THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics
MATH2040A (First Term, 2022-23)
Linear Algebra II
Course Review Notes

August 29, 2022

Contents

Topic 1: Vector space

Let F be a field. In this course we always assume either $F = \mathbb{R}$ or $F = \mathbb{C}$, but keep in mind that F can be other fields, for instance, $\mathbb{Q}$ or a finite field $\mathbb{F}_q$.

Definition 1.1. *A **vector space over a field** F is a set V equipped with two operations:*

$$\text{addition: } V \times V \to V,\ (\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} + \mathbf{y},$$
$$\text{scalar multiplication: } F \times V \to V,\ (a, \mathbf{x}) \mapsto a\mathbf{x}$$

satisfying the following axioms:

VS 1 *(Commutativity of addition)* $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ *for any* $\mathbf{x}, \mathbf{y} \in V$.

VS 2 *(Associativity of addition)* $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ *for any* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$.

VS 3 *(Existence of zero vector) There exists a vector* $\mathbf{0} \in V$ *satisfying* $\mathbf{x} + \mathbf{0} = \mathbf{x}$ *for any* $\mathbf{x} \in V$.

VS 4 *(Existence of additive inverse) For any* $\mathbf{x} \in V$*, there exists* $-\mathbf{x} \in V$ *such that* $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.

VS 5 $1\mathbf{x} = \mathbf{x}$ *for any* $\mathbf{x} \in V$.

VS 6 $(ab)\mathbf{x} = a(b\mathbf{x})$ *for any* $a, b \in F$ *and* $\mathbf{x} \in V$.

VS 7 *(Distributive law I)* $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$ *for any* $a \in F$ *and* $\mathbf{x}, \mathbf{y} \in V$.

VS 8 *(Distributive law II)* $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$ *for any* $a, b \in F$ *and* $\mathbf{x} \in V$.

*Elements of F are called **scalars** and elements of V are called **vectors**.*

Remark 1.2. *(VS 1) - (VS 4) say that $(V, +)$ is an abelian group.*

Here are some examples of vector spaces:

- Basic examples are given by F^n (as space of row/column vectors) and $M_{m\times n}(F)$, which are vector spaces over F under the usual addition and scalar multiplication of matrices.
- The set $P(F)$ of polynomials with coefficients in a field F is a vector space over F under the usual addition and scalar multiplication of polynomials.
- The set $F^\infty := \{(x_1, x_2, \dots) : x_j \in F \text{ for } j = 1, 2, \dots\}$ of sequences of elements in F is a vector space over F under componentwise addition and scalar multiplication.
- Let S be a nonempty set. Then the set $\mathcal{F}(S, F)$ of functions from S to F is a vector space over F under the usual addition and scalar multiplication of functions.
- The set of complex numbers $\mathbb{C} = \{a + b\mathbf{i} : a, b \in \mathbb{R}\}$ is a vector space over $\mathbb{R}$, and at the same time a vector space over $\mathbb{C}$.

Proposition 1.3. *Let V be a vector space over a field F. Then*

1. *The zero vector* $\mathbf{0} \in V$ *is unique.*

2. *The additive inverse* $-\mathbf{x}$ *of any vector* $\mathbf{x} \in V$ *is unique.*

3. *(Cancellation law) If* $\mathbf{x} + \mathbf{z} = \mathbf{y} + \mathbf{z}$*, then* $\mathbf{x} = \mathbf{y}$*.*

4. $0\mathbf{x} = \mathbf{0}$ *for any* $\mathbf{x} \in V$*.*

5. $(-a)\mathbf{x} = -(a\mathbf{x}) = a(-\mathbf{x})$ *for any* $a \in F$ *and any* $\mathbf{x} \in V$*.*

6. $a\mathbf{0} = \mathbf{0}$ *for any* $a \in F$*.*

Topic 2: Subspace

Definition 2.1. *A subset* W *of a vector space* V *over a field* F *is called a* ***subspace*** *of* V *if* W *is a vector space over* F *under the addition and scalar multiplication inherited from* V*.*

Proposition 2.2. *Let* V *be a vector space over* F*. A subset* $W \subset V$ *is a subspace if and only if the following conditions hold for the operations defined on* V*:*

1. $\mathbf{0} \in W$ *(i.e.* W *contains the zero vector of* V*).*

2. $\mathbf{x} + \mathbf{y} \in W$ *for any* $\mathbf{x}, \mathbf{y} \in W$ *(i.e.* W *is closed under addition).*

3. $a\mathbf{x} \in W$ *for any* $a \in F$ *and any* $\mathbf{x} \in W$ *(i.e.* W *is closed under scalar multiplication).*

Examples (and non-examples) of subspaces:

- For any vector space V, the subsets $\{0\}$ and V are subspaces of V; $\{\mathbf{0}\}$ is called the **zero vector space** or **trivial vector space**. A subspace $\{0\} \subsetneqq W \subset V$ is called **nontrivial**, and a subspace $W \subsetneqq V$ is called **proper**.

- In F^n, any hyperplane is defined by a linear equation

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

  where $a_1, a_2, \ldots, a_n, b \in F$. A hyperplane is a subspace if and only if it passes through the origin $\mathbf{0} \in F^n$, i.e. when $b = 0$.

  More generally, the solution set of a system of linear equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

  is a subset in F^n, which is a subspace if and only if the system is *homogeneous*; in this case, the subspace is nothing but the *null space*

$$N(A) := \{\mathbf{x} \in F^n : A\mathbf{x} = \mathbf{0} \text{ in } F^m\}$$

  of the coefficient matrix $A = (a_{ij}) \in M_{m\times n}(F)$.

- In $V = M_{n\times n}(F)$, we have the following subspaces:

$$W_1 := \{A \in M_{n\times n}(F) : A \text{ is diagonal}\},$$
$$W_2 := \{A \in M_{n\times n}(F) : \text{tr}(A) = 0\},$$

  where the **trace** of matrix $A \in M_{n\times n}(F)$ is defined by $\text{tr}(A) := \sum_{i=1}^{n} A_{ii}$.

- Given a positive integer n, let $P_n(F)$ be the set of all polynomials with coefficients in F and of degree less than or equal to n. Then $P_n(F)$ is a vector space over F.

  Remark 2.3. *Note that the set of polynomials with coefficients in F and a* fixed *degree n is* not *a vector space over F.*

- Given an open interval $I = (a, b) \subset \mathbb{R}$ and $n \in \mathbb{N} \cup \{\infty\}$, let $C^n(I)$ be the set of all functions $f : I \to \mathbb{R}$ which have a continuous nth derivative (for $n = 0$, $C^0(I)$ is simply the set of all continuous real-valued functions on I). Then

$$C^\infty(I) \subset \cdots \subset \cdots \subset C^n(I) \subset \cdots \subset C^1(I) \subset C^0(I) \subset \mathcal{F}(I, \mathbb{R})$$

  is an infinite chain of subspaces.

Proposition 2.4. *Any intersection of subspaces of a vector space V is also a subspace of V.*

Topic 3: Span and linear (in)dependence

Definition 3.1. *Let V be a vector space over F and $S \subset V$ a nonempty subset.*

- *We say a vector $\mathbf{x} \in V$ is a* ***linear combination*** *of vectors of S if there exist $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in S$ and $a_1, a_2, \ldots, a_n \in F$ such that*

$$\mathbf{x} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n.$$

  In this case, we also say $\mathbf{x}$ is a linear combination of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ and $a_1, a_2, \ldots, a_n$ the ***coefficients*** *of the linear combination.*

- *The* ***span*** *of S, denoted as span(S), is the set of all possible linear combinations of vectors of S, i.e.*

$$\textit{span}(S) := \{a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n : \textit{each } a_j \in F, \textit{each } \mathbf{x}_j \in S, n = 1, 2, \cdots\}.$$

  By convention, we set span$(\emptyset) = \{\mathbf{0}\}$.

