Revision notes - MA2101

Ma Hongqiang

April 27, 2017

Contents

1 Vector Spaces over a Field

Definition 1.1 (Field, Rings, Groups)**.**
Let $\mathbb{F}$ be a set containing at least two elements and equipped with the following two binary operations $+$(the addition, or plus) and $\times$(the multiplication, or times), where $\mathbb{F} \times \mathbb{F} := \{(x, y) \mid x, y \in \mathbb{F}\}$ is the **product set** of $\mathbb{F}$ with iteself:

$$\begin{aligned} + :& \mathbb{F} \times \mathbb{F} \to \mathbb{F} \\ & (x, y) \mapsto x + y; \\ \times :& \mathbb{F} \times \mathbb{F} \to \mathbb{F} \\ & (x, y) \mapsto x \times y; \end{aligned}$$

Axiom(0) Addition and multiplication are well defined on $\mathbb{F}$ in the sense that:

$$\forall x \in \mathbb{F}, \forall y \in \mathbb{F} \Rightarrow x + y \in \mathbb{F}$$
$$\forall x \in \mathbb{F}, \forall y \in \mathbb{F} \Rightarrow xy \in \mathbb{F}$$

Namely, the operation map $+$(resp. $\times$) takes element (x, y) in the domain $\mathbb{F} \times \mathbb{F}$ to some element $x+y$(resp. xy) in the codomain $\mathbb{F}$. The quintuple $(\mathbb{F}, +, 0; \times, 1)$ with two distinguished elements 0 (the additive identity) and 1 (the multiplicative identity), is called a **field** if the following **Eight Axioms** (and also **Axiom (0)**) are satisfied.

(1) Existence of an **additive identity** $0_{\mathbb{F}}$ or simply 0:

$$x + 0 = x = 0 + x, \forall x \in \mathbb{F}$$

(2) (Additive) Associativity:

$$(x + y) + z = x + (y + z), \forall x, y, z \in \mathbb{F}$$

(3) Additive Inverse
for every $x \in \mathbb{F}$, there is an **additive inverse** $-x \in \mathbb{F}$ of x such that

$$x + (-x) = 0 = (-x) + x$$

(4) Existence of a **multiplicative identity** $1_{\mathbb{F}}$ or simply 1:

$$x1 = x = 1x, \forall x \in \mathbb{F}$$

(5) (Multiplicative) Associativity:

$$(xy)z = x(yz), \forall x, y, z \in \mathbb{F}$$

(6) Multiplicative Inverse for nonzero element:
for every $0 \neq x \in \mathbb{F}$, there is a **multiplicative inverse** $x^{-1} \in \mathbb{F}$ such that

$$xx^{-1} = 1 = x^{-1}x.$$

(7) Distributive Laww:

$$(x+y)z = xz + yz, \forall x, y, z \in \mathbb{F}$$
$$z(x+y) = zx + zy, \forall x, y, z \in \mathbb{F}$$

(8) Commutativity for addition and multiplication:

$$x + y = y + x, \quad \forall x, y \in \mathbb{F}$$
$$xy = yx, \qquad \forall x, y \in \mathbb{F}$$

The triplet $(\mathbb{F}, +, \times)$ with only the Axioms (1)—(5) and (7)—(8) satisfied is called a (commutative) ring.
The pair $(\mathbb{F}, +)$ with only Axioms (1)—(3) satisfied by its binary operation $+$, is called an (additive) group.

Notation 1.1 (about $\mathbb{F}^\times$)**.** For a field $(\mathbb{F}, +, 0; \times, 1)$, we use $\mathbb{F}^\times$ to denote the set of nonzero elements in $\mathbb{F}$:

$$\mathbb{F}^\times := \mathbb{F} \setminus \{0\}$$

Definition 1.2 (Polynomial ring)**.**
Let $(\mathbb{F}, +, 0; \times, 1)$ be a field (or a ring), e.g. $\mathbb{F} = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$.

$$g(x) = \sum_{i=0}^{n} a_i x^i = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

with the **leading coefficient** $a_n \neq 0$, is called a **polynomial of degree** $n \geq 0$, in one variable x and with coefficients $a_i \in \mathbb{F}$. Let

$$\mathbb{F}[x] := \{\sum_{j=0}^{d} b_j x^j \mid d \geq 0, b_j \in \mathbb{F}\}$$

be the set of all polynomials in one variable x and with coefficients in $\mathbb{F}$.

Theorem 1.1 (Uniqueness of identity and inverse)**.** Let $\mathbb{F}$ be a field.

(1) $\mathbb{F}$ has only one additive identity 0.

(2) $\mathbb{F}$ has only one multiplicative identity 1.

(3) Every $x \in \mathbb{F}$ has only one additive inverse $-x$.

(4) Every $x \in \mathbb{F}^\times$ has only one multiplicative inverse x^{-1}.

Theorem 1.2 (Properties of $\mathbb{F}$)**.**

(1) (Cancelation Law) Let $b, x, y \in \mathbb{F}$. Then

$$b + x = b + y \Rightarrow x = y$$

(2) (Killing Power of 0)

$$0x = 0 = x0, \forall x \in \mathbb{F}$$

(3) In $\mathbb{F}$, we have

$$0_{\mathbb{F}} \neq 1_{\mathbb{F}}$$

(4) If $x \in \mathbb{F}^{\times}$, then its multiplicative inverse

$$x^{-1} \in \mathbb{F}^{\times}$$

(5) If $x + x' = 0$, then x and x' are multiplicative inverse to each other.

(6) If $xx'' = 1$ then x and x'' are multiplicative inverse to each other.

Definition 1.3 (Vector Space)**.**
Let $\mathbb{F}$ be a field and V a non-empty set, with a binary vector addition operation

$$\begin{aligned} + : V \times V &\to V \\ (\mathbf{v}_1, \mathbf{v}_2) &\mapsto \mathbf{v}_1 + \mathbf{v}_2 \end{aligned}$$

and scalar multiplication operation

$$\begin{aligned} \times : V \times V &\to V \\ (c, \mathbf{v}) &\mapsto c\mathbf{v} \end{aligned}$$

Axiom(0) These two operations are well defined in the sense that

$$\begin{aligned} \forall \mathbf{v}_i \in V &\Rightarrow \mathbf{v}_1 + \mathbf{v}_2 \in V \\ \forall c \in \mathbb{F}, \forall \mathbf{v} \in V &\Rightarrow c\mathbf{v} \in V \end{aligned}$$

$(V, +)$ is a **vector space over the field** $\mathbb{F}$ if the following **Seven Axioms** are satisfied.

(1) Existence of **zero vector** 0_V:

$$\mathbf{v} + \mathbf{0} = \mathbf{v} = \mathbf{v} + \mathbf{0}, \forall \mathbf{v} \in V$$

(2) (Additive) Associativity:

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}), \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$$

(3) Additive Inverse:
for every $\mathbf{v} \in V$, there is an additive inverse $-\mathbf{v}$ of $\mathbf{v}$ such that

$$\mathbf{v} + (-\mathbf{v}) = \mathbf{0} = (-\mathbf{v}) + \mathbf{v}$$

(4) The effect of $1 \in \mathbb{F}$ on V:

$$1\mathbf{v} = \mathbf{v}, \forall \mathbf{v} \in V$$

(5) (Multiplicative) Associativity:

$$(ab)\mathbf{v} = a(b\mathbf{v}), \forall a, b \in \mathbb{F}, \mathbf{v} \in V$$

(6) Distributive Law:

$$(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}, \forall a, b \in \mathbb{F}, \mathbf{v} \in V$$
$$a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}, \forall a \in \mathbb{F}, \mathbf{u}, \mathbf{v} \in V$$

(7) Commutativity for the vector addition:

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}, \forall \mathbf{u}, \mathbf{v} \in V$$

Remark: Every vector space V contains the zero vector $\mathbf{0}_V$.

2 Vector Subspaces

Definition 2.1 (Subspace)**.**
Let V be a vector space over a field $\mathbb{F}$. A non-empty subset $W \subseteq V$ is called a **vector subspace of** V if the following two conditions are satisfied:

(CA) (**C**losed under vector **A**ddition)

$$\forall \mathbf{w}_i \in W \Rightarrow \mathbf{w}_1 + \mathbf{w}_2 \in W$$

(CS) (**C**losed under **S**calar Multiplication)

$$\forall a \in \mathbb{F}, \forall \mathbf{w} \in W \Rightarrow a\mathbf{w} \in W$$

Obvious subspace of V are V and $\{\mathbf{0}\}$.
Remark: Every vector subspace W of V contains the zero vector $\mathbf{0}_W$.

Definition 2.2 (Linear Combination)**.**
Let V be a vector space over a field $\mathbb{F}$ and $W \subseteq V$ a non-empty subset. Then the following are equivalent:

(i) W is a vector subspace of V, i.e. W is closed under vector addition and scalar multiplication, in the sense of Definition of Subspace.

(ii) W is closed under linear combination:

$$\forall a_o \in \mathbb{F}, \forall \mathbf{w}_i \in W \Rightarrow a_1\mathbf{w}_1 + a_2\mathbf{w}_2 \in W$$

(iii) W together with the vector addition $+$ and the scalar multiplication $\times$, becomes a vector space.

Theorem 2.1 (Intersection of Subspace being a subspace)**.**
Let V be a vector space over a field $\mathbb{F}$ and let $W_\alpha \subseteq V(\alpha \in I)$ be vector subspaces of V. Then the intersection

$$\cap_{\alpha \in I} W_\alpha$$

is again a vector subspace of V

Remark: Union of subspaces may not be a subspace. However, union of subspaces is closed under scalar multiplication.

3 Linear Spans and Direct Sums of Subspaces

Definition 3.1 (Linear combination and Linear Span)**.**
Let V be a vector space over a field $\mathbb{F}$. A vector $\mathbf{v} \in V$ is called a **linear combination of some vectors** $\mathbf{v}_i \in V(1 \leq i \leq s)$ if

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_s\mathbf{v}_s$$

for some scalars $a_i \in \mathbb{F}$. Let $S \subseteq V$ be a non-empty subset. The subset $\text{Span}(S) :=$

$$\{\mathbf{v} \in V \mid \mathbf{v} \text{ is a linear combination of some vectors in } S\}$$

of V is called the **vector subspace of** V**spanned by the subset** S.

Theorem 3.1 (Span being a subspace)**.**

(i) The subset $\text{Span}(S)$ of V is indeed a vector subspace of V.

(ii) $\text{Span}(S)$ is the smallest vector subspace of V containing the set S:
firstly, $\text{Span}(S)$ is a vector subspace of V containing S;
secondly, if W is another vector subspace of V containing S, then $W \supseteq \text{Span}(S)$

Definition 3.2 (Sum of subspaces)**.**
Let V be a vector space over a field $\mathbb{F}$, and let U and W be vector subspaces of V. The subset

$$U + W = \mathbf{u} + \mathbf{w} \mid \mathbf{u} \in U, \mathbf{w} \in W$$

is called the **sum of the subspaces** U and W.

Theorem 3.2 (Sum being a subspace)**.**
Let U and W be vector subspaces of a vector space V over a field $\mathbb{F}$. For the sum $U + W$, we have:

1. $U + W = \text{Span}(U \cup W)$
2. $U + W$ is indeed a vector subspace of V.
3. $U + W$ is the smallest vector subspace of V containing both U and W:
   first, $U + W$ is a vector subspace of V containing both U and W; secondly, if T is another vector subspace of V containing both U and W then, $T \supseteq U + W$.

Note 1: Let U and W be two vector subspaces of a vector space V over a field $\mathbb{F}$. Then the following are equivalent.

1. The union $U \cup W$ is a vector subspace of V.
2. Either $U \subseteq W$ or $W \subseteq U$.

Definition 3.3 (Sum of many subspaces).
Let V be a vector space over a field $\mathbb{F}$ and let $W_i(1 \leq i < s)$ be vector subspaces of V. The subset

$$\begin{aligned}\sum_{i=1}^{s} W_i &= W_1 + \cdots + W_s \\ &= \{\sum_{i=1}^{s} \mathbf{w}_i \mid \mathbf{w}_i \in W\} \\ &= \{\mathbf{w}_1 + \cdots + \mathbf{w}_s \mid \mathbf{w}_i \in W_i\}\end{aligned}$$

Theorem 3.3 (Sum of many being a subspace).
Let $W_i(1 \leq i < s)$ be vector subspaces of a vector space V over a field $\mathbb{F}$. For the sum $\sum_{i=1}^{s} W_i$, we have

1. $\sum_{i=1}^{s} W_i = \text{Span}(\cup_{i=1}^{s} W_i)$
2. $\sum_{i=1}^{s} W_i$ is indeed a vector subspace of V.
3. $\sum_{i=1}^{s} W_i$ is the smallest vector subspace of V containing all W_i.

Definition 3.4 (Direct Sum of Subspaces).
Let V be a vector space over a field $\mathbb{F}$ and let W_1, W_2 be vector subspace of V. We say that the sum $W_1 + W_2$ is a **direct sum** of two vector subspaces W_1, W_2 if the intersection

$$W_1 \cap W_2 = \{\mathbf{0}\}$$

In this case, we denote $W_1 + W_2$ as $W_1 \oplus W_2$.
We write $W = W_1 \oplus W_2$ if W is a direct sum of W_1 and W_2.

Theorem 3.4 (Equivalent Direct Sum Definition).
Let W_1 and W_2 be two vector subspaces of a vector space V over a field $\mathbb{F}$. Set $W := W_1 + W_2$. Then the following are equivalent.

1. We have
$$W_1 + W_2 = W_1 \oplus W_2$$
i.e., $W_1 + W_2$ is a direct sum of W_1, W_2,i.e., $W_1 \cap W_2 = \{\mathbf{0}\}$

2. (Unique expression condition) Every vector $\mathbf{w} \in W$ can be expressed as
$$\mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2$$
for some $\mathbf{w}_i \in W_i$ and such expression of $\mathbf{w}$ is unique.

