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 some field $\mathbb{F}$, 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 $F \in \text{Hom}(V, W)$. 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

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

By extending a basis of the kernel of F to a basis of the whole V, one can prove the **Rank-Nullity Theorem**

$$\dim(\text{Ker}\, F) + \text{rank}(F) = \dim(V).$$

We use V^* to denote the vector space of linear map from V to $\mathbb{F}$, the **dual space** of V. Let $F : V \to W$ be a linear map. Its **dual map** $F^\dagger : W^* \to V^*$ is defined by

$$F^\dagger(\varphi) \triangleq \varphi \circ F$$

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

$$\varphi_i(v_j) \triangleq \delta^i_j$$

and there exists a natural isomorphism between V and its double dual $(V^*)^*$ by identifying $v \in (V^*)^*$ 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 $F^\dagger$ are transpose of each other. This together with the observation that $\text{rank}(F) + \dim(\ker F^*) = \dim(W)$ 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

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

If one identify the space of n-by-n square matrix over $\mathbb{F}$ 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 n-by-n square matrix A over $\mathbb{F}$ 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 $F \in \text{End}(V)$, 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.

1.2   Jordan-Chevalley Decomposition

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

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

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

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

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

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

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

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

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

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

Theorem 1.2.1. (Linear Independence of Generalized Eigenspaces) 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

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

Because v_1 is a generalized eigenvector, we may let k be the largest non-negative integer such that $(T - \lambda_1 I)^k v_1 \neq 0$, so that

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

Noting that the collection $\{(T - \lambda_1 I)^k, (T - \lambda_2 I)^{\dim V}, \ldots, (T - \lambda_n I)^{\dim V}\}$ of endomorphisms indeed 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

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

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

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

As we will later see, Decomposition 1.3 plays a central role in decomposition of V into generalized eigenspace of T if $\mathbb{F}$ is algebraically closed. See Theorem 1.2.2.

Theorem 1.2.2. (Decomposition into Generalized Eigenspaces) If $\mathbb{F}$ is algebraically closed, then clearly T has some eigenvalues. If we let $\{\lambda_1, \ldots, \lambda_m\}$ be the set of eigenvalues of T, we have

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

Proof. This 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. Now, suppose that such a decomposition always exists for complex vector spaces of strictly smaller dimension than V and their endomorphisms. By Equation 1.3, we may decompose V into

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

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 $T|_U \in \operatorname{End}(U)$. By inductive hypothesis, we may 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 $2 \le n \le m$. 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 Theorem 1.2.1 that $a_1 = a_3 = \cdots = a_m = 0$. We have shown $v \in G(\lambda_2, T|_U)$. ∎

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

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

(b) S and N commute.

If V is over an algebraically closed field $\overline{\mathbb{F}}$, then Theorem 1.2.2 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 F and

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

This together with the decomposition we obtained from Theorem 1.2.2:

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

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 Theorem 1.2.1, Theorem 1.2.2, nor any result that depends on them. The argument proceeds as follows: Because $\text{End}(V)$ is finite-dimensional, the minimal polynomial m exists. By the Fundamental Theorem of Algebra, we can factor m into linear terms:

$$m(x) = (x - \xi_1)^{r_1} \cdots (x - \xi_k)^{r_k}.$$

Consider the polynomials f_i defined by

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

Since $\{f_1, \ldots, f_k\}$ are coprime, Bézout's identity says that there exist polynomials $q_i \in \mathbb{C}[x]$ such that

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

8

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**:

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

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[1] . This now implies $S - S' = 0$, since no units are zero-divisor.

---

[1]Two commuting 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 Form

Theorem 1.3.1. (Jordan forms of Nilpotent Operators) Let V be a finite-dimensional vector space, and let $N \in \text{End}(V)$ be nilpotent. There exists some $v_1, \ldots, v_m \in V$ and nonnegative integer $n_1, \ldots, n_m$ such that

(a) $N^{n_1} v_1, \ldots, N v_1, v_1, \ldots, N^{n_m} v_m, \ldots, N v_m, v_m$ form a basis for V.

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

Proof. This is also 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.