Proposition 3.2. *Let $S \subset V$ be a subset of a vector space V over F. Then span(S) is the smallest subspace of V containing S.*

Definition 3.3. *We say a subset S of a vector space V* ***spans*** *(or* ***generates****) V if span$(S) = V$. In this case, we also say S is a* ***spanning set*** *(or* ***generating set****) for V.*

Definition 3.4. *Let V be a vector space over F. A subset $S \subset V$ is said to be **linearly dependent** if there exist distinct $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in S$ and scalars $a_1, a_2, \ldots, a_n \in F$, not all zero, such that*

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n = \mathbf{0};$$

*Otherwise, the subset $S \subset V$ is said to be **linearly independent**.*

Note: $\emptyset$ is linearly independent; if $0 \in S$ then S is linearly dependent, namely, a nonempty linearly independent subset cannot contain 0; $\{\mathbf{x}\}$ is linearly independent iff $\mathbf{x} \neq 0$; any set containing a linearly dependent subset is also linearly dependent; any subset of a linearly independent set is also linearly independent.

Proposition 3.5. *Given a nonempty subset S of a vector space V, then the following are equivalent:*

1. *S is linearly independent.*
2. *Each $\mathbf{x} \in span(S)$ can be expressed in a unique way as a linear combination of distinct vectors of S.*
3. *The only representation of $\mathbf{0}$ as a linear combination of distinct vectors of S is trivial, i.e. if*

   $$\mathbf{0} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n$$

   for some distinct vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in S$ and scalars $a_1, a_2, \ldots, a_n \in F$, then we must have $a_1 = a_2 = \cdots = a_n = 0$.

Proposition 3.6. *Let S be a linearly dependent subset of a vector space V. Then there is $\mathbf{x} \in S$ such that $span(S) = span(S \setminus \{\mathbf{x}\})$.*

Proposition 3.7. *Let S be a linearly independent subset of a vector space V, and assume $\mathbf{x} \in V \setminus S$. Then $\mathbf{x} \in span(S)$ if and only if the set $S \cup \{\mathbf{x}\}$ is linearly dependent.*

Topic 4: Basis and dimension

Definition 4.1. *A **basis** for a vector space V is a subset $\beta \subset V$ which is linearly independent and spans V (i.e. $V = span(\beta)$).*

Proposition 4.2. *Let V be a vector space and $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \subset V$ be a finite subset. Then β is a basis for V if and only if for any $\mathbf{x} \in V$, there exist unique $a_1, a_2, \ldots, a_n \in F$ such that $\mathbf{x} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$.*

Theorem 4.3. *Suppose S is a finite spanning set for a vector space V. Then there exists a subset $\beta \subset S$ which forms a basis for V, namely, any finite spanning set can be reduced to a basis.*

Theorem 4.4 (Replacement Lemma). *Let V be a vector space. Let $G \subset V$ be a spanning set for V consisting of n vectors, and $L \subset V$ be a linearly independent subset consisting of m vectors. Then, $m \leq n$ and there exists a subset $H \subset G$ consisting of exactly $n - m$ vectors such that $L \cup H$ spans V.*

Corollary 4.5. *(a) Any linearly independent subset of a vector space having a finite spanning set must be also finite.*

(b) Let V be a vector space having a finite basis. Then every basis of V contains the same number of vectors.

Definition 4.6. *(a) A vector space V is called* ***finite-dimensional*** *if it has a finite spanning set; otherwise, it is called* ***infinite-dimensional****.*

(b) For a finite-dimensional vector space V, the ***dimension*** *of V, denoted as dim(V), is the number of vectors in a finite basis for V (where the existence of a finite basis is assured by Theorem 4.3, and the number is well defined in a unique way by Corollary 4.5). For an infinite-dimensional vector space V, we write dim$(V) = \infty$.*

Corollary 4.7. *Let V be an n-dimensional vector space. Then the following statements hold:*

1. *Any finite spanning set for V has at least n vectors, and a spanning set with exactly n vectors is a basis for V.*
2. *Any linearly independent subset of V consisting of n vectors is a basis for V.*
3. *Every linearly independent subset of V can be extended to a basis for V.*

Theorem 4.8. *Let V be a finite-dimensional vector space and W be a subspace of V. Then W is finite-dimensional with dim$(W) \leq$ dim(V). Moreover, if dim$(W) =$ dim(V), then $W = V$.*

Corollary 4.9. *If W is a subspace of a finite-dimensional vector space V, then any basis of W can be extended to a basis of V.*

Remark 4.10. *This corollary implies that for any subspace $W \subset V$, there exists another subspace $Q \subset V$ such that $V = W \oplus Q$; see the definition of the direct sum of two subspaces in Exercises of Sect 1.3 (page 22 of the textbook).*

Examples:

- The *standard basis* of F^n consists of the vectors

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

  So F^n has dimension n over F.

- The *standard basis* of $M_{m\times n}(F)$ is given by

$$\{E_{ij} : 1 \leq i \leq m, 1 \leq j \leq n\},$$

  where $E_{ij} \in M_{m\times n}(F)$ is the matrix whose (i,j)-th entry is 1 and all other entries are 0. So $M_{m\times n}(F)$ has dimension $n \cdot m$ over F.

- In $M_{n\times n}(F)$ (whose dimension is n^2), we have the following subspaces:

- $W_1 := \{A \in M_{n\times n}(F) : A \text{ is diagonal}\}$ is a subspace of $M_{n\times n}(F)$ of dimension n.
- $W_2 := \{A \in M_{n\times n}(F) : \text{tr}(A) = 0\}$ is a subspace of $M_{n\times n}(F)$ of dimension $n^2 - 1$.
- $W_3 := \{A \in M_{n\times n}(F) : A^t = A\}$ is a subspace of $M_{n\times n}(F)$ of dimension $\frac{n(n+1)}{2}$.

- The **standard basis** of $P_n(F)$ is given by the subset $\{1, x, x^2, \ldots, x^n\}$, so $P_n(F)$ has dimension $n+1$ over F. Another basis for $P_n(F)$ is given by $\{1, x-a, (x-a)^2, \ldots, (x-a)^n\}$ where $a \in F$ is any scalar.

- The set of complex numbers $\mathbb{C}$ has dimension 2 as a vector space over $\mathbb{R}$ (a basis is given by $\{1, \mathbf{i}\}$), but dimension 1 as a vector space over $\mathbb{C}$ (any nonzero complex number z gives a basis $\{z\}$).

- The set of real numbers $\mathbb{R}$ has dimension 1 as a vector space over $\mathbb{R}$ (any nonzero complex number a gives a basis $\{a\}$), but is infinite dimensional as a vector space over $\mathbb{Q}$ (because $\mathbb{R}$ is uncountable).

Topic 5: Linear transformation

Definition 5.1. *(a) Let V and W be vector spaces over the same field F. A function $T : V \to W$ is linear if $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ and $T(a\mathbf{x}) = aT(\mathbf{x})$ for any $\mathbf{x}, \mathbf{y} \in V$ and $a \in F$.*

*(b) A function $T : V \to W$ which is linear is called a **linear transformation from** V **to** W. A linear transformation T from a vector space V into itself is called a **linear operator** on V.*

(c) We always use $\mathcal{L}(V, W)$ to denote the set of all linear transformations from V to W. In case $W = V$, we write $\mathcal{L}(V) := \mathcal{L}(V, V)$ to be the set of all linear operators on V.