Definition 3.5 (Direct Sum of Many Subspaces).
Let V be a vector space overa field $\mathbb{F}$ and let $W_i(1 \leq i \leq s; s \geq 2)$ be vector subspaces of V. We say that the sum $\sum_{i=1}^{s} W_i$ is a **direct sum of vector subspaces** W_i if the intersection

$$\left(\sum_{i=1}^{k-1} W_i\right) \cap W_k = \{\mathbf{0}\} \quad (2 \leq \forall k \leq s)$$

Theorem 3.5 (Equivalent Direct Multiple Sum Definition)**.**
Let $W_i (1 \leq i \leq s, s \geq 2)$ be vector subspaces of a vector space V over a field $\mathbb{F}$. Set $W := \sum_{i=1}^{s} W_i$. Then the following are equivalent.

1. We have
$$W_1 + \cdots + W_s = W_1 \oplus \cdots \oplus W_s$$
i.e., $\sum_{i=1}^{s} W_i$ is a direct sum of W_i.

2. $$\left(\sum_{i \neq j} W_i \right) \cap W_l = \{\mathbf{0}\} \quad (\forall 1 \leq l \leq s)$$

3. (Unique expression condition) Every vector $\mathbf{w} \in W$ can be expressed as
$$\mathbf{w} = \mathbf{w}_1 + \cdots + \mathbf{w}_s$$
for some $\mathbf{w}_i \in W_i$ and such expression of $\mathbf{w}$ is unique.

4 Linear Independence, Basis and Dimension

Definition 4.1 (Linear (in)dependence)**.**
Let V be a vector space over a field $\mathbb{F}$. Let T be a (not necessarily finite) subset of V and let

$$S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$$

be a finite subset of V.

(1) We call S a **linear independent set** or **L.I.**, if the vector equation below

$$x_1\mathbf{v}_1 + \cdots + x_m\mathbf{v}_m = \mathbf{0}$$

has only the so called **trivial solution**

$$(x_1, \dots, x_m) = (0, \dots, 0)$$

(2) We call S a **linear dependent set** or **L.D.** if there are scalars $a_1, \dots, a_m$ in $\mathbb{F}$ which are not all zero (i.e. $(a_1, \dots, a_m) \neq (0, \dots, 0)$) such that

$$a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m = \mathbf{0}$$

(3) The set T is a **linearly independent set** if every non-empty finite subset of T is linearly independent. The set T is a **linearly dependent set** if at least one non-empty finite subset of T is linearly dependent.

Theorem 4.1 (L.D./L.I. Inheritance)**.**

(1) Let $S_1 \subseteq S_2$. If the smaller set S_1 is linearly dependent then so is the larger set S_2. Equivalently, if the larger set S_2 is linearly independent then so is the smaller set S_1.

(2) $\{\mathbf{0}\}$ is a linearly dependent set.

(3) If $\mathbf{0} \in S$, then S is a linearly dependent set.

Definition 4.2 (Equivalent L.I./L.D. Definitions)**.**
Let $S = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ be a finite subset of a vector space V over a field $\mathbb{F}$. Then we have:

(1) Let $|S| \geq 2$. Then S is a linear dependent set if and only if some $\mathbf{v}_k \in S$ is a linear combination of the others, i.e. there are scalars

$$a_1, \dots, a_{k-1}, a_{k+1}, \dots, a_m$$

in $\mathbb{F}$ (with all these scalars vanishing allowed) such that

$$\mathbf{v}_k = \sum_{i \neq k} a_i\mathbf{v}_i = a_1\mathbf{v}_1 + \cdots + a_{k-1}\mathbf{v}_{k-1} + a_{k+1}\mathbf{v}_{k+1} + \cdots + a_m\mathbf{v}_m$$

(2) Let $|S| \geq 2$. Then S is linearly independent if and only if no $\mathbf{v}_k \in S$ is a linear combination of others.

(3) Suppose that $S = \{\mathbf{v}_1\}$ (a single vector). Then S is linearly dependent if and only if $\mathbf{v}_1 = \mathbf{0}$. Equivalently, S is linearly independent if and only if $\mathbf{v}_1 \neq \mathbf{0}$.

(4) Suppose that $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ (two vectors). Then S is linearly dependent if and only if one of $\mathbf{v}_1, \mathbf{v}_2$ is a scalar multiple of the other. Equivalently, S is linearly independent if and only if neither one of $\mathbf{v}_1, \mathbf{v}_2$ is a scalar multiple of the other.

Definition 4.3 (Basis, (in)finite demension)**.**
Let V be a nonzero vector space over a field $\mathbb{F}$. A subset B of V is called a **basis** if the following two conditions are satisfied.

(1) (Span) V is spanned by B: $V = \mathrm{Span}(B)$

(2) (L.I.) B is a linearly independent set.

If V has a basis B with **cardinality** $|B| < \infty$ we say that V is **finite dimensional** and define the **dimension** of V over the field $\mathbb{F}$ as the cardinality of B:

$$\dim_{\mathbb{F}} V := |B|$$

Otherwise, V is called **infinite-dimensional**.

If V equals the zero vector space $\{\mathbf{0}\}$, we define

$$\dim\{\mathbf{0}\} = 0$$

Theorem 4.2 (Equivalent Basis Definition I)**.**
Let $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$(with $\mathbf{v}_i \neq \mathbf{0}_V$) be a finite subset of a vector space V over a field $\mathbb{F}$. Then the following are equivalent.

(1) B is a basis of V.

(2) (Unique expression condition) Every vector $\mathbf{v} \in V$ can be expressed as

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$$

for some scalars $a_i \in \mathbb{F}$ and such expression of $\mathbf{v}$ is unique.

(3) V has the following direct sum decomposition:

$$V \begin{aligned} &= \mathrm{Span}\{\mathbf{v}_1\} \oplus \cdots \oplus \mathrm{Span}\{\mathbf{v}_n\} \\ &= \mathbb{F}\mathbf{v}_1 \oplus \cdots \oplus \mathbb{F}\mathbf{v}_n \end{aligned}$$

Theorem 4.3 (Deriving a basis from a spanning set)**.**
Suppose that a nonzero vector space V over a field $\mathbb{F}$ is spanned by a finite subset $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_s\}$, then we have:

(1) There is a subset $B_1 \subseteq B$ such that B_1 is a basis of V. In particular

$$dim_{\mathbb{F}} V = B_1 \leq B$$

(2) Let B_2 be a maximal linearly independent subset of B: first B_2 is L.I. and secondly every subsets B_3 of B larger than B_2 is L.D. Then B_2 is a basis of $V = \mathrm{Span}(B)$.

Theorem 4.4 (Dimension being well Defined)**.**
Let $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis of a vector space V over a field $\mathbb{F}$. Then we have:

(1) Suppose that S is a subset of V with $|S| > n = |B|$. Then S is L.D.

(2) Suppose that T is a subset of V with $|T| < n$. Then T does not span V.

(3) Suppose that B' is another basis of V. Then $|B'| = |B|$. So the dimension $\dim_{\mathbb{F}} V (= |B|)$ of V depends only on V, but not on the choice of its basis.
In other words, $\dim_{\mathbb{F}} V$ is well defined.

Theorem 4.5 (Expanding an L.I. set)**.**
Let B be a L.I. subset of a vector space V over a field. Then exactly one of the following two cases is true.

(1) B spans V and hence B is a basis of V.

(2) Let $\mathbf{w} \in V \setminus \mathrm{Span}(B)$(and hence $\mathbf{w} \notin B$). Then

$$B \cup \{\mathbf{w}\}$$

is a L.I. subset of V.

In particular, if V is of finite dimension n, then one can find $n - |B|$ vectors

$$\mathbf{w}_{|B|+1}, \cdots, \mathbf{w}_n$$

in V Span(B) such that

$$B \coprod \{\mathbf{w}_{|B|+1}, \cdots, \mathbf{w}_n\}$$

is a basis of V.

Theorem 4.6 (Equivalent Basis Definition II)**.**
Let B be a subset of vector space V of finite dimension $\dim_m athbbFV = n \geq 1$. Then the following are equivalent.

(1) B is a basis of V.

(2) B is L.I. and $|B| = n$.

(3) B spans V and $|B| = n$.

Theorem 4.7 (Basis of a direct sum)**.**
Let V be a (not necessarily finite-dimensional) vector space over a field $\mathbb{F}$.

(1) Suppose that B is a basis of V. Decompose it as a disjoint union

$$B = B_1 \coprod B_2 \cdots \coprod B_s$$

of non-empty sets B_i. Then B_i is a basis of $W_i := \mathrm{Span}(B_i)$ and

$$V = W_1 \oplus \cdots \oplus W_s$$

is a direct sum of nonzero vector subspaces W_i of V.

(2) Conversely, suppose that

$$V = W_1 \oplus \cdots \oplus W_s$$

is a direct sum of nonzero vector subspaces W_i of V. Let B_i be a basis of W_i. Then

$$B = B_1 \coprod B_2 \cdots \coprod B_s$$

is a basis of V and a disjoint union of non-empty sets B_i.

(3) In particular, if

$$V = W_1 \oplus \cdots \oplus W_s$$

is a direct sum, then

$$\dim_m athbbFV = \sum_{i=1}^{s} \dim_{\mathbb{F}} W_i$$

5 Row Space and Column Space

Definition 5.1 (Column/Row Space, Nullspace, Nullity, Range of A)**.**
Let

$$A = (a_i j) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

be an $m \times n$ matrix with entries in a field $\mathbb{F}$. Let

$$\mathrm{Col}(A) := \mathrm{Span}\{\mathbf{c}_1 := \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, \mathbf{c}_n := \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}\}$$

be the **column space** of A, and let

$$\mathrm{R}(A) := \mathrm{Span}\{\mathbf{r}_1 := (a_{11}, \dots, a_{1n}), \dots, \mathbf{r}_m := (a_{m1}, \dots, a_{mn})\}$$

be the **row space** of A so that we can write

$$A = (\mathbf{c}_1, \dots, \mathbf{c}_m) = \begin{pmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_m \end{pmatrix}$$

The **range** of A is defined as

$$R(A) = \{AX \mid X \in \mathbb{F}_c^n\}$$

The **nullity** of A is defined as the dimension of the **nullspace** or **kernel**

$$\mathrm{Ker}(A) := \mathrm{Null}(A) := \{X \in \mathbb{F}_c^n \mid AX = \mathbf{0}\}$$

i.e.

$$\mathrm{nullity}(A) = \dim \mathrm{Null}(A)$$

Theorem 5.1 (Rank of Matrix, Matrix Dimension Theorem)**.**

(1) The range equals the column space

$$R(A) = \mathrm{Col}(A)$$

(2) Column and row spaces have the same dimension

$$\dim_{\mathbb{F}} \mathrm{Col}(A) = \dim_{\mathbb{F}} \mathrm{R}(A) := \mathrm{rank}(A)$$

which is called the **rank** of A.

(3) There is a dimension theorem

$$\mathrm{rank}(A) + \mathrm{nullity}(A) = n$$

where n is the number of columns in A.

Theorem 5.2 (L.I. vs L.D.).
In previous theorem, suppose that $m = n$ so that A is a square matrix of order n. Then the following are equivalent.

(1) A is an invertible matrix, i.e. A has a so called **inverse** $A^{-1} \in M_n(\mathbb{F})$ such that

$$AA^{-1} = I_n = A^{-1}A$$

(2) A has nonzero determinant

$$\det(A) = |A| \neq 0$$

(3) The column vectors

$$\mathbf{c}_1, \ldots, \mathbf{c}_n$$

of A form a basis of the column vector n-space $\mathbb{F}_c^n$.

(4) The row vectors

$$\mathbf{r}_1, \ldots, \mathbf{r}_n$$

of A form a basis of the row vector n-space $\mathbb{F}_c^n$.

(5) The column vectors

$$\mathbf{c}_1, \ldots, \mathbf{c}_n$$

of A are linearly indepedent in $\mathbb{F}_c^n$.

(6) The row vectors

$$\mathbf{r}_1, \ldots, \mathbf{r}_n$$

of A are linearly indepedent in $\mathbb{F}_c^n$.

(7) The matrix equation

$$AX = \mathbf{0}$$

has the trivial solution only: $X = \mathbf{0}$.

Theorem 5.3 (Row operation preserves columns relations).
Suppose that A and B are row equivalent. Then we have:

(1) If the column vectors

$$\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_s}, \mathbf{a}_{j_1}, \ldots, \mathbf{a}_{j_t}$$

of A satisfies a relation

$$c_{i_1}\mathbf{a}_{i_1} + \cdots + c_{i_s}\mathbf{a}_{i_s} = c_{j_1}\mathbf{a}_{j_1} + \cdots + c_{j_t}\mathbf{a}_{j_t}$$

for some scalars $c_{i_k} \in \mathbb{F}$, then the corresponding column vectors

$$\mathbf{b}_{i_1}, \ldots, \mathbf{b}_{i_s}, \mathbf{b}_{j_1}, \ldots, \mathbf{b}_{j_t}$$

of B satisfies exactly the same relation

$$c_{i_1}\mathbf{b}_{i_1} + \cdots + c_{i_s}\mathbf{b}_{i_s} = c_{j_1}\mathbf{b}_{j_1} + \cdots + c_{j_t}\mathbf{b}_{j_t}$$

The converse is also true.

(2) The column vectors
$$\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_s}, \mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_t}$$
of A are linearly dependent if and only if the corresponding column vectors
$$\mathbf{b}_{i_1}, \dots, \mathbf{b}_{i_s}, \mathbf{b}_{j_1}, \dots, \mathbf{b}_{j_t}$$
of B are linearly dependent.

(3) The column vectors
$$\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_s}, \mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_t}$$
of A are linearly independent if and only if the corresponding column vectors
$$\mathbf{b}_{i_1}, \dots, \mathbf{b}_{i_s}, \mathbf{b}_{j_1}, \dots, \mathbf{b}_{j_t}$$
of B are linearly independent.

