

MAT 442: Advanced Linear Algebra

Arizona State University

Fall 2020

Contents

1	Vector Spaces	3
1.1	Fields	3
1.2	Vector Spaces	4
1.3	Subspaces	6
1.4	Linear combinations and systems of linear equations	8
1.5	Linear dependence and linear independence	9
1.6	Bases and dimension	10
1.7	Maximal linearly independent subsets	13
2	Linear Transformations and Matrices	14
2.1	Linear transformations, null spaces, and ranges	14
2.2	The matrix representation of a linear transformation	18
2.3	Composition of linear transformations and matrix multiplication	19
2.4	Invertibility and isomorphisms	23
2.5	The change of coordinate matrix	24
2.6	Dual spaces	25
3	Elementary matrix operations and systems of linear equations	26
3.1	Elementary matrix operations and elementary matrices	26
3.2	The rank of a matrix and matrix inverse	27
3.3	Systems of linear equations (theoretical aspect)	28
4	Determinants	29
4.1	Determinants of order 2	29
4.2	Determinants of order n	29
4.3	Properties	31
5	Diagonalization	31
5.1	Eigenvalues and eigenvectors	31
5.2	Diagonalizability	32
5.4	Invariant subspaces and the Cayley-Hamilton theorem	33
6	Inner product spaces	34
6.1	Inner products and norms	34
6.2	The Gram-Schmidt orthogonalization	36
6.3	The adjoint of a linear operator	37
6.4	Normal and self-adjoint operators	38
6.5	Unitary and orthogonal operators and their matrices	39

1 Vector Spaces

1.1 Fields

Definition 1 A *field* is a triple $(F, +, \cdot)$ with $+: F \times F \rightarrow F$ and $\cdot: F \times F \rightarrow F$ that satisfies the following axioms.

- (1) For all $a, b, c \in F$, $(a + b) + c = a + (b + c)$ (Associativity of addition)
- (2) For all $a, b, c \in F$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (Associativity of multiplication)
- (3) For all $a, b \in F$, $a + b = b + a$ (Commutativity of addition)
- (4) For all $a, b \in F$, $a \cdot b = b \cdot a$ (Commutativity of multiplication)
- (5) For all $a, b, c \in F$, $a \cdot (b + c) = a \cdot b + a \cdot c$ (Distributive law)
- (6) There exists $0 \in F$ such that $a + 0 = a$ for all $a \in F$ (Identity element of addition)
- (7) There exists $0 \in F$ such that $a \cdot 0 = a$ for all $a \in F$ (Identity element of multiplication)
- (8) For all $a \in F$, there exists $(-a) \in F$ such that $a + (-a) = 0$ (Additive inverse)
- (9) For all $a \in F$, there exists $a^{-1} \in F$ such that $a \cdot a^{-1} = 0$ (Multiplicative inverse)

Examples of Fields

- (1) $(\mathbb{R}, +, \cdot)$
- (2) $(\mathbb{C}, +, \cdot)$
- (3) $(\mathbb{Q}, +, \cdot)$
- (4) $(\{a + b\sqrt{2} : a, b \in \mathbb{Q}\}, +, \cdot)$
- (5) $(\mathbb{Z}_2, +, \cdot)$
- (6) $(\mathbb{Z}_p, +, \cdot)$, where p prime

1.2 Vector Spaces

Definition 1 A *vector space* V over a field F is a triple $(V, +, \cdot)$ where $+: V \times V \rightarrow V$ (addition), $\cdot: F \times V \rightarrow V$ (scalar multiplication) and the following conditions hold.

- (VS 1) For $x, y \in V$, $x + y = y + x$ (Commutativity of addition)
- (VS 2) For $x, y, z \in V$, $(x + y) + z = x + (y + z)$ (Associativity of addition)
- (VS 3) There exists an element $0_V \in V$ such that $0_V + x = x$ for all $x \in V$ (Identity element of addition)
- (VS 4) For $x \in V$, there is $y_x \in V$ such that $x + y_x = 0_V$ (Inverse elements of addition)
- (VS 5) For $x \in V$, $1 \cdot x = x$ (Identity element of scalar multiplication)
- (VS 6) For $a, b \in F$, $x \in V$, $(ab)x = a(bx)$ (Compatibility of scalar multiplication)
- (VS 7) For $a \in F$, $x, y \in V$, $a(x + y) = ax + ay$ (Distributivity of scalar multiplication with respect to vector addition)
- (VS 8) For $a, b \in F$, $x \in V$, $(a + b)x = ax + bx$ (Distributivity of scalar multiplication with respect to field addition)

Conventions

- We will often identify the set V with the vector space V .
- We will write av instead of $a \times v$.
- We will often write 0 for 0_V when V is clear from the context.

An $m \times n$ matrix with entries from a field F is a function $A: \{1, \dots, m\} \times \{1, \dots, n\} \rightarrow F$. We often organize the values in a rectangular array with m rows and n columns and use A_{ij} for $A(i, j)$.

Examples of Vector Fields

- (1) $\mathbb{F}^n = \{(a_1, \dots, a_n) : a_i \in F\}$
- (2) $\{f: S \rightarrow F : S \neq \emptyset\}$. Also notated $\mathcal{F}(S, F)$.
- (3) Set of all sequences over \mathbb{F}
- (4) $P(F)$. Set of all polynomials with coefficients in F .
- (5) $M_{m \times n}(\mathbb{F})$

Example: Is the empty set a vector space?

Answer. No. By rule (VS 3), there must exist $0_V \in V$, such that $0_V + x = x$ for all $x \in V$, but since $V = \emptyset$, $\nexists 0_V \in V$.

Example: Decide if V is a vector space.

$$(1) \quad V = \{(a_1, a_2) : a_1, a_2 \in \mathbb{R}\}, (a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 - b_2), c(a_1, a_2) = (ca_1, ca_2)$$

Answer. No. V violates axiom (VS 1).

$$(2) \quad V = \{(a_1, a_2) : a_1, a_2 \in \mathbb{R}\}, (a_1, a_2) + (b_1, b_2) = (a_1 + b_1, 0), c(a_1, a_2) = (ca_1, 0)$$

Answer. No. V violates axiom (VS 5).

$$(3) \quad V = \{(a_1, a_2) : a_1, a_2 \in \mathbb{R}\}, (a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 b_2), c(a_1, a_2) = (ca_1, a_2)$$

Answer. No. V violates axiom (VS 4).

Theorem 1.1. (Cancellation Law for Vector Addition) Let $x, y, z \in V$. If $x + z = y + z$, then $x = y$.

Proof. Suppose $x + z = y + z$. Let w be such that $z + w = 0_V$. Then

$$\begin{aligned} x &= x + 0_V \\ &= x + (z + w) \\ &= (x + z) + w \\ &= (y + z) + w \\ &= y + (z + w) \\ &= y + 0_V \\ &= y. \end{aligned}$$

□

Corollary 2

- There is unique $0_V \in V$ such that $0_V + x = x$ for all $x \in V$.

Proof. Let $v \in V$. Suppose that 0_V and $0'_V$ are zero elements.

$$v + 0_V = v = v + 0'_V.$$

Thus, by cancellation law, $0_V = 0'_V$.

□

- For every $x \in V$, there is unique $y_x \in V$ such that $x + y_x = 0_V$. *Proof.* Let $x \in V$. Suppose that y_x and y'_x are inverses of x . Then

$$x + y_x = 0 = x + y'_x.$$

Thus, by cancellation law, $y_x = y'_x$. □

Theorem 1.2. Let $x \in V$, $a \in F$. Then

- $0x = 0_V$;

Proof.

$$\begin{aligned} 0x + 0x &= (0 + 0)x \\ &= 0x \\ &= 0x + 0_V. \end{aligned}$$

Thus, by cancellation law, $0x = 0_V$. □

- $(-a)x = -(ax) = a(-x)$;

Proof. We can show

$$\begin{aligned} (-a)x + (ax) &= (-a + a)x \\ &= 0x \\ &= 0_V. \end{aligned}$$

Thus, the inverse of ax is unique, which implies $-ax = (-a)x$. We have $(-x) = (-1)x$, which implies $a(-x) = (a(-1))x = (-a)x$. □

- $a0_V = 0_V$.

Proof.

$$a0_V = a(0_V - 0_V) = a0_V - a0_V = 0_V.$$

□

1.3 Subspaces

Definition 2 Let $(V, +, \cdot)$ be a vector space and let $W \subseteq V$. Then W is called a *subspace* if $(W, +, \cdot)$ is a vector space.

Note: W is a subspace if

- $x + y \in W$ for $x, y \in W$

- $cx \in W$ for $x \in W, c \in F$
- W has a zero vector 0_W
- For $x \in W$ there is $y_x \in W$ such that $x + y_x = 0_W$

Theorem 1.3. Let W be a subset of V a vector space $(V, +, \cdot)$. Then W is a subspace of V if and only if the following hold.

- (1) $0_V \in W$
- (2) if $x, y \in W$, then $x + y \in W$
- (3) if $c \in F$ and $x \in W$, then $cx \in W$

Proof. (\implies): Suppose W is a subspace. Then (2) and (3) are satisfied. Then, $0_W + 0_W = 0_W = 0_V$. Thus, by cancellation law, $0_W = 0_V$.

(\impliedby): Suppose $x \in W$. We have $-x = (-1)x$ and $-1 \in F, x \in W$. Thus, $-x \in W$. \square

The *transpose* of an $m \times n$ matrix A is the $n \times m$ matrix A^t such that $(A^t)_{ij} = A_{ji}$. Note, $(A + B)^t = A^t + B^t$. Also, $(cA^t) = cA^t$. This is proven by,

$$\begin{aligned} ((A + B)^t)_{ij} &= (A + B)_{ji} \\ &= A_{ji} + B_{ji} \\ &= (A^t)_{ij} + (B^t)_{ij}. \end{aligned}$$

A is called *symmetric* if $A^t = A$.

An $n \times n$ matrix A is called *diagonal* if $A_{ji} = 0$ whenever $j \neq i$.

Theorem 1.4. Any intersection of subspaces of a vector space V is a subspace of V .

Definition 3

- Let $S_1, S_2 \subseteq V$. Then $S_1 + S_2$ is the set $\{x + y : x \in S_1, y \in S_2\}$
- V is called a direct sum of W_1 and W_2 if W_1, W_2 are subspaces of V such that $W_1 \cap W_2 = \{0\}$ and $W_1 + W_2 = V$. We then write $V = W_1 \oplus W_2$.

Example: Let $V = \mathbb{R}^2$, $S_1 = \{(a, 0) : a \in \mathbb{R}\}$, $S_2 = \{(0, b) : b \in \mathbb{R}\}$. Then $S_1 + S_2$ generates V .