Examples of linear transformations:

- Let $A \in M_{m\times n}(F)$. Regarding F^n and F^m as spaces of column vectors, then the map

$$L_A : F^n \to F^m,\ \mathbf{x} \mapsto A\mathbf{x}$$

  is a linear transformation, called the **left multiplication by** A.

- The transpose map $T : M_{m\times n}(F) \to M_{n\times m}(F)$, $A \mapsto A^t$ is a linear transformation.

- Differentiation and integration are linear operators, so they define linear maps such as:

$$T : P_n(\mathbb{R}) \to P_{n-1}(\mathbb{R}),\ f(x) \mapsto f'(x)$$

  and

$$T : P_{n-1}(\mathbb{R}) \to P_n(\mathbb{R}),\ f(x) \mapsto \int_0^x f(t)dt.$$

- For any vector space V and W, we have the **zero transformation** $T_0 : V \to W$ defined by $T_0(\mathbf{x}) = \mathbf{0}$ for any $\mathbf{x} \in V$, and the **identity transformation** $I_V : V \to V$ defined by $I_V(\mathbf{x}) = \mathbf{x}$ for any $\mathbf{x} \in V$.

Proposition 5.2. *Let $T : V \to W$ be a linear transformation. Then*

1. $T(\mathbf{0}_V) = \mathbf{0}_W$.
2. $T(\sum_{i=1}^n a_i \mathbf{x}_i) = \sum_{i=1}^n a_i T(\mathbf{x}_i)$ *for any* $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in V$ *and any* $a_1, a_2, \ldots, a_n \in F$ *(i.e. T preserves linear combinations).*

Theorem 5.3. *Let V and W be vector spaces, and let $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for V. Then given any $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n \in W$, there exists a unique linear transformation $T : V \to W$ such that $T(\mathbf{v}_i) = \mathbf{w}_i$ for $i = 1, \ldots, n$.*

Remark:

(a) In the above theorem T is constructed as follows: For $v = \sum_{i=1}^n a_i \mathbf{v}_i$ (where a_i are uniquely defined, since β is a basis for V), $T(v) = \sum_{i=1}^n a_i \mathbf{w}_i$.

(b) Let V be a vector space with a finite basis $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Then a linear transformation from V to another vector space W is uniquely determined by its values on β, namely, if $U, T : V \to W$ are linear and $U(\mathbf{v}_i) = T(\mathbf{v}_i)$ for $i = 1, \ldots, n$, then $U = T$.

Topic 6: Null space, range, and dimension theorem

Definition 6.1. *Let V and W be vector spaces and $T : V \to W$ be a linear transformation. The* ***null space*** *(or* ***kernel****) of T is defined as*

$$N(T) := \{\mathbf{x} \in V : T(\mathbf{x}) = \mathbf{0}\}.$$

The ***range*** *(or* ***image****) of T is defined as*

$$R(T) := \{T(\mathbf{x}) \in W : \mathbf{x} \in V\}.$$

Proposition 6.2. *Let $T : V \to W$ be a linear transformation. Then $N(T)$ is a subspace of V, and $R(T)$ is a subspaces of W.*

Proposition 6.3. *Let $T : V \to W$ be a linear transformation. Then, T maps a spanning set of V to a spanning set of $R(T)$, namely, if $V = span(S)$ for a subset S of V then $R(T) = span(T(S))$. Particularly, if β is a basis for V then $R(T) = span(T(\beta))$. Moreover, if β is a basis for V and $N(T) = \{\mathbf{0}\}$ then $T(\beta)$ is a basis for $R(T)$.*

Note: The proposition above implies that if V is finite-dimensional then the range space of a linear transformation $T : V \to W$ is a finite-dimensional subspace of W.

Definition 6.4. *Let $T : V \to W$ be a linear transformation such that $N(T)$ and $R(T)$ are finite-dimensional. Then we define the* ***nullity*** *of T, denoted nullity(T), and the* ***rank*** *of T, denoted rank(T), to be the dimensions of $N(T)$ and $R(T)$ respectively.*

Theorem 6.5 (Dimension Theorem)**.** *Let V and W be vector spaces such that V is finite-dimensional. Then for any linear transformation $T : V \to W$, we have*

$$\textit{nullity}(T) + \textit{rank}(T) = \textit{dim}(V).$$

Note: The proof of theorem above tells that a basis $\{v_1, v_2, \cdots, v_k\}$ for $N(T)$ can be extended to a basis $\{v_1, \cdots, v_k, v_{k+1}, \cdots, v_n\}$ for V such that $\{T(v_{k+1}), \cdots, T(v_n)\}$ is a basis for $R(T)$.

Proposition 6.6. *A linear transformation $T : V \to W$ is one-to-one if and only if $N(T) = \{\mathbf{0}\}$.*

Remark 6.7. *By definition, a linear transformation $T : V \to W$ is onto if and only if $R(T) = W$.*

Corollary 6.8. *Let V and W be vector spaces of equal finite dimensions, and let $T : V \to W$ be a linear transformation. Then the following things are equivalent:*

(a) *T is one-to-one.*

(b) *T is onto.*

(c) *rank$(T) =$ dim(V).*

(d) *nullity$(T) = 0$.*

Remark 6.9. *This corollary is not true in the infinite-dimensional case.*

Topic 7: Matrix representation of a linear transformation

Definition 7.1. *Let V be a finite-dimensional vector space and $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an ordered basis for V (meaning a basis with a specified order). Then for $\mathbf{x} \in V$, there exist unique $a_1, a_2, \ldots, a_n \in F$ such that $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{v}_i$. The **coordinate vector of** $\mathbf{x}$ **relative to** β, denoted $[\mathbf{x}]_\beta$, is the column vector*

$$[\mathbf{x}]_\beta := \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in F^n.$$

Remark 7.2. *This defines a map $\phi_\beta := [\cdot]_\beta : V \to F^n$, which is linear.*

Definition 7.3. *Let V and W be finite-dimensional vector spaces with ordered bases*

$$\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \text{ and } \gamma = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$$

respectively. Let $T : V \to W$ be a linear transformation. Then for each $1 \leq j \leq n$, there exist unique scalars $a_{ij} \in F$ for $1 \leq i \leq m$ such that

$$T(\mathbf{v}_j) = \sum_{i=1}^m a_{ij} \mathbf{w}_i \quad \text{for } 1 \leq j \leq n.$$

*The **matrix representation of** T **in the ordered bases** β **and** γ is defined as the matrix*

$$[T]_\beta^\gamma := (a_{ij}) \in M_{m \times n}(F).$$

If $W = V$ and $\gamma = \beta$, then we can write $[T]_\beta$ instead of $[T]_\beta^\beta$.

Note: From the definition above, equivalently we may write

$$[T(v_1)]_\gamma = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} \in F^m,\ [T(v_2)]_\gamma = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} \in F^m, \cdots, [T(v_n)]_\gamma = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \in F^m,$$

and

$$[T]_\beta^\gamma = ([T(v_1)]_\gamma, [T(v_2)]_\gamma, \cdots, [T(v_n)]_\gamma),$$

meaning that the jth column of $[T]_\beta^\gamma$ is given by the coordinate vector of $T(v_j)$ for the jth vector of β relative to γ.