(4) The column vectors
$$\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_s}$$
of A forms a basis of the column space
$$\mathrm{Col}(A) = \mathrm{Span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$$
if and only if the corresponding column vectors
$$\mathbf{b}_{i_1}, \dots, \mathbf{b}_{i_s}$$
form a basis of the column space
$$\mathrm{Col}(B) = \mathrm{Span}\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$$

(5) If
$$B_1 := \{\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_s}\}$$
is a maximal L.I. subset of the set
$$C = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$$
of all column vectors then B_1 is a basis of the column space $\mathrm{Col}(A)$ of A.

(6) Suppose that B is in row-echelon form with leading entries at columns
$$i_1, \dots, i_s$$
Then
$$\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_s}$$
forms a basis of the column space $\mathrm{Col}(A)$ of A.

(7) The row space of A and B are identical

$$\mathrm{R}(A) = \mathrm{R}(B)$$

But the column spaces of A and B may not be the same.

(8) Suppose that

$$B = (b_{ij}) = \begin{pmatrix} \mathbf{b}'_1 \\ \vdots \\ \mathbf{b}'_m \end{pmatrix}$$

is in row-echelon form with leading entries at rows

$$j_1, \dots j_t$$

Then

$$\mathbf{b}'_{j_1}, \dots, \mathbf{b}'_{j_t}$$

form a basis of the row space $\mathrm{R}(A) = \mathrm{R}(B)$ of A.

6 Quotient Spaces and Linear Transformations

Definition 6.1 (Sum of subsets of a space)**.**
Let V be a vector space over a field $\mathbb{F}$, and let S and T be subsets (which are not necessarily subspaces) of V. Define the **sum** of S and T as

$$S + T := \{s + t \mid s \in S, t \in T\}$$

In general, given subsets $S_i (1 \leq i \leq r)$ of V, we can define the **sum** of S_i as

$$\sum_{i=1}^{r} S_i = \{\sum_{i=1}^{r} \mathbf{x}_i \mid \mathbf{x}_i \in S_i\}$$

Theorem 6.1 (Inclusion and sum for subsets)**.** *(1) Associativity*

$$(S_1 + S_2) + S_3 = S_1 + (S_2 + S_3)$$

(2) Commutativity

$$S_1 + S_2 = S_2 + S_1$$

(3) If $S_1 \subseteq S_2$ and $T_1 \subseteq T_2$ then

$$S_1 + T_1 \subseteq S_2 + T_2$$

(4) If W is a subspace of V, then

$$W + \{\mathbf{0}\} = W, \quad W + W = W$$

(5) Suppose that W is a subspace of V, THen

$$S + W = W \Leftrightarrow S \subseteq W$$

Definition 6.2 (Coset $\bar{\mathbf{v}}$)**.**
Let V be a vector space over a field $\mathbb{F}$ and W a subspace of V. For any given $\mathbf{v} \in V$, the subset

$$\mathbf{v} + W := \{\mathbf{v} + \mathbf{w} \mid \mathbf{w} \in W\}$$

of V is called the **coset** of W containing $\mathbf{v}$. This subet is often denoted as

$$\bar{\mathbf{v}} := \mathbf{v} + W$$

The vector $\mathbf{v}$ is a **representative** of the coset $\bar{\mathbf{v}}$.

Theorem 6.2 (Coset Relations)**.**
Let W be a subspace of a vector space V. The following are equivalent

(1) $\mathbf{v} + W = W$,i.e.,$\bar{\mathbf{v}} = \bar{\mathbf{0}}$

(2) $\mathbf{v} \in W$

(3) $\mathbf{v} + W \subseteq W$

(4) $W \subseteq \mathbf{v} + W$

Theorem 6.3 (To be the same coset).
Let W be a subspace of V. Then for $\bar{\mathbf{v_i}} = \mathbf{v_i} + W$,

$$\bar{\mathbf{v_1}} = \bar{\mathbf{v_2}} \Leftrightarrow \mathbf{v_1} - \mathbf{v_2} \in W$$

Remark:Suppose that $V = U \oplus W$ is a direct sum of subspaces U and W. Then the map below is a bijection(and indeed an isomorphism)

$$\begin{aligned} f : U &\to V/W \\ \mathbf{u} &\mapsto \bar{\mathbf{u}} = \mathbf{u} + W \end{aligned}$$

Definition 6.3 (Quotient Space).
Let W be a subspace of V. Let

$$V/W := \{\bar{\mathbf{v}} = \mathbf{v} + W \mid \mathbf{v} \in V\}$$

be the set of all cosets of W. It is called the **quotient space** of V **modulo** W.
We define a binary addition operation on V/W:

$$\begin{aligned} + : V/W \times V/W &\to V/W \\ (\bar{\mathbf{v}}_1, \bar{\mathbf{v}}_2) &\mapsto \bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_2 := \overline{\mathbf{v}_1 + \mathbf{v}_2} \end{aligned}$$

and a scalar multiplication operation

$$\begin{aligned} \times : \mathbb{F} \times V/W &\to V/W \\ (a, \bar{\mathbf{v}}_1) &\mapsto a\bar{\mathbf{v}}_1 := a\bar{\mathbf{v}}_1 \end{aligned}$$

Theorem 6.4 (Quotient Space being Well Defined).
Let V be a vector space over a field $\mathbb{F}$ and W a vector subspace of V. Then we have:

(1) The binary addition operation and scalar multiplication operation on V/W is well defined.

(2) V/W together with these binary addition and scalar multiplication operations, becomes a vector space over the same field $\mathbb{F}$, with the zero vector

$$\mathbf{0}_{V/W} = \bar{\mathbf{0}}_V = \bar{\mathbf{w}}$$

for any $\mathbf{w} \in W$.

Definition 6.4 (Linear transformation, and its Kernel and Image; Isomorphism).
Let V_i be two vector spaces over the same field $\mathbb{F}$. A map

$$\varphi : V_1 \to V_2$$

is called a **linear transformation** from V_1 to V_2 if φ is compatible with the vector addition and scalar multiplication on V_1 and V_2 in the sense below:

$$\varphi(\mathbf{v}_1 + \mathbf{v}_2) = \varphi(\mathbf{v}_1) + \varphi(\mathbf{v}_2)$$
$$\varphi(a\mathbf{v}) = a\varphi(\mathbf{v})$$

When $\varphi : V \to V$ is a linear transformation from V to itself, we call φ a linear operator on V.
A linear transformation is called an **isomorphism** if it is a bijection. In this case, we denote

$$V_1 \simeq V_2$$

Remark: If $T : V \to W$ is a linear transformation, then $T(\mathbf{0}_V) = \mathbf{0}_W$.
Remark:(Direct sum vs quotient space)
Let $V = U \oplus W$, where U, W are subspaces of V. Then the map below is an isomorphism.

$$f : U \to V/W$$
$$\mathbf{u} \mapsto \bar{\mathbf{u}} = \mathbf{u} + W$$

Theorem 6.5 (Equivalent Linear Transformation definition)**.**
Let $\varphi : V_1 \to V_2$ be a map between two vector spaces V_i over the same field $\mathbb{F}$. The the following are equivalent.

(1) φ is a linear transformation.

(2) φ is compatible with taking linear combination in the sense below:

$$\varphi(a_1\mathbf{v}_1 + a_2\mathbf{v}_2) = a_1\varphi(\mathbf{v}_1) + a_2\varphi(\mathbf{v}_2)$$

for all $a_i \in \mathbb{F}, \mathbf{v}_i \in V$.

Theorem 6.6 (Evaluate T at a basis)**.**
Let V be a vector space over a field $\mathbb{F}$ and with a basis $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots\}$. Let $T : V \to W$ be a linear transformation.
Then T is uniquely determined by its valuations $T(\mathbf{u}_i) \quad (i = 1, 2, \ldots)$ at the basis B.
Namely, if $T' : V \to W$ is another linear transformation such that $T'(\mathbf{u}_i) = T(\mathbf{u}_i)\forall i$, then they are equal: $T' = T$.

Theorem 6.7 (Quotient map)**.**
Let V be a vector space over a field $\mathbb{F}$. Let W be a subspace V and V/W. One verifies that γ is surjective and

$$\ker(\gamma) = W$$

Theorem 6.8 (Image being a vector subspace)**.**
Let

$$\varphi : V \to W$$

be a linear transformation between two vector spaces over the same field $\mathbb{F}$. Let V_1 be a vector subspace of V. Then the image of V_1:

$$T(V_1) = \{T(\mathbf{u} \mid \mathbf{u} \in V_1\}$$

is a vector subspace of W.
In particular, $T(V)$ is a vector subspace of W.

Theorem 6.9 (Subspace vs. Kernel)**.**
Let V be a vector space over a field $\mathbb{F}$.

(1) Suppose that

$$\varphi : V \to U$$

is a linear transformation. Then the *kernel* $\ker(\varphi)$ is a vector subspace of V.

(2) Conversely, suppose W is a vector subspace of V. Then there is a linear transformation

$$\varphi : V \to U$$

such that

$$W = \ker(\varphi)$$

Theorem 6.10 (To be injective)**.** *Let*

$$\varphi : V \to W$$

be a linear transformation. Show that φ is injective if and only if $\ker(\varphi) = \{\mathbf{0}\}$.

Theorem 6.11 (Equivalent Isomorphism Definition)**.**
Let $\varphi : V \to W$ be a linear transformation. Then there is an isomorphism

$$\begin{aligned} \bar{\varphi} : V/\ker(\varphi) &\simeq \varphi(V) \in U \\ \bar{\mathbf{v}} &\mapsto \varphi(\mathbf{v}) \end{aligned}$$

such that

$$\varphi = \bar{\varphi} \circ \gamma$$

where

$$\gamma : V \to V/\ker(\varphi)\mathbf{v} \quad \mapsto \bar{\mathbf{v}}$$

is the quotient map, a linear transformation.
In particular, when φ is surjective, we have an isomorphism

$$\bar{\varphi} : V/\ker(\varphi) \simeq U$$

Theorem 6.12 (Finding basis of the quotient)**.**
Let V be a vector space over a field $\mathbb{F}$ of finite dimension n. Let W be a subspace with a basis $B_1 = \{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$.

(1) B_1 extends to a basis

$$B := B_1 \coprod \{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$$

of V.

(2) The cosets

$$\{\bar{\mathbf{w}}_1, \ldots, \bar{\mathbf{w}}_r\}$$

is a basis of the quotient space V/W. In particular,

$$\dim_{\mathbb{F}} V/W = \dim_{\mathbb{F}} V - \dim_{\mathbb{F}} W$$

(3)

$$B_1 \coprod \{\mathbf{u}_{r+1}, \ldots, \mathbf{u}_n\}$$

is a basis of V if and only if the cosets

$$\{\bar{\mathbf{u}_{r+1}}, \ldots, \bar{\mathbf{u}_n}\}$$

is a basis of V/W.

Theorem 6.13 (Goodies of Isomorphism)**.**
Let $\varphi : V \to W$ be an isomorphism and let B be a subset of V. Then we have:

(1) If there is a relation

$$\sum_{i=1}^{r} a_i \mathbf{v}_i = \sum_{i=r+1}^{s} a_i \mathbf{v}_i$$

among vectors $\mathbf{v}_i \in V$, then exactly the same relation

$$\sum_{i=1}^{r} a_i \varphi(\mathbf{v}_i) = \sum_{i=r+1}^{s} a_i \varphi(\mathbf{v}_i)$$

holds among vectors $\varphi(\mathbf{v}_i) \in W$. The converse is also true.

(2) B is linearly dependent if and only if so is $\varphi(B)$.

(3) B is linearly independent if and only if so is $\varphi(B)$;

(4) We have

$$\varphi(\mathrm{Span}(B)) = \mathrm{Span}(\varphi(B))$$

In particular,

$$\varphi(\mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_s\}) = \mathrm{Span}\{\varphi(\mathbf{v}_1), \ldots, \varphi(\mathbf{v}_s)\}$$

(5) B spans V if and only if $\varphi(B)$ spans W.

(6) B is a basis of V if and only if $\varphi(B)$ is a basis of W. In particular,

$$\dim V = \dim W$$

Theorem 6.14 (To be isomorphic finite-dimensional spaces)**.**
Let V and W be finite-dimensional vector spaces over the same field $\mathbb{F}$. Then the following are equivalent.

(1) $\dim_{\mathbb{F}} V = \dim_{\mathbb{F}} W = n$.

(2) There is an isomorphism

$$\varphi : V \simeq W$$

(3) For some n, we have:

$$V \simeq \mathbb{F}^n \simeq W$$

Theorem 6.15 (Dimension Theorem)**.**
Let $\varphi : V \to W$ be a linear transformation between vector spaces over a field $\mathbb{F}$. Then

$$\dim_{\mathbb{F}} \ker(\varphi) + \dim_{\mathbb{F}} \varphi(V) = \dim_{\mathbb{F}} V$$

Theorem 6.16 (2nd Isomorphism Theorem)**.**
Let W_1, W_2 be vector subspaces of a vector space V.

(1) The map

$$\varphi : W + 1/(W_1 + W_2)) \to (W_1 + W_2)/W_2$$
$$\mathbf{w} + W_1 \cap W_2 = \overline{\mathbf{w}} \mapsto \bar{(}\mathbf{w}) = \mathbf{w} + W_2$$

is a well difined isomorphism between vector spaces.

(2) A dimension formula:

$$\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2)$$

Theorem 6.17 (Equivalent isomorphism definition)**.**
Let

$$\varphi : V \to W$$

be a linear transformation between vector spaces over a field $\mathbb{F}$ and of the same finite dimension n. Then the following are equivalent.