Examples of subspaces

- (1) $\mathcal{C} = \{f \in \mathcal{F}(\mathbb{R}, \mathbb{R}) : f \text{ is continuous}\}$ is a subspace of $\mathcal{F}(\mathbb{R}, \mathbb{R})$.
- (2) $P_n(\mathbb{R}) = \{f : f \text{ is a polynomial and } \deg(f) \leq n\}$ is a subspace of $P(\mathbb{R})$.
- (3) $D = \{A \in M_{n \times n}(\mathbb{R}) : A \text{ is diagonal}\}$ is a subspace of $M_{n \times n}(\mathbb{R})$.
- (4) $W = \{A \in M_{n \times n}(\mathbb{R}) : \text{tr}(A) = 0\}$ is a subspace of $M_{n \times n}(\mathbb{R})$.

1.4 Linear combinations and systems of linear equations

Definition 4 Let S be a subset of a vector space V and let $v \in V$. Then v is called a linear combination of vectors of S if there exists $u_1, \dots, u_n \in S$ and $a_1, \dots, a_n \in F$ such that

$$v = \sum_{i=1}^n a_i v_i.$$

When solving a system of linear equations we reduced to what is called the reduction echelon form by performing the following three operations.

- Interchange two equations
- Multiply an equation by a non-zero element from F
- Add a scalar multiple of one equation to another

Example: $v = (2, 6, 8) \in \mathbb{R}^3$ is a linear combination of $v_1 = (1, 2, 1)$, $v_2 = (-2, -4, -2)$, $v_3 = (0, 2, 3)$, $v_4 = (2, 0, -3)$, $v_5 = (-3, 8, 16)$. Specifically, $v = a_1 v_1 + \dots + a_5 v_5$ where $(a_1, a_2, a_3, a_4, a_5) = (-4, 0, 7, 3, 0)$ is one solution.

Definition 5 Let S be a subset of a vector space V . If S is non-empty, we let $\text{span}(S)$ be the set of all linear combinations of vectors from S , and we set $\text{span}(\emptyset) = \{0_V\}$.

Theorem 1.5. Let S be a subset of a vector space V . Then $\text{span}(S)$ is a subspace of V . Moreover, if W is a subspace of V such that $S \subseteq W$, then $\text{span}(S) \subseteq W$.

Proof.

- (1) (Show $\text{span}(S)$ is a subspace): If $S = \emptyset$, then $\text{span}(S) = \{0_V\}$, which is a subspace. Assume $S \neq \emptyset$. Let $v \in S$. Then,

- (a) $0v = 0_V \in \text{span}(S)$
- (b) Suppose $x, y \in \text{span}(S)$. Then $x = \sum_{i=1}^n a_i v_i$, $y = \sum_{i=1}^m b_i w_i$ where $v_1, \dots, v_n, w_1, \dots, w_m \in S$ and $a_1, \dots, a_n, b_1, \dots, b_m \in F$. Let

$$u_i = \begin{cases} v_i & i \leq n \\ w_{i-n} & i > n \end{cases}, \quad c_i = \begin{cases} a_i & i \leq n \\ b_{i-n} & i > n. \end{cases}$$

Then $x + y = \sum_{i=1}^{n+m} c_i u_i$ and $u_i \in S$, $c_i \in F$ for $i = 1, \dots, n+m$. Thus, $x + y \in \text{span}(S)$.

- (c) Let $x \in \text{span}(S)$ and let $c \in F$. Then $x = \sum_{i=1}^n a_i v_i$ where $v_1, \dots, v_n \in S$ and $a_1, \dots, a_n \in F$. Then $cx = \sum_{i=1}^n (ca_i) v_i$ and $ca_i \in F$.

Therefore $\text{span}(S)$ is a subspace.

(2) Let $x \in \text{span}(S)$. Then, $x = \sum_{i=1}^n a_i v_i$, where $a_i \in F$ and $v_i \in S$. We will use induction on n .

- If $n = 1$, $x = a_1 v_1$, and $v_1 \in W$. Thus, $x \in W$.
- Suppose $x = \sum_{i=1}^n a_i v_i$. Then $x = \sum_{i=1}^{n-1} (a_i v_i) + a_n v_n$. Then $\sum_{i=1}^{n-1} (a_i v_i) \in W$ by inductive hypothesis and $a_n v_n \in W$. Thus $x \in W$.

Therefore, $\text{span}(S) \subseteq W$.

□

Note: In particular, $\text{span}(\text{span}(S)) = \text{span}(S)$.

Proof.

- (a) $\text{span}(S) \subseteq \text{span}(\text{span}(S))$
- (b) Let $x \in \text{span}(\text{span}(S))$. Then $\text{span}(S) \subseteq \text{span}(S)$ implies $\text{span}(S) \subseteq x$, which gives us $\text{span}(\text{span}(S)) \subseteq \text{span}(S)$. Therefore, $(\text{span}(\text{span}(S))) = \text{span}(S)$.

□

Definition 6 A subset S of V *generates* (spans) V if $\text{span}(S) = V$.

Example: Let $V = M_{2 \times 2}(\mathbb{R})$ and $S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\}$. Then $\text{span}(S) = V$.

1.5 Linear dependence and linear independence

Definition 7 A subset S of a vector space V is called *linearly dependent* if there exists distinct vectors $u_1, \dots, u_n \in S$ and $a_1, \dots, a_n \in F$ not all zero such that $a_1 u_1 + \dots + a_n u_n = 0$.

Notes:

- The empty set is linearly dependent.
- $S = \{u\}$ is linearly dependent if and only if $u \neq 0$.
- A set is called linearly independent if and only if the only representation of 0 as linear combinations of its vectors are trivial.

Theorem 1.6. Let $S_1 \subseteq S_2 \subseteq V$. If S_1 is linearly dependent, then so is S_2 .

Example: Let $S = \{(1, 0, 0, -1), (0, 1, 0, -1), (0, 0, 1, -1), (0, 0, 0, 1)\} \subseteq \mathbb{R}^4$. We'll show S is linearly independent. Suppose

$$a_1(1, 0, 0, -1) + a_2(0, 1, 0, -1) + a_3(0, 0, 1, -1) + a_4(0, 0, 0, 1) = (0, 0, 0, 0).$$

We get $a_1 = 0, a_2 = 0, a_3 = 0, -a_1 - a_2 - a_3 + a_4 = 0$, which implies $a_1 = a_2 = a_3 = a_4 = 0$.

Example: $P_k(x) = x^k + \cdots + x^n$ where $1 \leq k \leq n$. We'll show that $\{P_0(x), \dots, P_n(x)\}$ is linearly independent. Suppose $a_0P_0(x) + \cdots + a_nP_n(x) = 0$. The coefficient of x^i on the left hand side is $a_0 + \cdots + a_i$. Then we have

$$\begin{aligned} a_0 + \cdots + a_n &= 0 \\ a_0 + \cdots + a_{n-1} &= 0 \\ &\vdots \\ a_0 &= 0. \end{aligned}$$

Thus, $a_0 = \cdots = a_n = 0$.

Corollary 8 If $S_1 \subseteq S_2 \subseteq V$ and S_2 is linearly independent, then so is S_1 .

Theorem 1.7. Let S be a linearly independent subset of V and let $v \in V \setminus S$. Then $S \cup \{v\}$ is linearly dependent if and only if $v \in \text{span}(S)$.

Proof. Let S be linearly independent, $v \notin S$. We'll show $S \cup \{v\}$ is linearly dependent if and only if $v \in \text{span}(S)$.

(\implies): Suppose $T = S \cup \{v\}$ is linearly dependent. Then, there exists a_1, \dots, a_n not all zero and there exists $u_1, \dots, u_n \in T$ such that $a_1u_1 + \cdots + a_nu_n = 0_V$. Since S is linearly independent, $u_i = v$ for some $1 \leq i \leq n$. Say $u_1 = v$ and $a_i \neq 0$. Then $a_1u_1 = -a_2u_2 - \cdots - a_nu_n$. Thus $u_1 = -\frac{a_2}{a_1}u_2 - \cdots - \frac{a_n}{a_1}u_n$. Also $u_2, \dots, u_n \in S$. Therefore $u_1 \in \text{span}(S)$.

(\impliedby): Let $v \in \text{span}(S)$. Then there exists a_1, \dots, a_n not all zero and there exists $v_1, \dots, v_n \in S$ such that $v = a_1v_1 + \cdots + a_nv_n$. Then $v - a_1v_1 - \cdots - a_nv_n = 0_V$. The coefficient on v is $1 \neq 0$. Thus, $\{v\} \cup S$ is linearly dependent. \square

1.6 Bases and dimension

Definition 8 A *basis* β for a vector space V is a linearly independent subset of V that generates V .

Theorem 1.8. Let V be a vector space and let $\beta = \{u_1, \dots, u_n\}$. Then β is a basis of V if and only if each v can be uniquely written as

$$v = a_1u_1 + \cdots + a_nu_n$$

where $a_1, \dots, a_n \in F$.

Proof. (\Rightarrow): Suppose $\beta = \{u_1, \dots, u_n\}$ is a basis for a vector space V . Then $\text{span}(\beta) = V$ and every $v \in V$ can be written as a linear combination of u_1, \dots, u_n . Suppose $v = \sum_{i=1}^n a_i u_i$ and $v = \sum_{i=1}^n b_i u_i$. Then $0_V = \sum_{i=1}^n (a_i - b_i) u_i$. Thus $a_i = b_i = 0$ for every $i = 1, \dots, n$.

(\Leftarrow): Suppose every $v \in V$ can be written uniquely as $v = \sum_{i=1}^n a_i u_i$. Then $\text{span}(\{u_1, \dots, u_n\}) = V$. Suppose $a_1 u_1 + \dots + a_n u_n = 0_V$. Note, $0_V = u_1 \cdot 0 + \dots + u_n \cdot 0$. Since every vector has a unique representation as a linear vector, then $a_1 = \dots = a_n = 0$. Thus, $\{u_1, \dots, u_n\}$ is linearly independent and a basis for V . \square

Theorem 1.9. Let S be a finite set that generates V . Then there is a subset of S which is a basis for V .

Proof. If $S = \emptyset$ or $S = \{0_V\}$, then $V = \{0_V\}$ and \emptyset is a basis. Assume S contains a finite set of at least one nonzero vector v , which generates V . Then $\{v\}$ is linearly independent. Let $\{u_1, \dots, u_k\}$ is a maximal linearly independent subset of S . Then $S \subseteq \text{span}(\{u_1, \dots, u_k\})$. Then $\text{span}(S) \subseteq \text{span}(\text{span}(\{u_1, \dots, u_k\}))$. So, $\text{span}(S) = V \subseteq \text{span}(\{u_1, \dots, u_k\})$. Therefore $\{u_1, \dots, u_k\}$ is a basis for V . \square

Theorem 1.10. Let $G \subset V$, $|G| = n$ and $V = \text{span}(G)$. Suppose further that $L \subseteq V$, $|L| = m$ and L is linearly independent. Then $m \leq n$ and there exists a subset H of G such that $|H| = n - m$ and $L \cup H$ generates V .