Lemma 7.4. *Let V and W be finite-dimensional vector spaces with ordered bases β and γ respectively and let $T, U : V \to W$ be linear transformations. Then*

1. $[T + U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma$.
2. $[aT]_\beta^\gamma = a[T]_\beta^\gamma$.

Proposition 7.5. *Let V and W be vector spaces over F. Then the set $\mathcal{L}(V, W)$ of linear transformations from V to W is a vector space over F equipped with addition and scalar multiplication as follows:*

- *For any $T, U \in \mathcal{L}(V, W)$, we define $T + U \in \mathcal{L}(V, W)$ by $(T + U)(\mathbf{x}) = T(\mathbf{x}) + U(\mathbf{x})$ for any $\mathbf{x} \in V$.*
- *For any $a \in F$ and $T \in \mathcal{L}(V, W)$, we define $aT \in \mathcal{L}(V, W)$ by $(aT)(\mathbf{x}) = aT(\mathbf{x})$ for any $\mathbf{x} \in V$.*

Theorem 7.6. *Let V, W, Z be vector spaces over the same field F, and let $T : V \to W$ and $U : W \to Z$ be linear transformations.*

1. *Then the composition function $UT : V \to Z$ is a linear transformation.*
2. *If V, W, Z are all finite-dimensional with ordered bases α, β, γ respectively, then*

$$[UT]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta.$$

Corollary 7.7. *Let V and W be finite-dimensional vector spaces with ordered bases β and γ respectively and let $T : V \to W$ be a linear transformation. Then for any $\mathbf{x} \in V$, we have*

$$[T(\mathbf{x})]_\gamma = [T]_\beta^\gamma [\mathbf{x}]_\beta.$$

Topic 8: Invertibility and isomorphism

Definition 8.1. *Let V and W be vector spaces over the same field F. A function $T : V \to W$ is **invertible** if there exists a function $U : W \to V$ such that $TU = I_W$ and $UT = I_V$.*

Note:

(a) $T : V \to W$ is invertible if and only if $T : V \to W$ is bijective.

(b) If T is invertible, then such U is unique, called the inverse of T and denoted by T^{-1}.

(c) If $\dim(V) = \dim(W) < \infty$ then a linear transformation $T : V \to W$ is invertible if and only if $\text{rank}(T) = \dim(V)$. You may replace the statement $\text{rank}(T) = \dim(V)$ by any other equivalent statements in Corollary 6.8.

(d) If $T : V \to W$ is invertible then $T^{-1} : W \to V$ is invertible with $(T^{-1})^{-1} = T$.

(e) If $T : V \to W$ and $U : W \to Z$ are invertible then $UT : V \to Z$ is invertible with $(UT)^{-1} = T^{-1}U^{-1}$.

Proposition 8.2. *Let the linear transformation $T : V \to W$ be invertible, then the inverse $T^{-1} : W \to V$ is also a linear transformation.*

Lemma 8.3. *Suppose $T : V \to W$ is an invertible linear transformation. Then V is finite-dimensional if and only if W is finite-dimensional, and in such a case, we have dim$(V) =$ dim(W).*

Proposition 8.4. *Let V and W be finite-dimensional vector spaces with ordered bases β and γ respectively. Let $T : V \to W$ be a linear transformation. Then T is invertible if and only if $[T]_\beta^\gamma$ is an invertible matrix, and in such a case, $[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}$.*

Corollary 8.5. *Let V be a finite-dimensional vector space with an ordered basis β. Let $T : V \to V$ be a linear operator on V. Then T is invertible if and only if $[T]_\beta$ is an invertible matrix, and in such a case, $[T^{-1}]_\beta = ([T]_\beta)^{-1}$.*

Corollary 8.6. *Let $A \in M_{n\times n}(F)$. Then A is invertible if and only if L_A is invertible, and in such a case, $(L_A)^{-1} = L_{A^{-1}}$.*

Definition 8.7. *Let V and W be vector spaces. We say V is **isomorphic** to W if there exists an invertible linear transformation $T : V \to W$. In this case, T is called an **isomorphism** from V onto W.*

Theorem 8.8. *Let V and W be finite-dimensional vector spaces. Then V is isomorphic to W if and only if dim$(V) =$ dim(W).*

Corollary 8.9. *Let V be a vector space over F. Then V is isomorphic to F^n if and only if dim$(V) = n$.*

Theorem 8.10. *Let V and W be finite-dimensional vector spaces over F with dimensions n and m respectively. Then the map $\Phi : \mathcal{L}(V, W) \to M_{m\times n}(F)$ defined by mapping a linear transformation $T \in \mathcal{L}(V, W)$ to its matrix representation $[T]_\beta^\gamma \in M_{m\times n}(F)$ is an isomorphism. In particular, dim$(\mathcal{L}(V, W)) =$ dim$(V) \cdot$ dim(W).*

Definition 8.11. *Let β be an ordered basis for an n-dimensional vector space V over a field F. The map $\phi_\beta : V \to F^n$ defined by mapping $\mathbf{x} \in V$ to its coordinate vector $[\mathbf{x}]_\beta \in F^n$ is called the **standard representation of V with respect to** β.*

Proposition 8.12. *ϕ_β is an isomorphism.*

Given vector spaces V and W of dimensions n and m and equipped with ordered bases β and γ respectively, we have the following commutative diagram for any linear transformation $T : V \to W$:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\phi_\beta"'] & W \arrow[d, "\phi_\gamma"] \\ F^n \arrow[r, "L_A"'] & F^m \end{tikzcd}$$

Topic 9: Change of coordinates

Proposition 9.1. *Let β and β' be two ordered bases for a finite-dimensional vector space V, and let*

$$Q := [I_V]_{\beta'}^{\beta},$$

i.e. the matrix representation of the identity operator I_V on V in the ordered bases β' and β. Then

1. *Q is invertible.*
2. *For any $\mathbf{x} \in V$, $[\mathbf{x}]_\beta = Q[\mathbf{x}]_{\beta'}$.*

Definition 9.2. *The matrix $Q = [I_V]_{\beta'}^{\beta}$ is called the **change of coordinate matrix** which changes β'-coordinates into β-coordinates.*

To compute $Q = [I_V]_{\beta'}^{\beta}$, note that if $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $\beta' = \{\mathbf{v}'_1, \mathbf{v}'_2, \ldots, \mathbf{v}'_n\}$, then

$$Q = \left([\mathbf{v}'_1]_\beta \mid [\mathbf{v}'_2]_\beta \mid \cdots \mid [\mathbf{v}'_n]_\beta\right).$$

Proposition 9.3. *Let T be a linear operator on a finite-dimensional vector space V, and let β and β' be two ordered bases for V. Suppose that $Q = [I_V]_{\beta'}^{\beta}$ is the change of coordinate matrix which changes β'-coordinates into β-coordinates. Then*

$$[T]_{\beta'} = Q^{-1}[T]_\beta Q.$$

Corollary 9.4. *Let $A \in M_{m\times n}(F)$ and let $\gamma = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be an ordered basis for F^n. Then*

$$[L_A]_\gamma = Q^{-1}AQ,$$

where $Q = [I_V]_\gamma^\beta = (\mathbf{x}_1 \mid \mathbf{x}_2 \mid \cdots \mid \mathbf{x}_n)$ (where β is the standard ordered basis for F^n).

