \triangleq Av$. Clearly, $C(A)$ is isomorphic to $\mathrm{End}_{\mathbb{F}[T]}(V)$ as a $\mathbb{C}$-vector space. Because

$$\mathrm{End}_{\mathbb{F}[T]}(V) \cong \bigoplus_{(\lambda,d),(\mu,e)} \mathrm{Hom}_{\mathbb{F}[T]}\left(J(\lambda, d), J(\mu, e)\right)$$

as $\mathbb{F}[T]$-module, we only have to show:

(I) $\mathrm{Hom}_{\mathbb{F}[T]}\left(J(\lambda, d), J(\mu, e)\right) = 0, \quad$ if $\lambda \neq \mu$.

(II) $\dim_{\mathbb{F}}\left(\mathrm{Hom}_{\mathbb{F}[T]}\left(J(\lambda, d), J(\lambda, e)\right)\right) = \min\{d, e\}$.

(III) $\sum_{i=1}^{k_l} \sum_{j=1}^{k_l} \min\{d_{l,i}, d_{l,j}\} = \sum_{j=1}^{k_l}(2j-1)(d_{l,j}-1), \quad$ for all $l \in \{1, \ldots, t\}$.

The purely combinatorial fact (III) is easy to check. We now prove (I). Let $v \in J(\lambda, d)$ and $f \in \mathrm{Hom}_{\mathbb{F}[T]}\left(J(\lambda, d), J(\mu, e)\right)$. We are required to prove $f(v) = 0$. Let $\{w_1, \ldots, w_e\} \subseteq V$ be the Jordan basis for $J(\mu, e)$, i.e.,

$$\left[A|_{J(\mu,e)}\right]_{\{w_1, \ldots, w_e\}} = \begin{pmatrix} \mu & 1 & & \\ & \mu & 1 & \\ & & \ddots & \ddots \\ & & & \mu \end{pmatrix}$$

23

We may uniquely write $f(v) = \sum c_i w_i$. Because f is a $\mathbb{F}[T]$-module homomorphism, we know:

$$\sum c_i (A - \lambda I)^d w_i = f((A - \lambda I)^d v) = f(0) = 0 \tag{1.8}$$

Let $\mathcal{M}$ denote the $e \times e$-matrix:

$$\left[(A - \lambda I)^d |_{J(\mu, e)} \right]_{\{w_1, \ldots, w_e\}}$$

Because $\mathcal{M}$ is upper triangular whose diagonals are all nonzero[5], equation 1.8 implies $c_e = 0$. Repeating the same process, one see $c_i = 0$ for all i, proving (I). It remains to prove (II). Denote $t \triangleq \min\{d, e\}$, and let $\{v_1, \ldots, v_d\}$ and $\{w_1, \ldots, w_e\}$ respectively be the Jordan bases of $J(\lambda, d)$ and $J(\lambda, e)$. Let $\mathcal{N} \subseteq M_{d \times e}(\mathbb{F})$ be the $\mathbb{F}$-vector subspace of matrices of the form:

$$\begin{pmatrix} \mathcal{O}_1 & \mathcal{K} \\ \mathcal{O}_2 & \mathcal{O}_3 \end{pmatrix}$$

where t-by-$(d-t)$ matrix $\mathcal{O}_1$, $(e-t)$-by-$(d-t)$ matrix $\mathcal{O}_2$, and $(e-t)$-by-t matrix $\mathcal{O}$ are all zero, while $t \times t$-matrix $\mathcal{K}$ is upper triangular with $\mathcal{K}_{i,j} = \mathcal{K}_{i+k, j+k}$ for all i, j, k.

To show $\dim_{\mathbb{F}} \left(\mathrm{Hom}_{\mathbb{F}[T]}(J(\lambda, d)), J(\lambda, e) \right) = t$, we only have to show that the $\mathbb{F}$-vector space homomorphism $\phi : \mathrm{Hom}_{\mathbb{F}[T]} \left(J(\lambda, d), J(\lambda, e) \right) \to M_{d \times e}(\mathbb{F})$ defined by

$$\phi(f) \triangleq [f]_{\{v_1, \ldots, v_d\}}^{\{w_1, \ldots, w_e\}}$$

forms a $\mathbb{F}$-vector space isomorphism between $\mathrm{Hom}_{\mathbb{F}[T]} \left(J(\lambda, d), J(\lambda, e) \right)$ and $\mathcal{N}$. This requires proving three statements:

(S_1) ϕ is injective.

(S_2) $\mathrm{Im}(\phi) \subseteq \mathcal{N}$.

(S_3) $\mathrm{Im}(\phi) \supseteq \mathcal{N}$.

The injectivity of ϕ is clear from the definition. We now show $\mathrm{Im}(\phi) \subseteq \mathcal{N}$. Let $f \in \mathrm{Hom}_{\mathbb{F}[T]} \left(J(\lambda, d), J(\lambda, e) \right)$. Write $f(v_d) \triangleq \sum b_i w_i$. Because f is a $\mathbb{F}[T]$-module homomorphism, we know for all $i \leq d - 1$ that:

$$f(v_{d-i}) = f\left((A - \lambda I)^i v_d \right) = (A - \lambda I)^i f(v_d) = \sum_{j=1}^{e} b_j (A - \lambda I)^i w_j = \sum_{j=i+1}^{e} b_j w_{j-i} \tag{1.9}$$

---

[5]They are all $(\mu - \lambda)^d$.

24

Moreover, since $(A - \lambda I)^d v_d = 0$, we also know

$$\sum_{i=1}^{e-d} b_{i+d} w_i = (A - \lambda I)^d f(v_d) = f\left((A - \lambda I)^d v_d\right) = 0$$

which implies $b_{t+i} = 0$ for all i. This together with the fact that equation 1.9 holds for all $i \leq d - 1$ shows that $\text{Im}(\phi) \subseteq \mathcal{N}$. Let $f \in \text{Hom}_{\mathbb{F}}\left(J(\lambda, d), J(\lambda, e)\right)$ satisfies:

$$[f]_{\{v_1, \dots, v_d\}}^{\{w_1, \dots, w_e\}} \in \mathcal{N}$$

To show $\text{Im}(\phi) \supseteq \mathcal{N}$, we only have to show that f lies in $\text{Hom}_{\mathbb{F}[T]}\left(J(\lambda, d), J(\lambda, e)\right)$, i.e., f is moreover a $\mathbb{F}[T]$-module homomorphism. To prove that f forms a $\mathbb{F}[T]$-module homomorphism, clearly we only have to prove $f \circ A|_{J(\lambda,d)} = A|_{J(\lambda,e)} \circ f$, which can be easily checked by direct matrix computation. ∎

1.10 Cyclic subspace

Let V be a finite dimensional $\mathbb{F}$-vector space and $T \in \text{End}(V)$. Clearly, we may give V a $\mathbb{F}[x]$-module structure by

$$p(x)v \triangleq p(T)v$$

Clearly, $\mathbb{F}[x]$-submodules of V then correspond to T-invariant subspaces of V. Let $v \in V$. The **T-cyclic subspace of V generated by** v is then the $\mathbb{F}[x]$-submodule of V generated by v. We say v is a **T-cyclic vector** if V as a $\mathbb{F}[x]$-module is generated by v, and we say V is **T-cyclic** if V has some T-cyclic vector.

Example 1.10.1. Consider

$$T \triangleq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_2(\mathbb{R})$$

Such space is T-cyclic, but clearly not any nontrivial vector is T-cyclic.

Recall that **structure theorem for finitely generated modules over PID** says that for any finitely generated module M over a PID R, we have a unique R-module isomorphism:

$$M \cong R/(a_1) \oplus \cdots \oplus R/(a_n), \quad \text{where } a_1 \mid \cdots \mid a_n \in R \text{ are non-units}$$

Clearly, if $\{a_1, \ldots, a_n\}$ contains 0, then they must locate at the end of the sequence. However, in our case, where $M \triangleq V$ and $R \triangleq \mathbb{F}[x]$:

$$V \cong \mathbb{F}[x]/(a_1) \oplus \cdots \oplus \mathbb{F}[x]/(a_n), \quad \text{where } a_1 \mid \cdots \mid a_n \in \mathbb{F}[x] \text{ are not nonzero constant}$$

we see that if $a_n = 0$, then $R/(a_n)$ as a $\mathbb{F}$-vector space must be infinite dimensional. Therefore, every a_i has positive degree. Noting that

(i) Restriction of $f(T)$ onto the $\mathbb{F}$-vector space $R/(a_1)$ is 0.

(ii) $f \in \mathbb{F}[x]$ annihilate $R/(a_1)$

(iii) $a_1 \mid f$.

are equivalent, we see that the minimal polynomial of T is exactly the least common multiple of a_i, i.e., a_n. Because each $R/(a_i)$ is a T-invariant subspace, we know the characteristic polynomial of T is the product of the characteristic polynomials of the restriction of T onto $R/(a_i)$.

Let v correspond to $1 \in R/(a_1)$. We see that $\{v, \ldots, T^{(\deg a_1)-1}v\}$ must form a basis for $R/(a_1)$. Considering v, clearly a_1 must be the minimal polynomial of restriction of T. It

then follows from Cayley-Hamilton that a_1 is also the characteristic polynomial of restriction of T.

When written with respect to this basis, we see that T has the matrix representation of the companion matrix $C(a_1)$:

$$
C \triangleq \begin{pmatrix}
0 & 0 & \cdots & \cdots & -c_d \\
1 & 0 & 0 & \cdots & -c_{d-1} \\
0 & 1 & \ddots & \ddots & \vdots \\
\vdots & \vdots & \ddots & 0 & \vdots \\
0 & \cdots & \cdots & 1 & -c_1
\end{pmatrix}
$$

if we take

$$
a_1 = x^d + c_1 x^{d-1} + \cdots + c_d
$$

We have just shown that the companion matrix has the desired minimal polynomial and characteristic polynomial.

1.11   Primary decomposition theorem

Theorem 1.11.1. (Primary decomposition theorem) Let V be a finite dimensional $\mathbb{F}$-vector space, and $T \in \mathrm{End}(V)$. Consider its primary decomposition

$$V = V_1 \oplus \cdots \oplus V_n$$

There projection $E_i : V \to V_i$ onto V_i along the rest of the summands is a polynomial in T, and for every T-invariant subspace W, we have

$$W = W \cap V_1 \oplus \cdots \oplus W \cap V_n$$

Proof. For the first part, let the minimal polynomial of T be $p_1^{e_1} \cdots p_n^{e_n}$ with p_i irreducible. By Chinese remainder theorem, there exists a polynomial f_1 such that

$$f_1 \equiv 1 \pmod{p_1^{e_1}} \quad \text{and} \quad f_1 \equiv 0 \pmod{p_i^{e_i}}, \quad \text{for all } i \geq 2$$

Such f_i then clearly suffice as the projection. The second part is then the corollary of the first part. $\blacksquare$

1.12 Equivalent definition for cyclic

Equivalent Definition 1.12.1. (cyclic vector space) Let V be a finite dimensional $\mathbb{F}$-vector space, and $T \in \mathrm{End}(V)$. Then the followings are equivalent:

(i) T is cyclic.

(ii) $m_T = c_T$.

(iii) $C(T) = \mathbb{F}[T]$

Proof. (i) $\Longleftrightarrow$ (ii) follows from invariant form of structure theorem for finitely generated modules over PID. Clearly we have (i) $\Longrightarrow$ (iii). We now prove (iii) $\Longrightarrow$ (i). Assume T isn't cyclic for a contradiction. Then we may write

$$V \cong \mathbb{F}[x]/(a_1) \oplus \cdots \oplus \mathbb{F}[x]/(a_n) \oplus \mathbb{F}[x]/(m_T)$$

where $a_n \mid m_T$. Consider the $\mathbb{F}[x]$-module endomorphism that maps $(0, \ldots, 1)$ to $(0, \ldots, 1)$. Because such $\mathbb{F}[x]$-module homomorphism is a polynomial in T, we know it must be of the form $q m_T + 1$. However, if we require it to maps $(0, \ldots, 1, 0)$ to $(0, \ldots, 0, 0)$, then we see a contradiction. $\blacksquare$

1.13 Another proof for Jordan form

Let V be a finite dimensional $\mathbb{F}$-vector space and $T \in \mathrm{End}_{\mathbb{F}}(V)$. Give V the $\mathbb{F}[x]$-module structure by letting x acts T.

Theorem 1.13.1. (Structure theorem for finitely generated modules over PID, primary decomposition form) Let M be a finitely generated modules over PID R. Then M can be uniquely decomposed into sum of R quotient out primary ideals:

$$M \cong R/(f_1) \oplus \cdots R \oplus /(f_n), \quad \text{where } (f_i) \text{ are primary}$$

Proof. Omitted. $\blacksquare$

By structure theorem for finitely generated modules over PID, we may write

$$V \cong \bigoplus \mathbb{F}[x]/(f_i), \quad \text{as } \mathbb{F}[x]\text{-modules}$$

where $(f_i) \subseteq \mathbb{F}[x]$ are primary. Because (f_i) is primary, we know each f_i is a power of an irreducible. Clearly the least common multiples of all f_i is the minimal polynomial of T. Therefore if the minimal polynomial of T splits (over $\mathbb{F}$), then f_i are all linear. In such case, we may write

$$V \cong \bigoplus_{\lambda} \bigoplus \mathbb{F}[x]/\left((x-\lambda)^n\right)$$

One then see that

$$\left\{(x-\lambda)^{n-1}, \ldots, 1\right\} \text{ forms the Jordan basis of each block}$$

Let each eigenvalue λ_i has index d_i. Moreover, we see that after getting the Jordan Chevalley decomposition $T = D + N$ from Jordan block, because by Chinese remainder theorem, there exists a unique $g \in \mathbb{F}[x]$ with $\deg g < \sum d_i$ such that

$$g \equiv \lambda_i \pmod{(x-\lambda_i)^{d_i}}, \quad \text{for all } i$$

We see $g(T) = D$. Moreover we may show the existence of some f that makes $f(T) = 0$ and have zero constant term. If 0 is an eigenvalue, then $g \equiv 0 \pmod{x^d}$ so g already suffices. If 0 is not an eigenvalues, adding appropriate amount of the minimal polynomial to g suffices.

Note that the uniqueness of the Jordan-Chevalley decomposition follows from showing that two commuting diagonalizable matrix are simultaneously diagonalizable. Note that the uniqueness of Jordan form is guaranteed by the uniqueness in the structure theorem since we may reverse engineer a Jordan form $J(\lambda, d)$ back to showing that it is isomorphic as a $\mathbb{F}[x]$-module to $\mathbb{F}[x]/(x-\lambda)^d$ by invoking structure theorem for finitely generated modules

over PID on the $\mathbb{F}[x]$-module $J(\lambda, d)$ again and consider the minimal polynomial.

By structure theorem for finitely generated modules over PID, we may write

$$V \cong \mathbb{F}[x]/(f_1^{n_1}) \oplus \cdots \oplus \mathbb{F}[x]/(f_1^{n_{m_1}}) \oplus \cdots \oplus \mathbb{F}[x]/(f_k^{n_k}) \oplus \cdots \oplus \mathbb{F}[x]/(f_k^{n_{m_k}})$$

where f_i are distinct irreducible polynomials. We call

$$\mathbb{F}[x]/(f_i^{n_1}) \oplus \cdots \oplus \mathbb{F}[x]/(f_i^{n_{m_i}})$$

the primary part of V. Let n_{m_i} be maximal. By Chinese remainder theorem, there exists some polynomial f such that

$$f \equiv 1 \pmod{f_1^{n_{m_1}}} \quad \text{and} \quad f \equiv 0 \pmod{f_i^{n_{m_i}}} \text{ for } i \geq 2$$

Then $f(T)$ is a projection of V onto its first primary part along the direct sum of the other primary parts.

1.14 Simultaneous Triangulation

Theorem 1.14.1. (triangulable if and only if minimal polynomial splits if and only if Jordan form exists if and only if one step extension property) Let V be a finite dimensional $\mathbb{F}$-vector space, and let $T \in \text{End}(V)$. Then the followings are equivalent:

(i) T is triangulable.

(ii) minimal polynomial of T splits.

(iii) For every proper subspace W, there exists some $v \in V - W$ such that Tv is in the span of W and v.

Proof. (i) $\iff$ (ii) is clear. (iii) $\implies$ (i) is also clear. We now show (i) $\implies$ (iii). Let $\{e_1, \ldots, e_n\}$ be a basis that upper triangularize T. Because W is proper, we may let $i \in \{1, \ldots, n\}$ be the smallest index such that $e_i \notin W$. Such e_i then clearly suffice. ∎

Theorem 1.14.2. (One step extension property of commuting family of triangulable operators) Let V be a finite dimensional $\mathbb{F}$-vector space, and let elements of $\mathcal{F} \subseteq \text{End}(V)$ all be triangulable. If $\mathcal{F}$ commute and W is a proper $\mathcal{F}$-invariant subspace, then there is some $v \in V - W$ such that Tv is in the span of W and v for all $T \in \mathcal{F}$.

Proof. Clearly, we may suppose $\mathcal{F}$ is finite $\{T_1, \ldots, T_n\}$. By theorem 1.14.1, there exists some $v_1 \in V - W$ and $c_1 \in \mathbb{F}$ such that $(T_1 - c_1 I)v_1 \in W$. Therefore, the vector space $V_1 \triangleq \{v \in V : (T_1 - c_1 I)v \in W\}$ is strictly larger than W.

Because W is $\mathcal{F}$-invariant, we see V_1 is also $\mathcal{F}$-invariant. Clearly $T_2|_{V_1}$ is also triangulable. Therefore by theorem 1.14.1, there exists some $v_2 \in V_1 - W$ and $c_2 \in \mathbb{F}$ such that $(T_2 - c_2 I)v_2 \in W$. Let $V_2 \triangleq \{v \in V_1 : (T_1 - c_2 I)v \in W\}$. We now see that any vector in $V_n - W$ suffices. ∎

Corollary 1.14.3. (Commuting family of triangulable operators is simultaneously triangulable) Let V be a finite dimensional $\mathbb{F}$-vector space, and let elements of $\mathcal{F} \subseteq \text{End}(V)$ all be triangulable. If $\mathcal{F}$ commute it is simultaneously triangulable.

1.15 T is a constant multiple of the identity

Theorem 1.15.1. (When T preserves every subspace) Let V be a finite dimensional $\mathbb{F}$-vector space, and let $T \in \text{End}(V)$. If every subspace of V is T-invariant, then T is a constant multiple of identity.

Proof. Clearly every vector is an eigenvector. If T has distinct eigenvalues, then clearly we have a contradiction. ∎

Theorem 1.15.2. (When T commutes with every projection) Let V be a finite dimensional $\mathbb{F}$-vector space, and let $T \in \text{End}(V)$. If T commutes with every projection, then T is a constant multiple of identity.