Because N is nilpotent, we know $\dim(\text{Im } N) < \dim V$. Noting that $\text{Im } N$ is stable under N, we see $N|_{\text{Im } N} \in \text{End}(\text{Im } N)$. From inductive hypothesis, we have $v_1, \ldots, v_n \in \text{Im } N$ and nonnegative integer $k_1, \ldots, k_n$ such that

$$\{N^{k_1} v_1, \ldots, N v_1, v_1, \ldots, N^{k_n} v_n, \ldots, N v_n, v_n\} \text{ form a basis for Im } N.$$

and

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

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

$$\{N^{k_1+1} u_1, \ldots, N u_1, u_1, \ldots, N^{k_n} u_n, \ldots, N u_n, u_n\} \text{ is linearly independent,} \qquad (1.4)$$

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

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

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 $\{N^{k_1} v_1, \ldots, N^{k_n} v_n\}$ is linearly independent in the first place. We may now expand Set 1.4 to a basis

$$\{N^{k_1+1} u_1, \ldots, N u_1, u_1, \ldots, N^{k_n+1} u_n, \ldots, N u_n, u_n, w_1, \ldots, w_p\} \qquad (1.5)$$

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

$$\text{span}\{N^{k_1} u_1, \ldots, N u_1, u_1, \ldots, N^{k_n+1} u_n, \ldots, N u_n, u_n\}$$

so that Set 1.5 form a desired basis for V. ∎

10

Let V be some n-dimensional $\mathbb{F}$-vector space and $F \in \text{End}(V)$, and let $[F] \in M_n(\mathbb{F})$ be some matrix representation of F with respect to some basis. Let $\overline{\mathbb{F}}$ be the algebraically closure of $\mathbb{F}$, and consider the Jordan-Chevalley decomposition $[F] = S + N$, where $S, N \in M_n(\overline{\mathbb{F}})$. By Theorem 1.2.2, we have

$$(\overline{\mathbb{F}})^n = G(\lambda_1, F) \oplus \cdots \oplus G(\lambda_m, F)$$

Because both $[F]$ and S map $G(\lambda_j, F)$ into themselves, we know N also maps $G(\lambda_j, F)$ into themselves. Therefore, we may apply Theorem 1.3.1 to $N|_{G(\lambda_j, F)}$ and get a **Jordan basis**, here an invertible matrix $P \in M_n(\overline{\mathbb{F}})$ such that $P^{-1}[F]P$ is in the **Jordan form**. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of $[F]$, and let r_i be the dimension of the largest Jordan block with eigenvalue λ_i. It is clear that the **minimal polynomial**[2] of F, some polynomial m of smallest degree such that $m(F) = 0$, 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.

---

[2]It is worth mentioning that we didn't prove that the minimal polynomial of F is invariant under field extension. To prove such, one need to show that the minimal polynomial m of $[F]$ 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.

1.4   Spectral Theorem

Let V be a $\mathbb{F}$-vector space where $\mathbb{F}$ is a subfield of $\mathbb{C}$. 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

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

With an inner product equipped on V, we may discuss the **orthogonality** of vectors. Given a countable set $\{v_1, v_2, v_3, \dots\}$ 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 $\{e_1, e_2, e_3, \dots\}$ become orthogonal while

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

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**.

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, and its matrix representation is always the complex conjugation of that of F, regardless of the choices of bases. It is 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.[3]

Theorem 1.4.1. (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

---

[3]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.

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. ∎

1.5   SVD

Let V, W be two finite-dimensional inner product space over $\mathbb{F}$, where $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$. Given linear $F : V \to W$, clearly $F^\dagger F \in \text{End}(V)$ is self-adjoint. Therefore, by Spectral Theorem, there exists some orthonormal eigenbasis α of V with respect to $F^\dagger F$. Clearly, for every pair $v_1, v_2 \in \alpha$, we have $\langle F(v_1), F(v_2) \rangle = 0$ whether $F(v_1), F(v_2)$ are both nonzero or not. This implies that by deleting the zeros and negation of some elements, we may extend $\{F(v) \in W : v \in \alpha\}$ 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 lie in the diagonal. This process is called the **singular value decomposition of** F.