Definition 9.5. *Given two matrices $A, B \in M_{n\times n}(F)$. We say B is **similar** to A if there exists an invertible matrix Q such that $B = Q^{-1}AQ$.*

Topic 10: Eigenvalues and eigenvectors

Definition 10.1. *Let T be a linear operator on a vector space V over F. A nonzero vector $\mathbf{x} \in V$ is called an **eigenvector** of T if there exists $\lambda \in F$ such that $T(\mathbf{x}) = \lambda\mathbf{x}$; in this case, the scalar $\lambda \in F$ is called the **eigenvalue** of T corresponding to the eigenvector $\mathbf{x}$.*

*For a square matrix $A \in M_{n\times n}(F)$, a nonzero vector $\mathbf{x} \in F^n$ is called an **eigenvector** of A if it is an eigenvector of L_A, i.e. if there exists $\lambda \in F$ such that $A\mathbf{x} = \lambda\mathbf{x}$, and we call $\lambda \in F$ the **eigenvalue** of A corresponding to the eigenvector $\mathbf{x}$.*

Definition 10.2. *A linear operator T on a finite-dimensional vector space V is called **diagonalizable** if there exists an ordered basis β for V such that $[T]_\beta$ is a diagonal matrix.*

Proposition 10.3. *A linear operator T on a finite-dimensional vector space V is diagonalizable if and only if there exists an ordered basis β for V which consists of eigenvectors of T; in such a case, if we write $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, then*

$$[T]_\beta = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where λ_j is the eigenvalue of T corresponding to the eigenvector $\mathbf{v}_j$ (i.e. $T(\mathbf{v}_j) = \lambda_j \mathbf{v}_j$).

Proposition 10.4. *Let $A \in M_{n\times n}(F)$. Then $\lambda \in F$ is an eigenvalue of A if and only if $\det(A - \lambda I) = 0$.*

Definition 10.5. *The **characteristic polynomial** of a square matrix $A \in M_{n\times n}(F)$ is the polynomial $f_A(t) := \det(A - tI_n) \in P_n(F)$. Let T be a linear operator on an n-dimensional vector space V over F. Choose an ordered basis β for V. Then the **characteristic polynomial** of T is defined as the characteristic polynomial of $[T]_\beta$, i.e. $f_T(t) := \det([T]_\beta - tI_n) \in P_n(F)$.*

Proposition 10.6. *The characteristic polynomial $f_T(t)$ of a linear operator T on a finite-dimensional vector space V is well-defined, i.e. independent of the choice of the ordered basis β for V.*

Proposition 10.7. *Let $A \in M_{n\times n}(F)$. Then*

1. *The characteristic polynomial $f_A(t)$ is of degree n and has leading coefficient $(-1)^n$.*

2. *A has at most n distinct eigenvalues.*

Proposition 10.8. *Let T be a linear operator on a vector space V, and let λ be an eigenvalue of T. Then $\mathbf{x} \in V$ is an eigenvector of T corresponding to λ if and only if $\mathbf{v} \in N(T - \lambda I_V) \setminus \{\mathbf{0}\}$.*

Topic 11: Diagonalizability

Theorem 11.1. *A linear operator on an n-dimensional vector space having n distinct eigenvalues is diagonalizable.*

The converse of the above theorem may not be true, for instance, the identity operator is diagonalizable but has only one eigenvalue. The proof of this theorem is based on the following

Lemma 11.2. *Let T be a linear operator on a vector space V, and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be distinct eigenvalues of T. If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are eigenvectors of T corresponding to $\lambda_1, \lambda_2, \ldots, \lambda_k$ respectively, then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a linearly independent subset of V.*

We are going to look for necessary conditions on a diagonalizable linear operator.

Definition 11.3. *We say a polynomial $f(t) \in P(F)$ **splits over** F if there exist $c, a_1, a_2, \ldots, a_n \in F$ such that $f(t) = c(t - a_1)(t - a_2)\cdots(t - a_n)$.*

Proposition 11.4. *The characteristic polynomial of a diagonalizable linear operator on a finite-dimensional vector space V over F splits over F.*

Definition 11.5. *Let T be a linear operator on a vector space V, and let λ be an eigenvalue of T. The subspace*

$$E_\lambda := N(T - \lambda I_V) = \{\mathbf{x} \in V : T(\mathbf{x}) = \lambda \mathbf{x}\}$$

of V is called the ***eigenspace*** *of T corresponding to λ. Eigenspaces of a square matrix $A \in M_{n\times n}(F)$ are defined as those of the linear operator L_A.*

Definition 11.6. *Let λ be an eigenvalue of a linear operator on a finite-dimensional vector space (or a square matrix) with characteristic polynomial $f(t)$. The* ***algebraic multiplicity*** *of λ is defined to be the multiplicity of λ as a zero of $f(t)$, i.e. the largest positive integer j such that $(t-\lambda)^j \mid f(t)$.*

Proposition 11.7. *Let T be a linear operator on a finite-dimensional vector space V and let λ be an eigenvalue of T with algebraic multiplicity m. Then we have*

$$1 \leq dim(E_\lambda) \leq m.$$

We call $dim(E_\lambda)$ the ***geometric multiplicity*** *of λ.*

We now may figure out a test for diagonalizability of a linear operator T on a finite-dimensional vector space, as well as the construction of an ordered basis β of eigenvectors such that $[T]_\beta$ is diagonal.

Theorem 11.8. *Let T be a linear operator on an n-dimensional vector space V whose characteristic polynomial $f_T(t)$ splits. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be all the distinct eigenvalues of T with the algebraic multiplicities $m_1, m_2, \ldots, m_k$, respectively, (thus $1 \leq m_i \leq n$, $m_1 + m_2 + \cdots + m_k = n$). Then*

1. *T is diagonalizable if and only if $m_i = dim(E_{\lambda_i})$ for each $1 \leq i \leq k$.*
2. *If T is diagonalizable and β_i is an ordered basis for E_{λ_i} for each $1 \leq i \leq k$, then*

$$\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$$

with a specified order is an ordered basis for V which consists of eigenvectors of T (so that $[T]_\beta$ is diagonal).

Remark 11.9. *When T is diagonalizable, we have the* ***eigenspace decomposition*** *of V:*

$$V = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_k}.$$

The proof of Theorem 11.8 is based on the following

Lemma 11.10. *Let T be a linear operator on a vector space V, and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be distinct eigenvalues of T. For each $1 \leq i \leq k$, let $S_i \subset E_{\lambda_i}$ be a finite linearly independent subset. Then $S = S_1 \cup S_2 \cup \cdots \cup S_k$ is a linearly independent subset of V.*

An application of Theorem 11.8: Let $A \in M_{n\times n}(F)$ be diagonalizable. Thus the characteristic polynomial $f_A(t)$ splits over F, and one may write that

$$f_A(t) = (-1)^n (t-\lambda_1)^{m_1} \cdots (t-\lambda_k)^{m_k},$$

where λ_i are distinct eigenvalues and m_i are the associated (algebraic) multiplicities with $m_i = \dim(E_{\lambda_i})$ and $m_1 + \cdots + m_k = n$. Let $\beta_1, \cdots, \beta_k$ be the ordered bases of eigenspaces $E_{\lambda_1}, \cdots, E_{\lambda_k}$ respectively, so $\beta := \beta_1 \cup \cdots \cup \beta_k$ is an ordered basis of eigenvectors of A for F^n. By item 2 of Theorem 11.8,

$$[L_A]_\beta = \mathrm{diag}(\underbrace{\lambda_1, \cdots, \lambda_1}_{m_1 \text{ terms}}, \underbrace{\lambda_2, \cdots, \lambda_2}_{m_2 \text{ terms}}, \cdots, \underbrace{\lambda_k, \cdots, \lambda_k}_{m_k \text{ terms}}).$$