Proof. If T commute with projection $E \in \text{End}(V)$, then clearly both image and kernel of E is T-invariant. This together with theorem 1.15.1 suffices. ∎

Theorem 1.15.3. (When T commute with every diagonalizable operator) Let V be a finite dimensional $\mathbb{F}$-vector space, and let $T \in \text{End}(V)$. If T commutes with every diagonalizable operator, then T is a constant multiple of identity.

Proof. This follows from theorem 1.15.2 and noting that projections are diagonalizable. ∎

1.16   Commutator

Let $A \in M_n(\mathbb{F})$. Consider $L_A \in \text{End}(M_n(\mathbb{F}))$ defined by

$$L_A(B) \triangleq AB$$

Clearly, L_A has the same minimal polynomial as A. So L_A is diagonalizable if and only if A is diagonalizable. If A is diagonalizable, says, $A = PDP^{-1}$, then L_A has the eigenvectors:

$$L_A(PE_{i,j}) = \lambda_i(PE_{i,j})$$

Consider $\text{ad}_A \in \text{End}(M_n(\mathbb{F}))$ defined by

$$\text{ad}_A(B) \triangleq AB - BA$$

then ad_A has eigenvectors:

$$\text{ad}_A(PE_{i,j}P^{-1}) = (\lambda_i - \lambda_j)(PE_{i,j}P^{-1})$$

Let V be a finite dimensional vector space and $T \in \text{End}(V)$. If T is cyclic, then we know

$$V \cong \mathbb{F}[x]/m_T, \quad \text{where } m_T \in \mathbb{F}[x] \text{ is the minimal polynomial of } T$$

Let $G \in \text{End}(V)$ commutes with T, and suppose G maps $1 \in \mathbb{F}[x]/m_T$ to f, then $G = f(T)$.

Let V be a finite dimensional vector space and $T \in \text{End}(V)$. Suppose every $G \in \text{End}(V)$ that commutes with T is a polynomial in T. Assume for a contradiction that T doesn't has a cyclic vector. Then by structure theorem for finitely generated modules over PID, we may write

$$V \cong \mathbb{F}[x]/(h_1) \oplus \cdots \oplus \mathbb{F}[x]/(h_k) \oplus \mathbb{F}[x]/(m_T) \quad \text{with } h_k \mid m_T$$

Consider the module endomorphism G that maps $(0, \ldots, 0, 1, 0)$ to $(0, \ldots, 0, 2, 0)$ and maps $(0, \ldots, 0, 1)$ to $(0, \ldots, 0, 1)$. Because G maps $(0, \ldots, 1)$ to $(0, \ldots, 1)$, we know that if $G = m_T q + 1$ for some $q \in \mathbb{F}[x]$. However, this then implies G doesn't map $(0, \ldots, 0, 1, 0)$ to $(0, \ldots, 0, 2, 0)$

Chapter 2

What you probably need to know

2.1   Tensor Product

In the most general sense, the **tensor product** of two $\mathbb{F}$-vector space V_1, V_2 is merely a $\mathbb{F}$-vector space X associated with a $\mathbb{F}$-bilinear map $g : V_1 \times V_2 \to X$ that satisfies the **universal property** of tensor product: For every $\mathbb{F}$-bilinear map $f : V_1 \times V_2 \to W$, there exists a unique $\mathbb{F}$-linear map $\widetilde{f} : X \to W$ such that $f = \widetilde{f} \circ g$.

It is routine to check that the universal property of tensor product defines vector space up to an isomorphism compatible with itself. More precisely, if $Y, h : V_1 \times V_2 \to Y$ also satisfies the universal property, then there exists an isomorphism $f : X \to Y$ such that $f \circ g = h$.

Because of such, we often write X simply as $V_1 \otimes V_2$ and denote $g(v_1, v_2)$ by $v_1 \otimes v_2$. One common construction of the tensor product is to define $V_1 \otimes V_2$ as the space of $\mathbb{F}$-bilinear maps from $V_1^\vee \times V_2^\vee$ to $\mathbb{F}$ and to define the $\mathbb{F}$-bilinear map by $(v_1 \otimes v_2)(\varphi_1, \varphi_2) \triangleq \varphi_1(v_1)\varphi_2(v_1)$.

At this point, one may already observe that we may restate the universal property of tensor product as follows. Given a finite collection $\{V_1, \ldots, V_n\}$ of $\mathbb{F}$-vector spaces, we say the $\mathbb{F}$-vector space $V_1 \otimes \cdots \otimes V_n$ and the $\mathbb{F}$-multilinear map $\otimes : V_1 \times \cdots \times V_n \to V_1 \otimes \cdots \otimes V_n$ satisfies the universal property if for every $\mathbb{F}$-multilinear map $f : V_1 \times \cdots \times V_n \to W$, there exists a unique $\mathbb{F}$-linear map $\widetilde{f} : V_1 \otimes \cdots \otimes V_n$ such that $\widetilde{f}(v_1 \otimes \cdots \otimes v_n) = f(v_1, \ldots, v_n)$ for every $(v_1, \ldots, v_n) \in V_1 \times \cdots \times V_n$.

Even thought the construction of the tensor product is deemed unimportant by general consensus, here, we make the remark that the construction we gave do show that if B_j are finite bases for V_j, then $\{v_1 \otimes \cdots \otimes v_n \in V_1 \otimes \cdots \otimes V_n : v_j \in B_j \text{ for all } j\}$ forms a basis for $V_1 \otimes \cdots \otimes V_n$. In fact, this together with the universal property may be the only things about tensor product one will ever need to develop other properties of tensor products.

2.2 Exterior Algebra

Let V be some finite-dimensional $\mathbb{F}$-vector space. In this note, by the **exterior n-th power** $\bigwedge^n V$ of V, we mean the space of the antisymmetric multilinear maps from $(V^*)^n$ to $\mathbb{F}$, together with a multilinear map $\wedge : V^n \to \bigwedge^n V$ defined by

$$(v_1 \wedge \cdots \wedge v_n)(\varphi_1, \ldots, \varphi_n) \triangleq \begin{vmatrix} \varphi_1 v_1 & \cdots & \varphi_1 v_n \\ \vdots & \ddots & \vdots \\ \varphi_n v_1 & \cdots & \varphi_n v_n \end{vmatrix}$$

One may check that

(a) $\wedge$ is itself antisymmetric.

(b) If $\{v_1, \ldots, v_k\}$ forms a basis for V, then $\{v_{i_1} \wedge \cdots \wedge v_{i_n} \in \bigwedge^n V : 1 \le i_1 < \cdots < i_n \le k\}$ forms a basis for $\bigwedge^n V$.

(c) $\bigwedge^n V$ also satisfies its **universal property**: For every antisymmetric multilinear map $f : V^n \to W$, there exists a unique linear map $\widetilde{f}$ such that $\widetilde{f} = f \circ \wedge$.

Because of the universal property, after some tedious construction, we see that there exists some unique bilinear map $g : \bigwedge^n V \times \bigwedge^k V$ such that $g(v_1 \wedge \cdots \wedge v_n, v_{n+1} \wedge \cdots \wedge v_{n+k}) = v_1 \wedge \cdots \wedge v_{n+k}$ for all $(v_1, \ldots, v_{n+k}) \in V^{n+k}$. Therefore, when we define the **exterior algebra of V** by

$$\bigwedge V \triangleq \bigoplus_{n=0}^{\infty} \bigwedge^n V$$

where $\bigwedge^0 V = \mathbb{F}$ by convention, we see that $\bigwedge V$ indeed have the structure of an associative $\mathbb{F}$-algebra where the vector multiplication $\wedge : \bigwedge V \times \bigwedge V \to \bigwedge V$ satisfies

$$(v_1 \wedge \cdots \wedge v_n) \wedge (v_{n+1} \wedge \cdots \wedge v_{n+k}) = v_1 \wedge \cdots \wedge v_{n+k} \tag{2.1}$$

for all $(v_1, \ldots, v_{n+k}) \in V^{n+k}$. Again, since both exterior power and exterior algebra are categorical concept, the fore tedious construction doesn't really matter. What matter is the property they always have, for example Equation 2.1, the fact wedge products is antisymmetric, or the fact $\{v_{i_1} \wedge \cdots \wedge v_{i_n} \in \bigwedge^n V : 1 \le i_1 < \cdots < i_n \le k\}$ forms a basis for $\bigwedge^n V$. For future references, we close this section by giving two other important properties of the exterior algebra.

(a) For arbitrary $\sigma \in S_n$ and arbitrary $e_1, \ldots, e_n \in V$, we have

$$e_{\sigma(1)} \wedge \cdots \wedge e_{\sigma(n)} = \text{sgn}(\sigma) e_1 \wedge \cdots \wedge e_n$$

(b) For arbitrary $e_1, \ldots, e_n \in V$ and arbitrary $A \in M_n(\mathbb{F})$, we have

$$f_i = \sum_j A_{i,j} e_j \text{ for all } i \implies f_1 \wedge \cdots \wedge f_n = \det(A) e_1 \wedge \cdots \wedge e_n$$

2.3 Pfaffian

Given an $2n \times 2n$ skew-symmetric matrix A over $\mathbb{F}$, we define its **Pfaffian** to be

$$\text{pf}(A) \triangleq \frac{1}{2^n n!} \sum_{\sigma \in S_{2n}} \text{sgn}(\sigma) \prod_{k=1}^{n} A_{\sigma(2k-1),\sigma(2k)}$$

Theorem 2.3.1. (Equivalent Definition of Pfaffian) If we define $\omega \in \bigwedge^{2n} \mathbb{F}$ by

$$\omega \triangleq \sum_{i,j} A_{i,j} e_i \wedge e_j$$

then

$$\omega^n = 2^n n! \, \text{pf}(A) e_1 \wedge e_2 \wedge \cdots \wedge e_{2n}$$

Proof. Let X be the set of functions that maps $\{1, \ldots, n\}$ into $\{1, \ldots, 2n\} \times \{1, \ldots, 2n\}$ such that when we write $f(k) = (f_1(k), f_2(k))$ for $f \in X$, we have

$$\{f_1(k) : 1 \le k \le n\} \cup \{f_2(k) : 1 \le k \le n\} = \{1, \ldots, 2n\}$$

Because the wedge product is antisymmetric, we have

$$\omega^n = \sum_{f \in X} \left(\prod_{k=1}^{n} A_{f_1(k),f_2(k)} \right) e_{f_1(1)} \wedge e_{f_2(1)} \wedge \cdots \wedge e_{f_n(1)} \wedge e_{f_n(2)}$$

It remains to show

$$\sum_{f \in X} \left(\prod_{k=1}^{n} A_{f_1(k),f_2(k)} \right) e_{f_1(1)} \wedge e_{f_2(1)} \wedge \cdots \wedge e_{f_n(1)} \wedge e_{f_n(2)}$$

$$= \sum_{\sigma \in S_{2n}} \text{sgn}(\sigma) \left(\prod_{k=1}^{n} A_{\sigma(2k-1),\sigma(2k)} \right) e_1 \wedge \cdots \wedge e_{2n}$$

Define $\phi : S_{2n} \to X$ by

$$(\phi(\sigma))(k) \triangleq (\sigma(2k-1), \sigma(2k))$$

Clearly, ϕ is a bijection, and

$$\prod_{k=1}^{n} A_{\sigma(2k-1),\sigma(2k)} = \prod_{k=1}^{n} A_{\phi(\sigma)_1(k),\phi(\sigma)_2(k)}$$

and

$$e_{\phi(\sigma)_1(1)} \wedge e_{\phi(\sigma)_2(1)} \wedge \cdots \wedge e_{\phi(\sigma)_1(n)} \wedge e_{\phi(\sigma)_2(n)} = e_{\sigma(1)} \wedge e_{\sigma(2)} \wedge \cdots \wedge e_{\sigma(2n)} = \text{sgn}(\sigma) e_1 \wedge \cdots \wedge e_{2n}$$

Above equation need to be explained in the section before. ∎

Chapter 3

NTU Math M.A. Program Entrance Exam

3.1 Year 114

Question 1: Smith normal form and Jordan form

Let

$$A \triangleq \begin{pmatrix} 2 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & -1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

Find the Smith normal form of $xI - A \in M_4(\mathbb{R}[x])$ and the Jordan form of A.

Proof. The solution is that $\mathrm{SNF}_{\mathbb{R}[x]}(xI - A) = \mathrm{diag}\left\{1, 1, x - 1, (x-1)^2(x-2)\right\}$ and that

$$J \triangleq \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ and } P \triangleq \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & -1 & -1 & 0 \\ -1 & 0 & 1 & 0 \end{pmatrix}$$

makes $A = PJP^{-1}$. ∎

Let

$$A \triangleq \begin{pmatrix} 3 & -2 & -2 \\ -2 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix}$$

(i) Find the maximum of $\mathbf{x}^t A \mathbf{x}$ among all $\mathbf{x} \in \mathbb{R}^3$ that makes $|\mathbf{x}| = 1$.

(ii) Find the minimum of $Y^t A Y$ among all $Y \in M_{3 \times 2}(\mathbb{R})$ that makes $Y^t Y = I$.

Proof. Because A is symmetric, by spectral theorem for finite dimensional real vector space, we know there exists an orthonormal eigenbasis

$$\left\{ \frac{1}{\sqrt{6}} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$$

Denote them in order by $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, and let $\mathbf{x} = c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 + c_3 \mathbf{e}_3$. Because they form an orthonormal eigenbasis, we know

$$\mathbf{x}^t A \mathbf{x} = 5 c_1^2 - c_2^2 - c_3^2$$

This implies that the solution for (i) is -5 with $\mathbf{x} = \mathbf{e}_1$ and the solution for (ii) is -2 with $Y = \begin{pmatrix} \mathbf{e}_2 & \mathbf{e}_2 \end{pmatrix}$. ∎

Let $A \in M_n(\mathbb{C})$ be invertible. Show that $A^{-1} = p(A)$ for some $p(x) \in \mathbb{C}[x]$.

Proof. We are required to prove the existence of $p(x) \in \mathbb{C}[x]$ such that $p(x) \cdot x - 1 \in \mathbb{C}[x]$ maps A to 0. Because A is invertible, by looking at the Jordan form of A, we know the minimal polynomial m of A has nonzero constant term. Let the constant term of m be -1. We then see $p \triangleq \frac{m+1}{x}$ suffices. ∎

Let V and W be two finite dimensional inner product space over $\mathbb{R}$.

(i) Prove that for each $w \in W$, there exists a unique vector in V denoted by $T^* w$ such that $\langle Tv, w \rangle = \langle v, T^* w \rangle$ for all $v \in V$. Show that T^* is linear.

(ii) Prove that $(\operatorname{Im} T)^\perp = \ker T^*$.

(iii) Prove that $\dim \ker T - \dim \ker T^* = \dim V - \dim W$.

Proof. Let $\{e_1, \ldots, e_n\}, \{u_1, \ldots, u_m\}$ be orthonormal basis of V, W with $\{u_1, \ldots, u_k\}$ be a basis for $\operatorname{Im} T$.

Fix $w \in W$. Write

$$T^* w = c_1 e_1 + \cdots + c_n e_n$$

Because e_i are orthonormal, by premise, we can only have $c_i = \langle Te_i, w \rangle$. We have shown the uniqueness of $T^* w$. To see such $T^* w$ suffices, simply compute

$$\left\langle T \sum d_i e_i, w \right\rangle = \sum d_i \langle Te_i, w \rangle = \sum d_i c_i = \left\langle \sum d_i e_i, \sum c_i e_i \right\rangle = \left\langle \sum d_i e_i, T^* w \right\rangle$$

To see T^* is linear, just observe

$$\langle Te_i, w_1 + w_2 \rangle = \langle Te_i, w_1 \rangle + \langle Te_i, w_2 \rangle$$

Clearly we have $(\operatorname{Im} T)^\perp \supseteq \ker T^*$. Fix $i > k$. To prove the other direction, we are required to prove $u_i \in \ker T^*$. Because $\{u_1, \ldots, u_k\}$ spans $\operatorname{Im} T$ and $\{u_1, \ldots, u_m\}$ forms an orthonormal basis, we may compute

$$T^* u_i = \langle Te_1, u_i \rangle e_1 + \cdots + \langle Te_n, u_i \rangle e_n = 0$$

as desired.

Fix $1 \leq i \leq m$. Compute

$$T^* u_i = \langle Te_1, u_i \rangle e_1 + \cdots + \langle Te_n, u_i \rangle e_n = \sum_j A_{i,j} e_j, \quad \text{where } T(e_j) = \sum_i A_{i,j} u_i$$

We see that T has the matrix representation the transposition of that of T. This implies T^* have the same rank as T. Part (iii) then follows from rank-nullity theorem. ∎

> **Question 5**
>
> Let $\mathbb{F}$ be infinite, V be a finite dimensional $\mathbb{F}$-vector space, and $T \in \operatorname{End}(V)$. Show that
>
> $$V \text{ is } T\text{-cyclic} \iff V \text{ has only finite number of } T\text{-invariant subspaces}$$

Proof. From now on, we denote $\mathbb{F}[x]$ by R and give V a R-module structure:

$$f \cdot w \triangleq f(T) \cdot w$$

Note that T-invariant subspaces of V then correspond to R-submodules of V.

($\implies$): Let $v \in V$ be T-cyclic, and let $m_v \in R$ be the minimal polynomial that annihilate v. Because v is T-cyclic, clearly we have a surjective R-module homomorphism $R \twoheadrightarrow V$ defined by $f \mapsto f \cdot v$. The kernel of such module homomorphism is clearly the ideal generated by m_v. Therefore, by first isomorphism theorem, we have $V \cong R/(m_v)$. By correspondence theorem, R-submodules of $R/(m_v)$ correspond to the ideals of R that contains (m_v). Because R is a PID, this correspond to the dividing class of R that divides m_v, which clearly must be finite.

($\impliedby$): By structure theorem for finitely generated modules over PID, we may write

$$V \cong R/(a_1) \oplus \cdots \oplus R/(a_n), \quad \text{where } a_1 \mid \cdots \mid a_n$$

If $V \cong R/(a_1)$, then the vector $v \in V$ correspond to 1 forms a cyclic vector. Therefore, we only have to prove $n = 1$.

Assume for a contradiction that $n \geq 2$. We claim that the R-submodules

$$M_\lambda \triangleq R \cdot (1 \pmod{a_1} \times \lambda \pmod{a_2})$$

of R-module $R/(a_1) \oplus R/(a_2)$ are distinct for each $\lambda \in \mathbb{F}$. Because $\mathbb{F}$ is infinite, this would lead to a contradiction.

Let $M_\lambda = M_\mu$. This implies

$$(0 \pmod{a_1} \times \lambda - \mu \pmod{a_2}) \in M_\lambda \cap M_\mu$$

Therefore, by definition, there exists $h \in R$ such that

$$h \equiv 0 \pmod{a_1} \quad \text{and} \quad h\lambda \equiv (\lambda - \mu) \pmod{a_2}$$

Write $h = a_1 p$ for some $p \in R$. The latter now give us

$$\lambda a_1 p = (\lambda - \mu) + a_2 r, \quad \text{for some } r \in R$$

Because $a_1 \mid a_2$, we now see $a_1 \mid \lambda - \mu$, which implies $\lambda - \mu = 0$. $\blacksquare$

3.2 Year 113

Let

$$A \triangleq \begin{bmatrix} -2 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Find the Jordan-Chevalley decomposition of A and compute e^A.

