# Linear Algebra

Notes from TAU Course with Additional Information

Gabriel Domingues

2020-02-03

## Contents

# 1 Sets

## 1.1 Set Theory (Succinctly) and Logic

There will be no definition of a set. Instead, we postulate the existence of a relation $\in$ (read as "is in").

It is axiom the existence of the empty set $\varnothing$, that is, $\exists \varnothing : \forall x\,,\ x \notin \varnothing$

**Definition 1.1.1** (Principle of Double Inclusion)**.** *We define the following symbols:*

*Inclusion:* $A \subseteq B \Leftrightarrow \forall x\,,\ x \in A \Rightarrow\ x \in B$

*Equality:* $A = B \Leftrightarrow A \subseteq B$ *and* $B \subseteq A \Leftrightarrow \forall x\,,\ x \in A \Leftrightarrow\ x \in B$

*Proper Inclusion:* $A \underset{\neq}{\subsetneq} B \Leftrightarrow A \subseteq B$ *and* $A \neq B$

It is also axiom the existence of the power set: Given a set $A$, there is a set $\mathcal{P}(A)$ so that: $x \in \mathcal{P}(A) \Leftrightarrow x \subseteq A$

We can create new sets by the Principle of Restricted Comprehension: Let $A$ be a set and $P$ a predicate (given an object, it is either True or False), then the following is a set: $\{x \in A \mid P(x)\}$

**Example 1.1.1** (Set Difference)**.** *Let $A$ and $B$ be sets. Construct:* $A \setminus B = \{x \in A \mid x \notin B\}$

We can construct the sets we will mainly use:

Natural Numbers $\mathbb{N} = \{0, 1, 2, \cdots\}$

Integer Numbers $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$

Rational Numbers $\mathbb{Q} = \left\{ \dfrac{m}{n} \;\middle|\; m \in \mathbb{Z} \text{ and } n \in \mathbb{N} \setminus \{0\} \right\}$

Real Numbers $\mathbb{R}$

## 1.2 Operations on Sets

**Definition 1.2.1** (Set Operations). *For sets $A$ and $B$ these are sets (by axiom):*

$$\text{Union}: A \cup B = \{x \mid x \in A \text{ or } x \in B\}$$

$$\text{Intersection}: A \cap B = \{x \mid x \in A \text{ and } x \in B\}$$

$$\text{Cartesian Product}: A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}$$

**Definition 1.2.2** (Operation). *An operation $*$ on a set $A$ is a map:*

$$\begin{aligned} * : A \times A &\to B \\ (a, b) &\mapsto a * b \end{aligned}$$

*This operation can have (or lack) multiple properties. These are the most important:*

| Properties | Definition |
|---|---|
| Closed | $\forall\, a, b \in A\,,\ a * b \in A$ |
| Commutative | $\forall\, a, b \in A\,,\ a * b = b * a$ |
| Associative | $\forall\, a, b, c \in A\,,\ (a * b) * c = a * (b * c)$ |
| Neutral Element | $\exists\, e \in A\ :\ \forall\, a \in A\,,\ a * e = e * a = a$ |
| Inverse Element | $\forall\, a \in A\,,\ \exists\, b \in A\ :\ a * b = b * a = e$ |

**Definition 1.2.3** (Equivalence Relation). *An equivalence relation on a set $X$ is a predicate of two variables (has the value True or False), denoted $x \sim y$, that has these three properties:*

| | |
|---|---|
| Reflexivity | $\forall\, x \in X\,,\ x \sim x$ |
| Symmetry | $\forall\, x, y \in X\,,\ x \sim y\ \Leftrightarrow\ y \sim x$ |
| Transitivity | $\forall\, x, y, z \in X\,,\ x \sim y\ \wedge\ y \sim z \Rightarrow x \sim z$ |

## 1.3 Axioms of Field

**Definition 1.3.1** (Field). *A field $F$ is a set with operations* $(+ : F \times F \to F, \cdot : F \times F \to F)$ *iff:*

| Properties | Definition |
|---|---|
| Commutative | $\forall\,\alpha, \beta \in F\,,\ \alpha + \beta = \beta + \alpha$ |
| Associative | $\forall\,\alpha, \beta, \gamma \in F\,,\ (\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ |
| Neutral Element | $\exists\,0 \in F\ :\ \forall\,\alpha \in F\,,\ \alpha + 0 = \alpha$ |
| Inverse Element | $\forall\,\alpha \in F\,,\ \exists\,\beta \in F\ :\ \alpha + \beta = 0$ |
| Associative | $\forall\,\alpha, \beta, \gamma \in F\,,\ \alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$ |
| Distributive Right | $\forall\,\alpha, \beta, \gamma \in F\,,\ \alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$ |
| Distributive Left | $\forall\,\alpha, \beta, \gamma \in F\,,\ (\alpha + \beta) \cdot \gamma = \alpha \cdot \gamma + \beta \cdot \gamma$ |
| Unital Element | $\exists\,1 \in F\ :\ \forall\,\alpha \in F\,,\ 1 \cdot \alpha = \alpha$ |
| Inverse Element | $\forall\,\alpha \in F \setminus \{0\}\,,\ \exists\,\beta \in F\ :\ \alpha \cdot \beta = 1$ |
| Commutative | $\forall\,\alpha, \beta \in F\,,\ \alpha \cdot \beta = \beta \cdot \alpha$ |

Examples include $\mathbb{Q}$ and $\mathbb{R}$ with addition and multiplication of numbers.

# 2 Linear Equations

## 2.1 Linear Equations over a Field

**Definition 2.1.1** (Linear Equation)**.** *A linear equation (over $F$) in the tuple* $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in F^n$*, where we define* $F^n = \overbrace{F \times F \times \cdots \times F}^{n \text{ times}}$*. is something of the form:*

$$E \ : \ \sum_{i=1}^{n} a_i \cdot x_i = a_1 \cdot x_1 + a_2 \cdot x_2 + \cdots + a_n \cdot x_n = b$$

*where $a_i, b \in F$. We denote the solution to $E$ as* $\mathrm{sols}(E)$*. That is,*

$$\mathrm{sols}(E) := \left\{ \underline{x} \in F^n \ \middle| \ \sum_{i=1}^{n} a_i \cdot x_i = b \right\}$$

**Definition 2.1.2** (Homogeneous Equation)**.** *We say a linear equation $H$ is homogeneous if $b = 0$.*

**Definition 2.1.3** (Linear Equation System)**.** *A system of $k$ linear equations $L$ is a sequence of linear equations $(E_1, E_2, \cdots, E_k)$, for which we need to find $\underline{x} \in F^n$ that satisfies all equations. That is,*

$$\mathrm{sols}(L) = \bigcap_{i=1}^{k} \mathrm{sols}(E_i) = \mathrm{sols}(E_1) \cap \mathrm{sols}(E_2) \cap \cdots \cap \mathrm{sols}(E_k)$$

## 2.2 Gaussian Elimination

**Definition 2.2.1** (Leading Variable)**.** *We say $x_i$ the leading variable of the linear equation* $E \ : \ \sum_{k=1}^{n} a_i \cdot x_i$ *iff $a_i \neq 0$ and $\forall\, j \in \{1, \cdots, i\}$, $a_j = 0$.*

**Definition 2.2.2** (Canonical Form/ Row Reduced Echelon Form). *A linear system is said to be in canonical form if:*

1. *The sequence of leading variables strictly increases*
2. *Equations without leading variables (zero equation) come at the end*
3. *Each leading variable appears only in one equation with coefficient* 1