On the other hand, letting α be the standard ordered basis for F^n, it holds that $[L_A]_\beta = [I_{F^n}]_\alpha^\beta [L_A]_\alpha^\alpha [I_{F^n}]_\beta^\alpha$, where $[L_A]_\alpha^\alpha = A$, $[I_{F^n}]_\beta^\alpha = (\beta_1|\beta_2|\cdots|\beta_k)$, and $[I_{F^n}]_\alpha^\beta = ([I_{F^n}]_\beta^\alpha)^{-1}$. Therefore,

$$\mathrm{diag}(\underbrace{\lambda_1, \cdots, \lambda_1}_{m_1 \text{ terms}}, \underbrace{\lambda_2, \cdots, \lambda_2}_{m_2 \text{ terms}}, \cdots, \underbrace{\lambda_k, \cdots, \lambda_k}_{m_k \text{ terms}}) = Q^{-1}AQ,$$

with $Q := (\beta_1|\beta_2|\cdots|\beta_k)$, or equivalently we may write

$$A = Q\mathrm{diag}(\underbrace{\lambda_1, \cdots, \lambda_1}_{m_1 \text{ terms}}, \underbrace{\lambda_2, \cdots, \lambda_2}_{m_2 \text{ terms}}, \cdots, \underbrace{\lambda_k, \cdots, \lambda_k}_{m_k \text{ terms}})Q^{-1},$$

which is a very useful form in many applications. For instance, it is convenient to use it to compute A^j for any integer $j \geq 0$, particularly when j is very large. Another example is to solve an ODE system $\dot{x} = Ax$ by making a change of variables $x = Qy$ and reducing the system to $\dot{y} = Q^{-1}AQy$ with the diagonal coefficient matrix $Q^{-1}AQ$. □

Topic 12: Invariant subspace and Cayley-Hamilton Theorem

Theorem 12.1 (Cayley-Hamilton)**.** *Let T be a linear operator on a finite-dimensional vector space V, and let $f(t) := f_T(t)$ be its characteristic polynomial. Then $f(T) = T_0$, the zero transformation, so T is a "zero" of $f(t)$.*

Corollary 12.2. *Let $A \in M_{n\times n}(F)$ and $f(t)$ be its characteristic polynomial. Then we have $f(A) = O$, the zero matrix.*

To show the main theorem above, we need to introduce the invariant subspaces.

Definition 12.3. *Let T be a linear operator on a vector space V. A subspace $W \subset V$ is called **T-invariant** if $T(W) \subset W$ (or equivalently, $T(\mathbf{w}) \in W$ for any $\mathbf{w} \in W$).*

Definition 12.4. *Given a linear operator T on a vector space V and a nonzero vector $\mathbf{x} \in V$, the subspace*

$$W := span(\{T^k(\mathbf{x}) : k \in \mathbb{N}\}) = span(\{\mathbf{x}, T(\mathbf{x}), T^2(\mathbf{x}), \ldots\})$$

*is called the **T-cyclic subspace of** V **generated by** $\mathbf{x}$.*

Proposition 12.5. *W is the smallest T-invariant subspace of V containing $\mathbf{x}$.*

Note that if $W \subset V$ is a T-invariant subspace, then the restriction of T to W gives a linear operator $T_W := T|_W$ on W.

Lemma 12.6. *$f_{T_W}(t)$ divides $f_T(t)$.*

Lemma 12.7. *Let T be a linear operator on a finite-dimensional vector space V, and let $W \subset V$ be the T-cyclic subspace of V generated by a nonzero vector $\mathbf{x} \in V$. Denote $k := dim(W)$. Then*

1. $\beta := \{\mathbf{x}, T(\mathbf{x}), T^2(\mathbf{x}), \ldots, T^{k-1}(\mathbf{x})\}$ *is a basis for W.*
2. *Note $T^k(\mathbf{x}) \in W = \operatorname{span}(\beta)$. Then*

$$T^k(\mathbf{x}) = -(a_0\mathbf{x} + a_1T(\mathbf{x}) + a_2T^2(\mathbf{x}) + \cdots + a_{k-1}T^{k-1}(\mathbf{x})),$$

for some $a_0, \cdots, a_{k-1} \in F$, and the characteristic polynomial of $T_W \in \mathcal{L}(W)$ is given by

$$f_{T_W}(t) = (-1)^k(a_0 + a_1t + a_2t^2 + \cdots + a_{k-1}t^{k-1} + t^k).$$

Topic 13: Inner product space

From this point on, we assume $F = \mathbb{R}$ or $\mathbb{C}$.

Definition 13.1. *Let V be a vector space over F. An **inner product** on V is a map*

$$\langle \cdot, \cdot \rangle : V \times V \to F$$

such that, for any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and $c \in F$, we have

1. $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$.
2. $\langle c\mathbf{x}, \mathbf{y} \rangle = c\langle \mathbf{x}, \mathbf{y} \rangle$.
3. $\overline{\langle \mathbf{x}, \mathbf{y} \rangle} = \langle \mathbf{y}, \mathbf{x} \rangle$.
4. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ *and equality holds only when* $\mathbf{x} = \mathbf{0}$.

*An **inner product space** is a vector space V together with an inner product. If $F = \mathbb{C}$, we call V a **complex inner product space**; while if $F = \mathbb{R}$, we call V a **real inner product space**.*

Remark 13.2. *Properties 1 and 2 together say that the inner product is **linear** in its first argument.*

Remark 13.3. *Property 3 reduces to $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ whenever $F = \mathbb{R}$.*

Examples of inner product spaces:

- For $\mathbf{x} = (a_1, a_2, \ldots, a_n), \mathbf{y} = (b_1, b_2, \ldots, b_n) \in F^n$, we have the *standard inner product*

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{i=1}^{n} a_i\overline{b_i}.$$

- If $\langle \cdot, \cdot \rangle$ is an inner product on V and $r \in \mathbb{R}_{>0}$, then $\langle \mathbf{x}, \mathbf{y} \rangle' = r\langle \mathbf{x}, \mathbf{y} \rangle$ defines another inner product on V.

- Let $V = C([a,b])$ be the vector space of real-valued continuous functions on the interval $[a,b]$. Then for $f, g \in V$,

$$\langle f, g\rangle := \int_a^b f(t)g(t)dt$$

defines an inner product on V.

- Let $V = M_{n\times n}(F)$. For $A, B \in V$,

$$\langle A, B\rangle := \text{tr}(B^*A),$$

where $B^* = \overline{B}^t$ is the **conjugate transpose** of B, defines an inner product on V.

Proposition 13.4. *Let V be an inner product space. Then for any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and $c \in F$, we have*

1. $\langle \mathbf{x}, \mathbf{y} + \mathbf{z}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{x}, \mathbf{z}\rangle$.
2. $\langle \mathbf{x}, c\mathbf{y}\rangle = \overline{c}\langle \mathbf{x}, \mathbf{y}\rangle$.
3. $\langle \mathbf{0}, \mathbf{x}\rangle = \langle \mathbf{x}, \mathbf{0}\rangle = 0$ *for any* $\mathbf{x} \in V$.
4. $\langle \mathbf{x}, \mathbf{x}\rangle = 0$ *if and only if* $\mathbf{x} = \mathbf{0}$.
5. *If* $\langle \mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{z}, \mathbf{y}\rangle$ *for all* $\mathbf{y} \in V$*, then* $\mathbf{x} = \mathbf{z}$.