Proof. (Outline)

- Induction on m . For the inductive step consider $L = \{v_1, \dots, v_{m+1}\}$
- Apply induction to $\{v_1, \dots, v_m\}$. Thus $m \leq n$ and there is a subset $H' \dots$
- H' can't be empty, say $H' = \{u_1, \dots, u_{n-m}\}$. Thus $m \leq n$.
- To find H for L , use the fact $H' \cup \{v_1, \dots, v_m\}$ generates V and substitute one of vectors from H' with v_{m+1} .

Proof. Let $|G| = n$, $\text{span}(G) = V$, $|L| = m$, L is linearly independent. Then, using induction on m ,

(Base Case): If $m = 0$, then $L = \emptyset$, $|L| = 0$ and $H = G$ and $|H| = n = n - 0$.

(Inductive Step): Suppose $L = \{v_1, \dots, v_{m+1}\}$ is linearly independent. Let $L' = \{v_1, \dots, v_m\}$. By inductive hypothesis, there exists $m \leq n$ and there exists $H' \subseteq G$, such that $|H'| = n - m$ and $\text{span}(L' \cup H') = V$. Note, $H' \neq \emptyset$. Otherwise, $\text{span}(L') = V$. In particular, $v_{m+1} \in \text{span}(\{v_1, \dots, v_m\})$, but L is linearly independent, contradiction. Therefore, $H' \neq \emptyset$ and so $n - m > 0$. Thus, $m + 1 \leq n$. Say $H' = \{u_1, \dots, u_{n-m}\}$. Since $\text{span}(L' \cup H') = V$,

$$v_{m+1} = \sum_{i=1}^m a_i v_i + \sum_{i=1}^{n-m} b_i u_i.$$

Now, at least one of scalars b_i is nonzero, say $b_1 \neq 0$. Then, $b_1 u_1 = v_{m+1} - \sum_{i=2}^m a_i v_i - \sum_{i=2}^{n-m} b_i u_i$. So,

$$u_1 = \frac{1}{b_1} v_{m+1} - \sum_{i=2}^m \frac{a_i}{b_1} v_i - \sum_{i=2}^{n-m} \frac{b_i}{b_1} u_i,$$

which implies $u_1 \in \text{span}(\{v_1, \dots, v_{m+1}\} \cup \{u_2, \dots, u_{n-m}\})$. Thus, $u_1 \in \text{span}(L \cup \{u_2, \dots, u_{n-m}\})$. Let $H = \{u_2, \dots, u_{n-m}\}$. Then $|H| = n - m - 1 = n - (m + 1)$ and $H \subseteq G$. We have $u_i \in \text{span}(L \cup H)$ for $i = 1, \dots, n - m$ and $L \subseteq \text{span}(L \cup H)$. Thus, $V \subseteq \text{span}(L \cup \{u_1, \dots, u_{n-m}\}) \subseteq \text{span}(L \cup H)$. \square

Corollary 13 Suppose V has finite basis. Then every basis for V has the same cardinality.

Proof. Let β and γ be bases for V .

- β is linearly independent, $\text{span}(\gamma) = V$, $|\beta| \leq |\gamma|$ by Theorem 1.10.
- So, γ is linearly independent and $\text{span}(\beta) = V$. Again, by Theorem 1.10, $|\gamma| \leq |\beta|$.

\square

Definition 9 A vector space is called *finite-dimensional* if it has a finite basis. The number of vectors in a basis, is called the *dimension* of V , notated $\dim(V)$. A vector space which is not finite-dimensional is called *infinite-dimensional*.

Corollary 14 Suppose $\dim(V) = n$.

- (a) If S is finite and $\text{span}(S) = V$, then $n \leq |S|$. If $|S| = n$, then S is a basis.

Proof. Let S be finite and $\text{span}(S) = V$. Let β be a basis for V . Then, β is linearly independent and $|\beta| = n$. By Theorem 1.10, $|S| \geq |\beta| = n$. Suppose $|S| = n$. Then S contains a subset T such that T is linearly independent and $\text{span}(T) = V$. Consequently, T is a basis for V . Thus, $|T| = \dim(V) = n$, $T \subseteq S$, $|T| = n$. Therefore, $S = T$, which implies S is a basis. \square

- (b) If $|S| = n$ and S is linearly independent, then S is a basis.

Proof. Suppose L is linearly independent and $|L| = \dim(V)$. Let β be a basis for V . Then $|L| \leq |\beta|$, in addition, there exists $H \subseteq \beta$ such that $\text{span}(L \cup H) = V$ and $|H| = |\beta| - |L| = \dim(V) - \dim(V) = 0$. Thus, $H = \emptyset$. Thus $\text{span}(L) = V$. \square

- (c) Every linearly independent set can be extended to a basis.

Proof. Let L be linearly independent. Suppose $|L| < n$. Let β be a basis. Then there exists $H \subseteq \beta$ such that $\text{span}(L \cup H) = V$ and $|H| = n - |L|$. So, $|L \cup H| \leq n$. As before, $L \cup H$ contains an independent subset T such that $\text{span}(T) = V$. Then T is a basis and so $|T| = n$. Therefore, $T = L \cup H$. Thus $L \cup H$ is a basis. \square

Example:

- (1) \mathbb{F}^n , $\dim(\mathbb{F}^n) = n$.
- (2) $V = M_{m \times n}(\mathbb{F})$, $\dim(V) = mn$.
- (3) $V = \{A \in M_{n \times n} : A \text{ is symmetric}\}$, $\dim(V) = \frac{n(n+1)}{2}$.

Theorem 1.11. Let W be a subspace of a finite-dimensional space V . Then W is finite-dimensional and $\dim(W) \leq \dim(V)$. Moreover, if $\dim(W) = \dim(V)$, then $W = V$.

1.7 Maximal linearly independent subsets

Definition 10 A collection \mathcal{C} of sets is called a *chain* if for every $A, B \in \mathcal{C}$, $A \subseteq B$ or $B \subseteq A$.

Maximal Principle: Let \mathcal{F} be a family of sets. If for every chain \mathcal{C} in \mathcal{F} there is a set in \mathcal{F} which contains all members of \mathcal{C} , then \mathcal{F} contains a maximal element.

Theorem 1.12. Every vector space has a basis.

Proof. (Outline)

- Start with an arbitrary (finite) linearly independent set S in V and consider the family \mathcal{F} of all independent sets contains S . Argue that Maximal Principle applies and take the element β in \mathcal{F} .
- β generates V .

Cauchy's functional equation:

$$f(x + y) = f(x) + f(y)$$

What type of functions can f be?

- (1) $f : \mathbb{Q} \rightarrow \mathbb{Q}$ where $f(x) = \alpha \cdot x$ and $\alpha = f(1)$.
- (2) $f : \mathbb{R} \rightarrow \mathbb{R}$ allows for other fairly exotic functions.

For example, \mathbb{R} is a vector space over \mathbb{Q} . Let H be a basis of this vector space, commonly called Hamel basis. Then for every element $x \in \mathbb{R}$ there exists unique $h_1, \dots, h_n \in H$ and unique scalars $c_1, \dots, c_n \in \mathbb{Q}$ with $c_1, \dots, c_n \neq 0$ such that $x = \sum_{i=1}^n c_i h_i$. Then, for any $g : H \rightarrow \mathbb{R}$, we can extend g to $\bar{g} : \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$\bar{g}(x) = \sum_{i=1}^n c_i g(h_i).$$

So,

$$x + y = \sum_{i=1}^n d_i h_i + \sum_{i=1}^n a_i h_i = \sum_{i=1}^n c_i h_i.$$

2 Linear Transformations and Matrices

2.1 Linear transformations, null spaces, and ranges

Definition 1 A function $T : V \rightarrow W$ is called a *linear transformation* from V to W if for all $x, y \in V$ and $c \in F$ the following hold.

- $T(x + y) = T(x) + T(y)$
- $T(cx) = cT(x)$

Observations:

- T is linear if and only if $x_1, \dots, x_n \in V, a_1, \dots, a_n \in F$

$$T\left(\sum_{i=1}^n a_i x_i\right) = \sum_{i=1}^n a_i T(x_i).$$

- T is linear if and only if $T(cx + y) = cT(x) + T(y)$ for $x, y \in V, c \in F$.

Example:

- (1) Define $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ to be $T(a_1, a_2) = (2a_1 + a_2, a_1)$. Then

$$\begin{aligned} T((a_1, a_2) + (b_1, b_2)) &= T(a_1 + b_1, a_2 + b_2) \\ &= (2(a_1 + b_1) + a_2 + b_2, a_1 + b_1) \\ &= (2a_1 + a_2, a_1) + (2b_1 + b_2, b_1) \\ &= T(a_1, a_2) + T(b_1, b_2). \end{aligned}$$

It's also easy to show,

$$T(c(a_1, a_2)) = cT(a_1, a_2).$$

- (2) Let $T_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ where $T_\theta(a_1, a_2)$ is the vector obtained by (a_1, a_2) by the angle θ .

- (3) Let $V = \mathcal{C}(\mathbb{R})$, and $a, b \in \mathbb{R}$, with $a < b$. Let $T : V \rightarrow \mathbb{R}$, where

$$T(f) = \int_a^b f(t) dt.$$

Then, $T(f + g) = T(f) + T(g)$. Also, $T(cf) = cT(f)$.

Definition 2 Let $T : V \rightarrow W$ be linear.

- The *null space* (kernel) $N(T) = \{x \in V : T(x) = 0\}$

- The *range* (image) $R(T) = \{T(x) : x \in V\}$

Theorem 2.1. Let $T : V \rightarrow W$ be linear. Then $N(T)$ and $R(T)$ are subspaces of V and W .

Proof.

- (1) We have $0_V \in N(T)$ because $T(0_V) = 0_W$.
- (2) Suppose $x, y \in N(T)$. Then $T(x) = 0_W$, $T(y) = 0_W$. Thus, $T(x + y) = T(x) + T(y) = 0_W + 0_W = 0_W$. Thus $x + y \in N(T)$.
- (3) Suppose $x \in N(T)$ and $c \in F$. then $T(x) = 0_W$ and so $T(cx) = cT(x) = 0_W$. Thus, $cx \in N(T)$.