*A system in the canonical form is very easy to solve since we can isolate the leading variables and make the others (if exist) free variables.*

**Example 2.2.1.**

$$\begin{cases} x_1 + x_2 + x_3 & = 3 \\ 2x_1 - x_2 + 5x_3 & = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 + 0 + 2x_3 & = 1 \\ 0 + x_2 - x_3 & = 2 \end{cases}$$

*We can solve:* $\operatorname{sols}(L) = \left\{ \begin{pmatrix} 1-2t \\ 2+t \\ t \end{pmatrix} \middle| \, t \in \mathbb{R} \right\}$

**Definition 2.2.3** (Elementary Operations). *We define three types of elementary operation:*

| Operations | Calculation |
|---|---|
| Reordering the equations | $E_i \leftrightarrow E_j$ |
| Multiplying one equation by a non-zero constant $t$ | $E_i \to t \cdot E_i$ |
| Add multiple of one equation to another | $E_i \to E_i - t \cdot E_j$ |

*These operations are reversible, so they conserve the solutions. That is, the LES* $M = \varphi(L)$ *that we get after doing one of the elementary operations, is equivalent to* $L$*.*

**Theorem 2.2.1** (Gaussian Elimination Algorithm). *Every LES is equivalent to a LES in canonical form.*

*Proof.* Let $E_j : \sum_{i=1}^{n} a_{i,j} \cdot x_i = b_j$ and $L = (E_1, \cdots, E_k)$. We prescribe the following algorithm to get into canonical form:

**Algorithm 1** Gaussian Elimination Algorithm

$r \leftarrow 0$
**for** $1 \leq j \leq n$ **do**
  $i \leftarrow r + 1$
  **while** $i \leq k$ **and** $a_{i,j} = 0$ **do**
    $i \leftarrow i + 1$
  **end while**
  **if** $i < k$ **then**
    $r \leftarrow r + 1$
    **do** $E_i \leftrightarrow E_r$
    **do** $E_j \rightarrow \dfrac{1}{a_{r,j}} \cdot E_j$
    **for** $1 \leq m \leq k$ **do**
      **if** $m \neq r$ **then**
        **do** $E_m \rightarrow E_m - a_{m,j} \cdot E_r$
      **end if**
    **end for**
  **end if**
**end for**

□

# 3 Linear Combinations

## 3.1 Sequence of Tuples

**Definition 3.1.1** (Tuple field operations)**.** *We define addition of tuple and multiplication by number as:*

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad \text{and} \quad t \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} t \cdot x_1 \\ \vdots \\ t \cdot x_n \end{pmatrix}$$

**Definition 3.1.2.** *(Linear Combination) Given tuples $\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n \in F^k$ and numbers $\alpha_1, \alpha_2, \cdots, \alpha_n \in F$ a linear combination is:*

$$\sum_{i=1}^{n} \alpha_i \cdot \underline{v}_i = \alpha_1 \cdot \underline{v}_1 + \alpha_2 \cdot \underline{v}_2 + \cdots + \alpha_n \cdot \underline{v}_n \in F^k$$

**Definition 3.1.3** (Span of Sequence)**.** *For a set/sequence of tuples $S = (\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n)$, with $\underline{v}_i \in F^k$, we define:*

$$\mathrm{Span}(S) = \left\{ \sum_{i=1}^{n} \alpha_i \cdot \underline{v}_i \;\middle|\; \alpha_i \in F \right\}$$

*the set of all linear combinations*

**Proposition 3.1.1** (Span is closed)**.** *For any sequence $S$, $\mathrm{Span}(S)$ is closed under addition and multiplication by number.*

*Proof.* Let $S = (\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n)$ then:

$$\left(\sum_{i=1}^{n} \alpha_i \cdot \underline{v}_i\right) + \left(\sum_{i=1}^{n} \beta_i \cdot \underline{v}_i\right) = \left(\sum_{i=1}^{n} (\alpha_i + \beta_i) \cdot \underline{v}_i\right) \in \mathrm{Span}(S)$$

$$\lambda \cdot \left(\sum_{i=1}^{n} \alpha_i \cdot \underline{v}_i\right) = \left(\sum_{i=1}^{n} (\lambda \cdot \alpha_i) \cdot \underline{v}_i\right) \in \mathrm{Span}(S)$$

$\square$

**Definition 3.1.4** (Linear Dependent and Independent sequences). *We say* $S = (\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n)$ *is:*

*Linear Dependent if:*

$$\exists\, (x_1, x_2, \cdots, x_n) \neq \underline{0} \in F^n \;:\; \sum_{i=1}^{n} x_i \cdot \underline{v}_i = \underline{0}$$

*Linear Independent if:*

$$\nexists\, (x_1, x_2, \cdots, x_n) \neq \underline{0} \in F^n \;:\; \sum_{i=1}^{n} x_i \cdot \underline{v}_i = \underline{0}$$

*that is,*

$$\forall\, (x_1, x_2, \cdots, x_n) \in F^n,\; \sum_{i=1}^{n} x_i \cdot \underline{v}_i = \underline{0} \Rightarrow (x_1, x_2, \cdots, x_n) = \underline{0}$$

**Example 3.1.1** (Proportionality Condition). *A sequence* $S = (\underline{u}, \underline{v})$ *is linear independent iff the two tuples are not proportional.*

**Definition 3.1.5** (Linear Dependency). *We write the linear dependency of* $S = (\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n)$*:*

$$\mathrm{LD}(S) = \mathrm{sols}\left(\sum_{i=1}^{n} \underline{v}_i \cdot x_i = \underline{0}\right)$$

*Comment: Always given:* $\underline{0} \subseteq \mathrm{LD}(S)$

**Proposition 3.1.2** (LI sequences have trivial LD). *A sequence is Linear Independent iff* $\mathrm{LD}(S) = \{\underline{0}\}$.

*Proof.* We prove both directions:

(⇒) By contrary, suppose there is $\underline{x} = (x_1, x_2, \cdots, x_n) \neq \underline{0} \in \mathrm{LD}(S)$, then $S$ is not Linearly Independent

(⇐) Let $\underline{x} = (x_1, x_2, \cdots, x_n) \in F^n \;:\; \sum_{i=1}^{n} x_i \cdot \underline{v}_i = \underline{0}$. Therefore, $\underline{x} \in \mathrm{LD}(S) = \{\underline{0}\} \Rightarrow \underline{x} = \underline{0}$

□

## 3.2 Linear Equation Systems on Tuples

**Definition 3.2.1** (System as Tuples)**.** *For a linear system:*

$$L : \begin{cases} E_1 \ : \ \sum_{i=1}^n a_{1,i} \cdot x_i = b_1 \\ E_2 \ : \ \sum_{i=1}^n a_{2,i} \cdot x_i = b_2 \\ \qquad \vdots \\ E_k \ : \ \sum_{i=1}^n a_{k,i} \cdot x_i = b_n \end{cases} \quad \Leftrightarrow \sum_{i=1}^n \begin{pmatrix} a_{1,i} \\ a_{2,i} \\ \vdots \\ a_{k,i} \end{pmatrix} \cdot x_i = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{pmatrix}$$

*If we define* $\underline{a}_i = \begin{pmatrix} a_{1,i} \\ a_{2,i} \\ \vdots \\ a_{k,i} \end{pmatrix}$ *and* $\underline{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{pmatrix}$ *so that* $L \ : \ \sum_{i=1}^n \underline{a}_i \cdot x_i = \underline{b}$.