1.6 Appendix: Chinese Remainder Theorem and Bezout's Identity

Chapter 2

What you probably need to know

2.1 Matrix Exponential

Given two $n \times n$ square $\mathbb{F}$-matrices A, B, where $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$, by Cauchy-Schwarz inequality, we see that the Frobenius norm

$$
\begin{aligned}
\|AB\|^2 &= \sum_{i,j} \left(\sum_k |A_{i,k} B_{k,j}| \right)^2 \\
&\leq \sum_{i,j} \left(\sum_k |A_{i,k}|^2 \right) \left(\sum_k |B_{k,j}|^2 \right) \\
&= \sum_{i,j,k,l} |A_{i,k}|^2 |B_{j,l}|^2 = \sum_{i,k} |A_{i,k}|^2 \sum_{j,l} |B_{j,l}|^2 = \|A\|^2 \|B\|^2
\end{aligned}
$$

is submultiplicative. Therefore, when we define **matrix exponential** by

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

we see that 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. From now on, suppose $A \in M_n(\mathbb{C})$.

Let $A \in M_n(\mathbb{F})$. The first step is to compute and factorize the characteristic polynomial of A, and
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

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

Here the term e^S and e^N all can be easily computed. We have

$$S = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \implies e^S = \begin{bmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{bmatrix} \text{ and } e^N = \sum_{k=0}^{n} \frac{N^k}{k!}$$

2.2 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 $V_1^* \times V_2^*$ 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.

2.3 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 $\{v_{i_1} \wedge \cdots \wedge v_{i_n} \in \bigwedge^n V : 1 \leq i_1 < \cdots < i_n \leq k\}$ 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 $\{v_{i_1} \wedge \cdots \wedge v_{i_n} \in \bigwedge^n V : 1 \leq i_1 < \cdots < i_n \leq k\}$ 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

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

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

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

2.4 Pfaffian

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

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

Theorem 2.4.1. (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

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

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

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

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

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

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

Above equation need to be explained in the section before. $\blacksquare$

Chapter 3

NTU Math M.A. Program Entrance Exam

3.1 Year 113

> **Question 1**
>
> 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

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

Question 2

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, 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

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

24

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

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

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

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

∎

Question 3

Let V be the vector space of all n-by-n square matrices over $\mathbb{R}$, and let $f : V \to \mathbb{R}$ be linear. Suppose

$$f(AB) = f(BA) \text{ and } f(I) = n$$

for all $A, B \in V$. 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,

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

25

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

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

$\blacksquare$

Question 4

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

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

And

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

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^*. Observe

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

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

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

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

Question 5

Let V be some finite-dimensional vector space over some field $\mathbb{F}$ and let $T \in \text{End}(V)$. Let f, g be two relatively prime polynomials. Prove

$$\text{Ker}(f(T)g(T)) = \text{Ker} f(T) \oplus \text{Ker} g(T).$$

Proof. To show that the two kernels form a direct sum, assume for contradiction that there exists a nonzero vector $v \in V$ such that v belongs to both kernels. Since f is a polynomial satisfying $f(T)v = 0$, there must exist a polynomial p of minimal degree such that $p(T)v = 0$. By polynomial division, we can write

$$f = pq + r, \quad \text{where } q, r \text{ are polynomials and } \deg r < \deg p.$$

Since r has a strictly smaller degree than p, the minimality of p implies that $r = 0$. Thus, f is divisible by p. WLOG, we may assume that g is also divisible by p. This contradicts the assumption that f and g are relatively prime polynomials. We have shown the two kernels indeed form a direct sum. It is clear that

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

We now prove the opposite. Let $v \in \text{Ker}(f(T)g(T))$. Because f, g are relatively prime, by **Bezout's identity**, there exists some polynomials a, b such that

$$af + bg = 1$$

The proof then follows from noting

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

■

3.2 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

$$\begin{cases} \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} \text{ is linearly independent.} \\ \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\} \text{ is linearly independent.} \end{cases}$$

Use Gauss elimination to show

$$3\mathbf{v}_1 + 2\mathbf{v}_2 + \mathbf{v}_3 + \mathbf{v}_4 = 0$$

And conclude

$$W_1 \cap W_2 = \text{span}\{\mathbf{v}_3 + \mathbf{v}_4\}$$

∎

Let

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

Find an invertible matrix $Q \in M_3(\mathbb{C})$ such that

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

Find an invertible matrix $P \in M_3(\mathbb{C})$ such that PAP^{-1} is diagonal. *Note that my wording differs slightly with the original.*

Proof. Noting that characteristic polynomial is 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

$$P = \begin{bmatrix} 1 & 7 & 0 \\ 1 & -1 & 1 \\ 1 & -29 & 1 \end{bmatrix} \text{ and } PAP^{-1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