Therefore $N(T)$ is a subspace of V . $R(T)$ can be shown to be a subspace of W in a similar manner. \square

Theorem 2.2. Let $T : V \rightarrow W$ be linear and let $\beta = \{v_1, \dots, v_n\}$ be a basis for V . Then $R(T) = \text{span}(T(\beta)) = \text{span}(\{T(v_1), \dots, T(v_n)\})$.

Proof.

- (1) Let $w \in R(T)$. Then $w = T(v)$ for some $v \in V$. Then $v = \sum_{i=1}^n c_i v_i$ for some $c_1, \dots, c_n \in F$. Thus, $T(v) = \sum_{i=1}^n c_i T(v_i)$. Therefore, $w = T(v) \in \text{span}(T(\beta))$.
- (2) Let $w \in \text{span}(T(\beta))$. Then $w = \sum_{i=1}^n c_i T(v_i)$ for some $c_1, \dots, c_n \in F$. Thus $w = T(\sum_{i=1}^n c_i v_i)$ and $\sum_{i=1}^n c_i v_i \in V$. Thus $w \in R(T)$.

\square

- $\text{nullity}(T) = \dim(N(T))$
- $\text{rank}(T) = \dim(R(T))$

Theorem 2.3. (Dimension Theorem) Let $T : V \rightarrow W$ be linear. If V is finite-dimensional, then

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

Proof. (Outline)

- Start with a basis for $N(T)$, $\{v_1, \dots, v_k\}$ and extend it to a basis of V , $\{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$.
- Prove that $\{T(v_{k+1}), \dots, T(v_n)\}$ is a basis for $R(T)$.

Proof. Let $\{v_1, \dots, v_k\}$ be a basis for $N(T)$. Extend this basis to a basis β for V , say $\beta = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$. Note, $k = \dim(N(T))$ and $n = \dim(V)$. We claim that $\{T(v_{k+1}), \dots, T(v_n)\}$ is a basis for $R(T)$.

- (1) Clearly, $\text{span}(\{T(v_{k+1}), \dots, T(v_n)\}) \subseteq R(T)$. Let $w \in R(T)$. Then $w = T(v)$ for some $v \in V$ and $v = \sum_{i=1}^n c_i v_i$ for some $c_1, \dots, c_n \in F$. Then $w = T(v) = \sum_{i=1}^n c_i T(v_i) = \sum_{i=k+1}^n c_i T(v_i)$ since $T(v_i) = 0$ for $i \leq k$. Therefore, $w \in \text{span}(\{T(v_{k+1}), \dots, T(v_n)\})$. So, $\text{span}(\{T(v_{k+1}), \dots, T(v_n)\}) = R(T)$.
- (2) We'll show $\{T(v_{k+1}), \dots, T(v_n)\}$ is linearly independent. Suppose $\sum_{i=k+1}^n c_i T(v_i) = 0$. Then, $T(\sum_{i=k+1}^n c_i v_i) = 0$. Thus, $\sum_{i=k+1}^n c_i v_i \in N(T)$. So, $\sum_{i=k+1}^n c_i v_i = \sum_{i=1}^k d_i v_i$ for some $d_1, \dots, d_k \in F$. Therefore, $\sum_{i=k+1}^n c_i v_i - \sum_{i=1}^k d_i v_i = 0$. Since β is linearly independent, $c_{k+1} = \dots = c_n = 0$. In addition, $d_1 = \dots = d_k = 0$. Thus, $\{T(v_{k+1}), \dots, T(v_n)\}$ is linearly independent. □

Theorem 2.4. Let $T : V \rightarrow W$ be linear. Then T is injective if and only if $N(T) = \{0\}$.

Proof.

(\implies): Suppose $T : V \rightarrow W$ is injective. If $T(x) = 0$ then since $T(0) = 0$, we have $T(x) = T(0)$ and so $x = 0$.

(\impliedby): Suppose $N(T) = \{0\}$. If $T(x) = T(y)$, then $T(x) - T(y) = 0$. Thus, $T(x - y) = 0$. Thus, $x - y \in N(T)$, which implies $x - y = 0$ and thus $x = y$. □

Theorem 2.5. Let V, W be finite-dimensional vector spaces such that $\dim(V) = \dim(W)$ and let $T : V \rightarrow W$ be linear. Then the following are equivalent:

- (a) T is injective;
- (b) T is surjective;
- (c) $\text{rank}(T) = \dim(V) = \dim(W)$.

Proof. Note, T being injective is equivalent to $N(T) = \{0\}$, which is equivalent to $\text{nullity}(T) = 0$. By the dimension theorem, $\text{rank}(T) = \dim(V)$. So, $\text{rank}(T) = \dim(W)$, which is equivalent to $\dim(R(T)) = \dim(W)$. This is equivalent to $R(T) = W$ since $R(T) \subseteq W$. Thus, T is surjective. □

Theorem 2.6. Suppose $\{v_1, \dots, v_n\}$ is a basis for V . For $w_1, \dots, w_n \in W$ there exists exactly one linear transformation $T : V \rightarrow W$ such that $T(v_i) = w_i$ for every i .

Proof. (Outline)

- For $x \in V$ we can write $x = \sum_{i=1}^n a_i v_i$ uniquely.
- Let $T : V \rightarrow W$ be $T(x) = \sum_{i=1}^n a_i w_i$.

Proof. Let $\{v_1, \dots, v_n\}$ be a basis for V and $w_1, \dots, w_n \in W$. Then there exists unique $T : V \rightarrow W$ defined by $T(v_i) = w_i$.

- (1) Let $x \in V$. Then x can be uniquely written as $x = \sum_{i=1}^n a_i v_i$ where $a_i \in \mathbb{F}$. Let $T(x) := \sum_{i=1}^n a_i w_i$. $T : V \rightarrow W$ is well-defined, and thus a function.
- (2) Let T be linear and $u, v \in V$. Let $c \in \mathbb{F}$. We will show $T(cu + v) = cT(u) + T(v)$. Then, $u = \sum_{i=1}^n a_i v_i$ and $v = \sum_{i=1}^n b_i v_i$ for unique $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{F}$. Then $cu + v = \sum_{i=1}^n (ca_i + b_i) v_i$. So,

$$\begin{aligned} T(cu + v) &= \sum_{i=1}^n (ca_i + b_i) w_i \\ &= c \sum_{i=1}^n a_i w_i + \sum_{i=1}^n b_i w_i \\ &= cT(u) + T(v). \end{aligned}$$

- (3) Suppose $U : V \rightarrow W$ is linear and $U(v_i) = w_i$. Let $x \in V$. Then $x = \sum_{i=1}^n a_i v_i$. Then

$$\begin{aligned} U(x) &= U\left(\sum_{i=1}^n a_i v_i\right) \\ &= \sum_{i=1}^n a_i U(v_i) \\ &= \sum_{i=1}^n a_i w_i \\ &= T(x). \end{aligned}$$

- (4) Note, $T(v_i) = w_i$ because $v_i = 1 \cdot v_i$. Thus, $T(v_i) = 1 \cdot w_i = w_i$.

□

Corollary 7 Let $\{v_1, \dots, v_n\}$ be a basis of V . If $U, T : V \rightarrow W$ are linear and $T(v_i) = U(v_i)$, then $T = U$.

2.2 The matrix representation of a linear transformation

Definition 3 Let V be a finite-dimensional vector space. An *ordered basis* for V is a basis equipped with an ordering.

Example: Let $\beta = \left\{ e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$. Let $x = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} = 5e_1 + 6e_2 + 7e_3$.

Then $[x]_\beta = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix}$.

Definition 4 Let $\beta = \{u_1, \dots, u_n\}$ be an ordered basis for V . Then for every $x \in V$, $x = \sum a_i u_i$ and a_1, \dots, a_n are unique. The coordinate vector of x relative to β , $[x]_\beta = (a_1, \dots, a_n)^T$.

Let $T : V \rightarrow W$, $\beta = \{v_1, \dots, v_n\}$, $\gamma = \{w_1, \dots, w_m\}$ be ordered bases for V and W . We have

$$T(v_j) = \sum_{i=1}^m a_{ij} w_i.$$

Then $A = [a_{ij}]$ is called the matrix representation of T in the ordered bases β and γ and we write

$$A = [T]_\beta^\gamma.$$

If $V = W$ and $\beta = \gamma$, we say $A = [T]_\beta$.

Observations:

- The j th column of A is $[T(v_j)]_\gamma$.
- If $T, U : V \rightarrow W$ and $[U]_\beta^\gamma = [T]_\beta^\gamma$, then $T = U$.

Example: Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $\beta = \{(1, 0), (0, 1)\}$ and $\gamma = \{(1, 1), (1, -1)\}$. Let $T((1, 0)) = (1, 0)$ and $T((0, 1)) = (1, 1)$. Find $[T]_\beta^\gamma$.

We can see $T((1, 0)) = (1, 0) = \frac{1}{2}(1, 1) + \frac{1}{2}(1, -1)$. Also, $T((0, 1)) = (1, 1) = 1(1, 1) + 0(1, -1)$. Therefore, $[T]_\beta^\gamma = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{pmatrix}$.

Theorem 2.7. Let $T, U : V \rightarrow W$ are linear and let $a \in F$. Then,

- $aT + U$ is linear.

- The set of all linear transformations from V to W (with addition of functions and scalar multiplication) forms a vector space over F .

We notate $\mathcal{L}(V, W)$ as the vector space of all linear transformations from V to W . Also, $\mathcal{L}(V) = \mathcal{L}(V, V)$.

Theorem 2.8. Let $T, U : V \rightarrow W$ be linear. Then,

- $[T + U]_{\beta}^{\gamma} = [T]_{\beta}^{\gamma} + [U]_{\beta}^{\gamma}$
- $[T]_{\beta}^{\gamma} = a [T]_{\beta}^{\gamma}$.

2.3 Composition of linear transformations and matrix multiplication

Theorem 2.9. Let $T : V \rightarrow W$, $U : W \rightarrow Z$ be linear. Then $UT : V \rightarrow Z$ is linear.

Proof. Let $T : V \rightarrow W$, $U : W \rightarrow Z$ be linear. Then,

$$\begin{aligned} UT(au + v) &= U(T(au + v)) \\ &= U(aT(u) + T(v)) \\ &= aU(T(u)) + U(T(v)) \\ &= aUT(u) + UT(v). \end{aligned}$$

□

Theorem 2.10. Let $T, U_1, U_2 \in \mathcal{L}(V)$. Then

- $T(U_1 + U_2) = TU_1 + TU_2$ and $(U_1 + U_2)T = U_1T + U_2T$
- $T(U_1U_2) = (TU_1)U_2$
- $TI = IT = T$
- $a(U_1U_2) = (aU_1)U_2 = U_1(aU_2)$

Definition 5 Let A be an $m \times n$ matrix, B be an $n \times p$ matrix. We define AB to be the $m \times p$ matrix such that

$$(AB)_{ij} = \sum_{k=1}^n A_{ik}B_{kj}.$$

Note: $(AB)^t = B^tA^t$.