**Proposition 3.2.1** (N&SC for solution)**.** $L \ : \ \sum_{i=1}^n \underline{a}_i \cdot x_i = \underline{b}$ *has a solution iff* $\underline{b} \in \mathrm{Span}\left(\underline{a}_1, \underline{a}_2, \cdots, \underline{a}_n\right)$

**Lemma 3.2.1** (Homogeneous solutions are closed)**.** *Let* $L \ : \ \sum_{i=1}^n \underline{a}_i \cdot x_i = \underline{0}$ *be a homogeneous linear system, that is, a system of homogeneous linear equations. Then,* $\mathrm{sols}(H)$ *is closed under (tuple) addition and multiplication by number.*

*Proof.* Let $\underline{x}, \underline{y} \in \mathrm{sols}(H)$ then:

$$\left(\sum_{i=1}^n \underline{a}_i \cdot (x_i + y_i)\right) = \left(\sum_{i=1}^n \underline{a}_i \cdot x_i\right) + \left(\sum_{i=1}^n \underline{a}_i \cdot y_i\right) = \underline{0} + \underline{0} = \underline{0}$$

$$\left(\sum_{i=1}^n \underline{a}_i \cdot (\lambda \cdot x_i)\right) = \lambda \cdot \left(\sum_{i=1}^n \underline{a}_i \cdot x_i\right) = \lambda \cdot \underline{0} = \underline{0}$$

$\square$

*Further,* $\text{sols}(H) = \text{LD}\left(\underline{a}_1, \underline{a}_2, \cdots, \underline{a}_n\right)$, *by definition. And,* $\underline{0} \subseteq \text{sols}(H)$, *so* $\text{sols}(H) \neq \varnothing$

**Corollary 3.2.1** (LD is closed)**.** *For any sequence* $S$, $\text{LD}(S)$ *is closed under (tuple) addition and multiplication by number.*

**Theorem 3.2.1** (General Solution of LES)**.** *Let* $L$ *be a LES and* $H$ *be the respective homogeneous system. Let* $\underline{p} \in \text{sols}(L)$ *then:*

$$\text{sols}(L) = \left\{\underline{p} + \underline{q} \;\middle|\; \underline{q} \in \text{sols}(H)\right\}$$

*Proof.* We use double inclusion:

$(\supseteq)$ $\forall\, \underline{q} \in \text{sols}(H)\,,\ \sum_{i=1}^{n} \underline{a}_i \cdot \left(p_i + q_i\right) = \sum_{i=1}^{n} \underline{a}_i \cdot p_i + \sum_{i=1}^{n} \underline{a}_i \cdot q_i = \underline{b} + \underline{0} = \underline{b} \Rightarrow$ $p + \underline{q} \in \text{sols}(L)$

$(\subseteq)$ $\underline{x} \in \text{sols}(L) \Leftrightarrow \sum_{i=1}^{n} \underline{a}_i \cdot x_i = \underline{b} = \sum_{i=1}^{n} \underline{a}_i \cdot p_i \Leftrightarrow \sum_{i=1}^{n} \underline{a}_i \cdot \left(x_i - p_i\right) = \underline{0} \Rightarrow$ $\underline{x} - \underline{p} = \underline{q} \in \text{sols}(H)$

$\square$

**Corollary 3.2.2** (Uniqueness of Solution)**.** *If* $\text{sols}(H) = \{\underline{0}\}$ *(trivial), then the solution of the linear system is either unique or empty.*

## 3.3 Basis and Subspaces of $F^k$

**Definition 3.3.1** (Basis of Tuples)**.** *We say* $S = \left(\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n\right)$ *is a basis for a set* $U \subseteq F^k$ *if every element in* $U$ *can be uniquely represented as linear combination of* $S$*, that is:*

$$\forall\, \underline{u} \in U\,,\ \exists!\ \left(\alpha_1, \alpha_2, \cdots, \alpha_n\right) \in \mathbb{R}^n \,:\, \sum_{i=1}^{n} \alpha_i \cdot \underline{v}_i = \underline{u}$$

**Theorem 3.3.1** (N&SC for a Basis). *A sequence is a basis of its span iff the sequence is linear independent.*

*Proof.* We know every element in the span can be represented as a linear combination of $S$. It rests to show uniqueness iff the sequence is linear independent.

($\Rightarrow$) Then, $\underline{0}$ has a unique representation, which is taking every coefficient 0. Hence, there is no other linear combination that gets $\underline{0}$, that is, $S$ is linearly independent.

($\Leftarrow$) For any $\underline{u} \in U = \mathrm{Span}(S)$, we want $\sum_i \underline{v}_i \cdot x_i = \underline{u}$ to have a unique solution. Since a solution exists ($\underline{u} \in \mathrm{Span}(S)$), by the previous theorem, it is necessary and sufficient that the homogenous system has only trivial solution, that is, $\sum_i \underline{v}_i \cdot x_i = \underline{0}$ has only trivial solution. Meaning, the only linear combination of $S$ that gives $\underline{0}$ is the one with all zeros. I.e. $S$ is linearly independent.

$\square$

*If $S$ is a basis of its span, we say the span is **exact**.*

**Definition 3.3.2** (Subspaces as Spans). *A subset $U \subseteq F^k$ is a (finitely spanned) subspace of $F^k$ if it is a span of some sequence.*

# 4 Functions

## 4.1 Basic Definitions

**Definition 4.1.1** (Function). *A function $f : A \to B$ is defined as three sets:*

- ***Domain**: $A$*
- ***Codomain/Range**: $B$*
- ***Graph/Table**: $f \subseteq A \times B$*

*Such that:* $\forall\, a \in A\,,\ \exists!\, b \in B\ :\ (a, b) \in f$

*Instead of writing $(a, b) \in f$, we write $f(a) = b$. We call $b$ the **image** of $a$ and $a$ the **source** of $b$.*

**Definition 4.1.2** (Image).

$$\operatorname{Im}(f) = \{b \in B \mid \exists\, a \in A\ :\ f(a) = b\} = \{f(a) \mid a \in A\}$$

**Definition 4.1.3** (Injectivity). *Given $f : A \to B$ is called injective iff:*

$$\forall\, a_1, a_2 \in A\,,\ a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)$$

*which is equivalent to:* $\forall\, a_1, a_2 \in A\,,\ f(a_1) = f(a_2) \Rightarrow a_1 = a_2$

**Lemma 4.1.1** (Injectivity as Unique Solution). *A function is injective iff:*

$$\forall\, b \in \operatorname{Im}(f)\,,\ \exists!\, a \in A\ :\ f(a) = b$$

*Proof.* We prove both directions: Let $b \in \operatorname{Im}(f)$.

($\Rightarrow$) By definition, $\exists\, a \in A\ :\ f(a) = b$. Let $a' \in A\ :\ f(a') = b = f(a) \Rightarrow a' = a$, so it is unique.

($\Leftarrow$) By contrary, if $\exists\, a_1, a_2 \in A, a_1 \neq a_2\ :\ f(a_1) = b = f(a_2)$, then the source $a$ of $b$ is not unique.

□

**Definition 4.1.4** (Surjectivity)**.** *Given* $f : A \to B$ *is called surjective iff:* $B = \mathrm{Im}(f)$

**Lemma 4.1.2** (Surjectivity as existence of solution)**.** *A function is surjective iff:*

$$\forall\, b \in B\,,\ \exists\, a \in A\ :\ f(a) = b$$

*Proof.* We prove both directions: Let $b \in B$.