Proof. Routine computation give us

$$J = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \text{ and } P = \begin{bmatrix} 0 & 1 & -1 \\ 1 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \text{ and } A = PJP^{-1}$$

Therefore, the Jordan-Chevalley decomposition of A is

$$A = D + N \text{ where } D = P \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} P^{-1} \text{ and } N = P \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} P^{-1}$$

And

$$e^A = P \begin{bmatrix} e & 0 & 0 \\ 0 & e^{-1} & e^{-1} \\ 0 & 0 & e^{-1} \end{bmatrix} P^{-1}$$

∎

Let V be the space of polynomial in x over $\mathbb{R}$ of degree not higher than 2. Define an inner product for V by

$$\langle f, g \rangle \triangleq \int_{-1}^{1} fg$$

Find a polynomial $k(x, t)$ such that

$$f(x) = \int_{-1}^{1} k(x, t) f(t) dt \quad \text{for all } f \in V \tag{3.1}$$

Define $T \in \text{End}(V)$ by

$$T(a_2 x^2 + a_1 x + a_0) \triangleq 2a_2 x + a_1$$

That is, $Tf \triangleq f'$. Find the adjoint of T.

Proof. Because of the linearity of the function:

$$f(x) \mapsto g(x) \triangleq \int_{-1}^{1} k(x,t) f(t) dt$$

an endomorphism of V, for k to satisfy equation 3.1 for all $f \in V$, it only has to satisfies

$$1 = \int_{-1}^{1} k\, dt \text{ and } x = \int_{-1}^{1} kt\, dt \text{ and } x^2 = \int_{-1}^{1} kt^2\, dt \tag{3.2}$$

If we write k in the form

$$k = f_0(x) + f_1(x)t + f_2(x)t^2 + \cdots$$

then equation 3.2 becomes

$$1 = \frac{2f_0(x)}{1} + \frac{2f_2(x)}{3} + \frac{2f_4(x)}{5} + \cdots$$
$$x = \frac{2f_1(x)}{3} + \frac{2f_3(x)}{5} + \frac{2f_5(x)}{7} + \cdots$$
$$x^2 = \frac{2f_0(x)}{3} + \frac{2f_2(x)}{5} + \frac{2f_4(x)}{7} + \cdots$$

Thus, k can be

$$k(x,t) = \left(\frac{9}{8} + \frac{-15}{8}x^2\right) + \left(\frac{3x}{2}\right)t + \left(\frac{-15}{8} + \frac{45}{8}x^2\right)t^2$$

Routine computation give us an orthonormal basis

$$\left\{\sqrt{\frac{1}{2}}, \sqrt{\frac{3}{2}}x, \sqrt{\frac{45}{8}}(x^2 - \frac{1}{3})\right\} \text{ for } V$$

With respect to this ordered basis, the matrix representation of T is

$$\begin{bmatrix} 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{15} \\ 0 & 0 & 0 \end{bmatrix}$$

44

Because the matrix representation of $T^\dagger$ with respect to the same basis is the conjugate transpose of that of T, we may now compute

$$T^\dagger(a_2 x^2 + a_1 x + a_0) = \frac{15a_1}{2}(x^2 - \frac{1}{3}) + (3a_0 + a_2)x$$

■

Question 8: equivalent definition of trace function

Let $f : M_n(\mathbb{R}) \to \mathbb{R}$ be linear. Show that if

(i) $f(AB) = f(BA)$ for all $A, B \in M_n(\mathbb{R})$.   (**symmetry constraint**)

(ii) $f(I) = n$.   (**initial value condition**)

then f is the trace function.

Remark: The hard part is to come up with method 3.3.

Proof. Let $E_{i,j} \in V$ be the matrix whose only nonzero entry is 1, located in the i-th row and j-th column. Because f is linear, we only have to prove

(i) If $i \neq j$, then $f(E_{i,j}) = 0$.   (**f depends only on diagonal**)

(ii) $f(E_{1,1}) = \cdots = f(E_{n,n}) = 1$.

The fact that f depends only on diagonal is only a consequence of the symmetry constraint. Let $i \neq j$. Clearly, we wish to find $A \in M_n(\mathbb{R})$ such that

$$A(E_{i,j}) = 0 \quad \text{and} \quad (E_{i,j})A = E_{i,j} \tag{3.3}$$

The proof then follows from computing:

$$A(E_{i,j}) = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} & A_{;,i} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}$$

$$(E_{i,j})A = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ A_{j,;} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}$$

where $A_{;,i}$ locates at the j-th column, and $A_{j,;}$ locates at the i-th row. Because we have the initial value condition $f(I) = n$, to show $f(E_{1,1}) = \cdots = f(E_{n,n}) = 1$, we only have to

show:

$$f(E_{i,i}) = f(E_{j,j}) \text{ for all } i \neq j$$

which relies on the existences of **permutation matrices**, *square matrices that has exactly one entry of 1 in each row and each column with all other entries 0.* Let $\sigma \in S_n$. Clearly, the permutation matrix:

$$C = \begin{bmatrix} \mathbf{e}_{\sigma(1)} \\ \vdots \\ \mathbf{e}_{\sigma(n)} \end{bmatrix}, \qquad \text{where } \mathbf{e}_i \text{ is the } i\text{-th row of the identity matrix } I$$

when put on left side *permutes rows*:

$$CA = \begin{bmatrix} A_{\sigma(1),;} \\ \vdots \\ A_{\sigma(n),;} \end{bmatrix}$$

Because of such, its inverse C^{-1} must be

$$C^{-1} = \begin{bmatrix} \mathbf{e}_{\sigma^{-1}(1)} \\ \vdots \\ \mathbf{e}_{\sigma^{-1}(n)} \end{bmatrix} = \begin{bmatrix} (\mathbf{e}_{\sigma(1)})^t & \cdots & (\mathbf{e}_{\sigma(n)})^t \end{bmatrix}$$

which when put on right side *permutes columns*:

$$AC^{-1} = \begin{bmatrix} A_{;,\sigma(1)} & \cdots & A_{;,\sigma(n)} \end{bmatrix}$$

Let $\sigma = (i,j) \in S_n$. We now see from the symmetry constraint that:

$$f(E_{i,i}) = f(CE_{i,i}C^{-1}) = f(E_{j,i}C^{-1}) = f(E_{j,j})$$

■

Question 9: dual map rank formula: equation 3.4

Let U, V be two finite dimensional $\mathbb{F}$-vector spaces. Let $T : U \to V$ be a linear map, and let $T^\vee : U^\vee \to V^\vee$ be its dual map. Prove

$$T \text{ is injective} \iff T^\vee \text{ is surjective}$$

And

$$T \text{ is surjective} \iff T^\vee \text{ is injective}$$

Proof. Let $\{T(u_1), \ldots, T(u_n)\}$ be a basis for the image of T. Extend this to a basis $\{T(u_1), \ldots, T(u_n), v_1, \ldots, v_m\}$ for V. Let $\{\xi_1, \ldots, \xi_{n+m}\}$ be its dual basis. It is clear that $\{\xi_{n+1}, \ldots, \xi_{n+m}\}$ belongs to the kernel of $T^\vee$. Observe

$$T^\vee \xi_1 v_1 = 1$$

to conclude that $\xi_i \notin \operatorname{Ker} T^\vee$ for all $1 \le i \le n$. We have shown $\{\xi_{n+1}, \ldots, \xi_{n+m}\}$ is a basis for the kernel of $T^\vee$. In other words,

$$\operatorname{rank} T + \dim(\operatorname{Ker} T^\vee) = \dim V. \tag{3.4}$$

This, together with Rank-Nullity Theorem, proves both propositions. ∎

Theorem 3.2.1. (Bezout's identity) If $f, g \in \mathbb{F}[x]$ makes $\gcd(f,g) = h$, then there exists $a, b \in \mathbb{F}[x]$ such that:

$$af + bg = h$$

> ## Question 10: Bezout's identity
>
> Let V be some finite-dimensional $\mathbb{F}$-vector space, $T \in \operatorname{End}(V)$, and $f, g \in \mathbb{F}[x]$ two relatively prime polynomials. Prove
>
> $$\operatorname{Ker}(f(T)g(T)) = \operatorname{Ker} f(T) \oplus \operatorname{Ker} g(T).$$

Proof. To see that two kernels form a direct sum, let $v \in V$ satisfies $f(T)v = g(T)v = 0$. Because f, g are relatively prime, by Bezout's identity, there exists some $a, b \in \mathbb{F}[x]$ such that $af + bg = 1$. This implies

$$v = (af + bg)(T)v = a(T)f(T)v + b(T)g(T) = 0$$

as desired.[1]

It is clear that

$$\operatorname{Ker} f(T) \oplus \operatorname{Ker} g(T) \subseteq \operatorname{Ker}(f(T)g(T))$$

The opposite follows from noting

$$(af)(T)v \in \operatorname{Ker} g(T) \text{ and } (bg)(T)v \in \operatorname{Ker} f(T)$$

∎

---

[1]The moral of the story here is that $(fg)(T) = f(T)g(T) = f(T) \circ g(T)$. In more precision and deeper length, it is that endomorphism $\operatorname{End}(V)$ forms an algebra with the vector multiplication being the composition.

3.3   Year 112

Let

$$\mathbf{v}_1 = (1, 2, 0, 4), \mathbf{v}_2 = (-1, 1, 3, -3), \mathbf{v}_3 = (0, 1, -5, -2), \mathbf{v}_4 = (-1, -9, -1, -4)$$

be vectors in $\mathbb{R}^4$. Let W_1 be the subspace spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$, and let W_2 be the subspace spanned by $\mathbf{v}_3$ and $\mathbf{v}_4$. Find a basis for $W_1 \cap W_2$.

Proof. Use Gauss elimination to show that the solution space of

$$a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3 + d\mathbf{v}_4 = 0$$

is

$$(a, b, c, d) \in \operatorname{span}\{(3, 2, 1, 1)\}$$

It then follows from

$$W_1 \cap W_2 \subseteq W_2 = \operatorname{span}\{\mathbf{v}_3, \mathbf{v}_4\}$$

that

$$\{\mathbf{v}_3 + \mathbf{v}_4\} \text{ is a basis for } W_1 \cap W_2$$

∎

Let

$$A = \begin{bmatrix} 0 & 1 & -1 \\ 3 & -4 & 1 \\ 3 & -8 & 5 \end{bmatrix}$$

Find an $Q \in \operatorname{GL}_3(\mathbb{C})$ such that

$$QAQ^{-1} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 12 \\ 0 & 1 & 1 \end{bmatrix}$$

Find an invertible matrix $P \in \operatorname{GL}_3(\mathbb{C})$ such that PAP^{-1} is diagonal. *Note that my wording differs slightly with the original.*

Proof. Because characteristic polynomials are defined uniquely up to change of basis:

$$\det(tI - QAQ^{-1}) = \det(Q(tI - A)Q^{-1}) = \det(tI - A)$$

We may quickly compute that the characteristic polynomial of A is $t(t-4)(t+3)$. Routine computation now give us:

$$P = \begin{bmatrix} 1 & \frac{-7}{28} & \frac{7}{28} \\ 0 & \frac{1}{28} & \frac{-1}{28} \\ -1 & \frac{36}{28} & \frac{-8}{28} \end{bmatrix} \quad \text{and } PAP^{-1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

Let

$$B \triangleq \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 12 \\ 0 & 1 & 1 \end{bmatrix}$$

To find $Q \in \mathrm{GL}_3(\mathbb{C})$ that makes $QAQ^{-1} = B$, we first look for $M \in \mathrm{GL}_3(\mathbb{C})$ such that

$$MBM^{-1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -3 \end{bmatrix} = PAP^{-1}$$

We then see $Q \triangleq M^{-1}P$ suffices. It is then routine to compute

$$M^{-1} = \begin{bmatrix} -12 & 0 & 0 \\ -1 & 3 & 4 \\ 1 & 1 & -1 \end{bmatrix} \quad \text{and} \quad Q = M^{-1}P = \begin{bmatrix} -12 & \frac{84}{28} & \frac{-84}{28} \\ -5 & \frac{154}{28} & \frac{-42}{28} \\ 2 & \frac{-42}{28} & \frac{14}{28} \end{bmatrix}$$

∎

Question 13

Define **matrix trigonometry** by

$$\sin A \triangleq \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} A^{2n+1}$$

Compute

$$\sin \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$$

Show that there exist no matrix $A \in M_2(\mathbb{R})$ such that

$$\sin A = \begin{pmatrix} 1 & 2022 \\ 0 & 1 \end{pmatrix}$$

Proof. Using the identity

$$\sin A = \frac{e^{iA} - e^{-iA}}{2i}$$

we may compute

$$\sin \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \sin(1) & 3\cos(1) \\ 0 & \sin(1) \end{pmatrix}$$

Because $\sin(PBP^{-1}) = P(\sin B)P^{-1}$ for all B and all invertible P, and because

$$\sin \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \sin \lambda_1 & 0 \\ 0 & \sin \lambda_2 \end{pmatrix} \text{ and } \sin \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} \sin \lambda & \cos \lambda \\ 0 & \sin \lambda \end{pmatrix}$$

We may finish the proof by computing the Jordan form

$$\begin{pmatrix} 1 & 2022 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & \frac{1}{2022} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & \frac{1}{2022} \end{pmatrix}^{-1}$$

that there exists no matrix $A \in M_2(\mathbb{R})$ such that

$$\sin A = \begin{pmatrix} 1 & 2022 \\ 0 & 1 \end{pmatrix}$$

∎

Question 14

Let $A = (a_{ij}) \in M_n(\mathbb{C})$, and let $\lambda_1, \ldots, \lambda_n$ be roots of characteristic polynomial of A counted with multiplicity. Show that

$$A \text{ is normal} \iff \|A\|_F = \sum_{k=1}^{n} |\lambda_k|^2$$

where $\|A\|_F$ is the Frobenius norm of A.

Proof. Because the underlying field is $\mathbb{C}$, we may apply **Schur's decomposition** to get an upper triangular matrix D and an unitary matrix Q such that

$$A = QDQ^{-1}$$

Because A and D has the same characteristic polynomial, we know the eigenvalues of A lie on the diagonal line of D counted with multiplicity. The proof now follows from computing

$$\|A\|_F = \text{tr}(A^*A) = \text{tr}(D^*D) = \|D\|_F$$

and **Spectral Theorem for complex finite-dimensional vector space**. ∎

Remark: In my opinion, the next and final question of this exam is remarkably difficult compared to the rest in this note. Someone (not me) have posted this question on MSE and received multiple great answers. As it turned out, there are multiple (distinct?) approaches. The following proof I provided is merely an expansion of one of them. The link to the post is here:

https://math.stackexchange.com/q/4774695/1225113

Question 15

Let $A, B \in M_n(\mathbb{C})$. Suppose that all of the eigenvalues of A and B are positive real numbers. Prove

$$A^4 = B^4 \implies A = B$$

Proof. Because eigenvalues of A and B are all positive real, we may compute the Jordan forms of powers of Jordan blocks to see that A and B must share the same Jordan form. This implies A and B are similar, which implies the existence of some $P \in GL_n(\mathbb{C})$ such that $A = PBP^{-1}$. Clearly, to prove $A = B$, we only have to show that P and B commute. Using $A = PBP^{-1}$, we see:

$$B^4 = A^4 = PB^4P^{-1}$$

which implies P commute with B^4. To prove that P and B commute, we only have to show:

$$C(B^4) = C(B)$$

Clearly, the "$\supseteq$" part holds. To prove the "$\subseteq$" part also holds true, we only have to prove $\dim C(B^4) = \dim C(B)$, which follows from dimension formula for centralizer over algebraically closed field and computation of powers of Jordan blocks. ∎

3.4   Year 110

Let V be a finite-dimensional complex inner product space. Let $d \in \text{End}(V)$ satisfy

$$d^2 = 0$$

Let δ be the adjoint of d. Define $\Delta \in \text{End}(V)$ by

$$\Delta \triangleq d\delta + \delta d$$

Prove

(a) $\text{Ker}\, d\delta \subseteq \text{Ker}\, \delta$ and $\text{Ker}\, \delta d \subseteq \text{Ker}\, d$.

(b) $\text{Ker}\, \Delta = \text{Ker}\, d \cap \text{Ker}\, \delta$.

(c) $V = \text{Ker}\, \Delta \oplus \text{Im}\, d \oplus \text{Im}\, \delta$ is an orthogonal decomposition.

(d) $\text{Ker}\, d = \text{Ker}\, \Delta \oplus \text{Im}\, d$ is an orthogonal decomposition.

Proof. Part (a) follows from noting

$$\langle \delta v, \delta v \rangle = \langle v, d\delta v \rangle$$

and the positive definiteness of inner product. The hypothesis $d^2 = 0$ is unnecessary here. We now prove (b). The "$\subseteq$" is clear. We now prove the converse containment, which is done by unfolding the definition and premise. Let $v \in \text{Ker}\, \Delta$. Because $d\Delta = d\delta d$, we know $v \in \text{Ker}\, d\delta d$. This implies $dv \in \text{Ker}\, d\delta$, which by part (a) implies $dv \in \text{Ker}\, \delta$, and therefore by part (a) implies $v \in \text{Ker}\, d$.

We now prove (c). To prove that we have the desired orthogonal decomposition, we only have to prove

(I) $\{\text{Ker}\, \Delta, \text{Im}\, d, \text{Im}\, \delta\}$ is orthogonal.

(II) $V = \text{Ker}\, \Delta + \text{Im}\, d + \text{Im}\, \delta$

Part (I) follows from part (b) and unfolding the definitions. Noting that Δ is the sum of operators composing their adjoints, we see that Δ is self adjoint. Therefore, by spectral theorem for finite dimensional complex vector space, we have orthogonal decomposition $V = \text{Ker}\, \Delta \oplus \text{Im}\, \Delta$. Because we clearly have $\text{Im}\, \Delta \subseteq \text{Im}\, d + \text{Im}\, \delta$, part (II) follows.