Theorem 2.11. Let $T : V \rightarrow W$. $U : W \rightarrow Z$ be linear and let α, β, γ be ordered bases in V, W, Z . Then

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

Proof. (Outline)

- $\alpha = \{v_1, \dots, v_m\}$, $\beta = \{w_1, \dots, w_n\}$, $\gamma = \{z_1, \dots, z_p\}$.
- $A = [U]_{\beta}^{\gamma}$, $B = [T]_{\alpha}^{\beta}$
- $U(T(v_j)) = U(\sum_{k=1}^n B_{kj} w_k) = \sum_{k=1}^n B_{kj} U(w_k) = \sum_{k=1}^n B_{kj} \sum_{i=1}^p A_{ik} z_i = \sum_{i=1}^p C_{ij} z_i$
- Thus, by definition, the i, j th entry in $[UT]_{\alpha}^{\gamma}$ is $C_{ij} = \sum_{k=1}^n A_{ik} B_{kj}$.

Proof. Let $\alpha = \{v_1, \dots, v_m\}$, $\beta = \{w_1, \dots, w_n\}$ and $\gamma = \{z_1, \dots, z_p\}$. Let $A = [U]_{\beta}^{\gamma}$ and $B = [T]_{\alpha}^{\beta}$. Consider $[UT]_{\alpha}^{\gamma}$. Then

$$\begin{aligned} (UT)(v_j) &= U(T(v_j)) \\ &= U\left(\sum_{i=1}^m B_{ij} w_i\right) \\ &= \sum_{i=1}^m B_{ij} U(w_i) \\ &= \sum_{i=1}^m B_{ij} \cdot \sum_{k=1}^p A_{ki} z_k \\ &= \sum_{k=1}^p \left(\sum_{i=1}^m A_{ki} B_{ij}\right) z_k. \end{aligned}$$

If $C = [UT]_{\alpha}^{\gamma}$, then

$$C_{kj} = \sum_{i=1}^m A_{ki} B_{ij}.$$

□

Example: Let $\alpha = \{1, x, x^2\}$ and $\beta = \{1, x, x^2, x^3\}$ be standard bases for $P_2(\mathbb{R})$ and $P_3(\mathbb{R})$. Let $U : P_3(\mathbb{R}) \rightarrow P_2(\mathbb{R})$ and $T : P_2(\mathbb{R}) \rightarrow P_3(\mathbb{R})$ defined by $U(f(x)) = f'(x)$ and $T(f(x)) = \int_0^x f(t) dt$. From calculus, $UT = I$. We can see

$$\begin{aligned} U(1) &= 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ U(x) &= 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ U(x^2) &= 2x = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 \\ U(x^3) &= 3x^2 = 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2. \end{aligned}$$

Also,

$$\begin{aligned} T(1) &= x = 0 \cdot 1 + 1 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 \\ T(x) &= \frac{1}{2}x^2 = 0 \cdot 1 + 0 \cdot x + \frac{1}{2}x^2 + 0 \cdot x^3 \\ T(x^2) &= \frac{1}{3}x^3 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + \frac{1}{3}x^3. \end{aligned}$$

Thus,

$$\begin{aligned} [UT]_{\alpha}^{\alpha} &= [U]_{\beta}^{\alpha} [T]_{\alpha}^{\beta} \\ &= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix} \\ &= I. \end{aligned}$$

Theorem 2.12. Let A be an $m \times n$ matrix, B and C be $n \times p$ matrices, D, E be $q \times m$ matrices. Then

- $A(B + C) = AB + AC$, $(D + E)A = DA + EA$
- $a(AB) = (aA)B = A(aB)$
- $I_m A = A = A I_n$
- If V is n -dimensional with ordered basis β , then $[I_V]_{\beta} = I_n$

Theorem 2.13. Let A be an $m \times n$ matrix, B be an $n \times p$ matrix. Let u_j, v_j denote the j th columns of AB and B . Then

- (a) $u_j = Av_j$
- (b) $v_j = Be_j$

Theorem 2.14. Let V, W be finite-dimensional vector spaces with ordered bases β and γ and let $T : V \rightarrow W$ be linear. Then for $u \in V$

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

Proof. Let V, W be finite-dimensional vector spaces with ordered bases β and γ and let $T : V \rightarrow W$ be linear. Fix $u \in V$. Let $f : \mathbb{R} \rightarrow V$ and $g : \mathbb{R} \rightarrow W$ where $f(a) = au$

and $g(a) = aT(u)$. Let $\alpha = \{1\}$ be the standard basis for \mathbb{R} . Note that, g and f are linear transformations. Then,

$$\begin{aligned}
 [T(u)]_\gamma &= [g(1)]_\gamma \\
 &= [g]_\alpha^\gamma \\
 &= [Tf]_\alpha^\gamma \\
 &= [T]_\beta^\gamma [f]_\alpha^\beta \\
 &= [T]_\beta^\gamma [f(1)]_\beta \\
 &= [T]_\beta^\gamma [u]_\beta.
 \end{aligned}$$

□

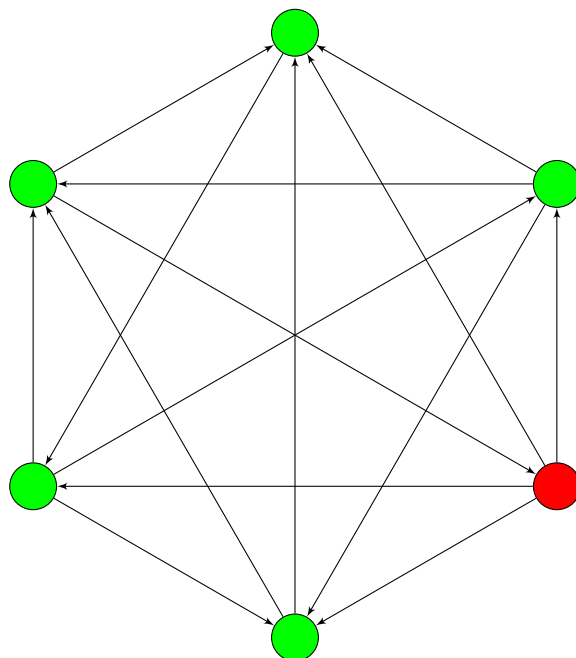
Let $L_A : F^n \rightarrow F^m$ be given by $L_A(x) = Ax$ where A is an $m \times n$ matrix.

Theorem 2.15. Let A be an $m \times n$ matrix. Then L_A is linear. Moreover, if $B \in M_{m \times n}(F)$ and β and γ are the standard ordered bases of F^n and F^m , then

- (a) $[L_A]_\beta^\gamma = A$
- (b) $L_A = L_B$ if and only if $A = B$
- (c) $L_{A+B} = L_A + L_B$, $L_{aA} = aL_A$
- (d) If $T : F^n \rightarrow F^m$ is linear, then there exists unique matrix C (namely $C = [UT]_\beta^\gamma$) such that $T = L_C$.
- (e) If E is an $n \times p$ matrix, then $L_{AE} = L_A L_E$
- (f) If $m = n$, $L_{I_n} = I_{F^n}$.

Theorem 2.16. Let A, B, C be such that $A(BC)$ is defined then $(AB)C = A(BC)$.

Tournaments: There are n players $\{1, \dots, n\}$ for every $i \neq j$ there is exactly one game between i and j which results in i winning or j winning. Let A be the incidence matrix of a tournament where we put $A_{ij} = 1$ if i wins with j and 0 otherwise. Show that $A^2 + A$ contains a column such that each entry but the diagonal is at least one.



2.4 Invertibility and isomorphisms

Definition 6 Let $T : V \rightarrow W$ be linear. A function $U : W \rightarrow V$ is called an *inverse* of T if $UT = I_V$, $TU = I_W$. If T has an inverse, then it's called *invertible*.

Note:

- If T is invertible, then the inverse is unique, denoted by T^{-1} .
- $(UT)^{-1} = T^{-1}U^{-1}$
- $(T^{-1})^{-1} = T$
- T is invertible if and only if T is a bijection.
- $T : V \rightarrow W$, V , W are finite-dimensional of equal dimensions.
- T is invertible if and only if $\text{rank}(T) = \dim(V)$.

Theorem 2.17. Let $T : V \rightarrow W$ be linear and invertible. Then $T^{-1} : W \rightarrow V$ is linear.

Definition 7 Let A be an $n \times m$ matrix. Then A is *invertible* if there exists an $n \times n$ matrix B such that $AB = BA = I$. Then B is called the inverse of A denoted by A^{-1} .

Lemma 19. Let T be an invertible transformation from V to W . Then V is finite-dimensional if and only if W is. In this case $\dim(V) = \dim(W)$.

Theorem 2.18. Let V, W be finite-dimensional with ordered bases β and γ . Let $T : V \rightarrow W$ be linear. Then T is invertible if and only if $[T]_{\beta}^{\gamma}$ is invertible. Furthermore $[T^{-1}]_{\gamma}^{\beta} = ([T]_{\beta}^{\gamma})^{-1}$.

Proof.

- If T is invertible, then by the lemma $\dim(V) = \dim(W)$ and the fact that $[T]_{\beta}^{\gamma}$ follows from previous facts and $T^{-1}T = I_V$.
- If $[T]_{\beta}^{\gamma}$ is invertible then it has an inverse B and so there is a transformation $U : W \rightarrow V$ with $B = [U]_{\gamma}^{\beta}$. Now check that $UT = I_V$.

□

Definition 8 Vector space V is isomorphic to W if there is an invertible linear transformation $T : V \rightarrow W$.

Theorem 2.19. Let V, W be finite-dimensional over F . Then V is isomorphic to W if and only if $\dim(V) = \dim(W)$.

Theorem 2.20. Let V, W be finite-dimensional over F of dimensions n and m . Let β, γ be ordered bases for V and W . Then $\phi : \mathcal{L}(V, W) \rightarrow M_{m \times n}(F)$ given by $\phi(T) = [T]_{\beta}^{\gamma}$ for $T \in \mathcal{L}(V, W)$ is an isomorphism.

Definition 9 Let β be an ordered basis for an n -dimensional vector space V over F . The *standard representation of V with respect to β* is the function $\phi_{\beta} : V \rightarrow F^n$ defined as $\phi_{\beta}(x) = [x]_{\beta}$.

Theorem 2.21. Let V be a finite-dimensional vector space with ordered basis β . Then ϕ_{β} is an isomorphism.