Note that one may find Q by solving a 9-by-9 linear equation. Here we give a smarter approach. By looking at QAQ^{-1}, we may reduce the problem into finding $v \in \mathbb{C}^3$ such that

$$A^3 v = 12Av + A^2 v \text{ and } \{v, Av, A^2 v\} \text{ is linear independent}$$

So that

$$Q \triangleq \begin{bmatrix} v & Av & A^2 v \end{bmatrix} \text{ suffices.}$$

Because the characteristic polynomial of A is $t(t-4)(t+3)$, by Cayley-Hamilton Theorem, we know all $v \in \mathbb{C}^3$ satisfy the first condition. To satisfy the linear independence, we write

$$v = c_1 e_1 + c_2 e_2 + c_3 e_3$$

where e_1, e_2, e_3 are the eigenvectors

$$e_1 \triangleq \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ and } e_2 \triangleq \begin{bmatrix} 7 \\ -1 \\ -29 \end{bmatrix} \text{ and } e_3 \triangleq \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

So that we may just looks for $c_1, c_2, c_3 \in \mathbb{C}$ that makes

$$\begin{bmatrix} c_1 & c_2 & c_3 \\ 0 & 4c_2 & -3c_3 \\ 0 & 16c_3 & 9c_3 \end{bmatrix} \text{ invertible.}$$

By computing the determinant, we see that an obvious choice is $c_1 = c_2 = c_3 = 1$. That is,

$$v = \begin{bmatrix} 8 \\ 1 \\ -27 \end{bmatrix} \text{ and } Q = \begin{bmatrix} 8 & 28 & 112 \\ 1 & -7 & -7 \\ -27 & -119 & -455 \end{bmatrix}$$

■

Question 8

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 9

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 10

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

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

3.3 Year 110

Question 11

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. Because the underlying field is $\mathbb{C}$, we may express d in its **Jordan form**. Because $d^2 = 0$, we know its Jordan blocks all have eigenvalue 0 and have size no greater than 2. Knowing that the matrix representation of δ with respect to the same Jordan basis is just the conjugate transpose of d, we may easily prove all four propositions with computation, except

$$\text{Ker}\, \Delta \perp \text{Im}\, d \quad \text{and} \quad \text{Im}\, d \perp \text{Im}\, \delta \quad \text{and} \quad \text{Im}\, \delta \perp \text{Ker}\, \Delta$$

Direct computation with the hint $\delta = d^\dagger$ and $d^2 = 0$ shows $\text{Im}\, d \perp \text{Im}\, \delta$. Direct computation with hint $\text{Ker}\, \Delta = \text{Ker}\, d \cap \text{Ker}\, \delta$ shows the other two orthogonal relationships. ∎

Question 12

Let $V = \mathbb{R}^n$ be the space of column vectors, and M a positive definite symmetric $n \times n$ real matrix. Suppose $A \in M_n(\mathbb{R})$ satisfies

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

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

$$P^t M P = I \text{ and } P^{-1} A P \text{ is diagonal}$$

Proof. Because M is symmetric and positive definite, by **real spectral theorem**, there

exists some symmetric positive definite **square root** $M^{\frac{1}{2}}$ of M. Define

$$B \triangleq M^{\frac{1}{2}} A M^{\frac{-1}{2}}$$

The fact that $M^{\frac{1}{2}}$ is symmetric and $MAM^{-1} = A^t$ allow us to do computation to directly show that B is symmetric. Therefore, by real spectral theorem, there exists some orthonormal Q such that QBQ^{-1} is diagonal. It is clear that $P \triangleq M^{\frac{-1}{2}} Q^{-1}$ makes $P^{-1}AP$ diagonal, and

$$P^t M P = (M^{\frac{-1}{2}} Q^{-1})^t M M^{\frac{-1}{2}} Q^{-1} = Q M^{\frac{-1}{2}} M M^{\frac{-1}{2}} Q^{-1} = I$$

■

Question 13

(a) Prove that invertible $M \in \mathrm{End}(\mathbb{C}^n)$ 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 $N^{n-1} \neq 0$. 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 $P \in \mathrm{End}(\mathbb{C}^n)$, $I + P$ 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 $B^{2n-2} \neq 0$. [1] ∎

Question 14

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 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 15

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

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

---

[1] 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.

that

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

Therefore,

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

∎

3.4   Year 109

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

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

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

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

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 $\text{rank}_{\mathbb{R}}(A) = \text{rank}_{\mathbb{C}}(A)^2$. Clearly, $\text{rank}_{\mathbb{R}}(A) \geq \text{rank}_{\mathbb{C}}(A)$. 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 $\text{rank}_{\mathbb{R}}(A) \leq \text{rank}_{\mathbb{C}}(A)$.