- ($\Rightarrow$) By definition, $b \in B = \mathrm{Im}(f) \Leftrightarrow \exists\, a \in A\ :\ f(a) = b$.
- ($\Leftarrow$) By contrary, if $\exists\, b \in B\ :\ \forall\, a \in A\,,\ f(a) \neq b$, then $B \setminus \mathrm{Im}(f) \neq \varnothing \Rightarrow B \neq \mathrm{Im}(f)$.

□

## 4.2 Composition and Inverses

**Definition 4.2.1** (Bijectivity)**.** *A function is bijective iff it is both injective and surjective.*

**Theorem 4.2.1** (Reverse Table)**.** *The reverse graph of* $f$*,* $g = \{(b, a) \in B \times A \mid (a, b) \in f\}$*, defines a function iff* $f$ *is bijective.*

*Proof.* We prove both directions:

- ($\Leftarrow$) $f$ is injective and surjective: $\forall\, b \in \mathrm{Im}(f) = B\,,\ \exists!\, a \in A\ :\ f(a) = b$. Putting $g(b) = a$ instead of $f(a) = b$, by definition, we get that $g$ is a function.
- ($\Rightarrow$) Using contrapositive, if it's not injective or not surjective, it is not invertible.
  - Not injective, $\exists\, a_1, a_2 \in A, a_1 \neq a_2\ :\ f(a_1) = f(a_2) = b$, hence $g(b)$ has two images.

- Not surjective $\exists\, b \in B \;:\; \forall\, a \in A\,,\; f(a) \neq b$, hence $g(b)$ has no image.

$\square$

**Definition 4.2.2** (Composition)**.** *Let $f : A \to B$ and $g : B \to C$ be two functions (s.t. $\mathrm{Im}(f) \subseteq \mathrm{Dom}(g)$), the composition $g \circ f$ is the function:*

$$\begin{aligned} h : A &\to C \\ a &\mapsto g\big(f(a)\big) \end{aligned}$$

*It is an associative operation.*

**Definition 4.2.3** (Commutative Diagram)**.** *A commutative diagram is a combination of maps as:*

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "h"'] & B \arrow[d, "g"] \\ & C \end{tikzcd}$$

*so that we can arrive at any point through any sequence of maps. In this simplest case, we must have $h = g \circ f$.*

**Theorem 4.2.2** (Character of Functions)**.** *Let $f : A \to B$ and $g : B \to C$.*

1. *$g \circ f$ injective $\Rightarrow f$ injective*
2. *$g \circ f$ surjective $\Rightarrow g$ surjective*

*Proof.* We use the contrapositive

1. $f$ is not injective: $\exists\, a_1, a_2 \in A, a_1 \neq a_2 \;:\; f(a_1) = f(a_2)$. By apply $g$ to both sides, we get $g \circ f$ is not injective.
2. $g$ is not surjective: $\exists\, c \in C \;:\; \forall\, b \in B\,,\; g(b) \neq c$. Letting $b = f(a)$, we get $g \circ f$ is not surjective.

$\square$

**Definition 4.2.4** (Identity). *For any set $A$ we define:*

$$\begin{aligned} \mathrm{Id}_A : A &\to A \\ a &\mapsto a \end{aligned}$$

*Let $f : A \to B$, then: $f \circ \mathrm{Id}_A = f = \mathrm{Id}_B \circ f$*

**Theorem 4.2.3** (Reverse Table as Inverses). *If $f$ is bijective and $g : B \to A$ its inverse table, then:*

$$\begin{aligned} g \circ f &= \mathrm{Id}_A \\ f \circ g &= \mathrm{Id}_B \end{aligned}$$

*Proof.* Let $f(a) = b$: $\forall\, a \in A\,,\ (g \circ f)(a) = g(f(a)) = g(b) = a$ and $\forall\, b \in B\,,\ (f \circ g)(b) = f(g(b)) = f(a) = b$. $\square$

**Lemma 4.2.1** (Uniqueness of Inverses). *For any associative operation $* : A \times A \to A$ with neutral element ($e$), the inverses are unique.*

*Proof.* Let $a, a'$ be inverses of $b$. Then:

$$a' = a' * e = a' * (b * a) = (a' * b) * a = e * a = a$$

$\square$

**Corollary 4.2.1** (Uniqueness of Function Inverses). *If $f$ has an inverse wrt composition, then it is unique. Heretofore, we denote the inverse $f^{-1}$.*

**Lemma 4.2.2** (Composition of Inverses). *For any associative operation $* : A \times A \to A$ with neutral element ($e$), if $a$ and $b$ are have inverses, then $a * b$ has an inverse. In particular, $(a * b)^{-1} = b^{-1} * a^{-1}$*

*Proof.* We show that $b^{-1} * a^{-1}$ is an inverse:

$$\begin{aligned} \left(b^{-1} * a^{-1}\right) * (a * b) &= b^{-1} * \left(a^{-1} * a\right) * b = b^{-1} * e * b = b^{-1} * b = e \\ (a * b) * \left(b^{-1} * a^{-1}\right) &= a * \left(b^{-1} * b\right) * a^{-1} = a * e * a^{-1} = a * a^{-1} = e \end{aligned}$$

$\square$

**Corollary 4.2.2** (Inverse of Composition). *The composition of bijective functions is again bijective.*

**Example 4.2.1** (Set of Functions). *Let ${}^AA = \{f \mid f : A \to A\}$, we have:*

1. *${}^AA$ is closed under composition*
2. *Composition is associative*
3. *There is a neutral element ($\mathrm{Id}_A$)*
4. *If $f : A \to A$ is a bijection, then it has an inverse*

*Let $G_A = \left\{f \in {}^AA \,\middle|\, f \text{ is bijective}\right\}$, so that every element of ${}^AA$ that has an inverse is in $G_A$. Then, $G_A$ is closed under composition.*

## 4.3 One-Sided Inverses

**Definition 4.3.1** (Left-Inverse). *Let $f : A \to B$ be a function. A left-inverse is $g : B \to A$ such that: $g \circ f = \mathrm{Id}_A$*

**Lemma 4.3.1** (N&SC for existence of left-inverses). *$f$ has left inverse iff $f$ is injective.*

*Proof.* We prove both directions:

($\Rightarrow$) Since $g \circ f = \mathrm{Id}_A$ is injective, we must have $f$ is injective.

($\Leftarrow$) Take any $a_0 \in A$. We define:

$$g(b) = \begin{cases} a & \text{if } b \in \mathrm{Im}(f) \text{ where } f(a) = b \\ a_0 & \text{if } b \notin \mathrm{Im}(f) \end{cases}$$

□

**Definition 4.3.2** (Right-Inverse). *Let $f : A \to B$ be a function. A right-inverse is $g : B \to A$ such that: $f \circ g = \mathrm{Id}_B$*

**Lemma 4.3.2** (N&SC for existence of right-inverses)**.** *$f$ has right inverse iff $f$ is surjective.*

*Proof.* We prove both directions:

(⇒) Since $f \circ g = \mathrm{Id}_B$ is surjective, we must have $f$ is surjective.

(⇐) We define $g(b)$ are any source $a$ of $b$ (there may be many, we pick an arbitrary one).