2.5 The change of coordinate matrix

Theorem 2.22. Let β, β' be two ordered bases for a finite-dimensional vector space V and let $Q = [I_V]_{\beta'}^{\beta}$. Then

- Q is invertible.
- For any $v \in V$, $[v]_{\beta} = Q[v]_{\beta'}$

Q is called the change of coordinate matrix. A linear operator on V is the linear transformation from V to V .

Theorem 2.23. Let T be a linear operator on a finite-dimensional vector space V . Let β, β' be ordered bases for V . Suppose $Q = [I_V]_{\beta'}^{\beta}$. Then

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

Corollary 26 Let $A \in M_{n \times n}$ and let $\gamma = \{u_1, \dots, u_n\}$ be an ordered basis for F^n . Then $[L_A]_{\gamma} = Q^{-1} A Q$ where Q is the matrix with the j th column equal to u_j .

2.6 Dual spaces

- Linear functional on V – linear transformation from V to F .
- $V^* = \mathcal{L}(V, F)$
- $V^{**} = (V^*)^*$

$$\dim(V^*) = \dim(V)$$

Let $\beta = \{x_1, \dots, x_n\}$ for $v \in V$ let $[v]_{\beta} = (a_1, \dots, a_n)^T$. Define

$$f_i(v) = a_i.$$

Theorem 2.24. Let V be a finite-dimensional vector space with ordered basis $\beta = \{x_1, \dots, x_n\}$ and let $\beta^* = \{f_1, \dots, f_n\}$. Then β^* is an ordered basis for V^* and for $f \in V^*$ where

$$f = \sum f(x_i) f_i.$$

Proof.

- Enough to show $f = \sum f(x_i) f_i$.
- Let $F := \sum f(x_i) f_i$. Then $F(x_j) = f(x_j)$ for every j .

□

Theorem 2.25. Let V, W be finite-dimensional vector spaces over F with ordered bases β and γ . Let $T : V \rightarrow W$ be linear. Then $T^t : W^* \rightarrow V^*$ given by $T^t(g) = gT$ is linear and

$$[T^t]_{\gamma^*}^{\beta^*} = ([T]_{\beta}^{\gamma})^t.$$

Proof.

- It's easy to see that T^t is a linear transformation from W^* to V^* .

- Let $\beta = \{x_1, \dots, x_n\}$, $\gamma = \{y_1, \dots, y_m\}$, $\beta^* = \{f_1, \dots, f_n\}$, $\gamma^* = \{g_1, \dots, g_m\}$, $A = [T]_{\beta}^{\gamma}$.
- The j th column of $[T^t]_{\gamma^*}^{\beta^*}$ is $T^t(g_j)$ which is

$$\sum_{k=1}^n (g_j Y)(x_k) f_k.$$

- Thus the i, j -th entry is $(T^t(g_j))(x_i)$ which is A_{ji} .

□

For $x \in V$ let $\hat{x} : V^* \rightarrow F$ given by $\hat{x}(f) = f(x)$.

Theorem 2.26. Let V be finite-dimensional and let $\psi : V \rightarrow V^{**}$ be given by $\psi(x) = \hat{\hat{x}}$. Then ψ is an isomorphism.

3 Elementary matrix operations and systems of linear equations

3.1 Elementary matrix operations and elementary matrices

Definition 1 Let A be a matrix. Elementary row operations:

- Interchange any two rows of A
- Add a scalar multiple of a row of A to another row
- Multiply any row of A by a non-zero scalar

Note: The same can be done for columns.

Definition 2 An $n \times n$ elementary matrix is a matrix obtained from I_n by an elementary operation. Its type is the type of the operation performed.

Theorem 3.1. Let $A \in M_{m \times n}(F)$ and suppose B is obtained by performing an elementary row (column) operation. Then there exists an $m \times m$ ($n \times n$) elementary matrix E such that $B = EA$ ($B = AE$).

Theorem 3.2. Elementary matrices are invertible and the inverse of an elementary matrix is an elementary matrix of the same type.

3.2 The rank of a matrix and matrix inverse

Definition 3 Let $A \in M_{m \times n}(F)$. The $\text{rank}(A)$ is the rank of $L_A : F^n \rightarrow F^m$.

Theorem 3.3. Let $T : V \rightarrow W$ be linear and let β, γ be ordered bases for V and W . Then

$$\text{rank}(T) = \text{rank}([T]_{\beta}^{\gamma}).$$

Theorem 3.4. Let A be an $m \times n$ matrix. If P and Q are invertible $m \times m$ and $n \times n$ matrices, then

- $\text{rank}(AQ) = \text{rank}(A)$
- $\text{rank}(PA) = \text{rank}(A)$

and so $\text{rank}(PAQ) = \text{rank}(A)$.

Proof.

- $R(L_{AQ}) = R(L_A)$ because $L_Q(F^n) = F^n$.
- $\dim(L_P(L_A(F^n))) = \dim(L_A(F^n))$ because $L_P : F^m \rightarrow F^m$ is an isomorphism.

□

Theorem 3.5. The rank of a matrix equals the maximum number of its linearly independent columns.

Proof.

- $R(L_A) = \text{span}(L_A(\{e_1, \dots, e_n\}))$ and $L_A(e_j)$ is the j th column of A .

□

Theorem 3.6. Let A be an $m \times n$ matrix of rank r . Then $r \leq \min\{m, n\}$ and A can be transformed to

$$D = \begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix}$$

using a finite number of elementary row and column operations.

Proof. Induction on m . In the inductive step use row and column operations to reduce to

$$\left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ 0 & & & \\ \hline & & B' & \\ 0 & & & \end{array} \right).$$

□

Corollary 7 Let A be an $m \times n$ matrix of rank r . Then there exist invertible matrices B and C of sizes $m \times m$ and $n \times n$ such that $D = BAC$.

Corollary 8 Let A be an $m \times n$ matrix. Then

- $\text{rank}(A^t) = \text{rank}(A)$
- $\text{rank}(A)$ is equal to the dimension of the row space of A
- dimension of the row space is equal to the dimension of the column space.

Corollary 9 Every invertible matrix is a product of elementary matrices.

Theorem 3.7. Let $T : V \rightarrow W$ and $U : W \rightarrow Z$ be linear transformations on finite-dimensional vector spaces. Let A, B be matrices such that AB is defined. Then

- (a) $\text{rank}(UT) \leq \text{rank}(U)$
- (b) $\text{rank}(UT) \leq \text{rank}(T)$
- (c) $\text{rank}(AB) \leq \text{rank}(A)$
- (d) $\text{rank}(AB) \leq \text{rank}(B)$

Proof.

- For (a), $R(UT) = U(R(T)) \subseteq U(W) = R(U)$
- (c) and (d) follow from (a) and discussion of the transpose
- (b) follows from the previous by considering matrix representations.

□

Observations: A is an invertible $n \times n$ matrix if and only if $(A|I_n)$ can be transformed into $(I_n|B)$ by elementary row operations, in this case $B = A^{-1}$.

3.3 Systems of linear equations (theoretical aspect)

- $Ax = b$ is consistent if it has at least one solution and inconsistent otherwise.
- $Ax = 0$ is called a homogeneous system and $Ax = b$ for $b \neq 0$ a nonhomogeneous system.

Theorem 3.8. Let $Ax = 0$ be a homogeneous system over F in n unknowns and let K be the set of solutions. Then $K = N(L_A)$ and so K is a subspace of F^n and $\dim(K) = n - \text{rank}(A)$.

Corollary 12 If $m < n$, then $Ax = 0$ has a non-zero solution.

Theorem 3.9. Let K be the solution set to $Ax = b$ and K_H be the solution set to $Ax = 0$. Then for some $s \in K$, $K = \{s\} + K_H$.

Theorem 3.10. Let $Ax = b$ be a system of n linear equations in n unknowns. Then A is invertible if and only if the system has exactly one solution.

Theorem 3.11. The system $Ax = b$ is consistent if and only if $\text{rank}(A) = \text{rank}(A|b)$.

4 Determinants

4.1 Determinants of order 2

Definition 1 If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then the determinant of A is $ad - bc$.

Notation: The determinant of A will be denoted by $\det(A)$ or $|A|$.

Theorem 4.1. For $u, v, w \in F^2$ and $k \in F$

$$\begin{aligned} \det \begin{pmatrix} u + kv \\ w \end{pmatrix} &= \det \begin{pmatrix} u \\ w \end{pmatrix} + k \det \begin{pmatrix} v \\ w \end{pmatrix} \\ \det \begin{pmatrix} w \\ u + kv \end{pmatrix} &= \det \begin{pmatrix} w \\ u \end{pmatrix} + k \det \begin{pmatrix} w \\ v \end{pmatrix}. \end{aligned}$$

Theorem 4.2. Let $A \in M_{2 \times 2}(F)$. Then the determinant of A is non-zero if and only if A is invertible. If A is invertible, then $A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}$.

4.2 Determinants of order n

Let \tilde{A}_{ij} be the matrix obtained from A by deleting the i th row and the j th column.

Definition 2 Let $A \in M_{n \times n}(F)$.

- If $n = 1$, then $\det(A) = A_{11}$.
- If $n \geq 2$, then $\det(A) = \sum_{j=1}^n (-1)^{1+j} A_{1j} \det(\tilde{A}_{1j})$.

The cofactor of the i, j th entry of A ,

$$c_{ij} = (-1)^{i+j} \det(\tilde{A}_{ij}).$$

Theorem 4.3. Let $a_1, \dots, a_n \in F^n$, let $k \in F$ and suppose $a_r = u + kv$ for some $u, v \in F^n$. Then

$$\begin{vmatrix} a_1 \\ a_{r-1} \\ a_r \\ a_{r+1} \\ a_n \end{vmatrix} = \begin{vmatrix} a_1 \\ a_{r-1} \\ u \\ a_{r+1} \\ a_n \end{vmatrix} + k \begin{vmatrix} a_1 \\ a_{r-1} \\ v \\ a_{r+1} \\ a_n \end{vmatrix}.$$

Corollary 4 If A has a row consisting of zeroes, then $\det(A) = 0$.

Lemma 5. Let $B \in M_{n \times n}(F)$ and $n \geq 2$. Suppose that the i th row of B is e_k for some $1 \leq k \leq n$. Then $\det(B) = (-1)^{i+k} \det(\tilde{B}_{ik})$.

Theorem 4.4. For $A \in M_{n \times n}(F)$ and $i \in \{1, \dots, n\}$

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} A_{ij} \det(\tilde{A}_{ij}).$$

Corollary 7 If $A \in M_{n \times n}(F)$ has two identical rows, then $\det(A) = 0$.

Proof. Induction on n . □

Theorem 4.5. If $A \in M_{n \times n}(F)$ and B is obtained from A by interchanging two rows, then $\det(B) = -\det(A)$.