(1) φ is an isomorphism

(2) φ is an injection

(3) φ is an surjection

7 Representation Matrices of Linear Transformations

Definition 7.1 (Coordinate vector)**.**
Let V be vector space of dimension $n \geq 1$ over a field $\mathbb{F}$. Let

$$B = B_V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$$

be a basis of V. Every vector $\mathbf{v} \in V$ can be expressed as a linear combination

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

and this expression is unique. We gather the coefficients c_i and form a column vector

$$[\mathbf{v}]_B := \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{F}_c^n$$

which is called the **coordinate vector** of $\mathbf{v}$ related to basis B.
One can recover $\mathbf{v}$ from its coordinate vector $[\mathbf{v}]_B$:

$$\mathbf{v} = B[\mathbf{v}]_B$$

Theorem 7.1 (Isomorphism $V \to \mathbb{F}_c^n$)**.**
Let V be an n-dimensional vector space over a field $\mathbb{F}$ and with a basis $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Show that the map

$$\begin{aligned} \varphi : V &\to \mathbb{F}_c^n \\ \mathbf{v} &\mapsto [\mathbf{v}]_B \end{aligned}$$

is an isomorphism between the vector space V and $\mathbb{F}_c^n$.

Theorem 7.2 (Representation matrix)**.**
Let

$$T : V \to W$$

be a linear transformation between vector spaces over a field $\mathbb{F}$. Let

$$B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$$

be a basis of V, and

$$B_W = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$$

be a basis of W. Let $A \in \mathbb{M}_{m\times n}(\mathbb{F})$. Then the following three conditions on A are equivalent.

$$[T(\mathbf{v})]_{B_W} = A[\mathbf{v}]_B$$

$$A = ([T(\mathbf{v}_1)]_{B_W}, \ldots, [T(\mathbf{v}_n)]_{B_W})$$

$$(T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)) = (\mathbf{w}_1, \ldots, \mathbf{w}_m)A$$

We denote the above matrix A as

$$[T]_{B,B_W} := A = ([T(\mathbf{v}_1)]_{B_W}, \ldots, [T(\mathbf{v}_n)]_{B_W})$$

and call it the **representation matrix** of T relative to B and B_W.

Theorem 7.3 (Linear transformation theory = Matrix theory)**.**
For every matrix

$$A \in \mathbb{M}_{m\times n}(\mathbb{F})$$

there is a unique linear transformation

$$T : V \to W$$

such that the representation matrix

$$[T]_{B,B_W} = A$$

Consequently, the map

$$\begin{aligned} \varphi : \mathrm{Hom}_{\mathbb{F}}(V, W) &\to \mathbb{M}_{m\times n}(\mathbb{F}) \\ T &\mapsto [T]_{B,B_W} \end{aligned}$$

is an isomorphism of vector spaces over $\mathbb{F}$.

Theorem 7.4 (Close relation between the space of vectors and space of their coordinates)**.**
Let

$$T : V \to W$$

be a linear transformation between the vector spaces V and W over the same field $\mathbb{F}$, of dimensions n and m, respectively. Let

$$B_V := (\mathbf{v}_1, \dots, \mathbf{v}_n)$$

be a basis of V, and

$$B_W := (\mathbf{w}_1, \dots, \mathbf{w}_m)$$

be a basis of W. Let

$$M_{m\times n}(\mathbb{F}) \ni A := [T]_{B_V,B_W} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$$

Then the following are isomorphisms:

$$\begin{aligned} \varphi : \mathrm{Ker}(T) &\to \mathrm{Null}(A) \\ \mathbf{v} &\mapsto [\mathbf{v}]_{B_V}, \\ \psi : \mathrm{Null}(A) &\to \mathrm{Ker}(T) \\ X &\mapsto B_V X, \\ \xi : \mathrm{R}(T) &\to \mathrm{R}(T_A) = \text{col.sp. of } A = \mathrm{Span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\} \\ \eta : \mathrm{R}(T_A) &\to \mathrm{R}(T) \\ Y &\mapsto B_W Y = (\mathbf{w}_1, \dots, \mathbf{w}_m)Y \end{aligned}$$

Below are some consequences of the isomorphism above

(1) The subset

$$\{X_1, \dots, X_s\}$$

of $\mathbb{F}_c^n$ is a basis of $\mathrm{Null}(A)$ if and only if the vectors

$$B_V X_1, \dots, B_V X_S$$

of V forms a basis of $\mathrm{Ker}(T)$.

(2) The subset
$$\{Y_1, \dots, Y_t\}$$
of $\mathbb{F}_c^m$ is a basis of $\mathrm{R}(T_A)$ if and only if the vectors
$$B_W Y_1, \dots, B_W Y_t$$
of W form a basis of $\mathrm{R}(T)$.

(3) The range of T is given by
$$\mathrm{R}(T) = \mathrm{Span}\{B_W \mathbf{a}_1, \dots, B_W \mathbf{a}_n\}$$

(4) $T : V \to W$ is an isomorphism if and only if its representation matrix $A = [T]_{B_V, B_W}$ is an invertible matrix in $M_n(\mathbb{F})$.

Theorem 7.5 (Representation Matrix of a Composite Map)**.**
Let V_1, V_2, V_3 be vector spaces of finite dimension over the same field $\mathbb{F}$ and let B_1, B_2, B_3 be theire respective bases. Let
$$T_1 : V_1 \to V_2$$
and
$$T_2 : V_2 \to V_3$$
be linear transformations. Then we have:
$$[T_2 \circ T_1]_{B_1,B_3} = [T_2]_{B_2,B_3}[T_1]_{B_1,B_2}$$

Theorem 7.6 (Representation matrix of inverse of an isomorphism)**.**
Let
$$T : V \to W$$
be an isomorphism between vector spaces over the same field $\mathbb{F}$ and of finite dimension. Let
$$T^{-1} : W \to V$$
be the inverse isomorphism of T. Let B_V(resp. B_W) be a basis of V(resp. W). Then
$$[T^{-1}]_{B_W,B_V} = [T_{B_V,B_W}]^{-1}$$

Theorem 7.7 (Representation matrix of map combination)**.**
Let
$$T_i : V \to W$$
be two linear transformations between finite-dimensional vector spaces over the same field $\mathbb{F}$. Let B(resp. B_W) be a basis of V(resp. W). Then for any $a_i \in \mathbb{F}$, the map linear combination $a_1T_1 + a_2T_2$ has the representation matrix
$$[a_1T_1 + a_2T_2]_{B,B_W} = a_1[T_1]_{B,B_W} + a_2[T_2]_{B,B_W}$$

Theorem 7.8 (Equivalent transition matrix definition)**.**
Let V be a vector space over a field $\mathbb{F}$ and of finite dimension $n \geq 1$. Let

$$B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$$

and

$$B' := (\mathbf{v}'_1, \dots, \mathbf{v}'_n)$$

be two bases of V. Let $P \in M_n(\mathbb{F})$. Then the following are equivalent.

(1)
$$P = ([\mathbf{v}'_1]_B, \dots, [\mathbf{v}'_n]_B)$$

(2)
$$B' = BP$$

(3) For any $\mathbf{v} \in V$, we have
$$P[\mathbf{v}]_{B'} = [\mathbf{v}]_B$$

This P is denoted as $P_{B' \to B}$ and called the **transition matrix** from basis B' to B. P is invertible.

Theorem 7.9 (Basis change theorem for representation matrix)**.**
Let V be a vector space over a field $\mathbb{F}$ and of finite dimension $n \geq 1$. Let

$$B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$$

and

$$B' := (\mathbf{v}'_1, \dots, \mathbf{v}'_n)$$

be two bases of V. Then

$$[T]_{B'} = P^{-1}[T]_B P$$

where

$$P = P_{B' \to B}$$

Definition 7.2 (Similar Matrices)**.**
Two square matrices (of the same order) $A_1, A_2 \in \mathbb{M}_n(\mathbb{F})$ are **similar** if there is an invertible matrix $P \in \mathbb{M}_n(\mathbb{F})$ such that

$$A_2 = P^{-1} A_1 P$$

In this case, we denote

$$A_1 \sim A_2$$

The similarity property is an equivalence relation.

Theorem 7.10. Similar matrices have the same determinant:

$$A_1 \sim A_2 \Rightarrow |A_1| = |A_2|$$

Definition 7.3 (Determinant/Trace of a linear operator)**.**
Let

$$T : V \to V$$

be a linear operator on a finite-dimensional vector space V.
We define the **determinant** $\det(T)$ of T as

$$\det(T) := \det([T]_B)$$

and the **trace** of T as

$$\mathrm{Tr}(T) = \mathrm{Tr}([T]_B)$$

where B is any basis of V.

Definition 7.4 (Characteristic polynomial $p_A(x), p_T(x)$)**.** (1) Let $A \in \mathbb{M}_n(\mathbb{F})$.

$$\begin{aligned} p_A(x) :&= |xI_n - A| \\ &= x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0 \end{aligned}$$

is called the **characteristic polynomial** of A, which is of degree n.

(2) Let

$$T : V \to V$$

be a linear operator on an n-dimensional vector space V. Set

$$A := [T]_B$$

where B is any basis of V. Then

$$\begin{aligned} p_T(x) :&= |xI_n - A| \\ &= x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0 \end{aligned}$$

is called the **characteristic polynomial** of T, which is of degree $n = \dim V$.

Theorem 7.11. Similar matrices have equal characteristic polynomial.

Theorem 7.12.
For $A \in \mathbb{M}_n(\mathbb{F})$, we have

$$\mathrm{Tr}(A) = -b_{n-1}$$

$$\det(A) = (-1)^n p_A(0)$$

8 Eigenvalue and Cayley-Hamilton Theorem

Definition 8.1 (Eigenvalue, eigenvector)**.**
Assume that

$$\lambda \in \mathbb{F}$$

(arabic*) Let V be a vector space over a field $\mathbb{F}$. Let

$$T : V \to V$$

be a linear operator. A *nonzero* vector $\mathbf{v}$ in V is called an **eigenvector** of T corresponding to the eigenvalue $\lambda \in \mathbb{F}$ of T if

$$T(\mathbf{v}) = \lambda \mathbf{v}$$

(arabic*) For an $n \times n$ matrix A in $\mathbb{M}_n(\mathbb{F})$, a nonzero column vector $\mathbf{u}$ in $\mathbb{F}_c^n$ is called an **eigenvector** of A corresponding to the eigenvalue $\lambda \in \mathbb{F}$ of A if

$$A\mathbf{u} = \lambda \mathbf{u}$$

Definition 8.2 (Equivalent definition of eigenvalue and eigenvector)**.**
Let V be a vector space of dimension n over a field $\mathbb{F}$ and with a basis B, Let

$$T : V \to V$$

be a linear operator. Assume that

$$\lambda \in \mathbb{F}$$

Then the following are equivalent:

(1) λ is an eigenvalue of T (corresponding to an eigenvector $\mathbf{0} \neq \mathbf{v} \in V$ of T, i.e. $T(\mathbf{v}) = \lambda \mathbf{v}$).

(2) λ is an eigenvalue of $[T]_B$ (corresponding to an eigenvector $\mathbf{0} \neq [\mathbf{v}]_B \in \mathbb{F}_c^n$ of $[T]_B$, i.e. $[T]_B[v]_B = \lambda[\mathbf{v}]_B$).

(3) The linear operator

$$\begin{aligned} \lambda I_V - T : V &\to V \\ \mathbf{x} &\mapsto \lambda \mathbf{x} - T(\mathbf{x}) \end{aligned}$$

is not an isomorphism, i.e. there is some

$$0 \neq \mathbf{v} \in \mathrm{Ker}(\lambda I_V - T)$$

(4) The matrix $\lambda I_n - [T]_B$ is not invertiblem i.e. the matrix equation

$$(\lambda I_n - [T]_B)X = 0$$

has a non=trivial solution.

(5) λ is a zero of the characteristic polynomial $p_T(x)$ of T

$$p_T(\lambda) = |\lambda I_n - [T]_B| = 0$$

Theorem 8.1 (Determinant $|A|$ as product of eigenvalues).
Let $A \in \mathbb{M}_n(\mathbb{F})$. Let $p(x)$ be the characteristic polynomial. Factorise

$$p(x) = (x - \lambda_1) \cdots (x - \lambda_n)$$

in some over field of $\mathbb{F}$. Then the determinant of A equals

$$\prod_{i=1}^{n} \lambda_i$$

Definition 8.3 (Eigenspace of an eigenvalue).
Let $\lambda \in \mathbb{F}$ be an eigenvalue of a linear operator

$$T : V \to V$$

on an n-dimensional vector space V over the field $\mathbb{F}$. The subspace (of all the eigenvectors corresponding to the eigenvalue λ, plus $\mathbf{0}_V$):

$$\begin{aligned} V_\lambda :&= V_\lambda(T) \\ :&= \operatorname{Ker}(\lambda I_V - T) \\ &= \{\mathbf{v} \in V \mid T(\mathbf{v}) = \lambda \mathbf{v}\} \end{aligned}$$

of V is called the **eigenspace** of T corresponding to the eigenvalue λ.

Definition 8.4 (Geometric/Algebraic Multiplicity).
Let $\lambda \in \mathbb{F}$ and $T : V \to V$ be as the previous definition.

(1) The dimension

$$\dim V_\lambda$$

of the eigenspace V_λ of T is called the **geometrix multiplicity** of the eigenvalue λ of T. We have

$$1 \leq \dim V_\lambda \leq n$$

(2) The **algebraic multiplicity** of the eigenvalue λ of T is defined to be the largest positive integer k such that $(x - \lambda)^k$ is a **factor** of the characteristic polynomial $p_T(x)$,i.e.

$$(x - \lambda)^k \mid p_T(x), \quad (x - \lambda)^{k+1} \nmid p_T(x)$$

We shall see that

$$\text{geometric multiplicity of } \lambda \leq \text{alg. multiplicity of } \lambda$$

Theorem 8.2 (Eigenspace of T and $[T]_B$)**.**
Let

$$T : V \to V$$

be a linear operator on an n-dimensional vector space V with a basis B. Set

$$A := [T]_B$$

The map

$$f : \mathrm{Ker}(T - \lambda I_V) \to \mathrm{Null}(A - \lambda I_n)$$
$$\mathbf{w} \mapsto [\mathbf{w}]_B$$

gives an isomorphism. In particular,

$$\dim V_lambda(T) = \dim V_\lambda(A)$$

Due to this isomorphism, the following are equivalent:

1. The subset
$$\{\mathbf{u}_1, \cdots, \mathbf{u}_s\}$$
of V is a basis of the eigenspace $V_\lambda(T)$ of T.

2. THe subset
$$\{[\mathbf{u}_1]_B, cdots, [\mathbf{u}_s]_B\}$$
of $\mathbb{F}_c^n$ is a basis of the eigenspace $V_\lambda([T]_B)$ of the representation matrix $[T]_B$ of T relative to a basis B of V.