□

**Theorem 4.3.1** (Conservation of Character)**.** *Let $f : A \to B$ arbitrary and $g : B \to C$ bijective.*

1. *$g \circ f$ injective $\Leftrightarrow$ $f$ injective*
2. *$g \circ f$ surjective $\Leftrightarrow$ $f$ surjective*

*Proof.* By the previous two lemmas:

1. $g \circ f$ injective $\Leftrightarrow \exists\, h_1 : C \to A \;:\; h_1 \circ (g \circ f) = \mathrm{Id}_A \Leftrightarrow f$ has left inverse $(h_2 = h_1 \circ g \Leftrightarrow h_1 = h_2 \circ g^{-1}) \Leftrightarrow f$ injective
2. $g \circ f$ surjective $\Leftrightarrow \exists\, h_1 : C \to A \;:\; (g \circ f) \circ h_1 = \mathrm{Id}_C$ eq. $f \circ h_1 = g^{-1}$ eq. $f \circ h_1 \circ g = \mathrm{Id}_A$ has right inverse $(h_2 = h_1 \circ g \Leftrightarrow h_1 = h_2 \circ g^{-1})$ $\Leftrightarrow f$ surjective

□

**Lemma 4.3.3** (Functions to Equations)**.**

$$\begin{aligned} &E : f(x) = g(x) \\ &F : h\big(f(x)\big) = h\big(g(x)\big) \end{aligned}$$

*For general $h$:* $\mathrm{sols}(E) \subseteq \mathrm{sols}(F)$*. If $h$ is injective:* $\mathrm{sols}(E) = \mathrm{sols}(F)$

# 5 Matrices

## 5.1 Products on Matrices

**Definition 5.1.1** (SumProd)**.** *Given two tuples $\underline{a}, \underline{x} \in F^n$, we define:*

$$\underline{a} \bullet \underline{x} := \sum_{i=1}^{n} a_i \cdot x_i \in F$$

*which is linear on both variables (distributive over addition and numbers move freely inside).*

**Definition 5.1.2** (SumProd with Sequences)**.** *Let $S = (\underline{a}_1, \underline{a}_2, \cdots, \underline{a}_n) \in (F^k)^n$, we can further define:*

$$S \bullet \underline{x} := \sum_{i=1}^{n} \underline{a}_i \cdot x_i \in F$$

*That way, we can write a linear system $L : \sum_{i=1}^{n} \underline{a}_i \cdot x_i = \underline{b}$ as $L : S \bullet \underline{x} = \underline{b}$ .*

**Definition 5.1.3** (Matrices)**.** *The set $M_{k\times n}(F)$ (also denoted $F^{k\times n}$) has elements which are rectangles of numbers with $k$ rows and $n$ columns. It is equivalent (isomorphic) to $(F^k)^n$. If we take a sequence $S = (\underline{a}_1, \underline{a}_2, \cdots, \underline{a}_n) \in (F^k)^n$, we write the matrix $A$ as:*

$$A = \begin{pmatrix} | & | & & | \\ \underline{a}_1 & \underline{a}_2 & \cdots & \underline{a}_n \\ | & | & & | \end{pmatrix}$$

*We define multiplication by tuples (written side-by-side) as:*

$$A\,\underline{x} = \begin{pmatrix} | & | & & | \\ \underline{a}_1 & \underline{a}_2 & \cdots & \underline{a}_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = S \bullet \underline{x} = \sum_{i=1}^{n} \underline{a}_i \cdot x_i$$

*We can also define with rows:*

$$A\,\underline{x} = \begin{pmatrix} — & \underline{r}_1 & — \\ — & \underline{r}_2 & — \\ & \vdots & \\ — & \underline{r}_k & — \end{pmatrix} \underline{x} = \begin{pmatrix} \underline{r}_1 \bullet \underline{x} \\ \underline{r}_2 \bullet \underline{x} \\ \vdots \\ \underline{r}_k \bullet \underline{x} \end{pmatrix}$$

*Note that is operation is linear on $\underline{x}$ (distributive over addition and numbers move freely inside). Also, we define $S_c(A) = (\underline{a}_1, \underline{a}_2, \cdots, \underline{a}_n)$, the column sequence, and $S_r(A) = (\underline{r}_1, \underline{r}_2, \cdots, \underline{r}_k)$, the row sequence.*

**Proposition 5.1.1** (LES as Matrix Equations)**.**

$$L : \sum_{i=1}^{n} \underline{a}_i \cdot x_i = \underline{b} \Leftrightarrow \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{pmatrix}$$

*We call the matrix $A$, we get: $L : A\,\underline{x} = \underline{b}$.*

**Definition 5.1.4** (Transpose)**.** *For $A \in M_{k\times n}(F)$ as a rectangle of numbers, we define:*

$$A = \begin{pmatrix} | & | & & | \\ \underline{a}_1 & \underline{a}_2 & \cdots & \underline{a}_n \\ | & | & & | \end{pmatrix} \Rightarrow A^t = \begin{pmatrix} — & \underline{a}_1 & — \\ — & \underline{a}_2 & — \\ & \vdots & \\ — & \underline{a}_n & — \end{pmatrix} \in M_{n\times k}(F)$$

**Lemma 5.1.1** (Transpose Changes Order)**.** $(AB)^t = B^t\,A^t$

## 5.2 Matricial Functions

**Definition 5.2.1** (Matricial Functions)**.** *Given $A \in M_{k\times n}(F)$, we define:*

$$\begin{aligned} T_A : F^n &\to F^k \\ \underline{x} &\mapsto A\,\underline{x} \end{aligned}$$

**Definition 5.2.2** (Linear Operation on Matrices)**.** *For $A, B \in M_{k\times n}(F)$, we define:*

$$\textit{Sum: } T_{A+B} = T_A + T_B, \textit{ that is, } T_{A+B} : \underline{x} \mapsto T_A(\underline{x}) + T_B(\underline{x})$$

$$\textit{Multiplication by number: } T_{\lambda A} = \lambda \cdot T_A \textit{ that is, } T_{\lambda A} : \underline{x} \mapsto \lambda \cdot T_A(\underline{x})$$

*In terms of the original rectangles of numbers, they are defined analogously to the tuple addition and multiplication by a number.*

**Definition 5.2.3** (Matrix Multiplication)**.** *We define: $T_{AB} = T_A \circ T_B$. Notice that, if $A \in M_{k\times n}(F)$ and $B \in M_{r\times m}(F)$, for $AB$ to be defined, we need $n = r$. Hence, $T_A : F^n \to F^k$ and $T_B : F^m \to F^n$.As a rectangle of numbers:*

$$T_A(T_B\,\underline{x}) = A\left(B\,\underline{x}\right) = A\begin{pmatrix} | & | & & | \\ \underline{b}_1 & \underline{b}_2 & \cdots & \underline{b}_n \\ | & | & & | \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A\left(\sum_{i=1}^{n} \underline{b}_i \cdot x_i\right)$$

$$= \sum_{i=1}^{n} A(\underline{b}_i) \cdot x_i = \begin{pmatrix} | & | & & | \\ A(\underline{b}_1) & A(\underline{b}_2) & \cdots & A(\underline{b}_n) \\ | & | & & | \end{pmatrix}\underline{x}$$

$$\Rightarrow A\,B = \begin{pmatrix} | & | & & | \\ A(\underline{b}_1) & A(\underline{b}_2) & \cdots & A(\underline{b}_n) \\ | & | & & | \end{pmatrix}$$