We now prove (d). Note that by (c), we already know that $\operatorname{Ker}\Delta+\operatorname{Im}d$ forms an orthogonal (therefore direct) sum with dimension same as $\operatorname{Ker}d$. Therefore, we only have to show that $\operatorname{Ker}\Delta\subseteq\operatorname{Ker}d$ and $\operatorname{Im}d\subseteq\operatorname{Ker}d$. The former is clear and mentioned in (b). The latter follows from observing the Jordan form of d. ∎

Question 17: Correspondence between PD and inner product

Let $M\in\mathrm{M}_n(\mathbb{R})$ be positive definite. Let $A\in\mathrm{M}_n(\mathbb{R})$ satisfies

$$MAM^{-1}=A^t$$

Show that there exists some $P\in\mathrm{M}_n(\mathbb{R})$ such that

$$P^tMP=I \text{ and } P^{-1}AP \text{ is diagonal}$$

Proof. Recall that positive define matrix corresponds to inner product, i.e., we have the inner product $\langle\cdot,\cdot\rangle_M$ on $\mathbb{R}^n$ defined by

$$\langle\mathbf{x},\mathbf{y}\rangle_M\triangleq\mathbf{y}^tM\mathbf{x}$$

Denoting $P\triangleq\begin{bmatrix}\mathbf{e}_1 & \cdots & \mathbf{e}_n\end{bmatrix}$, because

$$(P^tMP)_{i,j}\triangleq(\mathbf{e}_i)^tM\mathbf{e}_j$$

We see that

$$P^tMP=I \iff \{\mathbf{e}_1,\cdots,\mathbf{e}_n\} \text{ are orthonormal with respect to } \langle\cdot,\cdot\rangle_M$$

We also have

$$P^{-1}AP \text{ is diagonal} \iff \{\mathbf{e}_1,\cdots,\mathbf{e}_n\} \text{ are eigenvectors of } A$$

The rest then follows from spectral theorem for finite dimensional real vector space and checking:

$$\langle A\mathbf{x},\mathbf{y}\rangle_M=\mathbf{y}^tMA\mathbf{x}=\mathbf{y}^tA^tM\mathbf{x}=\langle\mathbf{x},A\mathbf{y}\rangle_M$$

∎

Question 18

(a) Prove that invertible $M\in\operatorname{End}(\mathbb{C}^n)$ always has a square root.

(b) Let $n>2$, and let N be a $n\times n$ matrix over some field $\mathbb{F}$ such that $N^n=0$ and $N^{n-1}\neq0$. Prove that there is no $n\times n$ square matrix B over $\mathbb{F}$ such that

$$B^2 = N.$$

Proof. For (a), we only have to show that given nilpotent $P \in \text{End}(\mathbb{C}^n)$, $I + P$ must have a square root, and then the rest of the proof will follow from decomposing M into its Jordan form. Suppose P has the nilpotency index k. That is, k is the smallest integer such that $P^k = 0$. Computing

$$(a_0 I + \cdots + a_k P^{k-1})^2 = \sum_{j=0}^{k-1} \left(\sum_{i=0}^{j} a_i a_{j-i} \right) P^j$$

we see that the processing of finding $a_0, \ldots, a_k \in \mathbb{R}$ to satisfy $(a_0 I + \cdots + a_k P^{k-1})^2 = I + P$ boils down to solving the following equations

$$a_0^2 = 1 \longrightarrow 2a_0 a_1 = 1 \longrightarrow 2a_0 a_2 + a_1^2 = 0 \longrightarrow 2a_0 a_3 + 2a_1 a_2 = 0 \longrightarrow \cdots$$

in order, and they clearly have the solution

$$a_0 = 1, a_1 = \frac{1}{2}, a_2 = \frac{-1}{8}, a_3 = \frac{1}{16}, a_4 = \frac{-5}{128}, a_4 = \frac{7}{256}, \ldots$$

In fact, this solution is just the coefficients of Taylor expansion of $\sqrt{1+x}$ at $x = 0$, and it has the closed form

$$a_j = \frac{\frac{1}{2}(\frac{1}{2} - 1) \cdots (\frac{1}{2} - j + 1)}{j!}$$

For (b), assume for a contradiction that $B^2 = N$, so $B^{2n} = 0$ and $B^{2n-2} \neq 0$. Various different methods show that given arbitrary nilpotent $F \in \text{End}(V)$ on finite-dimensional V, the nilpotent index of F must not be greater than $\text{Dim}(V)$. In our case, we see that $B^n = 0$, which together with $n > 2$ cause a contradiction to $B^{2n-2} \neq 0$. [2] ∎

> ### Question 19
>
> Let V be a vector space over some field $\mathbb{F}$, and let $u_1, \ldots, u_n \in V$ be linearly independent. Show that, for any $v_1, \ldots, v_n \in V$, the set $\{u_1 + tv_1, \ldots, u_n + tv_n\}$ is linearly independent for all but finitely many $t \in \mathbb{F}$.

Proof. Extend $\{u_1, \ldots, u_n\}$ to some basis $\{u_i : i \in I\}$ for V. Because every vector can be expressed as a finite sum of basis element, by rearranging the index, we may write for some

---

[2]Equation 1.1 and its previous observation is one way. Another is to decompose F into its Jordan form, which is doable because every field have an algebraic closure.

large enough $r \in \mathbb{N}$ that

$$v_j = \sum_{i=1}^{r} b_{i,j} u_i \text{ for all } 1 \leq j \leq n$$

Clearly, $\{u_1 + tv_1, \ldots, u_n + tv_n\}$ is linearly dependent only if the following determinant

$$\begin{vmatrix} 1 + tb_{1,1} & tb_{1,2} & tb_{1,3} & \cdots & tb_{1,n} \\ tb_{2,1} & 1 + tb_{2,2} & tb_{2,3} & \cdots & tb_{2,n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ tb_{n,1} & tb_{n,2} & tb_{n,3} & \cdots & 1 + tb_{n,n} \end{vmatrix}$$

equals to 0. This determinant is a polynomial in t over $\mathbb{F}$, which only have finite number of roots, since $\mathbb{F}$ is a field. ∎

Question 20

Let P be some $n \times n$ matrix over some field $\mathbb{F}$. Show

$$\text{rank}(P) + \text{rank}(I - P) = n \implies P^2 = P$$

Proof. Observe

$$v = v - Pv = (I - P)v, \quad \text{for all } v \in \text{Ker}(P)$$

This implies $\text{Ker}(P) \subseteq \text{Im}(I - P)$. It now follows from **Rank-Nullity Theorem** and

$$\text{rank}(P) + \text{rank}(I - P) = n$$

that

$$\text{Ker}(P) = \text{Im}(I - P).$$

Therefore,

$$(P - P^2)v = P(I - P)v = 0, \quad \text{for all } v \in \mathbb{F}^n \text{ as desired.}$$

∎

3.5 Year 109

Question 21

Let A be a 4×4 real symmetric matrix. Suppose that 1 and 2 are eigenvalues of A and the eigenspace of 2 is 3-dimensional. Suppose $(1, -1, -1, 1)^t$ is an eigenvector with eigenvalue 1.

(a) Find an orthonormal basis for the eigenspace for the eigenvalue 2 of A.

(b) Find Av, where $v = (1, 0, 0, 0)^t$.

Proof. Let $G(2, A)$ be the eigenspace of A with eigenvalue 2. By applying **Real Spectral Theorem** to A, we know that

$$\mathbb{R}^4 = \operatorname{span}\{(1, -1, -1, 1)\} \oplus G(2, A)$$

forms an orthogonal decomposition. By computing the determinant, we know

$$\{(1, -1, -1, 1), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\} \text{ forms a basis for } \mathbb{R}^4$$

Applying Gram-Schmidt process, we now see that

$$\left\{\frac{1}{2\sqrt{3}}(1, 3, -1, 1), \frac{1}{\sqrt{6}}(1, 0, 2, 1), \frac{1}{\sqrt{2}}(-1, 0, 0, 1)\right\} \text{ forms an orthonormal basis for } G(2, A).$$

Because

$$(1, 0, 0, 0) \cdot \frac{1}{2}(1, -1, -1, 1) = \frac{1}{2}$$

we know the orthogonal decomposition of $(1, 0, 0, 0)$ is exactly

$$(1, 0, 0, 0) = \frac{1}{4}(1, -1, -1, 1) + (\frac{3}{4}, \frac{1}{4}, \frac{1}{4}, \frac{-1}{4}).$$

Therefore

$$A(1, 0, 0, 0)^t = \frac{1}{4}(1, -1, -1, 1) + 2(\frac{3}{4}, \frac{1}{4}, \frac{1}{4}, \frac{-1}{4}) = (\frac{7}{4}, \frac{1}{4}, \frac{1}{4}, \frac{-1}{4})$$

■

Let A be a real $n \times n$ matrix. Prove

$$2\operatorname{rank}(A^2) \le \operatorname{rank}(A^3) + \operatorname{rank}(A)$$

Proof. The question is immediately solved once we decompose A into its **Jordan form**:

$$A = PJP^{-1} \text{ where } P, J \in M_n(\mathbb{C})$$

We will leave the detail of why the question is immediately solved by decomposing A into its Jordan form later. What's more important here is to show $\operatorname{rank}_{\mathbb{R}}(A) = \operatorname{rank}_{\mathbb{C}}(A)$.[3] Clearly, $\operatorname{rank}_{\mathbb{R}}(A) \ge \operatorname{rank}_{\mathbb{C}}(A)$. Suppose $A_1, \ldots, A_k \in \mathbb{R}^n$ are $\mathbb{R}$-linearly independent. If $(a_1 + ib_1)A_1 + \cdots + (a_k + ib_k)A_k = 0$, then $a_1 = \cdots = a_k = b_1 = \cdots = b_k = 0$ by $\mathbb{R}$-linear independence of $A_1, \ldots, A_k$. We have shown $\operatorname{rank}_{\mathbb{R}}(A) \le \operatorname{rank}_{\mathbb{C}}(A)$.

We now go back to the obvious detail. Suppose

$$A = \bigoplus_{i=1}^{r} J(\lambda_i, m_i) \oplus \bigoplus_{i=1}^{q} J(0, k_i), \quad \text{where } \lambda_i \text{ are all non-zero.}$$

Let $p \triangleq \sum m_i$. We have

$$\operatorname{rank}(A) = p + \sum_{i=1}^{q} (k_i - 1)$$

$$\operatorname{rank}(A^2) = p + \sum_{i=1}^{q} \max\{k_i - 2, 0\}$$

$$\operatorname{rank}(A^3) = p + \sum_{i=1}^{q} \max\{k_i - 3, 0\}$$

The proof then follows from checking

$$2\max\{k_i - 2, 0\} \le k_i - 1 + \max\{k_i - 3, 0\}$$

$\blacksquare$

---

[3]The same procedure can be used to show for any field $\mathbb{F} \subseteq \mathbb{K}$ and $A \in M_n(\mathbb{F})$, we have $\operatorname{rank}_{\mathbb{F}}(A) = \operatorname{rank}_{\mathbb{K}}(A)$

Question 23

Let V be a finite-dimensional real vector space, and let S, T and U be subspaces of V. Prove or disprove

(a) $\dim(S + T) = \dim(S) + \dim(T) - \dim(S \cap T)$.

(b) $\dim(S + T + U) = \dim(S) + \dim(T) + \dim(U) - \dim(S \cap T) - \dim(T \cap U) - \dim(U \cap S) + \dim(S \cap T \cap U)$.

Proof. For (a), let B be a basis for $S \cap T$, and let $B_S \subseteq S$ and $B_T \subseteq T$ satisfy

$$B_S \sqcup B, B_T \sqcup B \text{ are respectively bases for } S \text{ and } T$$

where $\sqcup$ means disjoint union. A tedious proof by contradiction shows that $B_S \cap B_T = \varnothing$ and $B_S \cup B_T$ are linearly independent. It is easy to see that $B_S \sqcup B \sqcup B_T$ spans $S + T$. We have shown $B_S \sqcup B \sqcup B_T$ forms a basis for $S + T$. This give us proposition (a).

For (b), let $V \triangleq \mathbb{R}^2, S \triangleq \text{span}(1, 0), T \triangleq \text{span}(0, 1)$ and $U \triangleq \text{span}(1, 1)$, and we see a counter example. $\blacksquare$

Question 24

Let

$$M \triangleq \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$$

(a) Compute e^M.

(b) Prove $\det(e^A) = \exp(\text{tr } A)$ for $A \in M_n(\mathbb{C})$.

(c) Prove or disprove

$$A \in M_n(\mathbb{C}) \text{ is nilpotent} \implies e^A - I \text{ is nilpotent}$$

Remark: the original phrasing of part (c) of this question is absolutely horrible. They should have written $\exp(A) - I$ instead of $\exp A - I$.

Proof. Decompose M into its **Jordan form**

$$M = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1}.$$

We see

$$e^M = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \left(\begin{bmatrix} e & 0 \\ 0 & e \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & e \\ -e & 2e \end{bmatrix}.$$

For part (b), decompose A into its Jordan form

$$A = QJQ^{-1} = Q(S + N)Q^{-1} \text{ where } S \text{ is diagonal and } N \text{ is nilpotent.}$$

Because

$$e^{QJQ^{-1}} = \sum_{k=0}^{\infty} \frac{(QJQ^{-1})^k}{k!} = \sum_{k=0}^{\infty} \frac{QJ^kQ^{-1}}{k!} = Q\left(\sum_{k=0}^{\infty} \frac{J^k}{k!} \right) Q^{-1} = Qe^JQ^{-1}$$

we have

$$\det(e^A) = \det(e^{QJQ^{-1}}) = \det(Qe^JQ^{-1}) = \det(e^J)$$

Noting that e^J is upper triangular with diagonal entries e^{λ_i} where λ_i are the diagonal entries of J, we may conclude

$$\det(e^J) = e^{\sum \lambda_i} = e^{\operatorname{tr} J} = e^{\operatorname{tr}(Q^{-1}AQ)} = e^{\operatorname{tr}(A)}.$$

For part (c), let $A^m = 0$, so

$$e^A - I = \sum_{k=0}^{m-1} \frac{A^k}{k!} - I = \sum_{k=1}^{m-1} \frac{A^k}{k!}$$

We have shown $e^A - I$ is a polynomial in A with constant coefficient zero. This implies $(e^A - I)^m = 0$. ∎

Question 25

Let U, V be two finite dimensional space over the same field. Let $T : U \to V$ be a linear map, and let $T^\vee : U^\vee \to V^\vee$ be its dual map. Prove

$$T \text{ is injective } \iff T^\vee \text{ is surjective}$$

And

$$T \text{ is surjective } \iff T^\vee \text{ is injective}$$

Proof. Let $\{T(u_1), \ldots, T(u_n)\}$ be a basis for the image of T. Extend this to a basis $\{T(u_1), \ldots, T(u_n), v_1, \ldots, v_m\}$ for V. Let $\{\xi_1, \ldots, \xi_{n+m}\}$ be its dual basis. It is clear that $\{\xi_{n+1}, \ldots, \xi_{n+m}\}$ belongs to the kernel of $T^\vee$. Observe

$$T^*\xi_1 v_1 = 1$$

to conclude that $\xi_i \notin \operatorname{Ker} T^\vee$ for all $1 \le i \le n$. We have shown $\{\xi_{n+1}, \ldots, \xi_{n+m}\}$ is a basis for the kernel of $T^\vee$. In other words,

$$\operatorname{rank} T + \operatorname{Dim}(\operatorname{Ker} T^*) = \dim V.$$

This, together with Rank-Nullity Theorem, proves both propositions. $\blacksquare$

3.6   Year 108

Find all possible Jordan forms for 8×8 real matrices $x^2(x-2)^3$ as minimal polynomial.

Proof. This question requires us to show that the Jordan form

$$J = \bigoplus_{i_1} J(\lambda_1, n_{i_1}) \oplus \cdots \oplus \bigoplus_{i_m} J(\lambda_m, n_{i_m})$$

has the minimal polynomial

$$T(x) = (x - \lambda_1)^{\max_{i_1} n_{i_1}} \cdots (x - \lambda_m)^{\max_{i_m} n_{i_m}}.$$

It is clear that J is a root of T. To see that T is indeed minimal, utilize the algebraic closure of the underlying field to show that the minimal polynomial of J can not have roots other than $\lambda_1, \ldots, \lambda_m$, and directly show that any polynomial that have J as a root must have power on factor $(x - \lambda_j)$ greater than $\max_{i_j} n_{i_j}$. ∎

Let V be a vector space over a field $\mathbb{F}$ of infinite elements, and let $v_1, \ldots, v_n$ be vectors in V such that there exists infinite numbers of $z \in \mathbb{F}$ satisfying

$$v_0 + zv_1 + \cdots + z^n v_n = 0$$

where $n > 0$. Show that

$$v_0 = v_1 = \cdots = v_n = 0$$

Proof. Let $\{w_\alpha : \alpha \in I\}$ be a basis for V, by rearranging the index, we may write

$$v_i = \sum_{k=1}^{r} c_{k,i} w_k \text{ for all } 1 \leq i \leq n$$

Fix k. We must have

$$c_{k,n} z^n + \cdots + c_{k,1} z + c_{k,0} = 0 \text{ for infinite } z \in \mathbb{F}$$

Because $\mathbb{F}$ forms a field, we know for this to be true, we must have $c_{k,n} = c_{k,n-1} = \cdots = c_{k,0} = 0$. ∎

Question 28

Let $V \in M_{n \times n}(\mathbb{R})$ and let $f \in \text{Hom}(V, \mathbb{R})$ satisfy

$$f(AB) = f(BA), \quad \text{for all } A, B \in V \text{ and } f(I) = n$$

Show that f is the trace function.

Proof. Let $E_{i,j} \in V$ be the matrix whose only nonzero entry is 1, located in the i-th row and j-th column. If $i \neq j$, then

$$E_{i,i}E_{i,j} = E_{i,j} \text{ and } E_{i,j}E_{i,i} = 0.$$

This, together with the linearity of f, shows that f depends only on the diagonal entries. For each permutation $\sigma \in S_n$, there exists a **permutation matrix** C such that

The $\sigma(i)$-th row of CA is identical to the i-th row of A.

And

The $\sigma(i)$-th column of AC^{-1} is identical to the i-th column of A.

Therefore,

$$A_{i,i} = (CAC^{-1})_{\sigma(i),\sigma(i)}.$$

Observing that

$$f(CAC^{-1}) = f(A(CC^{-1})) = f(A),$$

we have shown that f is stable under the permutation of diagonal entries. Define $\sigma \in S_n$ by

$$\sigma(i) \triangleq \begin{cases} i+1 & \text{if } 1 \leq i < n, \\ 1 & \text{if } i = n. \end{cases}$$

Let C_k be the permutation matrix corresponding to σ^k for all $1 \leq k \leq n$. The proof now follows from computing

$$nf(A) = \sum_{k=1}^{n} f(C_k A C_k^{-1}) = f\left(\sum_{k=1}^{n} C_k A C_k^{-1}\right) = f((\text{tr } A)I) = n \text{ tr } A.$$

$\blacksquare$

Let V be an n-dimensional vector space over $\mathbb{R}$ and $B : V \times V \to \mathbb{R}$ be a symmetric bilinear form.

(a) Let W be a subspace of V, and define

$$W^{\perp} \triangleq \{v \in V : B(v, w) = 0 \text{ for all } w \in W\}$$

Show that $\dim(W^{\perp}) \geq \dim(V) - \dim(W)$.

(b) Prove that $V = W \oplus W^{\perp}$ if and only if B is nondegenerate on W.