Also, the following are equivalent:

1. The subset
$$\{X_1, \cdots, X_s\}$$
of $\mathbb{F}_c^n$ is a basis of the eigenspace $V_\lambda([T]_B)$ of the representation matrix $[T]_B$ of T relative to a basis B of V.

2. The subset
$$\{BX_1, \cdots, BX_s\}$$
of V is a basis of the eigenspace $V_\lambda(T)$ of T.

Theorem 8.3 (Eigenspaces of similar matrices)**.**
Let $A \in \mathbb{M}_n(\mathbb{F})$. Suppose that $P^{-1}AP = C$. Show that

$$\mathbb{F}_c^n \supseteq V_\lambda(A) = PV_\lambda(C) := \{PX \mid X \in V_\lambda(C)\}$$

Theorem 8.4 (Sum of eigenspaces)**.**
Let

$$\lambda_1, \ldots, \lambda_k$$

be some distinct eigenvalues of a linear operator T on a vector space V over a field $\mathbb{F}$. Then the sum of eigenspaces

$$\begin{aligned} W &:= \sum_{i=1}^{k} V_{\lambda_i}(T) \\ &= V_{\lambda_i}(T) + \cdots + V_{\lambda_k}(T) \end{aligned}$$

is a direct sum:

$$\begin{aligned} W &= \oplus_{i=1}^{k} V_{\lambda_i}(T) \\ &= V_{\lambda_1}(T) \oplus \cdots \oplus V_{\lambda_k}(T) \end{aligned}$$

Definition 8.5 (Multiplication of linear operators $S_1, \ldots, S_r$)**.**
Let

$$T : V \to V$$

be a linear operator on a vector space V over a field $\mathbb{F}$. Define

$$T^s := T \circ \cdots \circ T \quad (s \text{ times})$$

which is a linear operator on V.

$$\begin{aligned} T^s : V &\to V \\ \mathbf{v} &\mapsto T^s(\mathbf{v}) \end{aligned}$$

Be convention, set

$$T^0 := I_V = \mathrm{id}_V$$

More generally, for a polynomial

$$f(x) = \sum_{i=0}^{r} a_i x^i$$

Define

$$f(T) = \sum_{i=0}^{r} a_i T^i$$

Then $f(T)$ is a linear operator on V.

$$\begin{aligned} f(T) : V \& to V \\ \mathbf{v} &\mapsto f(T)(\mathbf{v}) \end{aligned}$$

Similarly, for linear operators

$$S_i : V \to V$$

Define

$$S_1 S_2 \cdots S_r := S_1 \circ S_2 \circ \cdots \circ S_r$$

which is a linear operator.

Theorem 8.5 (Polynomials in T)**.**
Let

$$T : V \to V$$

be a linear operator on a vector space V over a field $\mathbb{F}$ and with a basis

$$B = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$$

Let

$$f(x), g(x) \in \mathbf{F}[x]$$

be polynomials. We have

(1)
$$[f(T)]_B = f([T]_B)$$

(2) The multiplication $f(T)g(T)$ as polynomials in T equals the composite $f(T) \circ g(T)$ as linear operators:
$$f(T)g(T) = f(T) \circ g(T)$$

(3) Commutativity:
$$f(T)g(T) = g(T)f(T)$$

(4) If $P \in \mathbb{M}_n(\mathbb{F})$ is invertible, then

$$f(P^{-1}AP) = P^{-1}f(A)P$$

(5) If
$$S : V \to V$$
is an isomorphism with inverse isomorphism
$$S^{-1} : V \to V$$
Then
$$f(S^{-1}TS) = S^{-1}f(T)S$$

Definition 8.6 (T-invariant Subspace)**.**
Let

$$T : V \to V$$

be a linear operator on a vector space V. A subspace W of V is called T-**invariant** if the image of W under the map T is included in W:

$$T(W) := \{T(\mathbf{w}) \mid \mathbf{w} \in W\}$$

i.e.,

$$T(\mathbf{w}) \in W \ \ \forall \mathbf{w} \in W$$

In this case, define the **restriction** of T on W as:

$$\begin{aligned} T \mid W : W &\to W \\ \mathbf{w} &\mapsto T(\mathbf{w}) \end{aligned}$$

Theorem 8.6 (Kernels and Images of Commutative Operators)**.**
Let $T_i : V \to V$ be two linear operator commutative to each other, i.e.

$$T_1 \circ T_2 = T_2 \circ T_1$$

as maps. Namely,

$$T_1(T_2(\mathbf{v})) = T_2(T_1(\mathbf{v})) \quad (\forall \mathbf{v} \in V)$$

Both

$$\mathrm{Ker}(T_2), \ \ ima(T_2)$$

are T_1-invariant subspaces of V.

Theorem 8.7 (Evaluate T on a basis of a subspace)**.**
Let

$$T : V \to V$$

be a linear operator on a vector space V over a field $\mathbb{F}$. Let W be a subspace of V with a basis $B_W = \{\mathbf{w}_1, \mathbf{w}_2, \cdots\}$. Then W is T-invariant, if and only if

$$T(B_W) \subseteq W$$

Theorem 8.8 (T-cyclic subspace)**.**
Let

$$T : V \to V$$

be a linear operator on a vector space V over a field $\mathbb{F}$. Fix a vector

$$\mathbf{0} \neq \mathbf{w}_1 \in V$$

1. The subspace

   $$W := \mathrm{Span}\{T^s(\mathbf{w}_1 \mid s \geq 0\}$$

   of V is T-invariant.
   W is called the T-cyclic subspace of V generated by $\mathbf{w}_1$.

2. Suppose that V is finite-dimensional. Let s be the smallest positive integer such that

   $$T^s(\mathbf{w}_1) \in \mathrm{Span}\{\mathbf{w}_1, T(\mathbf{w}_1), \dots, T^{s-1}(\mathbf{w}_1)\}$$

   We have

   $$\dim_{\mathbb{F}} W = s$$

   and

   $$B := \{\mathbf{w}_1, T(\mathbf{w}_1), \dots, T^{s-1}(\mathbf{w}_1)\}$$

   is a basis of W.

3. In (2), if

   $$T^s(\mathbf{w}_1) = c_0\mathbf{w}_1 + c_1 T(\mathbf{w}_1) + \cdots + c_{s-1}T^{s-1}(\mathbf{w}_1)$$

   for some scalars $c_i \in \mathbb{F}$, then the characteristic polynomial of the restiction operator $T \mid W$ on W is

   $$p_{T|W}(x) = -c_0 - c_1 x - \cdots - c_{s-1}x^{s-1} + x^s$$

Theorem 8.9 (Characteristic Polynomial of the Restriction Operator)**.**
Let

$$T : V \to V$$

be a linear operator on a vector space V over a field $\mathbb{F}$ and of dimension $n \geq 1$. Let W be a T-invariant subspace of V. Then the characteristic polynomial $p_{T|W}(x)$ of the restriction operator $T \mid W$ on W is a factor of the characteristic polynomial $p_T(x)$ of T, i.e.

$$p_T(x) = q(x)p_{T|W}(x)$$

for some polynomial $q(x) \in \mathbb{F}[x]$.

Theorem 8.10 (To be T-invariant in terms of $[T]_B$)**.**
A subspace W of an n-dimensional space V is T-invariant for a linear operator T on V, if and only if every basis B_W of W can be extended to a basis

$$B = B_W \cup B_2$$

of V such that the representation matrix of T relative to B, is of the form:

$$[T]_B = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix}$$

for some square matrices A_1, A_3 (automatically with $A_1 = [T \mid W]_{B_W}$).
In this case, the matrix

$$A_2 = 0$$

if and only if

$$W_2 := \mathrm{Span}(B_2)$$

is a T-invariant subspace of V (automatically with $[T \mid W_2]_{B_2} = A_3$)

Theorem 8.11 (Upper Triangular Form of a Matrix)**.**
Let $T : V \to V$ be a linear operator on an n-dimensional vector space V over a field $\mathbb{F}$. Suppose that the characteristic polynomial $p(x)$ is factorised as

$$p(x) = (x - \lambda_1)^{n_1} \cdots (x - \lambda_k)^{n_k}$$

for some $\lambda_i \in \mathbb{F}$. Then there is an basis B of V such that the representation matrix $[T]_B$ is upper triangular.

Theorem 8.12 (Characteristic Polynomials of Direct Sums)**.**
Let

$$T : V \to V$$

be a linear operator on an n-dimensional vector space V over a field $\mathbb{F}$. Suppose that there are T-invariant subspaces

$$W_i \quad (1 \leq i \leq r)$$

of V such that V is the direct sum

$$V = \oplus_{i=1}^{r} W_i$$

of W_i.
Then the characteristic polynomial $p_T(x)$ of T is the product:

$$p_T(x) = \prod_{i=1}^{r} p_{T|W_i}(x)$$

of the characteristic polynomials of the restriction operators $T \mid W_i$ on W_i.

Theorem 8.13 (To be direct sum of T-invariant subspaces)**.**
An n-dimensional vector space V with a linear operator T is a direct sum

$$V = \oplus_{i=1}^{r} W_i$$

of some T-invariant subspaces W_i, if and only if every set of bases B_i of W_i gives rises to a basis

$$B = B_1 \coprod \cdots \coprod B_r$$

of V such that the representation matrix of T relative to B is in the form

$$[T]_B = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & A_r \end{pmatrix}$$

with A_i of order $|B_i| = \dim W_i$

Theorem 8.14 (Cayley-Hamilton Theorem)**.**
Let

$$p_T(x) = |xI_n - [T]_B| = \sum_{i=0}^{n} b_i x^i$$

be the characteristic polynomial of a linear operator

$$T : V \to V$$

on an n-dimensional vector space V over a field $\mathbb{F}$ and with a basis B. Then T satisfies the equation $p_T(x) = 0$, i.e.

$$p_T(T) = 0I_V$$

which is the zero map on V.

9 Minimal Polynomial and Jordan Canonical Form

Definition 9.1 (Minimal Polynomial)**.**
Let

$$T : V \to V$$

be a linear operator on an n-dimensional vector space over a field $\mathbb{F}$. A nonzero polynomial

$$m(x) \in \mathbb{F}[x]$$

is a **minimal polynomial of** T if it satisfies:

1. $m(x)$ is monic,
2. Vanishing condition:
$$m(T) = 0I_V$$
3. Minimality degree condition:
Whenever $f(x) \in \mathbb{F}[x]$ is another nonzero polynomial such that $f(T) = 0I_V$, we have
$$\deg(f(x)) \geq \deg(m(x))$$

We can define a **minimal polynomial of** a matrix $A \in \mathbb{M}_N(\mathbb{F})$.

Remark: The existence of minimal polynomial is proven by Cayley-Hamilton theorem.

Theorem 9.1 (Uniqueness of a minimal polynomial $m_T(x)$)**.**
Let $T : V \to V$ be a linear operator on an n-dimensional vector space V over a field $\mathbb{F}$. Let $m(x)$ be a minimal polynomial of T. Let $f(x) \in \mathbb{F}[x]$. Then the following are equivalent.

(1) $f(T) = 0I_V$.

(2) $m(x)$ is a factor of $f(x)$,i.e., $m(x) \mid f(x)$.

In particular, there is exactly one minimal polynomial of T and will be denoted as

$$m_T(x) = m(x)$$

Further, if $A = [T]_B$, then $m_T(x) = m_A(x)$.

Theorem 9.2 (Minimal polynomials of similar matrices)**.**
If two matrices A_i are simialr: $A_1 \sim A_2$, then they have the same minimal polynomial

$$m_{A_1}(x) = m_{A_2}(x)$$

Theorem 9.3 (Minimal polynomials of direct sums)**.**
Consider the matrix

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix}$$

where $A_i \in \mathbb{M}_{n_i}(\mathbb{F})$ are square matrices. The minimal polynomial $m_A(x)$ of A is equal to the **least common multiple** of the minimal polynomials $m_{A_i}(x)$ of A_i,i.e.,

$$m_A(x) = \operatorname{lcm}\{m_{A_1}(x), \ldots, m_{A_r}(x)\}$$

Theorem 9.4.
The set of zeros of $p_T(x)$ and that of $m_T(x)$ are identical.

Definition 9.2 (Jordan Block)**.**
Let λ be a scalar in a field $\mathbb{F}$. The matrix below

$$J := J_s(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix} \in \mathbb{M}_s(\mathbb{F})$$

is called the **Jordan Block** of order s with eigenvalue λ.
The characteristic polynomial and minimal polynomial of J are identical:

$$m_J(x) = (x-\lambda)^s = p_J(x)$$

The eigenspace

$$V_\lambda(J) = \operatorname{Span}\{\mathbf{e}_1\}$$

has dimension 1, i.e. the geometric multiplicity of λ is 1, but the algebraic multiplicity of λ is s.