*Equivalently, we calculate as follows:*

$$\begin{pmatrix} — & \underline{r}_1 & — \\ — & \underline{r}_2 & — \\ & \vdots & \\ — & \underline{r}_k & — \end{pmatrix}\begin{pmatrix} | & | & & | \\ \underline{b}_1 & \underline{b}_2 & \cdots & \underline{b}_n \\ | & | & & | \end{pmatrix} = \begin{pmatrix} \underline{r}_1 \bullet \underline{b}_1 & \underline{r}_1 \bullet \underline{b}_2 & \cdots & \underline{r}_1 \bullet \underline{b}_n \\ \underline{r}_2 \bullet \underline{b}_1 & \underline{r}_2 \bullet \underline{b}_2 & \cdots & \underline{r}_2 \bullet \underline{b}_n \\ \vdots & \vdots & \ddots & \vdots \\ \underline{r}_n \bullet \underline{b}_1 & \underline{r}_n \bullet \underline{b}_2 & \cdots & \underline{r}_n \bullet \underline{b}_n \end{pmatrix}$$

*Notice it is not commutative.*

**Definition 5.2.4** (Identity Matrix)**.** *Let $\underline{e}_i \in F^n$ be such that the i-th component is 1 and every other is 0. Then:*

$$A\,\underline{e}_i = \begin{pmatrix} | & | & & | \\ \underline{a}_1 & \underline{a}_2 & \cdots & \underline{a}_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \underline{a}_i$$

*The $(n \times n)$ identity matrix is defined:*

$$I_n = \begin{pmatrix} | & | & & | \\ \underline{e}_1 & \underline{e}_2 & \cdots & \underline{e}_n \\ | & | & & | \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

*the diagonal of all ones. Also, notice:* $\mathrm{Id}_{F^n} = T_{I_n}$.

**Lemma 5.2.1** (Matrix of Transformation)**.** *Let $T : F^n \to F^k$ be a matricial function. Its corresponding matrix is:*

$$[T] = \begin{pmatrix} | & | & & | \\ T(\underline{e}_1) & T(\underline{e}_2) & \cdots & T(\underline{e}_n) \\ | & | & & | \end{pmatrix}$$

*That is, the matrix is uniquely determined by the function.*

*Proof.* If $A$ and $B$ are two matrices that define the same matricial function, then, for $i \in \{1, 2, \cdots, n\}$:

$$\underline{a}_i = A\,\underline{e}_i = T(\underline{e}_i) = B\,\underline{e}_i = \underline{b}_i$$

hence, every column is the same, so $A = B$. $\square$

**Definition 5.2.5** (Kernel)**.** *Given a matricial function $T_A$, we define:*

$$\ker(T_A) = \left\{ \underline{x} \in F^n \;\middle|\; T_A\big(\underline{x}\big) = \underline{0} \right\}$$

*Notice* $\{\underline{0}\} \subseteq \ker(T_A)$*. We further write* $\mathrm{sols}(A) = \ker(T_A) = \mathrm{sols}(A\,\underline{x} = \underline{0})$*.*

**Lemma 5.2.2** (N&SC for Injectivity)**.** *$T_A$ is injective iff* $\ker(T_A) = \{\underline{0}\}$*.*

*Proof.* We prove both directions:

($\Rightarrow$) $T\big(\underline{x}\big) = \underline{0} = T\big(\underline{0}\big) \Rightarrow \underline{x} = \underline{0}$, that is, $\{\underline{0}\} \supseteq \ker(T_A)$.

($\Leftarrow$) $T(\underline{x}) = T(\underline{y}) \Rightarrow T(\underline{x} - \underline{y}) = \underline{0} \Rightarrow \underline{x} - \underline{y} = \underline{0} \Rightarrow \underline{x} = \underline{y}$

□

## 5.3 Invertible Matrices

**Definition 5.3.1** (Inverse of a Matrix)**.** *Given a matricial function $T_A : F^n \to F^k$, we seek to find its inverse (if it exists). We define the inverse matrix as the matrix of the matrical function $T_A^{-1} = T_{A^{-1}}$. That is, $A^{-1}$ is the unique matrix that $A^{-1}\,A = I$ and $A\,A^{-1} = I$*

**Definition 5.3.2** (General Linear Group)**.** *The set of invertible matrices is denoted* $\mathrm{GL}_n(F)$*.*

**Lemma 5.3.1** (Inverse of Transpose)**.** $(A^t)^{-1} = (A^{-1})^t$

*Proof.*

$$\begin{aligned} A\,A^{-1} = I &\Rightarrow (A\,A^{-1})^t = I^t = I \Rightarrow (A^{-1})^t\,A^t = I \\ A^{-1}\,A = I &\Rightarrow (A^{-1}\,A)^t = I^t = I \Rightarrow A^t\,(A^{-1})^t = I \end{aligned}$$

By uniqueness of inverses, we have the proof. □

**Corollary 5.3.1** (Transpose of Invertible is Invertible)**.** *$A^t$ is invertible $\Leftrightarrow$ $A$ is invertible.*

We proceed to apply the Gaussian Elimination to matrices, as any linear equation system can be written as something of the form $A\,\underline{x} = \underline{b}$.

**Definition 5.3.3** (Elementary Functions)**.** *We define the following types of (invertible) elementary functions* $\varphi : F^k \to F^k$

| Operations | $\varphi$ | $\varphi^{-1}$ |
|---|---|---|
| Reordering the variables | $x_i \leftrightarrow x_j$ | $x_i \leftrightarrow x_j$ |
| Multiplying one variable by a non-zero constant $t$ | $x_i \to t \cdot x_i$ | $x_i \to \frac{1}{t} \cdot x_i$ |
| Add multiple of one variable to another | $x_i \to x_i - t \cdot x_j$ | $x_i \to x_i + t \cdot x_j$ |

*Notice that those are matricial functions. We further denote* $\Phi = \left[\varphi\right]$

**Definition 5.3.4** (Leading Coefficient)**.** *We say* $x_i$ *the leading coefficient of the tuple* $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ *iff* $a_i \neq 0$ *and* $\forall\, j : 1 \leq j < i\,,\ a_j = 0.$

**Definition 5.3.5** (Canonical Form / Reduced Row Echelon Form)**.** *A matrix is said to be in canonical form or rref (row reduced echelon form) if:*

1. *The leading coefficient of a nonzero row is always strictly to the right of the leading coefficient of the row above it.*
2. $\underline{0}$ *rows come at the end.*
3. *The leading coefficient of each row is* $1$*.*
4. *Each column containing a leading 1 (called pivot column) has zeros everywhere else.*

*An important example of a matrix in rref is the identity matrix. We also say* $T_A$ *is in canonical form.*

**Proposition 5.3.1** (Gaussian Elimination on Matrices)**.** *For every matricial function $T_A$, $\exists\, \varphi_1, \varphi_2, \cdots, \varphi_r$ elementary : $\varphi_r \circ \cdots \circ \varphi_2 \circ \varphi_1 \circ T_A = T_R$ is in canonical form.*

*Therefore, $T_A$ has "same character" (i.e. injective or surjective) as $T_R$.*

*Proof.* Follows directly from the Gaussian Elimination Algorithm. □

**Theorem 5.3.1** (N&SC for Inverting RREF)**.** *Let $R \in M_{k\times n}(F)$ be in canonical form. Then, $R$ is invertible iff $n = k$ and $R = I$.*

*Proof.* First, notice that the number of leading coefficients is $\min(n, k)$. We have two cases:

1. If $n < k$, there is a row of $\underline{0}$.
2. If $n > k$, there is a column without a leading coefficient.

Now, we look:

1. If there is a row of $\underline{0}$, then $T_R$ is not surjective, because it doesn't map to $\underline{e}_k$ (last row).
2. Suppose that there is a column without a leading coefficient, say, $\underline{a}$, the $i$-th column. We apply $T_R$ to the tuple $\underline{x}$:

$$x_j = \begin{cases} -a_j & \text{if } j\text{-th column is pivot} \\ 1 & \text{if } j = i \\ 0 & \text{else} \end{cases}$$

   Hence, $T_R\big(\underline{x}\big) = \underline{0} \Rightarrow \underline{0} \neq \underline{x} \in \ker(T_A) \Rightarrow T_A$ is not injective.