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

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

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

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

The proof then follows from checking

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

■

Question 18

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

---

[2]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 $\text{rank}_{\mathbb{F}}(A) = \text{rank}_{\mathbb{K}}(A)$

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. ∎

Question 19

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

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

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

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

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 20

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

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

And

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

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^*. Observe

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

to conclude that $\xi_i \notin \operatorname{Ker} T^*$ 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^*. In other words,

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

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

3.5 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 $\max_{i_j} n_{i_j}$. ∎

Question 22

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 23

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

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

And

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

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

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

$\blacksquare$

Question 24

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 $\tilde{v}_1, \ldots, \tilde{v}_n$ by

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

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

3.6 Year 107

Question 25

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

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

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 26

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

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

44

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

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

■

Question 27

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 $\text{pf}(A)$ of A by

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

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

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

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

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

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

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

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

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 2.4.1 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 $\det(A) = \det(OAO^t) = m_1^2 \cdots m_n^2$ $\blacksquare$

Question 28

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

$P \in GL_n(\mathbb{C})$ such that PAP^{-1} and PAP^{-1} are both upper triangular)

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

$$\mathbb{C}^n = G(\lambda_1, A) \oplus \cdots \oplus G(\lambda_n, A), \quad \text{See Theorem 1.2.2}$$

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 29

Show that

$$\begin{vmatrix} X_0 & X_1 & X_2 & \cdots & X_{n-1} \\ X_{n-1} & X_0 & X_1 & \cdots & X_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X_1 & X_2 & X_3 & \cdots & X_0 \end{vmatrix} = \prod_{j=0}^{n-1} \left(\sum_{k=0}^{n-1} \xi^{jk} X_k \right) \tag{3.3}$$

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 3.3 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. ∎

3.7 Year 106

Question 30

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 $\mathrm{rank}(A) = k$.

(b) Suppose $A^n = 0$ and $A^{n-1} \neq 0$. Find $\max_{v \in \mathbb{F}^n} \dim(\mathrm{span}(v, Av, \ldots, A^{n-1}v))$.

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

$$\begin{pmatrix} A_{1,1} & \cdots & A_{1,k} \\ \vdots & \ddots & \vdots \\ A_{k,1} & \cdots & A_{k,k} \end{pmatrix} \tag{3.4}$$

Because the subamtrix 3.4 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 $\mathrm{rank}(A) \geq k$. It remains to show $\mathrm{rank}(A) \leq k$. Assume $\mathrm{rank}(A) > k$ 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

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

∎

(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 $T \in \operatorname{End}(V)$ 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 $\operatorname{ch}(t) = (t-2)^2(t-1)$. 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 32

(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.[3]

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:

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

$\blacksquare$

Question 33

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[4] 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.

---
[3]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.

[4]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. ∎

3.8   Year 105

Prove the **Cayley-Hamilton Theorem**: If V is a finite dimensional vector space and $T \in \mathrm{End}(V)$ 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 35

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 $\mathrm{rank}(A) = r$. 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,[5] 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

$$A^\vee(\psi_i) = \begin{cases} \varphi_i & \text{if } i \leq r \\ 0 & \text{if } i > r \end{cases} \tag{3.5}$$

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

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

---

[5]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 36**
>
> 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

$$\mathbb{F}^n = \bigoplus_{\lambda_i \in \Lambda_i} (E(\lambda_1, A_1) \cap \cdots \cap E(\lambda_m, A_m)) \tag{3.6}$$

We shall show Equation 3.6 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

$$(A_1 - \lambda I)(A_2 v) = A_2(A_1 - \lambda I)v = 0 \tag{3.7}$$

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

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

Fix $\lambda_i \in \Lambda_i$. Again for the same reason Equation 3.7, 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 3.6. ■

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 38

Let V be a finite-dimensional complex vector space, and let $T \in \mathrm{End}(V)$ 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 1.2.2 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.