Definition 9.3 (Jordan Canonical Form)**.**
Let λ be a nonzero scalar in a field $\mathbb{F}$. Let

$$s_1 \le s_2 \le \cdots \le s_e$$

The following **Block Diagonal**

$$A(\lambda) = \begin{pmatrix} J_{s_1}(\lambda) & 0 & 0 & \cdots & 0 \\ 0 & J_{s_2}(\lambda) & 0 & \cdots & 0 \\ 0 & 0 & J_{s_3}(\lambda) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & J_{s_e}(\lambda) \end{pmatrix}$$

is called a **Jordan canonical form with eigenvalue** λ.
The order of $A(\lambda)$ is

$$s = \sum_{i=1}^{e} s_i$$

The characteristic polynomial and minimal polynomial of A are

$$p_{A(\lambda)}(x) = (x-\lambda)^s, \quad m_{A(\lambda)}(x) = (x-\lambda)^{s_e}$$

where s is also called algebraic multiplicity of λ of $A(\lambda)$.
The eigenspace of A

$$V_\lambda(A(\lambda)) = \operatorname{Span}\{\mathbf{e}_1, \mathbf{e}_{1+s_1}, \ldots, \mathbf{e}_{1+s_1+\cdots+s_{e-1}}\}$$

has dimension equal to e.
The geometric multiplicity of the eigenvalue λ of $A(\lambda)$ is $\dim V_\lambda(A(\lambda)) = e$. And we have,

$$e \leq s$$

More generally, let

$$\lambda_1, \dots, \lambda_k$$

be distinct scalars in $\mathbb{F}$. WLOG, we assume $\lambda_1 < \lambda_2 < \cdots < \lambda_k$ when $\mathbb{F} = \mathbb{R}$. Then the block diagonal

$$J = \begin{pmatrix} A(\lambda_1) & 0 & 0 & \cdots & 0 \\ 0 & A(\lambda_2) & 0 & \cdots & 0 \\ 0 & 0 & A(\lambda_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & A(\lambda_k) \end{pmatrix}$$

is called a **Jordan canonical form** where $A(\lambda_i)$ is a Jordan canonical form with eigenvalue λ_i as shown above.
Each $A(\lambda_i)$ is of order

$$s(\lambda_i)$$

and $s(\lambda_i)$ is also the number of times the same scalar λ_i appears on the diagonal of J and also the algebraic multiplicity of the eigenvalue λ_i of J.
So J is of order euqal to

$$\sum_{i=1}^{k} s(\lambda_i)$$

There are exactly

$$e(\lambda_i)$$

Jordan blocks (with eigenvalue λ_i) in $A(\lambda_i)$, the largest of which is of order

$$s_e(\lambda_i)$$

This $s_e(\lambda_i)$ is also the multiplicity of λ_i in the minimal polynomial $m_J(x)$.
There are exactly

$$\sum_{i=1}^{k} e(\lambda_i)$$

Jordan blocks in J.
Now we have

$$p_J(x) = \prod_{i=1}^{k} (x - \lambda_i)^{s(\lambda_i)}$$

$$m_J(x) = \prod_{i=1}^{k} (x - \lambda_i)^{s_e(\lambda_i)}$$

The eigenspace $V_{\lambda_i}(J)$ has dimension $e(\lambda_i)$ and is spanned by the $e(\lambda_i)$ ectors corresponding to the first columns of the $e(\lambda_i)$ Jordan blocks in $A(\lambda_i)$.
We also have

$$\dim V_{\lambda_i}(J) = e(\lambda_i) \leq s(\lambda_i)$$

Sometimes, a block diagonal J below

$$J = \begin{pmatrix} J_{s_1}(\lambda_1) & 0 & 0 & \cdots & 0 \\ 0 & J_{s_2}(\lambda_2) & 0 & \cdots & 0 \\ 0 & 0 & J_{s_3}(\lambda_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & J_{s_r}(\lambda_r) \end{pmatrix}$$

is also called a **Jordan Canonical Form**, where each $J_{s_i}(\lambda_i)$ is a Jordan block with eigenvalue $\lambda_i \in \mathbb{F}$, but these λ_i's may not be distinct.
Assume that there are exactly k distinct elements in the set

$$\{\lambda_1, \ldots, \lambda_r\}$$

and we assume that

$$\lambda_{m_i}$$

are these k distinct ones. These k of λ_{m_i} are just the distinct eigenvalues of J.
Let

$$s(\lambda_{m_i})$$

be the number of times the same scalar λ_{m_i} appears on the diagonal of J. Let

$$e(\lambda_{m_i})$$

be the number of Jordan blocks (among the r such in J) with eigenvalue of the same λ_{m_i}; among these $e(\lambda_{m_i})$ Jordan blocks, the largest is of order say

$$s_e(\lambda_{m_i})$$

The eigenspace $V_{\lambda_{m_i}}(J)$ has dimension $e(\lambda_{m_i})$ and is spanned by $e(\lambda_{m_i})$ vectors corresponding to the first columns of these $e(\lambda_{m_i})$ Jordan blocks.
Also,

$$p_J(x) = \prod_{i=1}^{k} (x - \lambda_i)^{s(\lambda_{m_i})}$$

$$m_J(x) = \prod_{i=1}^{k} (x - \lambda_i)^{s_e(\lambda_{m_i})}$$

As in the case of $A(\lambda)$, for the matrix J, we have

$$\dim V_{\lambda_{m_i}}(J) = e(\lambda_{m_i}) \leq s(\lambda_{m_i})$$

Theorem 9.5 (Jordan Canonical Form of a Linear Operator).
Let V be a vector space of dimension n over a field $\mathbb{F}$ and

$$T : V \to V$$

a linear operator with characteristic polynomial $p_T(x)$ and minimal polynomial $m_T(x)$ as follows

$$p_T(x) = (x - \lambda_1)^{n_1} \cdots (x - \lambda_k)^{n_k}$$
$$m_T(x) = (x - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k}$$

where

$$\lambda_1, \ldots, \lambda_k$$

are distinct scalars in $\mathbb{F}$.
Then there is a basis B of V such that the representative matrix $[T]_B$ equals a Jordan canonical form $J \in \mathbb{M}_n(\mathbb{F})$, with

$$n_i = s(\lambda_i), \qquad m_i = s_e(\lambda_i)$$

Such a block diagonal J is called a Jordan canonical form of T. It is unique up to re-ordering of λ_i. The basis B of V is called a Jordan canonical basis of T.

Theorem 9.6 (A canonical form of T is a canonical form of $[T]_B$).
Let V be an n-dimensional vector space over a field $\mathbb{F}$ and

$$T : V \to V$$

a linear operator. Let

$$A = [T]_B'$$

be the representation matrix of T relative to a basis

$$B' = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$$

of V. Let J be a Jordan canonical form. Then the following are equivalent:

(1) There is an invertible matrix $P \in \mathbb{M}_n(\mathbb{F})$ such that

$$P^{-1}AP = J$$

(2) THere is an invertible matrix $P \in \mathbb{M}_n(\mathbb{F})$ such that the representation matrix $[T]_B$ relative to the new basis

$$B = B'P$$

is J, i.e.

$$[T]_B = J$$

Theorem 9.7 (Existence of Jordan Canonical Form).
Let $\mathbb{F}$ be a field. Let A be a matrix in $\mathbb{M}_n(\mathbb{F})$. Let $p(x) = p_A(x)$ be the characteristic polynomial. Then the following is equivalent.

(1) A has a Jordan canonical form $J \in \mathbb{M}_n(\mathbb{F})$.

(2) Every zero of the characteristic polynomial $p(x)$ belongs to $\mathbb{F}$.

(3) We can factor $p(x)$ as

$$p(x) = (x - \lambda_1) \cdots (x - \lambda_n)$$

where all $\lambda_i \in \mathbb{F}$.

In particular, if $\mathbb{F}$ is so called **algebraically closed**, then every matrix $A \in \mathbb{M}_n(\mathbb{F})$ and every T on an n-dimensional vector space V over $\mathbb{F}$ have a Jordan canonical form $j \in \mathbb{M}_n(\mathbb{F})$.

Theorem 9.8 (Consequences of Jordan canonical forms)**.**
Let A be a matrix in $\mathbb{M}_n(\mathbb{F})$. Set $p(x) = p_A(x)$ and $m(x) = m_A(x)$.

1. The characteristic polynomial $p(x)$ and the minimal polynomial $m(x)$ have the same zero sets.

$$\{\alpha \in \mathbb{F} \mid p(\alpha) = 0\} = \{\alpha \in \mathbb{F} \mid m(\alpha) = 0\}$$

Also, the multiplicity n_i and m_i of a zero λ_i of $p(x)$ and $m(x)$ satisfy

$$n_i \geq m_i \geq 1$$

2. If $J \in \mathbb{M}_n(\mathbb{F})$ is a Jordan canonical form of A, then we have

$$\dim V_{\lambda_i}(A) = \dim V_{\lambda_i}(J) = e(\lambda_i) \leq s(\lambda_i)$$

Theorem 9.9 (Canonical forms of similar matrices)**.**
Let $A_i \in \mathbb{M}_n(\mathbb{F})$ and $J_i \in \mathbb{M}_n(\mathbb{F})$ be its Jordan canonical form. Then the following are equivalent.

(1) A_1 and A_2 are similar.

(2) We have $J_1 = J_2$ after re-ordering of their Jordan Block.

Definition 9.4 (Diagonalisable Operator)**.**
Let V be an n-dimensional vector space over a field $\mathbb{F}$. A linear operator $T : V \to V$ is **diagonalisable** over $\mathbb{F}$, if the representation matrix $[T]_B$ relative to some basis B of V is a diagonal matrix in $\mathbb{M}_n(\mathbb{F})$:

$$[T]_b = J = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

where λ_i are scalars in $\mathbb{F}$. This J is then an automatically a Jordan canonical form of T. Clearly,

$$\lambda_1, \ldots, \lambda_n$$

exhaust all zeros of $p_T(x)$ and the characteristic polynomial of T is

$$p_T(x) = (x - \lambda_1) \cdots (x - \lambda_n)$$

A square matrix $A \in \mathbb{M}_n(\mathbb{F})$ is diagonalisable over $\mathbb{F}$, if A is similar to a diagonal matrix in $\mathbb{M}_n(\mathbb{F})$, i.e.

$$P^{-1}AP = J = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

for some invertible $P \in \mathbb{M}_n(\mathbb{F})$, where λ_i are scalars in $\mathbb{F}$. This J is then automatically a Jordan canonical form of A.
Write

$$P = (\mathbf{p}_1, \ldots, \mathbf{p}_n)$$

with $\mathbf{p}_j$ the jth column of P.
The diagonalisability condition on A is equivalent to

$$AP = P \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

i.e.

$$(A\mathbf{p}_1, \ldots, A\mathbf{p}_n) = (\lambda_1\mathbf{p}_1, \ldots, \lambda_n\mathbf{p}_n)$$

i.e. each $\mathbf{p}_i$ is an eigenvector of A corresponding to the eigenvalue λ_i.
Suppose that $A = [T]_{B'}$. Then the condition above is equivalent to

$$[T(\mathbf{v}_i)]_{B'} = [T]_{B'}[\mathbf{v}_i]_{B'} = \lambda_i[\mathbf{v}_i]_{B'}$$

where $\mathbf{v}_i = B'\mathbf{p}_i \in V$ with

$$[\mathbf{v}_i]_{B'} = \mathbf{p}_i$$

i.e.

$$T(\mathbf{v}_i) = \lambda_i\mathbf{v}_i$$

i.e.

$$T(\mathbf{v}_1, \ldots, \mathbf{v}_n) = (\mathbf{v}_1, \ldots, \mathbf{v}_n) \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

i.e.

$$[T]_B = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

where

$$B = (\mathbf{v}_1, \ldots, \mathbf{v}_n) = B'P$$

is a basis of V since

$$([\mathbf{v}_1]_{B'}, \ldots, [\mathbf{v}_n]_{B'}) = (\mathbf{p}_1, \ldots, \mathbf{p}_n)$$

is a basis of column vector space $\mathbb{F}_c^n$.

Theorem 9.10.

(1) T is diagonalisable if and only if the representation matrix $[T]_{B'}$ relative to every basis B' is diagonalisable.

(2) A matrix $A \in \mathbb{M}_n(\mathbb{F})$ is diagonalisable if and only if the matrix transformation T_A on the column n-space $\mathbb{F}_c^n$ is diagonalisable.

Theorem 9.11 (Equivalent Diagonalisable Condition)**.**
Let V be an n-dimensional vector space over a field $\mathbb{F}$, and

$$T : V \to V$$

a linear operator. Then the following are equivalent:

1. T i s diagonalisable over $\mathbb{F}$, i.e. the representation matrix of T relative to some basis B of V is a diagonal matrix in $\mathbb{M}_n(\mathbb{F})$.

$$[T]_B = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

2. $[T]_{B'}$ is diagonalisable over $\mathbb{F}$ for every basis B' of V, i.e. there exists an invertible $P \in \mathbb{M}_n(\mathbb{F})$ such that

$$P^{-1}[T]_{B'}P = \text{diag}[\lambda_1, \ldots, \lambda_n]$$

for some scalars $\lambda_i \in \mathbb{F}$(automatically being eigenvalues of T).

3. A basis

$$B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$$

of V is formed by eigenvectors $\mathbf{v}_i$ of T.

4. There are n linearly independent eigenvectors $\mathbf{v}_i$ of T.

5. For the representation matrix $[T]_{B'}$ relative to every basis B' of V, a basis

$$P = (\mathbf{p}_1, \ldots, \mathbf{p}_n)$$

of the column n-space $\mathbb{F}_c^n$ is formed by eigenvectors $\mathbf{p}_i$ of $[T]_{B'}$.

6. For the representation matrix $[T]_{B'}$ relative to every basis B' of V, there are n linearly independent eigenvectors $\mathbf{p}_i$ of $[T]_{B'}$.