Therefore, if $n \neq k$, $T_R$ is not bijective. However if $n = k$ and $R$ still has a row of $\underline{0}$ or a column without a leading coefficient, $T_R$ is still not bijective. The remaining case it exactly when $n = k$ and $R = I$, which is trivially invertible. □

**Corollary 5.3.2** (Invertibility from RREF). $\boxed{\textit{A matrix is invertible iff it's rref is } I}$

**Corollary 5.3.3** (Character of Dimension). *For every matricial function $T_A$:*

1. *If $n < k$, $T_A$ is not surjective.*
2. *If $n > k$, $T_A$ is not injective.*

In order to invert a matrix $A$ we write the augmented matrix $\left(A \mid I\right)$ and apply elementary functions until we get $\left(I \mid A^{-1}\right)$.

**Example 5.3.1.** $A = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 2 & 3 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 0 & 3 & -1 \\ 0 & 1 & -2 & 1 \end{array}\right)$$

*Hence,* $A^{-1} = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix}$

**Definition 5.3.6** (Row Equivalence). *Two matrices are **row equivalent**, written:*

$$A \leftrightarrow B \Leftrightarrow \exists\, \Phi_1, \Phi_2, \cdots, \Phi_r \text{ elementary} : \; A = \Phi_1\, \Phi_2 \cdots \Phi_r\, B$$

**Lemma 5.3.2** (N&SC for Row Equivalence). $A \leftrightarrow B \Leftrightarrow \exists\, M \in \mathrm{GL}_k(F) : A = MB$

*Proof.* By Gaussian Elimination Theorem, a matrix is invertible iff it is row equivalent to the identity. □

**Lemma 5.3.3** (Row Equivalence Relation). *Row equivalence is an equivalence relation:*

*Proof.* We choose the matrices so that $A \leftrightarrow B \Leftrightarrow \exists\, M \in \mathrm{GL}_k(F) : A = MB$. Reflexive: Take $I$; Symmetric: Take $M^{-1}$; Transitive: Take $M_1\, M_2$. □

## 5.4 Matrix Spaces

**Definition 5.4.1** (Sols, Cols and Rows)**.** *Given*

$$A = \begin{pmatrix} | & | & & | \\ \underline{a}_1 & \underline{a}_2 & \cdots & \underline{a}_n \\ | & | & & | \end{pmatrix} = \begin{pmatrix} — & \underline{r}_1 & — \\ — & \underline{r}_2 & — \\ & \vdots & \\ — & \underline{r}_k & — \end{pmatrix} \in M_{k\times n}(F)$$

*we define:*

$$\begin{aligned}
&\text{cols}(A) = \text{Span}(S_c(A)) = \text{Span}\left(\underline{a}_1, \underline{a}_2, \cdots, \underline{a}_n\right) \\
&\text{rows}(A) = \text{cols}(A^t) = \text{Span}(S_r(A)) = \text{Span}\left(\underline{r}_1, \underline{r}_2, \cdots, \underline{r}_k\right) \\
&\text{sols}(A) = \text{sols}(A\,\underline{x} = \underline{0}) = \text{LD}(S_c(A)) = \text{LD}((\underline{a}_1, \underline{a}_2, \cdots, \underline{a}_n))
\end{aligned}$$

**Lemma 5.4.1** (Fundamental Theorem on Gaussian Elimination)**.**

$$A \leftrightarrow B \Leftrightarrow \text{rows}(A) = \text{rows}(B) \Leftrightarrow \text{sols}(A) = \text{sols}(B)$$

*Proof.* We prove each one:

$\Leftrightarrow \text{rows}(A) = \text{rows}(B)$

($\Rightarrow$) $\text{rows}(A) = \text{cols}(A^t) = \text{cols}(B^t\, M^t) = \text{cols}(B^t) = \text{rows}(B)$

($\Leftarrow$) By contrary, if $\text{rows}(A) \neq \text{rows}(B) \Rightarrow \text{cols}(A^t) \neq \text{cols}(B^t)$. If $\text{rows}(A) \setminus \text{rows}(B) \neq \varnothing \Rightarrow \exists\, x \in \mathbb{R}^n \;:\; \begin{cases} \exists\, y \in \mathbb{R}^n \;:\; A^t\,\underline{y} = \underline{x} \\ \nexists\, z \in \mathbb{R}^n \;:\; B^t\,\underline{z} = \underline{x} \end{cases}$, then, $\nexists\, M \in \text{GL}_k(F) : A = MB \;:\;$ otherwise $B^t\,(M^t\underline{y}) = \underline{x}$. The same if $\text{rows}(B) \setminus \text{rows}(A) \neq \varnothing$.

$\Leftrightarrow \text{sols}(A) = \text{sols}(B)$

($\Rightarrow$) $\underline{x} \in \text{sols}(A) \Leftrightarrow \underline{0} = A\,\underline{x} = M\,B\,\underline{x} \Leftrightarrow B\,\underline{x} = \underline{0} \Leftrightarrow \underline{x} \in \text{sols}(B)$

($\Leftarrow$) By contrary, if $\text{sols}(A) \setminus \text{sols}(B) \neq \varnothing \Rightarrow \exists\, x \in \mathbb{R}^n : \begin{cases} A\,\underline{x} = \underline{0} \\ B\,\underline{x} \neq \underline{0} \end{cases}$, then, $\nexists\, M \in \text{GL}_k(F) : B = MA$ : otherwise $B\,\underline{x} = 0$. The same if $\text{sols}(B) \setminus \text{sols}(A) \neq \varnothing$.