(c) Prove that if B is nondegenerate on V, then there exists a nonnegative integer p with $p \leq \text{Dim}(V)$ and a basis $\{e_1, \ldots, e_n\}$ such that

$$B(e_i, e_j) = \begin{cases} 1 & \text{if } 1 \leq i = j \leq p \\ -1 & \text{if } p + 1 \leq i = j \leq n \\ 0 & \text{if } i \leq j \end{cases}$$

Proof. Let $\{w_1, \ldots, w_m\}$ be a basis for W. Consider the linear map $F : V \to \mathbb{R}^m$ defined by $v \mapsto (B(v, w_1), \ldots, B(v, w_m))$. By Rank-Nullity Theorem, we have

$$\dim(W^{\perp}) = \dim(\text{Ker } F) = \dim(V) - \text{rank}(F) \geq \dim(V) - m$$

We have proved part (a). Note that $W \cap W^{\perp} = 0$ if and only if B is degenerate on W. This together with part (a) prove part (b).

Part (c) is independent of part (a) and (b). Because B is nondegenerate, we may let $\{v_1, \ldots, v_n\}$ be a basis for V such that $B(v_i, v_i)$ is positive for $i \leq p$ and negative for $i > p$. Define $\widetilde{v}_1, \ldots, \widetilde{v}_n$ by

$$\widetilde{v}_j \triangleq v_j - \frac{B(v_j, \widetilde{v}_{j-1})}{B(\widetilde{v}_{j-1}, \widetilde{v}_{j-1})} \widetilde{v}_{j-1} - \cdots - \frac{B(v_j, \widetilde{v}_1)}{B(\widetilde{v}_1, \widetilde{v}_1)} \widetilde{v}_1, \quad \text{(Gram-Schmidt process)}$$

It is clear that $\{\widetilde{v}_1, \ldots, \widetilde{v}_n\}$ is orthogonal. It is easy to normalize them, and the result is the desired $\{e_1, \ldots, e_n\}$. ∎

3.7 Year 107

Let

$$A = \begin{bmatrix} -1 & 3 & -2 \\ 2 & 3 & 0 \\ 11 & -6 & 7 \end{bmatrix}$$

Find the lower triangular Jordan form of A. Compute e^{tA} and the general solutions of $x'(t) = Ax(t)$ where x is a 3-dimensional column vector.

Proof. Computing

$$\det(tI - A^T) = \det(tI - A) = (t-3)^3$$

and

$$\mathrm{Ker}(3I - A^T) = \mathrm{span}(\begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix})$$

we may compute

$$A^T = \begin{bmatrix} 4 & 3 & 1 \\ -3 & -3 & -2 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & \\ & 3 & 1 \\ & & 3 \end{bmatrix} \begin{bmatrix} 4 & 3 & 1 \\ -3 & -3 & -2 \\ 2 & 2 & 1 \end{bmatrix}^{-1}$$

Therefore,

$$A = \begin{bmatrix} 4 & -3 & 2 \\ 3 & -3 & 2 \\ 1 & -2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 3 & & \\ 1 & 3 & \\ & 1 & 3 \end{bmatrix} \begin{bmatrix} 4 & -3 & 2 \\ 3 & -3 & 2 \\ 1 & -2 & 1 \end{bmatrix}$$

This give us

$$
e^{tA} = \begin{bmatrix} 4 & -3 & 2 \\ 3 & -3 & 2 \\ 1 & -2 & 1 \end{bmatrix}^{-1} \left(\begin{bmatrix} e^{3t} & & \\ & e^{3t} & \\ & & e^{3t} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ t & 1 & 0 \\ \frac{t^2}{2} & t & 1 \end{bmatrix} \right) \begin{bmatrix} 4 & -3 & 2 \\ 3 & -3 & 2 \\ 1 & -2 & 1 \end{bmatrix}
$$

$$
= e^{3t} \begin{bmatrix} 4 & -3 & 2 \\ 3 & -3 & 2 \\ 1 & -2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ t & 1 & 0 \\ \frac{t^2}{2} & t & 1 \end{bmatrix} \begin{bmatrix} 4 & -3 & 2 \\ 3 & -3 & 2 \\ 1 & -2 & 1 \end{bmatrix}
$$

$$
= e^{3t} \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -2 \\ -3 & 5 & -3 \end{bmatrix} \begin{bmatrix} 4 & -3 & 2 \\ 4t+3 & -3t-3 & 2t+2 \\ 2t^2+3t+1 & \frac{-3}{2}t^2 - 3t - 2 & t^2+2t+1 \end{bmatrix}
$$

$$
= e^{3t} \begin{bmatrix} -4t+1 & 3t & -2t \\ -4t^2+2t & 3t^2+1 & -2t^2 \\ -6t^2+11t & \frac{9}{2}t^2 - 6t & -3t^2+4t+1 \end{bmatrix}
$$

The space of solutions of $x'(t) = Ax(t)$ is the column space of e^{tA}. (The above heavy computation is verified by ChatGPT o3-model. I am not going to manually verify it.) ∎

Question 31

Let V be an n-dimensional complex vector space, and let $T \in \text{End}(V)$ satisfy $T^2 = 1$.

(a) Show that T is diagonalizable.

(b) Let $S \subseteq \text{End}(V)$ be the space of endomorphisms that commute with T. Express $\dim(S)$ in term of $\text{tr}(T)$.

Proof. Clearly, T can only have eigenvalues ± 1. Computing

$$
\begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}^2 = \begin{bmatrix} \lambda^2 & 2\lambda & 1 & \\ & \lambda^2 & 2\lambda & \ddots \\ & & \lambda^2 & \ddots & 1 \\ & & & \ddots & 2\lambda \\ & & & & \lambda^2 \end{bmatrix}
$$

we see that to satisfies $T^2 = 1$, each Jordan block of T must have size 1, otherwise there exists some $v \in \mathbb{C}^n$ such that $T^2 v \neq v$. We have shown T is indeed diagonalizable with eigenvalues ± 1. Therefore, there exists some basis under which T have the expression

$$
\begin{bmatrix} I_m & O \\ O & -I_{n-m} \end{bmatrix}
$$

65

Let $X \in M_n(\mathbb{C})$, and divide X into submatrices:

$$X = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \text{ where } A \in M_m(\mathbb{C}) \text{ and } D \in M_{n-m}(\mathbb{C})$$

We see

$$X \in S \iff B = C = O$$

This give us

$$\dim(S) = m^2 + (n-m)^2$$

We may now compute

$$\mathrm{tr}(T) = 2m - n \implies m = \frac{n + \mathrm{tr}(T)}{2} \implies \dim(S) = \frac{n^2 + (\mathrm{tr}(T))^2}{2}$$

∎

Question 32

Let $A = (A_{ij})$ be a $2n \times 2n$ real invertible skew-symmetric matrix.

(a) Show that all eigenvalues of A are pure-imaginary.

(b) Define the Pfaffian $\mathrm{pf}(A)$ of A by

$$\mathrm{pf}(A) = \frac{1}{2^n n!} \sum_{\sigma \in S_{2n}} \mathrm{sgn}(\sigma) \prod_{k=1}^{n} A_{\sigma(2k-1),\sigma(2k)}$$

Show that for every $2n \times 2n$ real matrix B, we have

$$\mathrm{pf}(BAB^t) = \mathrm{pf}(A) \det(B)$$

(c) Suppose for some $2n \times 2n$ real orthogonal matrix O we have

$$OAO^t = \mathrm{diag}\left\{ \begin{pmatrix} 0 & m_1 \\ -m_1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & m_2 \\ -m_2 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & m_n \\ -m_n & 0 \end{pmatrix} \right\}$$

where $m_i \in \mathbb{R}$ for all i, Show that $\det(A) = \mathrm{pf}(A)^2$

Proof. For part (a), one simply have to note that the matrix representation of adjoint of linear map T over $\mathbb{C}$ is just the complex conjugation of that of T. Suppose $Av = \lambda v$ for some $v \neq 0 \in \mathbb{C}^{2n}$. Because $A^* = A^t = -A$, we have

$$\lambda v \cdot v = Av \cdot v = v \cdot A^* v = -v \cdot Av = -\bar{\lambda} v \cdot v$$

66

This implies the real part of λ must be zero. We now prove (b). Compute

$$(BAB^t)_{ij} = \sum_{kl} B_{il} A_{lk} B_{jk}$$

Let $e_1, \ldots, e_{2n}$ be the standard basis for $\mathbb{R}^{2n}$. If we define $\omega \in \bigwedge \mathbb{R}^{2n}$ by

$$\omega \triangleq \sum_{ij} (BAB^t)_{ij} e_i \wedge e_j$$

we now have

$$\omega = \sum_{ijkl} B_{il} A_{lk} B_{jk} e_i \wedge e_j = \sum_{lk} A_{lk} f_l \wedge f_k, \quad \text{where } f_p = \sum_q B_{qp} e_q$$

This together with Theorem 2.3.1 give us

$$2^n n! \operatorname{pf}(A) \det(B) e_1 \wedge \cdots \wedge e_{2n} = 2^n n! \operatorname{pf}(A) f_1 \wedge \cdots \wedge f_{2n}$$
$$= \omega^n = 2^n n! \operatorname{pf}(BAB^t) e_1 \wedge \cdots \wedge e_{2n}$$

The proof of part (b) now follows from noting $e_1 \wedge \cdots \wedge e_{2n} \neq 0$. We now prove (c). Again, define $\omega \in \bigwedge \mathbb{R}^{2n}$ by

$$\omega \triangleq \sum_{ij} (OAO^t)_{ij} e_i \wedge e_j$$

Because of the form of OAO^t, we see

$$\omega = 2m_1 e_1 \wedge e_2 + 2m_2 e_3 \wedge e_4 + \cdots + 2m_n e_{2n-1} \wedge e_{2n}$$

This allow us to compute

$$\omega^n = 2^n n! \left(\prod m_i \right) e_1 \wedge \cdots \wedge e_{2n}$$

It now follows from Equivalent Definition of Pfaffian and part (b) of this question that

$$\operatorname{pf}(A) = \operatorname{pf}(OAO^t) = m_1 \cdots m_n$$

The rest of part (c) then follows from computing $\det(A) = \det(OAO^t) = m_1^2 \cdots m_n^2$ ∎

Question 33

Let $A, B \in M_n(\mathbb{C})$ be $n \times n$ complex matrices. Show that if A and B commute, then they are simultaneously diagonalizable. (i.e., there exists an invertible matrix

Proof. Before presenting the proof, I will like to point out that this question can not be solved using Jordan decomposition: It is possible that two commuting matrices over $\mathbb{C}$ does not share any Jordan basis. For example, consider T and T^2, where T has some Jordan block of size greater than 2.

We prove by induction. The base case is trivial. Let $\{\lambda_1, \ldots, \lambda_n\}$ be the set of eigenvalues of A. Consider the generalized eigenspaces $G(\lambda_i, A) = \mathrm{Ker}(A - \lambda_i I)^n$. It is well known that

$$\mathbb{C}^n = G(\lambda_1, A) \oplus \cdots \oplus G(\lambda_n, A), \quad \text{See Theorem 1.2.3}$$

Because A, B commute, we may conclude $G(\lambda_i, A)$ are all stable under B. This allow us to apply inductive hypothesis on $G(\lambda_1, A) \oplus \cdots \oplus G(\lambda_{n-1}, A)$ and $G(\lambda_n, A)$ to finish the proof. ∎

Question 34

Show that

$$
\begin{vmatrix}
X_0 & X_1 & X_2 & \cdots & X_{n-1} \\
X_{n-1} & X_0 & X_1 & \cdots & X_{n-2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
X_1 & X_2 & X_3 & \cdots & X_0
\end{vmatrix} = \prod_{j=0}^{n-1} \left(\sum_{k=0}^{n-1} \xi^{jk} X_k \right) \tag{3.5}
$$

where ξ is a n-th primitive root of unity.

Proof. Because ξ is a n-th primitive root of unity, we may compute that the following product of $n \times n$ matrices

$$
\begin{pmatrix}
X_0 & X_1 & X_2 & \cdots & X_{n-1} \\
X_{n-1} & X_0 & X_1 & \cdots & X_{n-2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
X_1 & X_2 & X_3 & \cdots & X_0
\end{pmatrix}
\begin{pmatrix}
\xi^0 & \xi^0 & \xi^0 & \cdots & \xi^0 \\
\xi^0 & \xi^1 & \xi^2 & \cdots & \xi^{n-1} \\
\xi^0 & \xi^2 & \xi^4 & \cdots & \xi^{2(n-1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\xi^0 & \xi^{n-1} & \xi^{2(n-2)} & \cdots & \xi^{(n-1)^2}
\end{pmatrix}
$$

is exactly

$$
\begin{pmatrix}
\xi^0 Y_0 & \xi^0 Y_1 & \xi^0 Y_2 & \cdots & \xi^0 Y_{n-1} \\
\xi^0 Y_0 & \xi^1 Y_1 & \xi^2 Y_2 & \cdots & \xi^{n-1} Y_{n-1} \\
\xi^0 Y_0 & \xi^2 Y_1 & \xi^4 Y_2 & \cdots & \xi^{2(n-1)} Y_{n-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\xi^0 Y_0 & \xi^{n-1} Y_1 & \xi^{2(n-2)} Y_2 & \cdots & \xi^{(n-1)^2} Y_{n-1}
\end{pmatrix}
$$

where

$$
Y_j = \sum_{k=0}^{n-1} \xi^{jk} X_k \text{ for all } j
$$

Noting that the right hand side of Equation 3.5 is just $\prod Y_j$ and that

$$
\begin{vmatrix}
\xi^0 Y_0 & \xi^0 Y_1 & \xi^0 Y_2 & \cdots & \xi^0 Y_{n-1} \\
\xi^0 Y_0 & \xi^1 Y_1 & \xi^2 Y_2 & \cdots & \xi^{n-1} Y_{n-1} \\
\xi^0 Y_0 & \xi^2 Y_1 & \xi^4 Y_2 & \cdots & \xi^{2(n-1)} Y_{n-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\xi^0 Y_0 & \xi^{n-1} Y_1 & \xi^{2(n-2)} Y_2 & \cdots & \xi^{(n-1)^2} Y_{n-1}
\end{vmatrix}
= \left(\prod_{j=0}^{n-1} Y_j \right)
\begin{vmatrix}
\xi^0 & \xi^0 & \xi^0 & \cdots & \xi^0 \\
\xi^0 & \xi^1 & \xi^2 & \cdots & \xi^{n-1} \\
\xi^0 & \xi^2 & \xi^4 & \cdots & \xi^{2(n-1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\xi^0 & \xi^{n-1} & \xi^{2(n-2)} & \cdots & \xi^{(n-1)^2}
\end{vmatrix}
$$

we conclude that to close out the proof with the general formula $\det(AB) = \det(A)\det(B)$, it only remains to show the determinant of the **Vandermonde matrix**

$$
\begin{vmatrix}
\xi^0 & \xi^0 & \xi^0 & \cdots & \xi^0 \\
\xi^0 & \xi^1 & \xi^2 & \cdots & \xi^{n-1} \\
\xi^0 & \xi^2 & \xi^4 & \cdots & \xi^{2(n-1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\xi^0 & \xi^{n-1} & \xi^{2(n-2)} & \cdots & \xi^{(n-1)^2}
\end{vmatrix}
$$

is nonzero. ∎

69

3.8 Year 106

> **Question 35**
>
> Let $\mathbb{F}$ be some field, and let $A \in M_n(\mathbb{F})$.
>
> (a) Show that if k is the largest integer such that some $k \times k$ submatrix of A has a nonzero determinant, then $\text{rank}(A) = k$.
>
> (b) Suppose $A^n = 0$ and $A^{n-1} \neq 0$. Find $\max_{v \in \mathbb{F}^n} \dim(\text{span}(v, Av, \ldots, A^{n-1}v))$.

Proof. We first prove part (a). Because exchanging two columns or exchanging two rows do not change the rank, we may WLOG suppose the $k \times k$ submatrix is

$$\begin{pmatrix} A_{1,1} & \cdots & A_{1,k} \\ \vdots & \ddots & \vdots \\ A_{k,1} & \cdots & A_{k,k} \end{pmatrix} \tag{3.6}$$

Because the subamtrix 3.6 has nonzero determinant, we know that there exists no nonzero $(c_1, \ldots, c_k) \in \mathbb{F}^k$ such that $c_1 A_{:,1} + \cdots + c_k A_{:,k} = 0$, where $A_{:,j}$ stands for the j-th column of A. We have shown $\text{rank}(A) \geq k$. It remains to show $\text{rank}(A) \leq k$. Assume $\text{rank}(A) > k$ for a contradiction. WLOG, suppose $A_{:,1}, \ldots, A_{:,k+1}$ are linearly independent. Recall that for any matrix, square or not, the columns rank equals the row rank, so there exists some $1 \leq i_1 < \cdots < i_{k+1} \leq n$ such that

$$\left\{ (A_{i_j,1}, \ldots, A_{i_j,k+1}) \in \mathbb{F}^{k+1} : 1 \leq j \leq k+1 \right\} \text{ is linearly independent}$$

This implies

$$\begin{vmatrix} A_{i_1,1} & \cdots & A_{i_1,k+1} \\ \vdots & \ddots & \vdots \\ A_{i_{k+1},1} & \cdots & A_{i_{k+1},k+1} \end{vmatrix} \neq 0$$

causing a contradiction. For part (b), we note that the Jordan decomposition always exists for nilpotent endomorphism over finite dimensional space, whether the base field is algebraically closed or not. (Observe that both the statement and the proof for Theorem 1.3.1 do not mention whether the base field of V is closed or not). Therefore, we may conclude the solution is

$$\max_{v \in \mathbb{F}^n} \dim(\text{span}(v, Av, \ldots, A^{n-1}v)) = n$$

∎

(a) Let $A = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}$. Find the general solution of linear system of ODE: $x' = Ax$.

(b) Let V be the space of all real polynomial having degree strictly less than 4 with inner product $\langle f, g \rangle \triangleq \int_{-1}^{1} f(t)g(t)dt$. Define $T \in \text{End}(V)$ by $T(f) \triangleq f' + 3f$. Use the Gram-Schmidt process to replace $\beta = \{1, 1 + x, x + x^2, x^2 + x^3\}$ by an orthonormal basis, and find the matrix representation of the adjoint $T^\dagger$ with respect to this new basis.