7. Let

$$\lambda_{m_1}, \ldots, \lambda_{m_k}$$

be the only distinct eigenvalues of T and let B_i be a basis of the eigenspace $V_{\lambda_i}(T)$. Then

$$B = (B_1, \ldots, B_k)$$

is a basis of V, automatically with

$$[T]_B = \begin{pmatrix} \lambda_{m_1} I_{|B_1|} & 0 & 0 & \cdots & 0 \\ 0 & \lambda_{m_2} I_{|B_2|} & 0 & \cdots & 0 \\ 0 & 0 & \lambda_{m_3} I_{|B_3|} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_{m_k} I_{|B_k|} \end{pmatrix}$$

8. Let

$$\lambda_{m_1}, \ldots, \lambda_{m_k}$$

be the only distinct eigenvalues of T. Then V is a direct sum of the eigenspaces

$$V = V_{\lambda_{m_i}}(T) \oplus \cdots \oplus V_{\lambda_{m_k}}(T)$$

9. Let

$$\lambda_{m_1}, \ldots, \lambda_{m_k}$$

be the only distinct eigenvalues of T. Then

$$\sum_{i=1}^{k} \dim V_{\lambda_{m_i}}(T) = \dim V$$

10. T has a Jordan canonical form J which is diagonal.

Theorem 9.12 (Minimal polynomial and diagonalisability)**.**
Let $\mathbb{F}$ be a field. Let A be a matrix in $\mathbb{M}_n(\mathbb{F})$. Let $m(x) = m_A(x)$ be minimal polynomial of A. Then the following are equivalent:

(1) A is diagonalisable over $\mathbb{F}$.

(2) The minimal polynomial $m(x)$ is a product of distinct linear polynomials in $\mathbb{F}[x]$.

$$m(x) = (x - \lambda_1) \cdots (x - \lambda_k)$$

where λ_i are distinct scalars in $\mathbb{F}$

(3) We can factor $m(x)$ over $\mathbb{F}$ as

$$m(x) = (x - \lambda_1) \cdots (x - \lambda_k)$$

for some scalars $\lambda_i \in \mathbb{F}$ and $m(x)$ has only simple zeros.

(4) Let $p(x) = p_A(x)$ be the characteristic polynomial. Then we can factorise $p(x)$ over $\mathbb{F}$ as

$$p(x) = (x - \lambda_1)^{n_1} \cdots (x - \lambda_k)^{n_k}$$

where λ_i are distinct scalars in $\mathbb{F}$. The dimension of the eigenspace satisfies:

$$\dim V_{\lambda_i} = n_i$$

Theorem 9.13.
Let V be an n-dimensional vector space over a field $\mathbb{F}$.
A linear operator

$$T : V \to V$$

on V is **nilpotent** if

$$T^m = 0I_V$$

for some positive integer m.
Suppose that T has a Jordan canonical form $J \in \mathbb{M}_n(\mathbb{F})$. The following are equivalent.

(1) T is nilpotent.

(2) J equals some $A(\lambda)$ with $\lambda = 0$.

(3) Every eigenvalue of T is zero/

(4) The characteristic polynomial of T is $p_T(x) = x^n$.

(5) The minimal polynomial of T is $m_T(x) = x^s$ for some $s \geq 1$.

Theorem 9.14 (Additive Jordan Decomposition)**.**
Suppose a linera operator

$$T : V \to V$$

has a Jordan canonical form in $\mathbb{M}_n(\mathbb{F})$. There are linear operators

$$T_S : V \to V$$

and

$$T_n : V \to V$$

satisfying the following:

(1) A decomposition

$$T = T_s + T_n$$

(2) T_s is semi-simple.

(3) T_n is nilpotent.

(4) Commutativity

$$T_s \circ T_n = T_n \circ T_s$$

(5) There are polynomials $f(x), g(x)$ in $\mathbb{F}[x]$ such that

$$T_s = f(T) \qquad T_n = g(T)$$

This decomposition is unique. We call it Jordan decomposition.

10 Quadratic Forms, Inner Product Spaces and Conics

Definition 10.1 (Bilinear forms)**.**
Let V be a vector space over a field $\mathbb{F}$. Consider the map H below:

$$\begin{aligned} H : V \times V &\to \mathbb{F} \\ (\mathbf{x}, \mathbf{y}) &\mapsto H(\mathbf{x}, \mathbf{y}) \end{aligned}$$

(1) H is called a **bilinear form on** V if H is linear in both variables, i.e., for all

$$\mathbf{x}_i, \mathbf{y}_j, \mathbf{x}, \mathbf{y} \in V,\ a_i, b_i \in \mathbb{F}$$

we have

$$\begin{aligned} H(a_1\mathbf{x}_1 + a_2\mathbf{x}_2, \mathbf{y}) &= a_1 H(\mathbf{x}_1, y) + a_2 H(\mathbf{x}_2, y) \\ H(\mathbf{x}, b_1\mathbf{y}_1 + b_2\mathbf{y}_2) &= b_1 H(\mathbf{x}, \mathbf{y}_1) + b_2 H(\mathbf{x}, \mathbf{y}_2) \end{aligned}$$

(2) A bilinear form H on V is **symmetric** if

$$H(\mathbf{x}, \mathbf{y}) = H(\mathbf{y}, \mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in V$$

Theorem 10.1 (Representation Matrix)**.**
Suppose that

$$B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$$

is a basis of a vector space V over a field $\mathbb{F}$. Let

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

be a matrix in $mathbbM_n(\mathbb{F})$.
We define the function

$$\begin{aligned} H_A : V \times V &\to \mathbb{F} \\ (\sum_{i=1}^{n} x_i\mathbf{v}_i, \sum_{j=1}^{n} y_i\mathbf{v}_j) \mapsto H_A(\sum_{i=1}^{n} x_i\mathbf{v}_i, \sum_{j=1}^{n} y_i\mathbf{v}_j) \end{aligned}$$

where

$$\begin{aligned} &H_A(\sum_{i=1}^{n} x_i\mathbf{v}_i, \sum_{j=1}^{n} y_i\mathbf{v}_j) \\ :=&\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j \\ =&(x_1, \dots, x_n) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ =&X^tAY \end{aligned}$$

(1) Then H_A is a bilinear form on V and called the **bilinear form associated with** A (and relative to the basis B of V).

(2) Conversely, every bilinear form H on V is of the form H_A for some A in $\mathbb{M}_n(\mathbb{F})$. Indeed, just set

$$a_{ij} = H(\mathbf{v}_i, \mathbf{v}_j), \quad A := (a_{ij})$$

Then one can use the bilinearity of H, show that $H = H_A$.
The matrix A is called the representation matrix of H relative to the basis of B¿

(3) H_A is a symmetric bilinear form if and only if A is a **symmetric matrix**.

Definition 10.2 (Non-degenerate bilinear forms)**.**
A bilinear form H on V is **non-degenerate** if for every $\mathbf{y}_0 \in V$, we have:

$$H(\mathbf{x}, \mathbf{y}_0) = 0 (\forall \mathbf{x} \in V) \Rightarrow \mathbf{y}_0 = \mathbf{0}$$

A bilinear form $H = H_A$ is non-degenerate if and only if its representation matrix A is invertible.

Definition 10.3 (Congruent matrices)**.**
Two matrices A and B in $\mathbb{M}_n(\mathbb{F})$ are **congruent** if there is an invertible matrix $P \in \mathbb{M}_n(\mathbb{F})$ such that

$$B = P^t A P$$

Being congruent is an equivalent relation.
Consider the bilinear form

$$\begin{aligned} H : \mathbb{F}_c^n \times \mathbb{F}_c^n &\to \mathbb{F} \\ (X, Y) &\mapsto X^t A Y \end{aligned}$$

If we write

$$X = PY$$

with an invertible matrix $P \in \mathbb{M}_n(\mathbb{F})$ and introduce Y as a new coordinate system for $\mathbb{F}_c^n$, then

$$\begin{aligned} H(X_1, X_2) &= X_1^t A X_2 \\ &= (PY_1)^t A (PY_2) \\ &= Y_1^t (P^t A P) Y_2 \end{aligned}$$

Thus, the bilinear form above would have a simpler form in new coordinates Y, if $P^t AP$ (which is congruent to A) is simpler. This simplification is very useful in classfying all conics.

Theorem 10.2 (Weak version of Principle Axis Theorem)**.**
Let $A \in \mathbb{M}_n(\mathbb{F})$ be a symmetric matrix. Then there is an invertible matrix P in $\mathbb{M}_n(\mathbb{F})$ such that the matrix $P^t AP$ is diagonal:

$$P^t A P = \operatorname{diag}[d_1, \ldots, d_n] =: D$$

i.e. A is congruent to a diagonal matrix D.
In this case, the bilinear form

$$\begin{aligned}
H(X_1, X_2) &= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j \\
&= X_1^t A X_2 \\
&= Y_1^t D Y_2 \\
&= (y_{11}, \ldots, y_{1n}) \begin{pmatrix} d_1 & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & d_n \end{pmatrix} \begin{pmatrix} y_2 1 \\ \vdots \\ y_2 n \end{pmatrix} \\
&= \sum_{j=1}^{n} d_j y_{1j} y_{2j}
\end{aligned}$$

where we have used the substitution:

$$X_i = PY_i$$

Definition 10.4 (Inner Product, Orthogonal, Norm)**.**
We start with the real version.
Consider a function H:

$$\begin{aligned}
V \times V &\to \mathbb{R} \\
(\mathbf{x}, \mathbf{y}) &\mapsto H(\mathbf{x}, \mathbf{y})
\end{aligned}$$

on a vector space V over the field $\mathbb{R}$ of real numbers.
The function H is called a **real inner product** and V a **real inner product space**, if the following three conditions are satisfied, where we denote

$$\langle \mathbf{x}, \mathbf{y} \rangle := H(\mathbf{x}, \mathbf{y})$$

(1) H is a bilinear form, i.e., for all

$$\mathbf{x}_i, \mathbf{y}_j, \mathbf{x}, \mathbf{y} \in V, a_i, b_i \in \mathbb{R}$$

we have

$$\begin{aligned}
\langle a_1\mathbf{x}_1 + a_2\mathbf{x}_2, \mathbf{y} \rangle &= a_1 \langle \mathbf{x}_1, \mathbf{y} \rangle + a_2 \langle \mathbf{x}_2, \mathbf{y} \rangle \\
\langle \mathbf{x}, b_1\mathbf{y}_1 + b_2\mathbf{y}_2 \rangle &= b_1 \langle \mathbf{x}, \mathbf{y}_1 \rangle + b_2 \langle \mathbf{x}, \mathbf{y}_2 \rangle
\end{aligned}$$

(2) H is symmetric, i.e., for all $\mathbf{x}, \mathbf{y} \in V$, we have

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$$

(3) Positivity:
For all $\mathbf{0} \neq \mathbf{x} \in V$, we have

$$\langle \mathbf{x}, \mathbf{x} \rangle > 0$$

Next is the complex version. Consider a function H:

$$V \times V \to \mathbb{C}$$
$$(\mathbf{x}, \mathbf{y}) \mapsto H(\mathbf{x}, \mathbf{y})$$

on a vector space V over the field $\mathbb{C}$ of real numbers.
The function H is called a **complex inner product** and V a **complex inner product space**, if the following three conditions are satisfied, where we denote

$$\langle \mathbf{x}, \mathbf{y} \rangle := H(\mathbf{x}, \mathbf{y})$$

(1) H is a bilinear form, i.e., for all

$$\mathbf{x}_i, \mathbf{y}_j, \mathbf{x}, \mathbf{y} \in V, a_i, b_i \in \mathbb{R}$$

we have

$$\langle < a_1\mathbf{x}_1 + a_2\mathbf{x}_2, \mathbf{y} \rangle = a_1\langle \mathbf{x}_1, \mathbf{y} \rangle + a_2\langle \mathbf{x}_2, \mathbf{y} \rangle$$
$$\langle < \mathbf{x}, b_1\mathbf{y}_1 + b_2\mathbf{y}_2 \rangle = \bar{b}_1\langle \mathbf{x}, \mathbf{y}_1 \rangle + \bar{b}_2\langle \mathbf{x}, \mathbf{y}_2 \rangle$$

(2) H is symmetric, i.e., for all $\mathbf{x}, \mathbf{y} \in V$, we have

$$\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$$

(3) Positivity:
For all $\mathbf{0} \neq \mathbf{x} \in V$, we have

$$\langle \mathbf{x}, \mathbf{x} \rangle > 0$$

We have three more definitions:

(1) The **norm** of a vector $\mathbf{x} \in V$ is denoted and defined as:

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

We have

$$\|\mathbf{x}\| \geq 0$$

and

$$\|\mathbf{x}\| == 0 \Leftrightarrow \mathbf{x} = \mathbf{0}_V$$

(2) Two vectors $\mathbf{x}, \mathbf{y}$ in V are **orthogonal** to each other and denoted as

$$\mathbf{x} \perp \mathbf{y}$$

if their inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0$$

(3) Sometimes, we use

$$(V, \langle , \rangle)$$

to denote a vector space V with an inner product $\langle , \rangle$.

Definition 10.5 (Non-degenerate Inner Product)**.**
Let $(V, \langle, \rangle)$ be an inner product space over a field $\mathbb{F}$ with $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. Then the product $\langle, \rangle$ is **non-degenerate** in the sense:
for every $\mathbf{u}_0 \in V$

$$\langle \mathbf{u}_0, \mathbf{y} \rangle = 0 (\forall \mathbf{y} \in V) \Rightarrow \mathbf{u}_0 = \mathbf{0}_V$$

and for every $\mathbf{v}_0 \in V$,

$$\langle \mathbf{x}, \mathbf{v}_0 \rangle = 0 (\forall \mathbf{x} \in V) \Rightarrow \mathbf{v}_0 = \mathbf{0}_V$$

Definition 10.6 (Orthonormal basis)**.**
Let $(V, \langle, \rangle)$ be a real or complex inner product space. A basis $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n$ is called an orthonormal basis of the inner product space V, if it satisfies the following two conditions:

(1) **Orthogonality**:
for all $i \neq j$, we have:

$$\mathbf{v}_i \perp \mathbf{v}_j \quad \text{i.e.,} \quad \langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$$

(2) **Normalised**:
for all i, we have

$$\|\mathbf{v}_i\| = 1$$

Namely, $\mathbf{v}_i$ is a unit vector.