Proof. Say rows a_r and a_s are interchanged. Play with the matrix that has rows r and s equal to $a_r + a_s$. □

Theorem 4.6. If $A \in M_{n \times n}(F)$ and B is obtained from A by adding a multiple of one row to another, then $\det(B) = \det(A)$.

Proof. Use expansion formula from Theorem 4.4. □

Corollary 10 If $A \in M_{n \times n}(F)$ and $\text{rank}(A) < n$, then $\det(A) = 0$.

Proof. Rows of A are linearly dependent so say $a_1 = \sum_{i \geq 2} c_i a_i$. Use Theorem 4.6. □

4.3 Properties

Summary of properties

- Let $A, B \in M_{n \times n}(F)$. Then $\det(AB) = \det(A)\det(B)$.
- $A \in M_{n \times n}(F)$ is invertible if and only if $\det(A) \neq 0$ and $\det(A^{-1}) = \frac{1}{\det(A)}$.
- Let $A \in M_{n \times n}(F)$. Then $\det(A^t) = \det(A)$.
- (Cramer's Rule) Let $Ax = b$ where $A \in M_{n \times n}(F)$ and let M_k be the $n \times n$ matrix obtained from A by replacing column k with b . Then $x_k = \frac{\det(M_k)}{\det(A)}$.

5 Diagonalization

5.1 Eigenvalues and eigenvectors

Definition 1 Let V be finite-dimensional and let $T : V \rightarrow V$ be linear. T is called *diagonalizable* if there is an ordered basis β such that $[T]_\beta$ is a diagonal matrix.

Matrix A is diagonalizable if L_A is.

Definition 2 A non-zero vector $v \in V$ such that $T(v) = \lambda v$ for some $\lambda \in F$ is called an *eigenvector*. The scalar λ is called an *eigenvalue* of T .

Theorem 5.1. Let V be finite-dimensional. A linear operator on V is diagonalizable if and only if there is an ordered basis β for V consisting of eigenvectors.

Theorem 5.2. Let $A \in M_{n \times n}(F)$. Then $\lambda \in F$ is an eigenvalue of A if and only if $\det(A - \lambda I_n) = 0$.

Definition 3

- Let $A \in M_{n \times n}(F)$. Then $f(t) = \det(A - \lambda I_n)$ is called the *characteristic polynomial* of A .
- If T is a linear operator V , then the characteristic polynomial of T is the characteristic polynomial of $A = [T]_\beta$ where β is an ordered basis for V .

Note: The characteristic polynomial of T is well-defined.

Theorem 5.3. Let $A \in M_{n \times n}(F)$.

- $f_A(t)$ is a polynomial of degree n with the leading coefficient $(-1)^n$.
- A has at most n distinct eigenvalues.

Theorem 5.4. Let T be a linear operator on V and let λ be an eigenvalue of T . A vector $v \in V$ is an eigenvector of T corresponding to λ if and only if $v \neq 0$ and $v \in N(T - \lambda I)$.

5.2 Diagonalizability

Theorem 5.5. Let T be a linear operator with distinct eigenvalues $\lambda_1, \dots, \lambda_k$. If v_1, \dots, v_k are eigenvectors of T such that v_i corresponds to λ_i , then $\{v_1, \dots, v_k\}$ is linearly independent.

Corollary 6 If T has n distinct eigenvalues, then T is diagonalizable.

Definition 4 A polynomial $f(t) \in P(F)$ splits over F if $f(t) = c(t - a_1) \dots (t - a_n)$ for some $c, a_1, \dots, a_n \in F$.

Theorem 5.6. Let T be a linear operator. If T is diagonalizable, then its characteristic polynomial splits.

Definition 5 The (algebraic) *multiplicity* of an eigenvalue λ is the largest positive integer k such that $(t - \lambda)^k | f(t)$.

The eigenspace of T with respect to λ is

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

Theorem 5.7. Let λ be an eigenvalue of T of multiplicity m . Then

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

Proof. (Outline)

- Start with an ordered basis $\{v_1, \dots, v_p\}$ for E_λ and extend it to an ordered basis β for V .
- Then $A = [T]_\beta$ will have the following form

$$\begin{pmatrix} \lambda I_p & B \\ 0 & C \end{pmatrix}.$$

Lemma 9. Let $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of T and let $v_i \in E_{\lambda_i}$. If $v_1 + \dots + v_k = 0$, then $v_i = 0$ for every i .

Theorem 5.8. Let $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of T and let $S_i \subseteq E_{\lambda_i}$ be finite and linearly independent. Then $S_1 \cup \dots \cup S_k$ is linearly independent.

Proof. (Outline) Say $S_i = \{v_{i_1}, \dots, v_{i_{n_i}}\}$ and suppose $\sum_{i=1}^k \sum_{j=1}^{n_i} a_{ij} v_{ij} = 0$. Consider $w_i = \sum_{j=1}^{n_i} a_{ij} v_{ij}$.

Theorem 5.9. Let T be a linear operator on a finite-dimensional vector space V such that the characteristic polynomial of T splits. Let $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of T . Then

- (a) T is diagonalizable if and only if the multiplicity of each λ_i equals $\dim(E_{\lambda_i})$.
- (b) If T is diagonalizable and β_i is an ordered basis for E_{λ_i} , then $\beta = \beta_1 \cup \cdots \cup \beta_k$ is an ordered basis for V .

Proof. (Outline)

- Part (b) follows from the proof.
- (\implies): m_i - the multiplicity of λ_i , $d_i = \dim(E_{\lambda_i})$, $\beta_i = \beta \cap E_{\lambda_i}$, $n_i = |\beta_i|$. We have $n_i \leq d_i \leq m_i$ and $\sum n_i = n = \sum m_i$. It follows that $d_i = m_i$ for every i .
- (\impliedby): Let β_i be an ordered basis for E_{λ_i} ; let $\beta = \beta_1 \cup \cdots \cup \beta_k$. Then β is a basis for V consisting of eigenvectors.

Note: T is diagonalizable if and only if

- The characteristic polynomial splits and
- the multiplicity of λ_i is $\text{nullity}(T - \lambda_i I) = n - \text{rank}(T - \lambda_i I)$.

5.4 Invariant subspaces and the Cayley-Hamilton theorem

Definition 6

- Let $T : V \rightarrow V$. A subspace $W \subseteq V$ is called a T -invariant subspace of V if $T(W) \subseteq W$.
- The T -cyclic subspace of V generated by x is

$$\text{span}(\{x, T(x), T^2(x), \dots\}).$$

Note: T -cyclic subspace is a minimal subspaces which is T -invariant and contains x .

Theorem 5.21. Let T be a linear operator on a finite-dimensional vector space V and let W be a T -invariant subspace of V . Then the characteristic polynomial of the restriction of T to W , T_W , divides the characteristic polynomial of T .

Theorem 5.22. Let T be a linear operator on a vector space V of dimension n and let W be a T -cyclic subspace of V generated by a non-zero vector v of dimension k . Then

- (a) $\{v, T(v), \dots, T^{k-1}(v)\}$ is a basis for W .
- (b) If $a_0 + a_1 T(v) + \cdots + a_{k-1} T^{k-1}(v) + T^k(v) = 0$, then the characteristic polynomial of T_W is $f(t) = (-1)^k(a_0 + a_1 t + \cdots + a_{k-1} t^{k-1} + t^k)$.

Proof. (Outline)

- (a) Let j be the largest integer such that $\beta = \{v, T(v), \dots, T^{j-1}(v)\}$ is linearly independent. Then $\text{span}(\beta)$ is T -invariant and it follows that $j = k$.
- (b) Let $\beta = v, T(v), \dots, T^{k-1}(v)$ and suppose $a_0v + a_1T(v) + \dots + a_{k-1}T^{k-1}(v) + T^k(v) = 0$. Use the form of $[T_W]_\beta$ to notice that $f_{T_W}(t) = (-1)^k(a_0 + a_1t + \dots + a_{k-1}t^{k-1} + t^k)$.

Theorem 5.23, Cayley-Hamilton. Let V be finite-dimensional and let $T : V \rightarrow V$ be linear with the characteristic polynomial $f(t)$. Then $f(T) = T_0$.

Proof. Note that $f(T)$ is a linear operator. We will show $f(T)(v) = 0$ for every v .

- Take $v \neq 0$ and consider the T -cycle subspace generated by v , W .
- Let $k = \dim(W)$. Then $T^k \in \text{span}(\{v, T(v), \dots, T^{k-1}(v)\})$ by 5.22(a) and we are done by 5.22(b).

□

Corollary 15 Let A be an $n \times n$ matrix and let $f(t)$ be its characteristic polynomial. Then $f(A) = 0_n$, the $n \times n$ zero matrix.

6 Inner product spaces

6.1 Inner products and norms

Definition 1 Let V be a vector space over F . An *inner product* on V is a function $\langle \cdot, \cdot \rangle : V \times V \rightarrow F$ such that the following conditions hold.

- $\langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$
- $\langle cx, y \rangle = c\langle x, y \rangle$
- $\overline{\langle x, y \rangle} = \langle y, x \rangle$
- $\langle x, x \rangle > 0$ if $x \neq 0$ and $\langle 0, 0 \rangle = 0$.

Definition 2 The *adjoint* of an $m \times n$ matrix X is the $n \times m$ matrix A^* such that $(A^*)_{ij} = \overline{A_{ji}}$.

Theorem 6.1.

- $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$

- $\langle x, cy \rangle = \bar{c} \langle x, y \rangle$
- $\langle x, 0 \rangle = \langle 0, x \rangle = 0$
- $\langle x, x \rangle = 0$ if and only if $x = 0$
- If $\langle x, y \rangle = \langle x, z \rangle$ for every $x \in V$, then $y = z$.

Definition 3 Let V be an inner product space. For $x \in V$, the *norm* of x is $\|x\| = \sqrt{\langle x, x \rangle}$.

Theorem 6.2. Let V be an inner product space over F . For $x, y \in V$ and $c \in F$.

- (a) $\|cx\| = |c| \cdot \|x\|$
- (b) $\|x\| = 0$ if and only if $x = 0$ and $\|x\| \geq 0$ for any x .
- (c) (Cauchy-Schwarz Inequality) $|\langle x, y \rangle| \leq \|x\| \|y\|$
- (d) (Triangle Inequality) $\|x + y\| \leq \|x\| + \|y\|$

Proof. (Outline)

- (c) Expand $\|x - cy\|^2$ and apply with $c = \frac{\langle x, y \rangle}{\langle y, y \rangle}$
- (d) Note that $\langle x, y \rangle + \langle y, x \rangle = 2\operatorname{Re}\langle x, y \rangle$ and $\operatorname{Re}\langle x, y \rangle \leq |\langle x, y \rangle|$.