Proof. The characteristic polynomial of A is $\text{ch}(t) = (t-2)^2(t-1)$. Luckily, A is diagonalizable:

$$A = P \operatorname{diag}(1, 2, 2)P^{-1} \text{ where } P = \begin{pmatrix} -3 & 4 & 2 \\ 1 & 1 & 1 \\ -3 & 1 & 0 \end{pmatrix}$$

This implies

$$e^{At} = P \operatorname{diag}(e^t, e^{2t}, e^{2t})P^{-1} = \begin{pmatrix} -3e^t + 4e^{2t} & -6e^t - 6e^{2t} & 6e^t - 6e^{2t} \\ e^t - e^{2t} & 2e^t + 3e^{2t} & -2e^t + 2e^{2t} \\ -3e^t + 3e^{2t} & -6e^t - 6e^{2t} & 6e^t - 5e^{2t} \end{pmatrix}$$

We now do the computation for part (b). The new orthonormal basis is

$$\left\{ \frac{1}{\sqrt{2}}, \frac{\sqrt{3}}{\sqrt{2}}x, \frac{3\sqrt{5}}{2\sqrt{2}}(x^2 - \frac{1}{3}), \frac{5\sqrt{7}}{2\sqrt{2}}(x^3 - \frac{3}{5}x) \right\}$$

With respect to this basis, we have

$$[T] = \begin{pmatrix} 3 & \sqrt{3} & 0 & \sqrt{7} \\ 0 & 3 & \sqrt{15} & 0 \\ 0 & 0 & 3 & \sqrt{35} \\ 0 & 0 & 0 & 3 \end{pmatrix} \implies [T^\dagger] = \begin{pmatrix} 3 & 0 & 0 & 0 \\ \sqrt{3} & 3 & 0 & 0 \\ 0 & \sqrt{15} & 3 & 0 \\ \sqrt{7} & 0 & \sqrt{35} & 3 \end{pmatrix}$$

■

Question 37: polar decomposition and spectral theorem

(a) Let $A \in M_n(\mathbb{R})$. Show that there exists orthonormal $Q \in M_n(\mathbb{R})$ and positive semi-definite symmetric matrix $P \in M_n(\mathbb{R})$ such that $A = QP$.[a]

(b) Let V be a finite dimensional $\mathbb{C}$-vector space, and let $T \in \text{End}(V)$. Show

$$T \text{ is normal} \iff T^\dagger = p(T) \text{ for some } p \in \mathbb{C}[x]$$

---

[a]I was first confused by the original phrasing of the question, so I did some googling. According to Wikipedia, the term "orthogonal matrix" and the term "orthonormal matrix" are interchangeable. They mean a real square matrix P such that $PP^t = I$. The term "positive semi-definite matrix" means a real square matrix Q such that $v^t Q v \geq 0$ for all $v \in \mathbb{R}^n$

Proof. We first prove part (a). By SVD, there exists some orthonormal $U, V \in M_n(\mathbb{R})$ and diagonal $D \in M_n(\mathbb{R}_0^+)$ such that $A = UDV$. Clearly $Q \triangleq UV$ and $P \triangleq V^t DV$ suffice.[4]

We now prove part (b). The "if" part is clear. Suppose T is normal. By Spectral Theorem, there exists some orthonormal basis β of V such that $[T]_\beta = \text{diag}(z_1, \ldots, z_n)$ for some $z_1, \ldots, z_n \in \mathbb{C}$. Because $[T^\dagger]_\beta = \text{diag}(\overline{z_1}, \ldots, \overline{z_n})$. The polynomial p we are looking for only has to satisfy $p(z_i) = \overline{z_i}$ for each i. Write $\{z_1, \ldots, z_n\} = \{x_1, \ldots, x_k\}$ where $x_1, \ldots, x_k$ are distinct. We have reduced the original question into an interpolation problem with well known solution:

$$p(x) \triangleq \prod_{i=1}^{k} \frac{\overline{x_i}(x - x_1) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_k)}{(x_i - x_1) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - z_k)}$$

∎

Question 38

Let $T \in \text{End}(V)$ for finite dimensional $\mathbb{C}$-vector space V. Prove the existence and uniqueness of Jordan-Chevalley decomposition $T = S + N$. Show that there exists some polynomial $\widetilde{p} \in \mathbb{C}[x]$ such that $S = \widetilde{p}(T)$ and $\widetilde{p}(0) = 0$.

Proof. In my personal opinion, this is a shitty question for exam. There are multiple distinct[5] approaches to show the existence and uniqueness of Jordan-Chevalley decomposition, but there is essentially only one approach, i.e., the ones quoting the Chinese Remainder Theorem without giving a proof, that is short enough to fully write down in exam time.

---

[4]To see $v^t(V^t DV)v$ is always nonnegative, express v as the linear combination of columns of V and perform direct computation.

[5]distinct in the sense that how much ring theory one wish to directly employ.

Here, we offer a proof that use only Bezout's identity, i.e., the one we presented at the end of Section 1.2, which is short enough I guess.

The only part the question require us to prove that the argument presented at the end of Section 1.2 didn't show is that we may select some $\widetilde{p} \in \mathbb{C}[x]$ so that not only $S = \widetilde{p}(T)$ but also $\widetilde{p}$ has zero constant term. Let $p = \sum \xi_i q_i f_i$ where ξ_i, q_i, f_i are defined at the end of Section 1.2. To prove this, we first note that because $T - p(T)$ is nilpotent, we may find n such that the polynomial $(x - p)^n$ sends T to 0. If p has zero constant term, then $\widetilde{p} \triangleq p$ suffice. If p has nonzero constant term, then $(x - p)^n$ also has nonzero constant term and $\widetilde{p} \triangleq p - \frac{\text{constant term of } p}{\text{constant term of } (x-p)^n}(x - p)^n \in \mathbb{C}[x]$ is a polynomial that sends T to 0 and has zero constant term. ∎

3.9 Year 105

Question 39

Prove the **Cayley-Hamilton Theorem**: If V is a finite dimensional vector space and $T \in \text{End}(V)$ has characteristic polynomial $p(x)$, then $p(T) = 0$.

Proof. **Cayley-Hamilton Theorem** has been addressed at the end of the section on Jordan form. Fix some basis β of V, let $\mathbb{F}$ be the underlying field of V, and consider the Jordan form of $[T]_\beta \in M_{\dim(V)}(\mathbb{F})$, which is a matrix J in general with coefficients in $\overline{\mathbb{F}}$. By definition, $[T]_\beta$ have characteristic polynomial p. Because $J = P[T]_\beta P^{-1}$ for some $P \in M_{\dim(V)}(\overline{\mathbb{F}})$ and in general $\det(ABA^{-1}) = \det(B)$, J have the characteristic polynomial same as $[T]_\beta$, i.e., p. Clearly $p(J) = 0$ by direct computation, since J is in Jordan form. Now, because $0 = p(J) = Pp([T]_\beta)P^{-1}$, we see $p([T]_\beta) = 0$, which implies $p(T) = 0$. ∎

Question 40

Let $\mathbb{F}$ be arbitrary field and $A \in M_{m \times n}(\mathbb{F})$. Show that:

(i) The row ranks of A equals the column ranks of A.

(ii) If A have rank r, then AA^t also have rank r.

Proof. Suppose $\text{rank}(A) = r$. Let $\alpha \triangleq \{e_1, \ldots, e_n\}$ and $\beta \triangleq \{e'_1, \ldots, e'_m\}$ respectively be basis for $\mathbb{F}^n$ and $\mathbb{F}^m$ such that $\{A(e_1) = e'_1, \ldots, A(e_r) = e'_r\}$ form a basis for image of A,[6] and let γ and δ respectively be the standard basis for $\mathbb{F}^n$ and $\mathbb{F}^m$. Let $\alpha^\vee \triangleq \{\phi_1, \ldots, \phi_n\}, \beta^\vee \triangleq \{\psi_1, \ldots, \psi_m\}, \gamma^\vee$, and $\delta^\vee$ be the dual bases. Let $A^\vee : (\mathbb{F}^m)^* \to (\mathbb{F}^n)^*$ be the dual map of A. After setting these definitions, to get full credits of the first problem, one only need to show

$$[A^\vee]_{\delta^\vee}^{\gamma^\vee} = A^t$$

and show

$$A^\vee(\psi_i) = \begin{cases} \varphi_i & \text{if } i \leq r \\ 0 & \text{if } i > r \end{cases} \tag{3.7}$$

They are all routine to check. The second problem seems to be wrong. Consider the field $\mathbb{F} = \mathbb{Z}_2$, and

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

---

[6]To construct α and β, consider: $\{A(v_1), \ldots, A(v_r)\}$ is linearly independent then $v_1, \ldots, v_r$ are linearly independent.

Question 41

Let $\mathbb{F}$ be arbitrary field and let $\mathcal{C}$ be a possibly infinite collection of $n \times n$ $\mathbb{F}$-matrices that are all diagonalizable and commute. Prove that there exists some $P \in GL_n(\mathbb{F})$ such that PA_iP^{-1} are diagonal for all $i \in I$.

Proof. Because $\text{End}(\mathbb{F}^n)$ is finite dimensional, there exists some finite subset of $\mathcal{C}$ whose spans contains all $\mathcal{C}$. Clearly, if there exists a basis that contains only eigenvectors of the elements this finite subset, then every linear operator in $\mathcal{C}$ with respect to this basis is diagonal, since each of them is a finite linear combination of the elements of the finite subset.

We have reduced the case of $\mathbb{C}$ being finite. Write $\mathcal{C} \triangleq \{A_1, \ldots, A_m\}$. Let Λ_i be the set of eigenvalues of A_i. It remains to show

$$\mathbb{F}^n = \bigoplus_{\lambda_i \in \Lambda_i} \left(E(\lambda_1, A_1) \cap \cdots \cap E(\lambda_m, A_m) \right) \tag{3.8}$$

We shall show Equation 3.8 by induction on m. The base case is when $m = 2$. Let $\lambda \in \Lambda_1$. Clearly, for each $v \in E(\lambda, A_1)$, we have

$$(A_1 - \lambda I)(A_2 v) = A_2(A_1 - \lambda I)v = 0 \tag{3.9}$$

We have shown eigenspaces with respect to A_1 are all stable under A_2. Because restriction of diagonalizable linear operator is still diagonalizable, for each $\lambda \in \Lambda_1$, we have

$$E(\lambda, A_1) = \bigoplus_{\mu \in \Lambda_2} E(\mu, A_2|_{E(\lambda, A_1)}) = \bigoplus_{\mu \in \Lambda_2} E(\lambda, A_1) \cap E(\mu, A_2)$$

where the second equality hold trivially. The base case $m = 2$ now follows from $\mathbb{F}^n = \bigoplus_{\lambda \in \Lambda_1} E(\lambda, A_1)$. We now prove the inductive case. Suppose

$$\mathbb{F}^n = \bigoplus_{\lambda_i \in \Lambda_i} \left(E(\lambda_1, A_1) \cap \cdots \cap E(\lambda_{m-1}, A_{m-1}) \right)$$

Fix $\lambda_i \in \Lambda_i$. Again for the same reason Equation 3.9, for each $i \leq m-1$, eigenspace $E(\lambda_i, A_i)$ is stable under A_m, so clearly $V \triangleq E(\lambda_1, A_1) \cap \cdots \cap E(\lambda_{m-1}, A_{m-1})$ is stable under A_m. Again because restriction of diagonalizable linear operator is still diagonalizable, we have

$$V = \bigoplus_{\lambda \in \Lambda_m} E(\lambda, A_m|_V) = \bigoplus_{\lambda \in \Lambda_m} E(\lambda_1, A_1) \cap \cdots \cap E(\lambda_{m-1}, A_{m-1}) \cap E(\lambda, A_m)$$

where the second equality hold by definition of V as one can check. We have shown Equation 3.8. ■

75

Prove the **Spectral Theorem for finite dimensional complex vector space**.

Proof. This has been addressed in the section on spectral theorem. Long story short, you consider Jordan form, use Gram-Schmidt process to get Schur's form, and use the hypothesis of the linear operator being normal to deduce that its Schur's form is in fact diagonal, via direct computation. ∎

Question 43

Let V be a finite-dimensional complex vector space, and let $T \in \operatorname{End}(V)$ has characteristic polynomial p. Suppose $p = q_1 q_2$ for some q_1 and $q_2 \in \mathbb{C}[x]$ that have no common roots in $\mathbb{C}$. Show that there exists subspace W_1 and $W_2 \subseteq V$ that satisfies

(i) $V = W_1 \oplus W_2$.

(ii) W_1 and W_2 are both stable under T.

(iii) the characteristic polynomials of $T|_{W_1}$ and $T|_{W_2}$ are respectively q_1 and q_2.

Proof. Suppose q_1 and q_2 respectively have distinct roots $\{\lambda_1, \ldots, \lambda_r\}$ and $\{\lambda_{r+1}, \ldots, \lambda_m\}$. Recall that we prove in Theorem 1.2.3 that

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$$

We claim

$$W_1 \triangleq G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_r, T) \text{ and } W_2 \triangleq G(\lambda_{r+1}, T) \oplus \cdots \oplus G(\lambda_m, T)$$

suffice. Clearly they are stable under T. It remains to show $T|_{W_1}$ and $T|_{W_2}$ respectively have characteristic polynomial q_1 and q_2. Let $\alpha \subseteq W_1$ and $\beta \subseteq W_2$ be such that $\alpha \cup \beta$ form a Jordan basis for $V = W_1 \oplus W_2$. It is then routine to check (in order) that:

(i) α and β each form a basis respectively for W_1 and W_2.

(ii) $[T]_{\alpha \cup \beta} = [T|_{W_1}]_\alpha \oplus [T|_{W_2}]_\beta$. (the $\oplus$ here simply means putting a $n \times n$ matrix and a $m \times m$ matrix to form a $(n+m) \times (n+m)$ matrix the way you imagined).

(iii) α, β are respectively Jordan basis for $T|_{W_1}, T|_{W_2}$.

(iv) $T|_{W_1}$ and $T|_{W_2}$ respectively have characteristic polynomials q_1 and q_2.

∎

3.10 Year 104

> **Question 44**
>
> (i) For each $x \in \mathbb{R}$, let V_x be the subspace of $\mathbb{R}^4$ spanned by
>
> $$(x, 1, 1, 1), (1, x, 1, 1), (1, 1, x, 1), (1, 1, 1, x)$$
>
> Determine all $x \in \mathbb{R}$ such that $\dim(V_x) \leq 3$.
>
> (ii) Find the dimension and a basis for the space of $\mathbb{R}$-linear maps from $\mathbb{R}^5$ to $\mathbb{R}^3$ whose kernel contains $(0, 2, -3, 0, 1)$.

Proof. The first problem boils down to factorizing the determinant:

$$\begin{vmatrix} x & 1 & 1 & 1 \\ 1 & x & 1 & 1 \\ 1 & 1 & x & 1 \\ 1 & 1 & 1 & x \end{vmatrix} \tag{3.10}$$

into linear terms. Even though one may just perform direct computation and "guesses work" to find the solutions, we offer a different approach: Because Matrix 3.10 equals to $(x-1)I + J$ where all entries of $J \in M_4(\mathbb{R})$ are 1, by Theorem 3.10.1, the determinant of Matrix 3.10 is $(x-1)^3(x-1+4)$. This implies the solutions of the first problem is $\{1, -3\}$.

In the second problem, if we identify $\mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^5, \mathbb{R}^3)$ with $M_{3 \times 5}(\mathbb{R})$ the natural way, then the space in the question consists precisely of those matrix whose every row is orthogonal to $(0, 2, -3, 0, 1)$. Clearly, the subspace of $\mathbb{R}^5$ consisting of vectors orthogonal to $(0, 2, -3, 0, 1)$ is 4-dimensional, and says, if $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ is a basis, then the space in the question is exactly 12-dimensional with basis $\{E_{i,j} : 1 \leq i \leq 3, 1 \leq j \leq 4\}$ where the non i-th row of $E_{i,j}$ are all 0 and the i-th row of $E_{i,j}$ is $\mathbf{v}_j$. We have reduced the problem into finding $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$. Clearly, $\mathbf{v}_1 \triangleq (1, 0, 0, 0, 0), \mathbf{v}_2 \triangleq (0, 1, 1, 0, 1), \mathbf{v}_3 \triangleq (0, 1, 0, 0, -2), \mathbf{v}_4 \triangleq (0, 0, 0, 1, 0)$ suffices. $\blacksquare$

Theorem 3.10.1. (Determinant of special matrices) Let $\mathbb{F}$ be some field, $\alpha, \beta \in \mathbb{F}$, and every entry of $J \in M_n(\mathbb{F})$ be 1. The matrix $\alpha I + \beta J$ is $\mathbb{F}$-diagonalizable with eigenvalues:

$$\overbrace{\alpha, \alpha, \ldots, \alpha}^{n-1 \text{ copies}}, \quad \text{and} \quad \alpha + n\beta \tag{3.11}$$

Therefore,

$$\det(\alpha I + \beta J) = (\alpha)^{n-1}(\alpha + n\beta) \tag{3.12}$$

Proof. Noting $\text{rank}_{\mathbb{F}}(J) = 1$, we see that J is $\mathbb{F}$-diagonalizable with eigenvalues:[7]

$$\overbrace{0, 0, \ldots, 0,}^{n-1 \text{ copies}} \quad \text{and} \quad n$$

Let $\mathbf{v} \in \mathbb{F}^n$. Clearly, $J\mathbf{v} = \mathbf{0} \implies (\alpha I + \beta J)\mathbf{v} = \alpha\mathbf{v}$. Also, clearly, $J\mathbf{v} = n\mathbf{v} \implies (\alpha I + \beta J)\mathbf{v} = (\alpha + n\beta)\mathbf{v}$. This proves Equation 3.11. To see Equatoin 3.12, just consider the Jordan form. ∎

Question 45

Let

$$A \triangleq \begin{bmatrix} -1 & 4 & -2 \\ -2 & 5 & -2 \\ -1 & 2 & 0 \end{bmatrix}$$

(i) Compute the characteristic polynomial of A.

(ii) If $f(x) \triangleq (x-3)^2 + 5$, find the eigenvalue of $f(A)$.

(iii) Find an orthogonal matrix $P \in M_3(\mathbb{R})$ such that $P^{-1}AP$ is diagonal.

Proof. The characteristic polynomial of A is $(x-2)(x-1)^2$. By Spectral Theorem of polynomial (Deduced from our Theorem 1.6.1), the set of eigenvalues of $f(A)$ is then exactly $\{6, 9\}$. The third question is clearly wrong: There is no such P. By Spectral Theorem for finite dimensional real vector space, for such P to exist, A must be symmetric. ∎

Question 46

Let $A, B \in \text{GL}_n(\mathbb{R})$. Show that

(i) If $ABA^{-1}B^{-1} = cI$, then $c = \pm 1$.

(ii) If $AB - BA = cI$, then $c = 0$.

Proof. Note that $ABA^{-1}B^{-1} = cI$ implies $AB = cBA$, which implies $\det(AB) = c^n \det(BA)$, which implies $c^n = 1$, which implies $c \in \{-1, 1\}$.