∎

Chapter 4

NTU MATH M.A. Program Entrance Exam Round 2

4.1   NTU Round 2 Year 106

> **Question 39**
>
> 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.

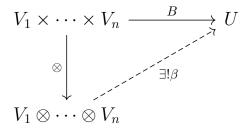*Proof.* ∎

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

$$
\begin{array}{ccc}
V_1 \times \cdots \times V_n & \xrightarrow{\ \ B\ \ } & U \\
\Big\downarrow{\scriptstyle \otimes} & \nearrow{\scriptstyle \exists! \beta} & \\
V_1 \otimes \cdots \otimes V_n & &
\end{array}
$$

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,

$$
\begin{aligned}
(c_1 H_{z_1} + H_{z_2})(x \otimes y) &= c_1 x \otimes y \otimes z_1 + x \otimes y \otimes z_2 \\
&= x \otimes y \otimes (c_1 z_1 + z_2) = h_{c_1 z_1 + z_2}(x, y)
\end{aligned}
$$

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

$$H((x \otimes y) \otimes z) = h(x \otimes y, z) = H_z(x \otimes y) = x \otimes y \otimes z$$
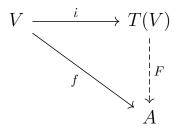
Let $\otimes : X \times Y \times Z \to X \otimes Y \otimes Z$ denotes the tensor product, we now have

$$\otimes = H \circ F \circ \otimes$$

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

$$f(k) \triangleq \begin{cases} w_1 \otimes \cdots \otimes w_n & \text{if } k = n \\ 0 & \text{if otherwise} \end{cases} \quad \text{and } g(k) \triangleq \begin{cases} w_{n+1} \otimes \cdots \otimes w_{n+l} & \text{if } k = l \\ 0 & \text{if otherwise} \end{cases}$$

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

$$T^k(V) = \text{span}\{v_{I(1)} \otimes \cdots \otimes v_{I(k)} : I \in S\}$$

If we define $I^k(V) \triangleq I(V) \cap T^k(V)$, one then have

$$I^k(V) = \text{span}\{v_{I(1)} \otimes \cdots \otimes v_{I(k)} : I(j) = I(j+1) \text{ for some } j\} \qquad (5.1)$$

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

$$x \sim y \iff x - y \in I(V)$$

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

$$[x] + [y] \triangleq [x + y] \text{ and } [x] \wedge [y] \triangleq [xy] \text{ and } c[x] \triangleq [cx]$$

61

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
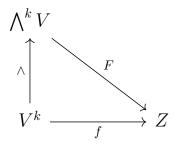
$$\text{span}\{v_{I(1)} \wedge \cdots \wedge v_{I(k)} : I \in S\}$$

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 $\{v_{I(1)} \wedge \cdots \wedge v_{I(k)} : I \in S\}$ which spans $\wedge^k(V)$. $\blacksquare$

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

$$\alpha \wedge \beta = \sum_{I \in S_k, J \in S_l} a_I b_J (v_{I(1)} \wedge \cdots \wedge v_{I(k)} \wedge v_{J(1)} \wedge \cdots \wedge v_{J(l)})$$

$$= \sum_{I \in S_k, J \in S_l} (-1) a_I b_J (v_{I(1)} \wedge \cdots \wedge v_{J(1)} \wedge v_{I(k)} \cdots \wedge \cdots \wedge v_{J(l)})$$

$$= \sum_{I \in S_k, J \in S_l} (-1)^k a_I b_J (v_{J(1)} \wedge v_{I(1)} \wedge \cdots \wedge v_{I(k)} \wedge v_{J(2)} \wedge \cdots \wedge v_{J(l)})$$

$$= \sum_{I \in S_k, J \in S_l} (-1)^{kl} a_I b_J (v_{J(1)} \wedge \cdots \wedge v_{J(l)} \wedge v_{I(1)} \wedge \cdots \wedge v_{I(k)}) = (-1)^{kl} \beta \wedge \alpha$$