Definition 4

- Vectors $x, y \in V$ are called *orthogonal* if $\langle x, y \rangle = 0$.
- A set of vectors $S \subseteq V$ is called *orthogonal* if any two distinct vectors are orthogonal.
- S is called *orthonormal* if it is orthogonal and $\|x\| = 1$ for every $x \in S$.

Old Town Problem: There are n people living in an odd town and they form clubs. A club must contain an odd number of members and for any two distinct clubs there must be an even (possibly zero) number of people in both of them. What is the maximum number of clubs that can be formed?

6.2 The Gram-Schmidt orthogonalization

Definition 5 An ordered basis which is orthonormal is called an *orthonormal basis*.

Theorem 6.3. Suppose $S = \{v_1, \dots, v_k\}$ is an orthogonal subset of V such that $v_i \neq 0$. For $y \in \text{span}(S)$,

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

Corollary 4 Any orthogonal set of non-zero vectors is linearly independent.

Theorem 6.4 (Graham-Schmidt algorithm). Let $S = \{w_1, \dots, w_n\}$ be a linearly independent subset. Define $S' = \{v_1, \dots, v_n\}$ as follows, $v_1 := w_1$ and for $k \geq 2$

$$v_k = w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j \rangle}{\|v_j\|^2} v_j.$$

Then S' is orthogonal and $\text{span}(S') = \text{span}(S)$.

Proof. This is induction on $|S|$. For the inductive step, first check that $v_k \neq 0$ and then compute $\langle v_k, v_i \rangle$ for $i < k$. Then $\dim(\text{span}(S'_k)) = \dim(\text{span}(S_k))$ because S'_k is linearly independent. \square

Theorem 6.5. Let V be a finite-dimensional inner product space and let $\beta = \{v_1, \dots, v_n\}$ be an orthonormal basis for V . Then for every $x \in V$, $x = \sum_{i=1}^n \langle x, v_i \rangle v_i$.

Corollary 7 Let $\beta = \{v_1, \dots, v_n\}$ be orthonormal, let $T : V \rightarrow V$ be linear, and let $A = [T]_\beta$. Then $A_{ij} = \langle T(v_j), v_i \rangle$.

Fourier coefficient of x relative to β is $\langle x, y \rangle$ where $y \in \beta$.

Definition 6 Let S be a non-empty subset of V . Define $S^\perp = \{x \in V : \langle x, y \rangle = 0 \text{ for every } y \in S\}$.

Note: S^\perp is a subspace of V .

Theorem 6.6. Let W be a finite-dimensional subspace of an inner product space V and let $y \in V$. Then there exist unique $u \in W$ and $z \in W^\perp$ such that $y = u + z$. Furthermore, if $\{v_1, \dots, v_k\}$ is an orthonormal basis for W , then $u = \sum \langle y, v_i \rangle v_i$.

Proof. (Outline)

- Let $u = \sum \langle y, v_i \rangle v_i$ and $z = y - u$. Check that $z \in W^\perp$.

- For the uniqueness $u - u' \in W$ and $z' - z \in W^\perp$.

Corollary 9 For any $x \in W$, $\|y - x\| \geq \|y - u\|$ and if $\|y - x\| = \|y - u\|$ then $x = u$.

Theorem 6.7. Suppose $S = \{v_1, \dots, v_k\}$ is an orthonormal set in an n -dimensional inner product space V . Then

- (a) S can be extended to an orthonormal basis $\{v_1, \dots, v_n\}$ for V .
- (b) $\{v_{k+1}, \dots, v_n\}$ is an orthonormal basis for $(\text{span}(S))^\perp$.
- (c) If W is a subspace of V , then $\dim(W) + \dim(W^\perp) = \dim(V)$.

6.3 The adjoint of a linear operator

Theorem 6.8. Let V be a finite-dimensional inner product space over F , and let $g : V \rightarrow F$ be a linear transformation. Then there exists a unique vector $y \in V$ such that $g(x) = \langle x, y \rangle$ for every $x \in V$.

Proof. (Outline) Take an orthonormal basis $\beta = \{v_1, \dots, v_n\}$ and define $y = \sum \overline{g(v_i)} v_i$.

Theorem 6.9. Let V be a finite-dimensional inner product space and let $T : V \rightarrow V$ be linear. Then there exists a unique function $T^* : V \rightarrow V$ such that $\langle T(x), y \rangle = \langle x, T^*(y) \rangle$ for all x, y . Furthermore, T^* is linear.

Proof. (Outline)

- Fix y . Check that $g(x) = \langle T(x), y \rangle$ is linear.
- Theorem 6.8 gives unique y' such that $g(x) = \langle x, y' \rangle$. Define $T^*(y) = y'$.
- Note that T^* is a function and check that it is linear.

Theorem 6.10. Let V be a finite-dimensional inner product space and let β be an orthonormal ordered basis for V . If T is a linear operator on V , then $[T^*]_\beta = [T]_\beta^*$.

Theorem 6.11. Let V be an inner-product space, and let T, U be linear operators on V . Then

- (a) $(T + U)^* = T^* + U^*$
- (b) $(cT)^* = \bar{c}T^*$
- (c) $(TU)^* = U^*T^*$
- (d) $T^{**} = T$
- (e) $I^* = I$

6.4 Normal and self-adjoint operators

Lemma 15. Let T be a linear operator on a finite-dimensional inner product space V . If T has an eigenvector, then so does T^* .

Theorem 6.14 (Schur). Let T be a linear operator on a finite-dimensional inner product space V and suppose that the characteristic polynomial of T splits. Then there exists an orthonormal basis β for V such that $[T]_\beta$ is upper triangular.

Proof. (Outline)

- Show that W^\perp is T -invariant and $\dim(W^\perp) = n - 1$.
- The characteristic polynomial of T_{W^\perp} divides $f_T(t)$ (so it splits) and by **IH** for some γ , $[T_{W^\perp}]_\gamma$ is upper-triangular.
- Let $\beta = \gamma \cup \{z\}$ Then $[T]_\beta$ is upper-triangular.

Note: If V has an orthonormal basis of eigenvectors of T , then $TT^* = T^*T$.

Definition 7 $T : V \rightarrow V$ is *normal* if $TT^* = T^*T$. $A \in M_{n \times n}(F)$ is *normal* if $AA^* = A^*A$.

Theorem 6.15. Let T be a normal operator on V . Then

- $\|T(x)\| = \|T^*(x)\|$
- $T - cI$ is normal for every $c \in F$.
- If x is an eigenvector of T , then x is an eigenvector of T^* .
- If λ_1, λ_2 are distinct eigenvalues of T with eigenvectors x_1, x_2 , then $\langle x_1, x_2 \rangle = 0$.

Theorem 6.16. 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 consisting of eigenvectors of T .

Proof. (Outline) Suppose T is normal.

- T splits over C and so apply Schur's lemma to get an orthonormal basis $\beta = \{v_1, \dots, v_n\}$.
- $A := [T]_\beta$ is upper-triangular and so $T(v_1) = A_{11}v_1$.
- Show that e_k is an eigenvector by induction on k using the fact that $A_{jk} = \langle T(v_k), v_j \rangle$.

The converse is easy.

Definition 8 $T : V \rightarrow V$ is *self-adjoint* (Hermitian) if $T = T^*$. $A \in M_{n \times n}(F)$ is *Hermitian* if $A = A^*$.

Lemma 19. Let T be a Hermitian operator on a finite-dimensional inner product space V . Then,

- (a) All eigenvalues of T are real.
- (b) If V is a real inner product space, then the characteristic polynomial splits.

Theorem 6.17. Let T be a linear operator on a finite-dimensional real inner-product space V . Then T is Hermitian if and only if there exists an orthonormal basis for V consisting of eigenvectors of T .

Proof. The characteristic polynomial splits and so we may apply Schur's lemma. $A := [T]_\beta$ is upper-triangular and so A^* . Thus it must be a diagonal matrix. \square

6.5 Unitary and orthogonal operators and their matrices

Definition 9 Let V be a finite-dimensional inner-product space over F and let $T : V \rightarrow V$ be linear. If $\|T(x)\| = \|x\|$ for every $x \in V$, then T is called *unitary* if $F = C$ and *orthogonal* if $F = R$.

Theorem 6.18. Let T be a linear operator on a finite-dimensional inner-product space V . Then the following statements are equivalent.

- (a) $TT^* = T^*T = I$
- (b) $\langle T(x), T(y) \rangle = \langle x, y \rangle$ for all $x, y \in V$
- (c) If β is an orthonormal basis, then so is $T(\beta)$.
- (d) There exists an orthonormal basis β such that $T(\beta)$ is orthonormal.
- (e) $\|T(x)\| = \|x\|$ for every x .

Proof. (Outline)

- (d) \implies (e) Let $\beta = \{v_1, \dots, v_n\}$ be orthonormal such that $T(\beta)$ is orthonormal. Take $x \in V$. Then $x = \sum a_i v_i$ and

$$\|x\|^2 = \sum |a_i|^2 = \sum \|T(x)\|^2.$$

- $(e) \implies (a)$ $\langle x, x \rangle = \langle x, T^*T(x) \rangle$ by (e). Thus $\langle x, (I - T^*T)(x) \rangle = 0$ for every x . Set $U := I - T^*T$. Then U is self-adjoint and so there is an orthonormal basis consisting of eigenvectors of U . Check that $U(x) = \lambda x$ implies $\lambda = 0$; $U = T_0$; $T^*T = I$; $TT^* = I$ as well because $[T]_\beta$ is a square matrix.

Definition 10 A square matrix A is called *orthogonal* if $A^t A = A A^t = I$ and *unitary* if $A^* A = A A^* = I$.

Note:

- $AA^* = I$ if and only if A are orthonormal.
- $A^*A = I$ if and only if columns of A are orthonormal.
- If λ is an eigenvalue of a unitary (orthogonal) matrix, then $|\lambda| = 1$.

Definition 11 $A, B \in M_{n \times n}(C)$ ($A, B \in M_{n \times n}(R)$) are *unitarily (orthogonally) equivalent* if there exists a unitary (orthogonal) matrix P such that $A = P^{-1}BP$.

Theorem 6.19. Let $A \in M_{n \times n}(C)$. Then A is normal if and only if A is unitarily equivalent to a diagonal matrix.

Theorem 6.20. Let $A \in M_{n \times n}(R)$. Then A is symmetric if and only if A is orthogonally equivalent to a diagonal matrix.

Theorem 6.21 (Schur). Let $A \in M_{n \times n}(F)$ and suppose $f_A(t)$ splits over F .

- If $F = C$, then A is unitarily equivalent to a complex upper-triangular matrix.
- If $F = R$, then A is orthogonally equivalent to a real upper-triangular matrix.