Note that $AB - BA = cI$ implies $AB = cI + BA$, which implies $\text{trace}(AB) = nc + \text{trace}(BA)$, which implies $nc = 0$, which implies $c = 0$. ∎

---

[7]Note that here, by $n \in \mathbb{F}$, we mean $\overbrace{1 + \cdots + 1}^{n \text{ copies}} \in \mathbb{F}$

Question 47

Given $A \in M_n(\mathbb{R})$, show

$$A^3 = A \implies \text{rank}(A) = \text{tr}(A^2)$$

Proof. $A^3 = A$ implies that $x^3 - x = x(x-1)(x+1)$ is a polynomial that send A to 0. By long division, minimal polynomial of A must divides $x(x-1)(x+1)$. This implies that the set of eigenvalues of A is contained by $\{-1, 0, 1\}$. It then follows from considering the Jordan form of A that $\text{rank}(A) = \text{tr}(A^2)$. ∎

Question 48

Given $A \in M_n(\mathbb{R})$, show

$$\text{rank}(A) + \text{rank}(I - A) = n \implies A^2 = A$$

Proof. Suppose A has distinct eigenvalues $\lambda_1 = 0, \lambda_1 = 1, \lambda_2, \ldots, \lambda_k \in \mathbb{C}$ with algebraic dimension d_i (So $d_i = 0$ if λ_i is not an eigenvalue), and suppose the Jordan form of A is

$$\bigoplus_{i=1}^{k} J(\lambda_i, r_1) \oplus \cdots \oplus J(\lambda_i, r_{m_i})$$

Compute

$$\text{rank}(A) = n - m_1 \text{ and } \text{rank}(I - A) = n - m_2$$

This by premise give us

$$m_1 + m_2 = n$$

which implies that the $d_2 = \cdots = d_k = 0$ and all Jordan blocks of A have size 1. This implies that the minimal polynomial of A is either $x, x - 1$, or $x(x - 1)$. Either way, this implies $A^2 = A$. ∎

Question 49

Consider $A, B \in M_n(\mathbb{Q})$ with $A^n = 0$ but $A^{n-1} \neq 0$. Show

$$AB = BA \implies B = F(A) \text{ for some } F \in \mathbb{Q}[x] \text{ such that } \deg(F) < n$$

Proof. Looking at the Jordan form of A, we know there exists some $\mathbf{e} \in \mathbb{Q}^n$ such that $\{\mathbf{e}, A\mathbf{e}, \ldots, A^{n-1}\mathbf{e}\}$ forms a basis for $\mathbb{Q}^n$. Therefore, there exists some $c_0, \ldots, c_{n-1} \in \mathbb{Q}$

such that if we define $F(x) \triangleq c_{n-1}x^{n-1} + \cdots + c_1 x + c_0$, then $F(A)$ and B have the same action on $\mathbf{e}$. To see such F suffices, we only have to show B and $F(A)$ moreover agree on rest of the $\{\mathbf{e}, A\mathbf{e}, \ldots, A^{n-1}\mathbf{e}\}$, which is a consequence of the commutation:

$$BA^i\mathbf{e} = A^i B\mathbf{e} = A^i F(A)\mathbf{e} = F(A)A^i\mathbf{e}$$

$\blacksquare$

3.11 Year 103

Question 50

Let $A \in \mathrm{M}_{m \times n}(\mathbb{F}), B \in \mathrm{M}_{n \times p}(\mathbb{F})$. Show that

(i) $\mathrm{rank}\, AB \leq \min\{\mathrm{rank}\, A, \mathrm{rank}\, B\}$.

(ii) If $n = p$ and B is invertible, then $\mathrm{rank}\, AB = \mathrm{rank}\, A$.

Proof. Clearly, $\mathrm{rank}\, AB \leq \mathrm{rank}\, A$. We also have:

$$\mathrm{rank}\, AB = \mathrm{rank}\, B^t A^t \leq \mathrm{rank}\, B^t = \mathrm{rank}\, B$$

Let $\{v_1, \ldots, v_n\} \subseteq \mathbb{F}^n$ be a basis such that $\{Av_1, \ldots, Av_t\}$ forms a basis for image of A. The fact that $\mathrm{rank}\, AB = \mathrm{rank}\, B$ then follows from noting $Av_i = ABB^{-1}v_i \in \mathrm{Im}\, AB$. $\blacksquare$

Question 51

Let $A \in M_{m \times n}(\mathbb{C})$. Show that

$$\mathrm{rank}\, A^*A = \mathrm{rank}\, A$$

Proof. By SVD, there exists $P \in U(m), Q \in U(n)$ and diagonal D with nonnegative spectrum such that $A = PDQ^*$. IT then follows from $A^*A = QD^2Q^*$ that $\mathrm{rank}\, A^*A = \mathrm{rank}\, D^2 = \mathrm{rank}\, D = \mathrm{rank}\, A$. $\blacksquare$

Question 52

Let $A \in \mathrm{M}_n(\mathbb{R})$ be symmetric. Show that

(i) If λ is an eigenvalue in $\mathbb{C}$, then $\lambda \in \mathbb{R}$.

(ii) A possess an orthonormal eigenbasis in $\mathbb{R}^n$.

Proof. $\blacksquare$

Chapter 4

NTU MATH M.A. Program Entrance Exam Round 2

4.1 NTU Round 2 Year 106

> **Question 53**
>
> Show that if
> $$\mathbb{R}^n = W_1 \cup W_2 \cup \cdots \cup W_k \cup \cdots$$
> then $\mathbb{R}^n = W_k$ for some k.

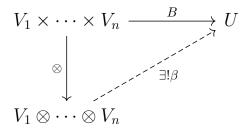*Proof.* ∎

Chapter 5

Archived

5.1 Tensor Algebra

Abstract

In this section, by the term **ring**, we mean a ring with a multiplication identity, and by the term **real algebra**, we mean a real vector space equipped with a vector multiplication compatible with both scalar multiplication and addition. In this definition, for a real algebra A to be a ring, A must be associative. By the term **ideal**, we mean a 2-sided ideal. If we say a multi-linear map $M : V^k \to Z$ is **alternating**, we mean that M maps $(v_1, \ldots, v_n)$ to 0 if two arguments coincide.

Given a finite collection $(V_1, \ldots, V_n)$ of finite dimensional real vector space, by the term **tensor product of** $V_1, \ldots, V_n$, we mean a real vector space usually denoted by $V_1 \otimes \cdots \otimes V_n$ and a multilinear map $\otimes : V_1 \times \cdots \times V_n \to V_1 \otimes \cdots \otimes V_n$ satisfying the universal property: If $B : V_1 \times \cdots \times V_n \to Z$ is a multilinear map, then there exists a unique linear map $\beta : V_1 \otimes \cdots \otimes V_n$ such that

$$B(v_1, \ldots, v_n) = \beta(v_1 \otimes \cdots \otimes v_n)$$

In other words, we have the commutative diagram

This approach indeed define a pair of vector space and multilinear map uniquely up to isomorphism, in the sense of Theorem 5.1.1, where we define the isomorphism between tensor product.

Theorem 5.1.1. (Uniqueness of Tensor product) Given a finite collection $(V_1, \ldots, V_n)$ of finite dimensional real vector space, if $V_1 \otimes \cdots \otimes V_n, V_1 \otimes' \cdots \otimes' V_n$ both satisfy the universal property, then there exists an linear isomorphism $T : V_1 \otimes \cdots \otimes V_n \to V_1 \otimes' \cdots \otimes' V_n$ such that

$$T(v_1 \otimes \cdots \otimes v_n) = v_1 \otimes' \cdots \otimes' v_n$$

Proof. Because $V_1 \otimes \cdots \otimes V_n$ satisfies the universal property, there exists a linear map $T : V_1 \otimes \cdots \otimes V_n \to V_1 \otimes' \cdots \otimes' V_n$ such that

$$\otimes' = T \circ \otimes$$

It remains to show T is bijective. Similarly, because $V_1 \otimes' \cdots \otimes' V_n$ satisfies the universal property, there exists a linear map $T' : V_1 \otimes' \cdots \otimes' V_n \to V_1 \otimes \cdots \otimes V_n$ such that

$$\otimes = T' \circ \otimes'$$

Composing the two equations, we have

$$\otimes' = T \circ T' \circ \otimes'$$

It then follows from uniqueness of the induced linear map in universal property that $T \circ T' = \mathbf{id} : V_1 \otimes' \cdots \otimes' V_n \to V_1 \otimes' \cdots \otimes' V_n$. This implies T is indeed bijective. ∎

We have shown that tensor products is unique up to isomorphism. A construction further shows that if B_i are bases for V_i, then

$$\{v_1 \otimes \cdots \otimes v_n : v_i \in B_i \text{ for all } 1 \leq i \leq n\} \text{ form a basis for } V_1 \otimes \cdots \otimes V_n$$

Theorem 5.1.2. (Associativity of the Tensor product) Given three finite-dimensional real vector spaces X, Y, Z, there exists a unique linear isomorphism $F : X \otimes Y \otimes Z \to (X \otimes Y) \otimes Z$ that satisfy

$$F(x \otimes y \otimes z) = (x \otimes y) \otimes z$$

Proof. Define $f : X \times Y \times Z \to (X \otimes Y) \otimes Z$ by

$$f(x, y, z) \triangleq (x \otimes y) \otimes z$$

It follows from the universal property that there exists a unique linear map $F : X \otimes Y \otimes Z \to (X \otimes Y) \otimes Z$ such that

$$F(x \otimes y \otimes z) = f(x, y, z) = (x \otimes y) \otimes z$$

It remains to show F is bijective. For all $z \in Z$, define $h_z : X \times Y \to X \otimes Y \otimes Z$ by

$$h_z(x, y) \triangleq x \otimes y \otimes z$$

If follows from the universal property that there exists a unique linear map $H_z : X \otimes Y \to X \otimes Y \otimes Z$ such that

$$H_z(x \otimes y) = h_z(x, y) = x \otimes y \otimes z$$

Define $h : (X \otimes Y) \times Z \to X \otimes Y \otimes Z$ by

$$h(v, z) \triangleq H_z(v)$$

It is clear that h in linear in $(X \otimes Y)$. We now show h is linear in Z, that is

$$H_{c_1 z_1 + z_2} = c_1 H_{z_1} + H_{z_2}$$

By definition,

$$(c_1 H_{z_1} + H_{z_2})(x \otimes y) = c_1 x \otimes y \otimes z_1 + x \otimes y \otimes z_2$$
$$= x \otimes y \otimes (c_1 z_1 + z_2) = h_{c_1 z_1 + z_2}(x, y)$$

It then follows from the uniqueness part of the universal property that $H_{c_1 z_1 + z_2} = c_1 H_{z_1} + H_{z_2}$. (done)

We have shown h is indeed bilinear. It follows from the universal property that there exists a unique linear map $H : (X \otimes Y) \otimes Z \to X \otimes Y \otimes Z$ such that

$$H((x \otimes y) \otimes z) = h(x \otimes y, z) = H_z(x \otimes y) = x \otimes y \otimes z$$
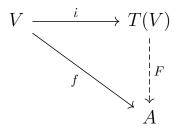
Let $\otimes : X \times Y \times Z \to X \otimes Y \otimes Z$ denotes the tensor product, we now have

$$\otimes = H \circ F \circ \otimes$$

It then follows from universal property that $H \circ F = \mathbf{id} : X \otimes Y \otimes Z \to X \otimes Y \otimes Z$. This implies F is indeed bijective. (done) ∎

Let V be a finite-dimensional real vector space. By its **tensor algebra**, we mean any real associative algebra $T(V)$ with an injective linear map $i : V \to T(V)$ that satisfies the universal property: If A is a real associative algebra and $f : V \to A$ is a linear map, then there exists a unique algebra homomorphism $F : T(V) \to A$ such that the

diagram

commutes. The proof that such definition is indeed unique up to isomorphism is similar to that of Theorem 5.1.1 and thus omitted. We now give the most useful construction.

Let V be finite-dimensional real vector space. We use the notation

$$T^n(V) \triangleq \overbrace{V \otimes \cdots \otimes V}^{n \text{ copies}}$$

and call $T^n(V)$ the n-**th tensor power** of V or the n-**fold tensor product** of V. Define

$$T(V) \triangleq \bigoplus_{n=0}^{\infty} T^n(V)$$
$$= \mathbb{R} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots$$

and define for all $f, g \in T(V)$ the multiplication

$$(fg)(n) \triangleq \sum_{k=0}^{n} f(k)g(n-k)$$

where

$$\left(\sum_I a_I v_{I(1)} \otimes \cdots \otimes v_{I(k)} \right) \left(\sum_J b_J v_{J(1)} \otimes \cdots \otimes v_{J(l)} \right)$$
$$\triangleq \sum_{I,J} a_I b_J v_{I(1)} \otimes \cdots \otimes v_{I(k)} \otimes v_{J(1)} \otimes \cdots \otimes v_{J(l)}$$

where $\{v_1, \ldots, v_m\}$ is some basis for V, I run through the set of function that maps $\{1, \ldots, k\}$ into $\{1, \ldots, m\}$ and J run through the set of function that maps $\{1, \ldots, l\}$ into $\{1, \ldots, m\}$. For example, given two elements

$$(5, 0, v_1 \otimes v_2, 0, 0, \ldots) \text{ and } (7, v_3, 0, 0, \ldots)$$

of $T(V)$, their product is defined to be

$$(35, 5v_3, 7v_1 \otimes v_2, v_1 \otimes v_2 \otimes v_3, 0, 0, \dots)$$

Tedious effort shows that our multiplication is consistent with abuse of notation in the sense that if $f, g \in T(V)$ is defined by

$$f(k) \triangleq \begin{cases} w_1 \otimes \cdots \otimes w_n & \text{if } k = n \\ 0 & \text{if otherwise} \end{cases} \quad \text{and} \quad g(k) \triangleq \begin{cases} w_{n+1} \otimes \cdots \otimes w_{n+l} & \text{if } k = l \\ 0 & \text{if otherwise} \end{cases}$$

then

$$(fg)(k) = \begin{cases} w_1 \otimes \cdots \otimes w_{n+l} & \text{if } n = k + l \\ 0 & \text{if otherwise} \end{cases}$$

does form an associative algebra with multiplication identity $1 \in \mathbb{R}$. Thus, $T(V)$ is in fact a ring. Let $I(V) \subseteq T(V)$ be the ideal generated by $\{v \otimes v : v \in V\}$. By definition, ideal $I(V)$ is a subgroup of $T(V)$. To see that $I(V)$ is closed under scalar multiplication, observe that for all $t \in \mathbb{R}$ and $x \in T(V)$, the scalar multiplication tx is identical to tx where t is treated as an element of $T(V)$, so it follows from definition of ideal that $I(V)$ is also a vector subspace of $T(V)$. Let $\{v_1, \dots, v_n\}$ be a basis for V, and let S be the set of function that maps $\{1, \dots, n\}$ into $\{1, \dots, k\}$. We know for a fact that

$$T^k(V) = \text{span}\left\{v_{I(1)} \otimes \cdots \otimes v_{I(k)} : I \in S\right\}$$

If we define $I^k(V) \triangleq I(V) \cap T^k(V)$, one then have

$$I^k(V) = \text{span}\left\{v_{I(1)} \otimes \cdots \otimes v_{I(k)} : I(j) = I(j+1) \text{ for some } j\right\} \qquad (5.1)$$

This is proved by showing $I^0(V) \oplus I^1(V) \oplus I^2(V) \oplus \cdots$ is indeed the smallest ideal containing $\{v \otimes v : v \in V\}$. Define an equivalence class on $T(V)$ by

$$x \sim y \overset{\triangle}{\Longleftrightarrow} x - y \in I(V)$$

Because ideal form a subgroup, we see that our definition indeed give an equivalence relation. We then can define on the set of equivalence class $T(V) \setminus I(V)$ addition, scalar multiplication and vector multiplication

$$[x] + [y] \triangleq [x+y] \text{ and } [x] \wedge [y] \triangleq [xy] \text{ and } c[x] \triangleq [cx]$$

which is well defined and form an algebra as one can check. We call this algebra $T(V) \setminus I(V)$ the **exterior algebra** $\wedge^*(V)$. Note that if we refer to $v, w \in T^k(V)$ as elements of $\wedge^*(V)$, we mean $[v], [w]$. Immediately, we see that the wedge product is **alternating** in the sense that if $v \in V$, then

$$v \wedge v = 0$$

and is **anti-symmetric** in the sense that if $v, w \in V$, then

$$v \wedge w = -w \wedge v$$

We use the notation

$$\wedge^k(V) \triangleq \left\{ [x] \in \wedge^*(V) : x \in \overbrace{V \otimes \cdots \otimes V}^{k \text{ copies}} \right\}$$

Immediately from Equation 5.1, we see that $\wedge^k(V)$ is the vector space
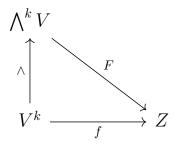
$$\text{span} \left\{ v_{I(1)} \wedge \cdots \wedge v_{I(k)} : I \in S \right\}$$

where $\{v_1, \ldots, v_n\}$ is a basis for V and S is the set of function that maps $\{1, \ldots, k\}$ into $\{1, \ldots, n\}$. If we define the vector subspace $I^k(V) \triangleq T^k(V) \cap I(V)$, there exists a natural vector space isomorphism

$$\wedge^k(V) \underset{\text{v.s.}}{\cong} T^k(V) \setminus I^k(V); [x] \leftrightarrow [x]$$

where $T^k(V) \setminus I^k(V)$ is the quotient vector space.

Theorem 5.1.3. (Universal mapping property for alternating k-linear map) For any vector space Z over $\mathbb{R}$ and any alternating k-linear map $f : V^k \to Z$, there is a unique linear map $F : \bigwedge^k V \to Z$ such that the diagram

commute, i.e.,

$$F(v_1 \wedge \cdots \wedge v_k) = f(v_1, \ldots, v_k) \text{ for all } v_1, \ldots, v_k \in V$$

Proof. By universal property of tensor product, there exists unique linear map $h : T^k(V) \to Z$ such that

$$h(v_1 \otimes \cdots \otimes v_k) = f(v_1, \cdots, v_k)$$

Because f is alternating, we see from the characterization of $I^k(V)$ given in Equation 5.1 that h vanishes on $I^k(V)$. We then can induce a linear map

$$F : \wedge^k(V) \cong \frac{T^k(V)}{I^k(V)} \to Z$$

by $F([x]) \triangleq h(x)$. This then give us the desired

$$F(v_1 \wedge \cdots \wedge v_k) = h(v_1 \otimes \cdots \otimes v_k) = f(v_1, \ldots, v_k)$$

Note that F is unique because all such linear map take the same values on $\{v_{I(1)} \wedge \cdots \wedge v_{I(k)} : I \in$ which spans $\wedge^k(V)$. ∎

Let $\{w_1, \ldots, w_l\} \subseteq V$ be linear independent. An immediate consequence of the universal mapping property for alternating k-linear map is that one may define alternating multilinear $f : V^l \to \mathbb{R}$ by

$$B(v_1, \ldots, v_l) \triangleq \det(M) \text{ where } v_i = \sum_j M_{i,j} w_j$$

and see that $F : \wedge^l(V) \to \mathbb{R}$ take $w_1 \wedge \cdots \wedge w_l$ to 1. This implies that

$$w_1 \wedge \cdots \wedge w_l \neq 0$$

Theorem 5.1.4. (Anti-symmetry of wedge product) If $\alpha \in \wedge^k(V), \beta \in \wedge^l(V)$, then $\alpha \wedge \beta = (-1)^{kl}(\beta \wedge \alpha)$.