Theorem 10.3 (Gram-Schmidt Process)**.**
Let $V = \mathbb{F}_c^n$ with $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Employ the standard inner product $\langle, \rangle$ for V.
Let

$$(\mathbf{u}_1, \ldots, \mathbf{u}_r)$$

be a basis of a subspace W of V. Then one can apply the following **Gram-Schmidt process** to get an orthonormal basis

$$(\mathbf{v}_1, \ldots, \mathbf{v}_r)$$

of W.

$$\mathbf{v}'_1 = \mathbf{u}_1$$

$$\mathbf{v}'_2 = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{v}'_1 \rangle}{\|\mathbf{v}'_1\|^2} \mathbf{v}'_1$$

$$\mathbf{v}'_k = \mathbf{u}_k - \sum_{i=1}^{k-1} \frac{\langle \mathbf{u}_k, \mathbf{v}'_i \rangle}{\|\mathbf{v}'_i\|^2}$$

$$\mathbf{v}_j = \frac{\mathbf{v}'_j}{\left\|\mathbf{v}'_j\right\|}$$

Definition 10.7 (Adjoint matrices A^*)**.**
For a matrix $A = (a_{ij}) \in \mathbb{M}_n(\mathbb{C})$, the **adjoint** of A is defined as

$$A^* = (\overline{A})^t = (\overline{a_{ij}})^t$$

i.e., the (i, j)-entry of A^* equals $\overline{a_{ji}}$. Note that

$$A^* = \overline{(A^t)}$$

Theorem 10.4 (Adjoint matrix A^* and inner product)**.**
Let $V = \mathbb{F})_c^n$ with $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Employ the standard inner product $\langle , \rangle$ for V. For a matrix $A \in \mathbb{M}_n(\mathbb{F})$, we have

$$\langle AX, Y\rangle = \langle X, A^*Y\rangle$$

Theorem 10.5 (Adjoint Linear Operator)**.**
Let $T : V \to V$ be a linear operator on an n-dimensional inner product space V over a field $\mathbb{F}$. Then we have:

(1) There is a **unique** linear operator

$$T^* : V \to V$$

on V such that

$$\langle T(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, T^*(\mathbf{v})\rangle$$

Such T^* is called the **adjoint linear operator** of T.

(2) Let $B = (\mathbf{w}_1, \ldots, \mathbf{w}_n)$ be an orthonormal basis of the inner product space V. Then

$$[T^*]_B = ([T]_B)^*$$

Theorem 10.6 (Adjoint of adjoint)**.** $(T^*)^* = T$.

Theorem 10.7 (Adjoint of linear map combinations)**.**

(1) Suppose that $T = \alpha I_V$ is a scalar map. Then

$$T^* = \overline{\alpha} I_V$$

(2)

$$(a_1T_1 + a_2T_2)^* = \overline{a_1}T_1^* + \overline{a_2}T_2^*$$

(3)

$$(T_1 \circ T_2)^* = T_2^* \circ T_1^*$$

Definition 10.8 (Orthogonal, Unitary, Self-adjoint, Normal linear operators)**.**
Let $A \in \mathbb{M}_n(\mathbb{C})$ (resp. let $T : V \to V$ be a linear operator on an n-dimensional inner product space over a field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and with an orthonormal basis B). Let A^* (resp. T^*) be the adjoint of A (resp. T).

(1) A linear operator T over a real inner product space is **orthogonal** if

$$TT^* = I_V$$

(2) A real matrix A in $\mathbb{M}_n(\mathbb{R})$ is **orthogonal** if

$$AA^t = I_n$$

(3) A linear operator T over a complex inner product space is **unitary** if

$$TT^* = I_V$$

(4) A complex matrix A in $mathbbM_n(\mathbb{C})$ is **unitary** if

$$AA^* = I_n$$

(5) T is **self-adjoint** if its adjoint T^* equals itself:

$$T = T^*$$

When the field $\mathbb{F} = \mathbb{R}$, a self adjoint operator is also called a **symmetric** operator.

(6) A complex matrix $A \in \mathbb{M}_n(\mathbb{C})$ is **self-adjoint** if the adjoint matrix of A equals itself:

$$A^* = A$$

(7) A linear operator T over a complex inner product space is **normal** if

$$TT^* = T^*T$$

(8) A complex matrix $A \in \mathbb{M}_n(\mathbb{C})$ is **normal** if

$$AA^* = A^*A$$

Orthogonal, Unitary, self-adjoint operators are normal.

Theorem 10.8. T is orthogonal, unitary, self-adjoint or normal if and only if its representation matrix $A := [T]_B$ (relative to one hence every orthonormal basis B) is respectively orthogonal, unitary, self-adjoint and normal.

Theorem 10.9 (Equivalent unitary matrix definition)**.**
For a real matrix P in $\mathbb{M}_n(\mathbb{C})$, the following are equivalent, if we employ the standard inner product on $\mathbb{C}^n_c$.

(1) P is unitary, i.e. $PP^* = I_n$.

(2) Write

$$P = (\mathbf{p}_1, \ldots, \mathbf{p}_n)$$

where the $\mathbf{p}_j$ are the column vectors of P. Then the column vectors $\mathbf{p}_1, \ldots, \mathbf{p}_n$ form an orthonormal basis of $\mathbb{C}^n_c$.

(3) The matrix transformation

$$\begin{aligned} T_P : \mathbb{C}^n_c &\to \mathbb{C}^n_c \\ X &\mapsto PX \end{aligned}$$

preserves the standard inner product, i.e., for all X, Y in $\mathbb{C}^n_c$, we have

$$\langle PX, PY \rangle = \langle X, Y \rangle$$

(4) The matrix transformation T_P preserves the distance, i.e. for all X, Y in $\mathbb{C}_c^n$, we have

$$\|PX - PY\| = \|X - Y\|$$

(5) The matrix transformation T_P preserves the norm, i.e. for all X in $\mathbb{C}_c^n$, we have

$$\|PX\| = \|X\|$$

(6) For one and hence every orthogonal basis

$$B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$$

of $\mathbb{C}_c^n$, the new basis

$$B' = BP$$

is again an orthonormal basis of $\mathbb{C}_c^n$.

Theorem 10.10 (Eigenvalues of orthogonal or unitary matrices)**.**

(1) If a real matrix $P \in \mathbb{M}_n(\mathbb{R})$ is orthogonal, then every zero of $p_P(x)$ has modulus equal to 1. In particular, the determinant

$$|P| = \pm 1$$

(2) If a complex matrix $P \in \mathbb{M}_n(\mathbb{C})$ is unitary, then every eigenvalue

$$\lambda_i = r_1 + r_2\sqrt{-1}$$

of P has **modulus**

$$|\lambda| = \sqrt{r_1^2 + r_2^2} = 1$$

In particular, the determinant $|P| \in \mathbb{C}$ has modulus 1.

Theorem 10.11 (Eigenvalue of self-adjoint linear operators)**.**

(1) Suppose that a real matrix $A \in \mathbb{M}_n(\mathbb{R})$ is symmetric. Every zero of $p_A(x)$ is a real number.

(2) Suppose a complex matrix $A \in \mathbb{M}_n(\mathbb{C})$ is self-adjoint. Every zero of $p_A(x)$ is a real number.

(3) More generally, suppose that T is a self adjoint linear operator. Every zero of $p_T(x)$ is a real number.

(4) Suppose that T is a self-adjoint linear operator. Let $\mathbf{v}_i (i = 1, 2)$ be two eigenvectors corresponding to two distinct eigenvalues λ_i of T. We have

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$$

Definition 10.9 (Positive/Negative definite linear operators)**.**
Let $A \in \mathbb{M}_n(\mathbb{C})$ (resp. let V be an n-dimensional inner product space which is over a field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$).

(1) T is **positive definite** if T is self-adjoint and

$$\langle T(\mathbf{v}), \mathbf{v}\rangle > 0$$

(2) T is **negative definite** if T is self-adjoint and

$$\langle T(\mathbf{v}), \mathbf{v}\rangle < 0$$

Thus T is negative definite if and only if $-T$ is positive definite.

(3) A is **positive definite** if A is self-adjoint and

$$(AX)^t\overline{X} = X^t A^t \overline{X} > 0$$

(4) A is **negative definite** if A is self-adjoint and

$$(AX)^t\overline{X} = X^t A^t \overline{X} < 0$$

Thus, A is negative definite if and only if $-A$ is positive definite.

Theorem 10.12 (Equivalent Positive-Definite Definition)**.**
Let

$$A = (a_{ij}) \in \mathbb{M}_n(\mathbb{R})$$

be a symmetric real matrix. Then A is positive definite if and only if all its **principal minors**

$$(a_{ij})_{1\le i,j\le r}(1 \le r \le n)$$

of order r have positive determinants.
Let T be a self-adjoint linera operator on an inner product space V which is over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and with an orthonormal basis B. Set

$$A := [T]_B \in \mathbb{M}_n(\mathbb{F})$$

Then the following are equivalent.

(1) T is positive definite.

(2) A is positive definite.

(3) Every eigenvalue of T is positive.

(4) Every eigenvalue of A is positive.

(5) One can write A as

$$A = C^*C$$

for some invertible complex matrix $C \in \mathbb{M}_n(\mathbb{C})$

Theorem 10.13. Let $A \in \mathbb{M}_n(\mathbb{C})$. The function H on $V := \mathbb{C}^n_c$

$$\begin{aligned} H : V \times V &\to \mathbb{C} \\ (X, Y) &\mapsto \langle X, Y \rangle := (AX)^t \overline{Y} \end{aligned}$$

defines an inner product on V if and only if A is positive definite.

Theorem 10.14 (Principle Axis Theorem)**.**

1. Let $T : V \to V$ be a linear operator on a real inner product space V of dimension n. Then T is self-adjoint (i.e., $T^* = T$) if and only if there is an orthonormal basis B such that
$$[T]_B$$
is a diagonal matrix in $\mathbb{M}_n(\mathbb{R})$.

2. A real matrix $A \in \mathbb{M}_n(\mathbb{R})$ is self-adjoint (i.e. $A^* = A$) if and only if there is an orthogonal matrix P such that
$$P^{-1}AP = P^tAP$$
is a diagonal matrix in $\mathbb{M}_n(\mathbb{R})$.

3. Let $T : V \to V$ be a linear operator on a complex inner product space V of dimension n. Then T is self-adjoint (i.e., $T^* = T$) if and only if there is an orthonormal basis B such that
$$[T]_B$$
is a diagonal matrix in $\mathbb{M}_n(\mathbb{C})$.

4. A complex matrix $A \in \mathbb{M}_n(\mathbb{C})$ is self-adjoint (i.e. $A^* = A$) if and only if there is an unitary matrix U such that
$$U^{-1}AU = U^*AU$$
is a diagonal matrix in $\mathbb{M}_n(\mathbb{C})$.

Theorem 10.15 (Orthogonal Complement)**.**
Let W be a subspace of an inner product space V. Take an orthogonal basis B_W of W.

(1) One can extend B_W to an orthonormal basis $B = (B_W, B_2)$ of V.

(2) B_2 is an orthonormal basis of so called **orthogonal complement** of W:
$$W^\perp := \{\mathbf{x} \in V \mid \langle \mathbf{x}, \mathbf{v} \rangle = 0, \forall \mathbf{w} \in W\}$$

(3)
$$V = W \oplus W^\perp$$

Definition 10.10 (Quadratic Form)**.**
Let V be a vector space over a field $\mathbb{F}$. A function

$$K : V \to \mathbb{F}$$

or simply $K(\mathbf{x})$ is a **quadratic form** if there is a symmetric bilinear form

$$H : V \times V \to \mathbb{F}$$

such that

$$K(\mathbf{x}) = H(\mathbf{x}, \mathbf{x})$$

Theorem 10.16 (Principle Axis Theorem of Quadratic Form)**.**
Let

$$f(x_1, \dots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

be a quadratic form in coordinates

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

with

$$A = (a_{ij}) \in \mathbb{M}_n(\mathbb{R})$$

a symmetric matrix. Then there is an orthogonal matrix P such that f has the following **standard form**

$$f(x_1, \dots, x_n) = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$

in the new coordinates

$$Y := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = P^{-1} X$$

where $\lambda_i \in \mathbb{R}$ are the eigenvalues of A.
This standard form is unique up to relabelling of $\lambda_i y_i^2$.

11 Problems

1 Let $A \in \mathbb{M}_n(\mathbb{C})$ be a complex matrix of order $n \geq 9$ and let

$$f(x) := (x-1)^2(x-2)^3(x-3)^4$$

Suppose that A is self-adjoint and $f(A) = 0$. Find all possible minimal polynomials $m_A(x)$ of A.

2 Let V be a finite-dimensional inner product space and $T : V \to V$ invertible. Prove that there exists a unitary operator U and a positive operator P on V such that $T = U \circ P$.

3 AY1314Sem2 Question 6(iii)

4 AY1415Sem2 Question 8(iv) –(vi)

5 Let V be a finite dimensional vector space over a field $\mathbb{F}$ and let T be a linear operator on V. Suppose there exists $v \in V$ such that $\{\mathbf{v}, T(\mathbf{v}), \ldots, T^{n-1}(\mathbf{v})\}$ is a basis for V where $n = \dim(V)$.

(a) Prove that the linear operators $I_V, T, \ldots, T^{n-1}$ are linearly independent. (Done)

(b) Let S be a linear operator on V such that $S \circ T = T \circ S$. Write

$$S(\mathbf{v}) = a_0\mathbf{v} + a_1 T(\mathbf{v}) + \cdots + a_{n-1}T^{n-1}(\mathbf{v})$$

where $a_0, a_1, \cdots, a_{n-1} \in \mathbb{F}$.
Prove that $S = p(T)$ where $p(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}$.

(c) Suppose $p_T(x) = (x-\lambda_1)^{r_1}(x-\lambda_2)^{r_2} \cdots (x-\lambda_k)^{r_k}$ where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of T. Find $m_T(x)$.

6 Let A be an invertible $n \times n$ matrix over a field $\mathbb{F}$.

(a) Show that $c_{A^{-1}}(x) = x^n[c_A(0)]^{-1}c_A(\frac{1}{x})$ (Done)

(b) Show that $m_{A^{-1}}(x) = x^k[m_A(0)]^{-1}m_A(\frac{1}{x})$.