Remark 13.5. *Properties 1 and 2 here together say that the inner product is* ***conjugate linear*** *in its second argument.*

Definition 13.6. *The* ***norm*** *(or* ***length****) of a vector $\mathbf{x} \in V$ in an inner product space V is defined by* $\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}$.

Proposition 13.7. *Let V be an inner product space. The norm $\|\cdot\|$ induced by the inner product satisfies the following three properties:*

1. $\|\mathbf{x}\| \geq 0, \forall\, \mathbf{x} \in V$; $\|\mathbf{x}\| = 0$ *if and only if* $\mathbf{x} = \mathbf{0}$.
2. $\|c\mathbf{x}\| = |c|\|\mathbf{x}\|, \forall\, \mathbf{x} \in V, \forall\, c \in F$.
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|, \forall\, \mathbf{x}, \mathbf{y} \in V$ *(Triangle Inequality).*

Note: The above three properties are usually used as a direct definition of a norm in the general situation. The proof of the triangle inequality is based on the Cauchy-Schwarz inequality: $|\langle \mathbf{x}, \mathbf{y}\rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|, \forall\, \mathbf{x}, \mathbf{y} \in V$.

Definition 13.8. *Let V be an inner product space. Two vectors $\mathbf{x}, \mathbf{y} \in V$ are said to be* ***orthogonal*** *(or* ***perpendicular****), denoted as $\mathbf{x} \perp \mathbf{y}$, if $\langle \mathbf{x}, \mathbf{y}\rangle = 0$. A subset $S \subset V$ is called* ***orthogonal*** *if any two distinct vectors in S are orthogonal. A* ***unit vector*** *in V is a vector $\mathbf{x} \in V$ with $\|\mathbf{x}\| = 1$. A subset $S \subset V$ is called* ***orthonormal*** *if S is orthogonal and all vectors in S are unit vectors.*

Topic 14: Gram-Schmidt orthogalization

Lemma 14.1. *Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be an orthogonal subset of V consisting of nonzero vectors. Then S is linearly independent.*

Note: If $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is an orthogonal subset of V consisting of nonzero vectors, then

$$\left\{ \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|^2}, \ldots, \frac{\mathbf{x}_n}{\|\mathbf{x}_n\|^2} \right\}$$

is an orthonormal subset of V; this process is called **normalization**.

Theorem 14.2 (Gram-Schmidt Process)**.** *Let $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ be a linearly independent subset of an inner product space V. Set $\mathbf{v}_1 = \mathbf{w}_1$ and*

$$\mathbf{v}_k := \mathbf{w}_k - \sum_{j=1}^{k-1} \frac{\langle \mathbf{w}_k, \mathbf{v}_j \rangle}{\|\mathbf{v}_j\|^2} \mathbf{v}_j$$

for $2 \leq k \leq n$. Then $S' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal subset of V consisting of nonzero vectors such that span(S') = span(S).

Corollary 14.3. *Any finite dimensional inner product space has an orthonormal basis (namely, an ordered basis which is orthonormal).*

Proposition 14.4. *Let V be an inner product space and $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be an orthogonal subset of V consisting of nonzero vectors. Then for any $\mathbf{y} \in$ span(S),*

$$\mathbf{y} = \sum_{i=1}^{k} \frac{\langle \mathbf{y}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \mathbf{v}_i.$$

Particularly, whenever $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is an orthonormal subset of V, it holds that for any $\mathbf{y} \in$ span(S),

$$\mathbf{y} = \sum_{i=1}^{k} \langle \mathbf{y}, \mathbf{v}_i \rangle \mathbf{v}_i.$$

Topic 15: Orthogonal complement

Definition 15.1. *Let W be a subspace of an inner product space V. The **orthogonal complement** of W is defined as the subspace*

$$W^{\perp} := \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{w} \rangle = 0 \textit{ for all } \mathbf{w} \in W\}.$$

Proposition 15.2. *Let V be an inner product space and $W \subset V$ be a finite-dimensional subspace. Then for any $\mathbf{y} \in V$, there exist unique $\mathbf{u} \in W$ and $\mathbf{z} \in W^{\perp}$ such that*

$$\mathbf{y} = \mathbf{u} + \mathbf{z}.$$

Furthermore, if $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *is an orthonormal basis for* W*, then*

$$\mathbf{u} = \sum_{i=1}^{k} \langle \mathbf{y}, \mathbf{v}_i \rangle \mathbf{v}_i.$$

The vector $\mathbf{u} \in W$ *is called the* ***orthogonal projection*** *of* $\mathbf{y}$ *on* W.

Corollary 15.3. *With notations as above, then*

$$\|\mathbf{y} - \mathbf{x}\| \geq \|\mathbf{y} - \mathbf{u}\|$$

for any $\mathbf{x} \in W$*, and equality holds if and only if* $\mathbf{x} = \mathbf{u}$*. In other words,* $\mathbf{u}$ *is the unique vector in* W *which is closest to* $\mathbf{y}$.

Proposition 15.4. *Suppose* $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *is an orthonormal subset in an* n*-dimensional inner product space* V*. Then*

1. S *can be extended to an orthonormal basis* $\{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ *for* V.
2. *If* $W =$ *span*(S)*, then* $S_1 := \{\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ *is an orthonormal basis for* $W^{\perp}$.
3. *For any subspace* W *of* V*, we have dim*$(V) =$ *dim*$(W) +$ *dim*$(W^{\perp})$.

Remark 15.5. *We have* $V = W \oplus W^{\perp}$.

Topic 16: Adjoint of a linear operator

Proposition 16.1. *Let* V *be a finite-dimensional inner product space over* F*. Then for any linear transformation (or* ***linear functional****)* $g : V \to F$*, there exists a unique* $\mathbf{y} \in V$ *such that* $g(\mathbf{x}) = \langle \mathbf{x}, \mathbf{y} \rangle$ *for any* $\mathbf{x} \in V$.

Theorem 16.2. *Let* V *be a finite-dimensional inner product space, and let* T *be a linear operator on* V*. Then there exists a unique linear operator* T^* *on* V *such that*

$$\langle T(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, T^*(\mathbf{y}) \rangle$$

for any $\mathbf{x}, \mathbf{y} \in V$*. We call* T^* *the* ***adjoint*** *of* T.

Remark 16.3. *In practice, if* $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ *be an orthonormal basis for* V*, then* T^* *is computed by the formula*

$$T^*(\mathbf{y}) = \sum_{i=1}^{n} \overline{\langle T(\mathbf{v}_i), \mathbf{y} \rangle} \mathbf{v}_i$$

for any $\mathbf{y} \in V$.

Proposition 16.4. *Let* V *be a finite-dimensional inner product space, and let* β *be an orthonormal basis for* V*. Then for any* $T \in \mathcal{L}(V)$*, we have* $[T^*]_\beta = [T]_\beta^*$.

Corollary 16.5. *Let* A *be an* $n \times n$ *matrix. Then* $L_{A^*} = (L_A)^*$.

Proposition 16.6. *Let V be an inner product space, and let $T, U \in \mathcal{L}(V)$. Then*

1. $(T+U)^* = T^* + U^*$.
2. $(cT)^* = \overline{c}T^*$ *for any* $c \in F$.
3. $(TU)^* = U^*T^*$.
4. $T^{**} = T$.
5. $I^* = I$.