$\blacksquare$

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

$$\wedge^k(A)(v_1 \wedge \cdots \wedge v_n) = Av_1 \wedge \cdots \wedge Av_n$$

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

$$\wedge^n A(v_1 \wedge \cdots \wedge v_n) = \left(\sum_j A_{j,1} v_j \right) \wedge \cdots \wedge \left(\sum_j A_{j,n} v_j \right)$$

$$= \sum_{\sigma \in S_n} A_{\sigma(1),1} \cdots A_{\sigma(n),n} v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(n)}$$

$$= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{\sigma(1),1} \cdots A_{\sigma(n),n} v_1 \wedge \cdots \wedge v_n$$

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

$$T \text{ is bounded} \quad \overset{\triangle}{\Longleftrightarrow} \quad \exists M \in \mathbb{R}, \forall x \in \mathcal{X}, \|Tx\| \le M\|x\|$$

The proof of equivalency is simple. For $(\longrightarrow)$, observe

$$\|Tx\| = \|x\| \cdot \|T\frac{x}{\|x\|}\| \le \Big(\sup\{\|Ty\| : \|y\| = 1\} \Big)\|x\|$$

For $(\longleftarrow)$, observe

$$\|Tx - Ty\| = \|T(x - y)\| \le M\|x - y\|$$

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

$$T \text{ is a bounded operator} \quad \Longleftrightarrow \quad T \text{ is continuous on } \mathcal{X}$$

Proof. If T is bounded, we see that T is Lipschitz.

$$\|Tx - Ty\| \le M\|x - y\|$$

Now, suppose T is linear and continuous at 0. Let ϵ satisfy

$$\sup_{\|y\| \le \epsilon} \|Ty\| \le 1$$

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

$$\text{finding } m, M \in \mathbb{R}^+ \text{ such that } m\|x\|_\infty \leq \|x\| \leq M\|x\|_\infty$$

66

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

$$m\|x\|_\infty = \|x\|_\infty (\min_{y \in S} \|y\|) \leq \|x\|_\infty \cdot \left\| \frac{x}{\|x\|_\infty} \right\| = \|x\| \text{ (done)} \quad \text{(done)}$$

$\blacksquare$

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

$$\text{to find } \delta \text{ such that } \forall h \in \mathcal{X} : \|h\| \leq \delta, \|T(x + h) - Tx\| \leq \epsilon$$

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

$$\|x\|_1 \leq M\|x\| \quad (x \in V)$$

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

$$\|T\|_{\text{op}} \triangleq \inf\{M \in \mathbb{R}^+ : \forall x \in \mathcal{X}, \|Tx\| \le M\|x\|\} \tag{5.3}$$

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

$$\|T\|_{\text{op}} = \sup_{x\in\mathcal{X}, x\ne 0} \frac{\|Tx\|}{\|x\|}$$

Proof. Define $J \triangleq \{M \in \mathbb{R}^+ : \forall x \in \mathcal{X}, \|Tx\| \leq M\|x\|\}$ and observe

$$J = \{M \in \mathbb{R}^+ : M \geq \frac{\|Tx\|}{\|x\|}, \forall x \neq 0 \in \mathcal{X}\}$$

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

$$\|T\|_{\mathrm{op}} = \sup_{x \in \mathcal{X}, x \neq 0} \frac{\|Tx\|}{\|x\|} \tag{5.4}$$

$$= \sup_{x \in \mathcal{X}, \|x\|=1} \|Tx\| \tag{5.5}$$

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 $\|(T_1 + T_2)x\| \leq \|T_1x\| + \|T_2x\|$.

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

$$\langle y, y \rangle = \langle z + \frac{\langle y, x \rangle}{\|x\|^2} x, z + \frac{\langle y, x \rangle}{\|x\|^2} x \rangle$$
$$= \langle z, z \rangle + \left| \frac{\langle y, x \rangle}{\langle x, x \rangle} \right|^2 \langle x, x \rangle$$
$$= \langle z, z \rangle + \frac{|\langle x, y \rangle|^2}{\langle x, x \rangle}$$

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

$$\|x + y\|^2 = \langle x + y, x + y \rangle$$
$$= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle$$
$$= \langle x, x \rangle + \langle y, y \rangle + 2 \operatorname{Re} \langle x, y \rangle$$
$$\leq \|x\|^2 + \|y\|^2 + 2 |\langle x, y \rangle|$$
$$\leq \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| = \left(\|x\| + \|y\| \right)^2$$

indeed holds true for space equipped with only positive semi-definite Hermitian form.