Proof. Let $v_1, \ldots, v_n$ be a basis of V. Let S_k be the space of function that maps $\{1, \ldots, k\}$ into $\{1, \ldots, n\}$, S_l be the space of function that maps $\{1, \ldots, l\}$ into $\{1, \ldots, n\}$. We may then write

$$\alpha = \sum_{I \in S_k} a_I(v_{I(1)} \wedge \cdots \wedge v_{I(k)}) \text{ and } \beta = \sum_{J \in S_l} b_J(v_{J(1)} \wedge \cdots \wedge v_{J(l)})$$

and compute

$$\alpha \wedge \beta = \sum_{I \in S_k, J \in S_l} a_I b_J (v_{I(1)} \wedge \cdots \wedge v_{I(k)} \wedge v_{J(1)} \wedge \cdots \wedge v_{J(l)})$$

$$= \sum_{I \in S_k, J \in S_l} (-1) a_I b_J (v_{I(1)} \wedge \cdots \wedge v_{J(1)} \wedge v_{I(k)} \cdots \wedge \cdots \wedge v_{J(l)})$$

$$= \sum_{I \in S_k, J \in S_l} (-1)^k a_I b_J (v_{J(1)} \wedge v_{I(1)} \wedge \cdots \wedge v_{I(k)} \wedge v_{J(2)} \wedge \cdots \wedge v_{J(l)})$$

$$= \sum_{I \in S_k, J \in S_l} (-1)^{kl} a_I b_J (v_{J(1)} \wedge \cdots \wedge v_{J(l)} \wedge v_{I(1)} \wedge \cdots \wedge v_{I(k)}) = (-1)^{kl} \beta \wedge \alpha$$

∎

Following from Theorem 5.1.4, Equation 5.1 and tedious effort, one can see that if $\{v_1, \ldots, v_n\}$ is a basis for V, then

$$\{v_{i_1} \wedge \cdots \wedge v_{i_k} : 1 \leq i_1 < \cdots < i_k \leq n\}$$

form a basis for $\wedge^k(V)$. If $A : V \to W$ is a linear map, we define linear map $\wedge^k A : \wedge^k(V) \to \wedge^k(W)$ by linear extension of

$$\wedge^k(A)(v_1 \wedge \cdots \wedge v_n) = A v_1 \wedge \cdots \wedge A v_n$$

Note that if $A : V \to V$ and $\dim(V) = n$, then $\wedge^n A : \wedge^n(V) \to \wedge^n(V)$ is given by the determinant since given basis $\{v_1, \ldots, v_n\}$, we have

$$\wedge^n A(v_1 \wedge \cdots \wedge v_n) = \left(\sum_j A_{j,1} v_j \right) \wedge \cdots \wedge \left(\sum_j A_{j,n} v_j \right)$$

$$= \sum_{\sigma \in S_n} A_{\sigma(1),1} \cdots A_{\sigma(n),n} v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(n)}$$

$$= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) A_{\sigma(1),1} \cdots A_{\sigma(n),n} v_1 \wedge \cdots \wedge v_n$$

5.2 Operator Norm

Abstract

This section introduces the concept of the operator norm and proves some fundamental results related operator norm and finite-dimensional normed spaces. For example, we establish results such as a linear operator being bounded if and only if it is continuous and the equivalence of all norms on finite-dimensional vector spaces.

In this section, and particularly in functional analysis, we say a function T between two metric space is a **bounded operator** if T always map bounded set to bounded set. In particular, if T is a linear transformation between two normed space, we say T is a **bounded linear operator**. Now, suppose $\mathcal{X}, \mathcal{Y}$ are two normed space over $\mathbb{R}$ or $\mathbb{C}$. In space $L(\mathcal{X}, \mathcal{Y})$, alternatively, we can define

$$T \text{ is bounded} \quad \overset{\triangle}{\iff} \quad \exists M \in \mathbb{R}, \forall x \in \mathcal{X}, \|Tx\| \leq M\|x\|$$

The proof of equivalency is simple. For ($\longrightarrow$), observe

$$\|Tx\| = \|x\| \cdot \|T\frac{x}{\|x\|}\| \leq \left(\sup\{\|Ty\| : \|y\| = 1\} \right)\|x\|$$

For ($\longleftarrow$), observe

$$\|Tx - Ty\| = \|T(x-y)\| \leq M\|x - y\|$$

We first show that a linear transformation is continuous if and only if it is bounded.

Theorem 5.2.1. (Liner Operator is Bounded if and only if it is Continuous)
Given two normed space $\mathcal{X}, \mathcal{Y}$ over $\mathbb{R}$ or $\mathbb{C}$ and $T \in L(\mathcal{X}, \mathcal{Y})$, we have

$$T \text{ is a bounded operator} \quad \iff \quad T \text{ is continuous on } \mathcal{X}$$

Proof. If T is bounded, we see that T is Lipschitz.

$$\|Tx - Ty\| \leq M\|x - y\|$$

Now, suppose T is linear and continuous at 0. Let ϵ satisfy

$$\sup_{\|y\| \leq \epsilon} \|Ty\| \leq 1$$

Observe that for all $x \in \mathcal{X}$, we have

$$\|Tx\| = \frac{\|x\|}{\epsilon}\left\|T\frac{\epsilon x}{\|x\|}\right\| \leq \frac{\|x\|}{\epsilon}$$

∎

Here, we introduce a new terminology, which shall later show its value. Given a set X, we say two metrics d_1, d_2 on X are **equivalent**, and write $d_1 \sim d_2$, if we have

$$\exists m, M \in \mathbb{R}^+, \forall x, y \in X, md_1(x, y) \leq d_2(x, y) \leq Md_1(x, y)$$

Now, given a fixed vector space V, naturally, we say two norms $\|\cdot\|_1, \|\cdot\|_2$ on V are **equivalent** if

$$\exists m, M \in \mathbb{R}^+, \forall x \in X, m\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1$$

We say two metric d_1, d_2 on X are **topologically equivalent** if the topology they induce on X are identical.

A few properties can be immediately spotted.

(a) Our definition of $\sim$ between metrics of a fixed X is an equivalence relation.

(b) Our definition of $\sim$ between norms on a fixed V is an equivalence relation.

(c) Equivalent norms induce equivalent metrics.

(d) Equivalent metrics are topologically equivalent.

We now prove if V is finite-dimensional, then all norms on V are equivalent. This property will later show its value, as used to prove linear map of finite-dimensional domain is always continuous

Theorem 5.2.2. (All Norms on Finite-dimensional space are Equivalent) Suppose V is a finite-dimensional vector space over $\mathbb{R}$ or $\mathbb{C}$. Then

$$\text{all norms on } V \text{ are equivalent}$$

Proof. Let $\{e_1, \ldots, e_n\}$ be a basis of V. Define ∞-norm $\|\cdot\|_\infty$ on V by

$$\left\|\sum \alpha_i e_i\right\|_\infty \triangleq \max |\alpha_i|$$

It is easily checked that $\|\cdot\|_\infty$ is indeed a norm. Fix a norm $\|\cdot\|$ on V. We reduce the problem into

$$\text{finding } m, M \in \mathbb{R}^+ \text{ such that } m\|x\|_\infty \leq \|x\| \leq M\|x\|_\infty$$

92

We first claim

$$M = \sum \|e_i\| \text{ suffices}$$

Compute

$$\|x\| = \left\|\sum \alpha_i e_i\right\| \leq \sum |\alpha_i| \|e_i\| \leq \|x\|_\infty \sum \|e_i\| = M\|x\|_\infty \text{ (done)}$$

Note that reverse triangle inequality give us

$$\left|\|x\| - \|y\|\right| \leq \|x - y\| \leq M\|x - y\|_\infty \tag{5.2}$$

Then we can check that

(a) $\|\cdot\| : \left(V, \|\cdot\|_\infty\right) \to \mathbb{R}$ is Lipschitz continuous because of Equation 5.2.

(b) $S \triangleq \{y \in V : \|y\|_\infty = 1\}$ is sequentially compact in $\|\cdot\|$ and non-empty.

Now, by EVT, we know $\min_{y \in S} \|y\|$ exists. Note that $\min_{y \in S} \|y\| > 0$, since $0 \notin S$. We claim

$$m = \min_{y \in S} \|y\| \text{ suffices}$$

Fix $x \in V$ and compute

$$m\|x\|_\infty = \|x\|_\infty(\min_{y \in S}\|y\|) \leq \|x\|_\infty \cdot \left\|\frac{x}{\|x\|_\infty}\right\| = \|x\| \text{ (done)} \text{ (done)}$$

∎

Theorem 5.2.3. (Linear map of Finite-dimensional Domain is always Continuous) Given a finite-dimensional normed space $\mathcal{X}$ over $\mathbb{R}$ or $\mathbb{C}$, an arbitrary normed space $\mathcal{Y}$ over $\mathbb{R}$ or $\mathbb{C}$ and a linear transformation $T : \mathcal{X} \to \mathcal{Y}$, we have

$$T \text{ is continuous}$$

Proof. Fix $x \in \mathcal{X}, \epsilon$. We wish

to find δ such that $\forall h \in \mathcal{X} : \|h\| \leq \delta, \|T(x + h) - Tx\| \leq \epsilon$

Let $\{e_1, \ldots, e_n\}$ be a basis of $\mathcal{X}$. Note that $\|\sum \alpha_i e_i\|_1 \triangleq \sum |\alpha_i|$ is a norm. Because $\mathcal{X}$ is finite-dimensional, we know $\|\cdot\|$ and $\|\cdot\|_1$ are equivalent. Then, we can fix $M \in \mathbb{R}^+$ such that

$$\|x\|_1 \leq M\|x\| \quad (x \in V)$$

93

We claim
$$\delta = \frac{\epsilon}{M(\max \|Te_i\|)} \text{ suffices}$$
Fix $\|h\| \le \delta$ and express $h = \sum \alpha_i e_i$. Compute using linearity of T
$$
\begin{aligned}
\|T(x+h) - Tx\| &= \|\sum \alpha_i Te_i\| \\
&\le \sum |\alpha_i| \, \|Te_i\| \\
&\le \|h\|_1 (\max \|Te_i\|) \\
&\le M\|h\|(\max \|Te_i\|) = \epsilon \text{ (done)}
\end{aligned}
$$
∎

We now see that, because Linear transformation is bounded if and only if it is continuous and Linear map of finite-dimensional domain is always continuous, if $\mathcal{X}$ is finite-dimensional, then all linear map of domain $\mathcal{X}$ are bounded. A counter example to the generalization of this statement is followed.

> **Example 1 (Differentiation is an Unbounded Linear Operator)**
> $$\mathcal{X} = \left(\mathbb{R}[x]|_{[0,1]}, \|\cdot\|_\infty\right), D(P) \triangleq P'$$
> Note that $\{x^n\}_{n \in \mathbb{N}}$ is bounded in $\mathcal{X}$ and $\{D(x^n)\}_{n \in \mathbb{N}}$ is not.

Now, suppose $\mathcal{X}, \mathcal{Y}$ are two fixed normed spaces over $\mathbb{R}$ or $\mathbb{C}$. We can easily check that the set $BL(\mathcal{X}, \mathcal{Y})$ of bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$ form a vector space over whichever field $\mathcal{Y}$ is over.

Naturally, our definition of boundedness of linear operator derive us a norm on $BL(\mathcal{X}, \mathcal{Y})$, as followed

$$\|T\|_{\text{op}} \triangleq \inf \left\{ M \in \mathbb{R}^+ : \forall x \in \mathcal{X}, \|Tx\| \le M\|x\| \right\} \tag{5.3}$$

Before we show that our definition is indeed a norm, we first give some equivalent definitions and prove their equivalency.

Theorem 5.2.4. (Equivalent Definitions of Operator Norm) Given two fixed normed space $\mathcal{X}, \mathcal{Y}$ over $\mathbb{R}$ or $\mathbb{C}$, a bounded linear operator $T : \mathcal{X} \to \mathcal{Y}$, and define $\|T\|_{\text{op}}$ as in Equation 5.3, we have

$$\|T\|_{\text{op}} = \sup_{x \in \mathcal{X}, x \neq 0} \frac{\|Tx\|}{\|x\|}$$

Proof. Define $J \triangleq \{M \in \mathbb{R}^+ : \forall x \in \mathcal{X}, \|Tx\| \leq M\|x\|\}$ and observe

$$J = \left\{ M \in \mathbb{R}^+ : M \geq \frac{\|Tx\|}{\|x\|}, \forall x \neq 0 \in \mathcal{X} \right\}$$

This let us conclude

$$\sup_{x \in \mathcal{X}, x \neq 0} \frac{\|Tx\|}{\|x\|} = \min J = \|T\|_{\mathrm{op}}$$

∎

It is now easy to see

$$\|T\|_{\mathrm{op}} = \sup_{x \in \mathcal{X}, x \neq 0} \frac{\|Tx\|}{\|x\|} \qquad (5.4)$$

$$= \sup_{x \in \mathcal{X}, \|x\|=1} \|Tx\| \qquad (5.5)$$

It is not all in vain to introduce the equivalent definitions. See that the verification of $\|\cdot\|_{\mathrm{op}}$ being a norm on $BL(\mathcal{X}, \mathcal{Y})$ become simple by utilizing the equivalent definitions.

(a) For positive-definiteness, fix non-trivial T and fix $x \in \mathcal{X} \setminus N(T)$. Use Equation 5.4 to show $\|T\|_{\mathrm{op}} \geq \frac{\|Tx\|}{\|x\|} > 0$.

(b) For absolute homogenity, use Equation 5.5 and $\|Tcx\| = |c| \cdot \|Tx\|$.

(c) For triangle inequality, use Equation 5.5 and $\|(T_1 + T_2)x\| \leq \|T_1 x\| + \|T_2 x\|$.

Naturally, and very very importantly, Equation 5.4 give us

$$\|Tx\| \leq \|T\|_{\mathrm{op}} \cdot \|x\| \qquad (x \in \mathcal{X})$$

This inequality will later be the best tool to help analyze the derivatives of functions between Euclidean spaces, and perhaps better, it immediately give us

$$\frac{\|T_1 T_2 x\|}{\|x\|} \leq \frac{\|T_1\|_{\mathrm{op}} \cdot \|T_2\|_{\mathrm{op}} \cdot \|x\|}{\|x\|} = \|T_1\|_{\mathrm{op}} \cdot \|T_2\|_{\mathrm{op}}$$

Then Equation 5.4 give us

$$\|T_1 T_2\|_{\mathrm{op}} \leq \|T_1\|_{\mathrm{op}} \cdot \|T_2\|_{\mathrm{op}}$$

5.3 Cauchy-Schwarz for Positive semi-definite Hermitian form

> Sometimes, we do not require $\langle \cdot, \cdot \rangle : V^2 \to \mathbb{F}$ to be positive-definite, but only **positive semi-definite**, i.e. $\langle v, v \rangle \geq 0$ for all $v \in V$. In this case, we say $\langle \cdot, \cdot \rangle : V^2 \to \mathbb{F}$ is a **positive semi-definite Hermitian form**. If we again induce a functional $\|\cdot\| : V \to \mathbb{R}$ from some positive semi-definite Hermitian form using Equation **??**, then the functional $\|\cdot\| : V \to \mathbb{R}$ in general is not positive-definite, and is called a **semi norm**, since, as we later show, it still satisfies triangle inequality and absolute homogeneity.

Theorem 5.3.1. (Basic Property of Positive semi-definite Hermitian form) Given a positive semi-definite Hermitian form $\langle \cdot, \cdot \rangle : V^2 \to \mathbb{F}$ and $x, y \in V$, we have

$$\langle x, x \rangle = 0 \implies \langle x, y \rangle = 0$$

Proof. Assume $\langle x, y \rangle \neq 0$. Fix $t > \frac{\|y\|^2}{2|\langle x, y \rangle|^2}$. Compute

$$\begin{aligned}
\|y - t\langle y, x \rangle x\|^2 &= \|y\|^2 + \|(-t)\langle y, x \rangle x\|^2 + \langle -t\langle y, x \rangle x, y \rangle + \langle y, -t\langle y, x \rangle x \rangle \\
&= \|y\|^2 + t^2 |\langle x, y \rangle|^2 \|x\|^2 - t\langle y, x \rangle \langle x, y \rangle - t\langle x, y \rangle \langle y, x \rangle \\
&= \|y\|^2 - 2t |\langle x, y \rangle|^2 < 0 \; \text{CaC}
\end{aligned}$$

$\blacksquare$

Theorem 5.3.2. (Cauchy-Schwarz Inequality) Given a positive semi-definite Hermitian form $\langle \cdot, \cdot \rangle : V^2 \to \mathbb{C}$ on vector space V over $\mathbb{C}$, we have

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$$

Moreover, the equality hold true if x, y are linearly independent. In addition, if $\langle \cdot, \cdot \rangle : V^2 \to \mathbb{F}$ is an inner product, then the equality hold true only if x, y are linearly independent.

Proof. We first prove

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\| \qquad (x, y \in V)$$

Fix $x, y \in V$. Theorem 5.3.1 tell us $\|x\| = 0 \implies \langle x, y \rangle = 0$. Then we can reduce the problem into proving

$$\frac{|\langle x, y \rangle|^2}{\|x\|^2} \leq \|y\|^2$$

Set $z \triangleq y - \frac{\langle y, x \rangle}{\|x\|^2} x$. We then have

$$\langle z, x \rangle = \langle y - \frac{\langle y, x \rangle}{\|x\|^2} x, x \rangle = \langle y, x \rangle - \frac{\langle y, x \rangle}{\|x\|^2} \langle x, x \rangle = 0$$

Then from $y = z + \frac{\langle y, x \rangle}{\|x\|^2} x$, we can now deduce

$$\begin{aligned} \langle y, y \rangle &= \langle z + \frac{\langle y, x \rangle}{\|x\|^2} x, z + \frac{\langle y, x \rangle}{\|x\|^2} x \rangle \\ &= \langle z, z \rangle + \left| \frac{\langle y, x \rangle}{\langle x, x \rangle} \right|^2 \langle x, x \rangle \\ &= \langle z, z \rangle + \frac{|\langle x, y \rangle|^2}{\langle x, x \rangle} \end{aligned}$$

Because $\langle z, z \rangle \geq 0$, we now have

$$\langle y, y \rangle = \langle z, z \rangle + \frac{|\langle x, y \rangle|^2}{\langle x, x \rangle} \geq \frac{|\langle x, y \rangle|^2}{\langle x, x \rangle} \quad \text{(done)}$$

The equality hold true if and only if $\langle z, z \rangle = 0$. This explains the other two statements regarding the equality. ∎

Now, with Cauchy-Schwarz Inequality, we can check the triangle inequality

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \langle x, x \rangle + \langle y, y \rangle + 2 \operatorname{Re} \langle x, y \rangle \\ &\leq \|x\|^2 + \|y\|^2 + 2 |\langle x, y \rangle| \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| = \left(\|x\| + \|y\| \right)^2 \end{aligned}$$

indeed holds true for space equipped with only positive semi-definite Hermitian form.