Corollary 16.7. *Let A, B be $n \times n$ matrices. Then*

1. $(A+B)^* = A^* + B^*$.
2. $(cA)^* = \overline{c}A^*$ *for any* $c \in F$.
3. $(AB)^* = B^*A^*$.
4. $A^{**} = A$.
5. $I^* = I$.

Topic 17: Normal operator and self-adjoint operator

Lemma 17.1. *Let T be a linear operator on a finite-dimensional inner product space V. If T has an eigenvector, then so does T^*.*

Theorem 17.2 (Schur)**.** *Let T be a linear operator on a finite-dimensional inner product space V. Suppose that the characteristic polynomial of T splits. Then there exists an orthonormal basis β for V such that $[T]_\beta$ is upper triangular.*

Definition 17.3. *Let V be an inner product space. We say that a linear operator T on V is* ***normal*** *if $T^*T = TT^*$. An $n \times n$ (real or complex) matrix A is called* ***normal*** *if $A^*A = AA^*$.*

We say that a linear operator $T \in \mathcal{V}$ is

- **unitary** if $T^*T = TT^* = I$;
- **self-adjoint** (or Hermitian) if $T^* = T$;
- **anti-self-adjoint** (or skew-Hermitian) if $T^* = -T$.

All of these are examples of normal operators. For example, the rotation $T : \mathbb{R}^2 \to \mathbb{R}^2$ by an angle θ in the counterclockwise direction, which is represented by

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

is unitary, and hence normal.

Proposition 17.4. *Let V be an inner product space, and let T be a normal operator on V. Then we have*

1. $\|T(\mathbf{x})\| = \|T^*(\mathbf{x})\|$ *for any* $\mathbf{x} \in V$.
2. $T - cI_V$ *is normal for any* $c \in F$.
3. *If* $T(\mathbf{x}) = \lambda\mathbf{x}$, *then* $T^*(\mathbf{x}) = \overline{\lambda}\mathbf{x}$.
4. *If* $\mathbf{x}_1$ *and* $\mathbf{x}_2$ *are eigenvectors of* T *corresponding to distinct eigenvalues, then* $\mathbf{x}_1$ *and* $\mathbf{x}_2$ *are orthogonal.*

Theorem 17.5 (Spectral Theorem for complex inner product spaces)**.** *Let T be a linear operator on a finite-dimensional complex inner product space V. Then T is normal if and only if there exists an orthonormal basis for V which consists of eigenvectors of T.*

Remark 17.6. *This theorem is not true in the infinite-dimensional case.*

Definition 17.7. *Let T be a linear operator on an inner product space V. We say T is* ***self-adjoint*** *if $T^* = T$. An $n \times n$ (real or complex) matrix A is called* ***self-adjoint*** *if $A^* = A$.*

Lemma 17.8. *Let T be a self-adjoint operator on a finite-dimensional inner product space V. Then*

1. *Every eigenvalue of T is real.*
2. *Suppose V is a real inner product space. Then the characteristic polynomial of T splits over $\mathbb{R}$.*

Theorem 17.9 (Spectral Theorem for real inner product spaces)**.** *Let T be a linear operator on a finite-dimensional real inner product space V. Then T is self-adjoint if and only if there exists an orthonormal basis for V which consists of eigenvectors of T.*

Topic 18: Unitary (orthogonal) operator

Definition 18.1. *Let T be a linear operator on a finite-dimensional inner product space V over F. If $\|T(\mathbf{x})\| = \|\mathbf{x}\|$ for any $\mathbf{x} \in V$, we call T a* ***unitary operator*** *(resp.* ***orthogonal operator****) when $F = \mathbb{C}$ (resp. $F = \mathbb{R}$).*

Lemma 18.2. *Let U be a self-adjoint operator on a finite-dimensional inner product space V. If $\langle \mathbf{x}, U(\mathbf{x}) \rangle = 0$ for all $\mathbf{x} \in V$, then $U = T_0$, the zero operator.*

Theorem 18.3. *For a linear operator T on a finite-dimensional inner product space V, the following are equivalent*

1. $T^*T = TT^* = I$.
2. *T preserves the inner product on V, i.e. $\langle T(\mathbf{x}), T(\mathbf{y}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ for any $\mathbf{x}, \mathbf{y} \in V$.*
3. *$T(\beta)$ is an orthonormal basis for V for any orthonormal basis β for V.*
4. *There exists an orthonormal basis β for V such that $T(\beta)$ is an orthonormal basis for V.*

5. $\|T(\mathbf{x})\| = \|\mathbf{x}\|$ *for any* $\mathbf{x} \in V$.

Definition 18.4. *A matrix $A \in M_{n\times n}(\mathbb{R})$ is called **orthogonal** if $A^tA = AA^t = I_n$. The set of $n \times n$ orthogonal matrices is denoted by $O(n)$. A matrix $A \in M_{n\times n}(\mathbb{C})$ is called **unitary** if $A^*A = AA^* = I_n$. The set of $n \times n$ unitary matrices is denoted by $U(n)$.*

Remark 18.5. *A linear operator T on an inner product space V is unitary (resp. orthogonal) if and only if there exists an orthonormal basis β for V such that $[T]_\beta$ is unitary (resp. orthogonal).*

Remark 18.6. *Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in F^n$. Then the matrix $A = (\mathbf{v}_1 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_n) \in M_{n\times n}(F)$ is unitary (resp. orthogonal) if and only if $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthonormal basis for $F^n = \mathbb{C}^n$ (resp. $F^n = \mathbb{R}^n$).*

Theorem 18.7. *Let $A \in M_{n\times n}(\mathbb{C})$. Then A is normal if and only if A is unitarily equivalent to a diagonal matrix, i.e. there exists $P \in U(n)$ such that P^*AP is diagonal.*

Theorem 18.8. *Let $A \in M_{n\times n}(\mathbb{R})$. Then A is self-adjoint if and only if A is orthogonally equivalent to a diagonal matrix, i.e. there exists $P \in O(n)$ such that P^tAP is diagonal.*

Topic 19: Spectral theorem

Proposition 19.1. *Let V be an inner product space and $W \subset V$ be a finite-dimensional subspace with an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$. Then the orthogonal projection $T : V \to V$ defined by*

$$T(\mathbf{y}) = \sum_{i=1}^{k} \langle \mathbf{y}, \mathbf{v}_i \rangle \mathbf{v}_i$$

is a linear operator on V such that

1. $N(T) = W^\perp$ *and* $R(T) = W$.
2. $T^2 = T$.
3. T *is self-adjoint.*

Remark 19.2. *Properties 1 and 2 above uniquely determine the orthogonal projection onto W.*

Theorem 19.3. *Let T be a linear operator on a finite-dimensional inner product space V over F with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$. Assume that T is normal (resp. self-adjoint) when $F = \mathbb{C}$ (resp. $F = \mathbb{R}$). For $1 \le i \le k$, let $E_i = E_{\lambda_i}$ be the eigenspace of T corresponding to λ_i and let T_i be the orthogonal projection onto E_i. Then*

1. $V = E_1 \oplus E_2 \oplus \cdots \oplus E_k$.
2. $E_i^\perp = \bigoplus_{j\neq i} E_j$ *for* $1 \le i \le k$.
3. $T_iT_j = \delta_{ij}T_j$ *for* $1 \le i, j \le k$.
4. $I_V = T_1 + T_2 + \cdots + T_k$ *(resolution of the identity operator induced by T).*
5. $T = \lambda_1T_1 + \lambda_2T_2 + \cdots + \lambda_kT_k$ *(spectral decomposition of T).*

—END—