$\square$

**Lemma 5.4.2.** *Let $R$ be in canonical form. Then, either $R = I$ or it is (up to an exchange of columns):*

$$\left(\begin{array}{c|c} I_r & L \\ \hline 0 & 0 \end{array}\right) \Rightarrow \text{sols}(R) = \text{cols}\left(\begin{array}{c} -L \\ \hline I_{n-r} \end{array}\right)$$

*(up to the same exchange of rows).*

**Example 5.4.1.** *We have:*

$$R = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 4 \end{pmatrix} \Rightarrow \text{sols}(R) = \text{cols}\begin{pmatrix} -2 & -3 \\ 1 & 0 \\ 0 & -4 \\ 0 & 1 \end{pmatrix}$$

*We had to switch columns 2 and 3 in $R$, we performed the same switch to rows in the shape.*

**Theorem 5.4.1** (Rank-Dimension)**.**

$$\dim \text{cols}(A) = \dim \text{rows}(A)$$

*Proof.* Let $R$ be the rref form of $A$. Since $R = M\,A$ for some $M \in \text{GL}_k(F)$, we get: $\text{rows}(A) = \text{rows}(R)$, by the lemma above and $\dim \text{cols}(A) = \dim \text{cols}(R)$ since $T_M$ is an isomorphism. Therefore, we only need to prove for the rref form. From the previous lemma, $\dim \text{rows}(A) = r$ (number of pivot columns). Also $\dim \text{cols}(A) = n - \dim \text{sols}(A) = n - (n-r) = r$. $\square$

**Definition 5.4.2** (Rank)**.** *We call*

$$\text{rank}(A) = \dim \text{cols}(A) = \dim \text{rows}(A) = \textit{number of pivot columns}$$

# 6 Determinants

## 6.1 Multilinear Alternating Function

**Definition 6.1.1** (Determinant)**.** *We define the determinant function* det : $(F^n)^n \to F$ *so that:*

| Operations | Calculation |
|---|---|
| Multilinearity | $\det(\alpha \cdot \underline{u} + \beta \cdot \underline{v}, \cdots) = \alpha \cdot \det(\underline{u}, \cdots) + \beta \cdot \det(\underline{v}, \cdots)$ |
| Alternating | $\det(\underline{u}, \cdots, \underline{v}, \cdots) = -\det(\underline{v}, \cdots, \underline{u}, \cdots)$ |
| Normalized | $\det(\underline{e}_1, \underline{e}_2, \cdots, \underline{e}_n) = 1$ |

*Moreover, we have:* $\det(\underline{u}, \underline{u}, \cdots) = 0$

*For a square matrix, we use the sequence of columns:*

$$\det(A) = \det(\underline{a}_1, \underline{a}_2, \cdots, \underline{a}_n) = \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix}$$

**Proposition 6.1.1** (Change in Determinants)**.** *For elementary operations, we have those relations:*

| Operations | $\varphi$ | $\det(\Phi\, A)$ |
|---|---|---|
| Reordering the equations | $x_i \leftrightarrow x_j$ | $\det(A)$ |
| Multiplying one equation by a non-zero constant $t$ | $x_i \to t \cdot x_i$ | $t\ \det(A)$ |
| Add multiple of one equation to another | $x_i \to x_i - t \cdot x_j$ | $\det(A)$ |

**Definition 6.1.2** (Permutation)**.** *A permutation is a bijective function* $\sigma : \{1, 2, \cdots, n\} \to \{1, 2, \cdots, n\}$*. A transpostion is the simplest type of permutation which consist of only switching two numbers, every permutation can be written as a composition of transpositions. We define* $\operatorname{sgn}(\sigma) = (-1)^{\#\ transpositions}$*.*

**Theorem 6.1.1** (Leibnitz Formula)**.**

$$\det(A) = \sum_{\sigma} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

*Proof.* By linearity:

$$\det(A) = \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_n=1}^{n} \det\left(\underline{e}_{k_1}, \underline{e}_{k_2}, \cdots, \underline{e}_{k_n}\right) \prod_{i=1}^{n} a_{i,k_i}$$

we get the formula by noticing:

$$\det\left(\underline{e}_{k_1}, \underline{e}_{k_2}, \cdots, \underline{e}_{k_n}\right) = \begin{cases} 0 & \text{if one of } k_i\text{'s are equal} \\ \operatorname{sgn}(\sigma) & \text{otherwise, where } \sigma(i) = k_i \end{cases}$$

□

**Corollary 6.1.1.** $\det(A\,B) = \det(A) \cdot \det(B)$

**Theorem 6.1.2** (Laplace Formula)**.** *The* $(i, j)$ *minor of* $A$*, denoted* $A_{i,j}$ *, is the matrix we get when we delete the* $i$*-th row and* $j$*-th column. For arbitrary column* $j$*:*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{i,j})$$

*Proof.* By linearity on the $j$-th column:

$$\det(A) = \sum_{i=1}^{n} a_{i,j} \det(\cdots, \underline{a}_{j-1}, \underline{e}_i, \underline{a}_{j+1}, \cdots)$$

And we calculate:

$$\begin{aligned} &\det(\cdots, \underline{a}_{j-1}, \underline{e}_i, \underline{a}_{j+1}, \cdots) = (-1)^{i+j} \det(\underline{e}_i, \cdots, \underline{a}_{j-1}, \underline{a}_{j+1}, \cdots) \\ &= (-1)^{i+j} \det(\underline{e}_i, \cdots, \underline{a}_{j-1} - a_{i,j-1} \cdot \underline{e}_i, \underline{a}_{j+1} - a_{i,j+1} \cdot \underline{e}_i, \cdots) \\ &= (-1)^{i+j} \det(A_{i,j}) \end{aligned}$$

□

**Corollary 6.1.2** (Determinant on Upper Triangular Matrix)**.** *If $U$ is an upper triangular matrix,* $\det(U)$ *is the product of the elements in the main diagonal.*

**Theorem 6.1.3** (N&SC for Invertibility)**.** *$A$ is invertible iff* $\det(A) \neq 0$

*Proof.* We prove both directions:

($\Rightarrow$) $\exists\, B \in M_n(F) \,:\, A\, B = I$, so that $\det(A) \cdot \det(B) = 1 \Rightarrow \det(A) \neq 0$

($\Leftarrow$) $A = \Phi_r \cdots \Phi_2\, \Phi_1\, R \Rightarrow \det(A) = t\, \det(R)$, $t \neq 0$. $R$ is upper diagonal and $\det(R) \neq 0$ hence, the diagonal has no zeros, i.e. $R = I$.

□

## 6.2 Cramer's Rule and Adjungate Matrix

**Definition 6.2.1** (Adjungate)**.** *Given $A \in M_n(F)$, we define* $\operatorname{adj}(A) \in M_n(F)$ *such that:*

$$\operatorname{adj}(A)_{i,j} = (-1)^{i+j} \det(A_{j,i})$$

**Lemma 6.2.1** (Adjungate Formula)**.**

$$\operatorname{adj}(A)\, A = \det(A)\, I$$

*Proof.* Sufficient to notice:

$$\sum_{i=1}^{n} (-1)^{i+j}\, a_{i,k} \det(A_{i,j}) = \begin{cases} \det(A) & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}$$

Since if $j \neq k$ we have the determinant with repeated columns. □

**Corollary 6.2.1** (Inverse Formula)**.** *If* $\det(A) \neq 0$*:*

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$$

**Theorem 6.2.1** (Cramer's Rule)**.** *If $A$ is invertible, then the solution $\underline{x}$ of $A\,\underline{x} = \underline{b}$ is such that:*

$$x_i = \frac{\det(B_i)}{\det(A)}$$

*where $B_i$ is the matrix we get from $A$ by replacing its $i$-th column by $\underline{b}$.*

*Proof.*

$$\underline{x} = A^{-1}\,\underline{b} = \frac{1}{\det(A)}\ \mathrm{adj}(A)\,\underline{b}$$

notice that: $\sum_{i=1}^{n}(-1)^{i+j}\,b_i\ \det(A_{i,j}) = \det(B_i)$. $\square$

# 7 Ring of Polynomials

## 7.1 Polynomials

**Definition 7.1.1** (Polynomial is $c_{00}$)**.** *A polynomial is a sequence $p \in F^\infty$ such that almost all of its components (i.e. except finitely many) are $0$. We will write with a dummy variable $X$:*

$$p = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \\ 0 \\ \vdots \end{pmatrix} \Rightarrow p(X) = a_0 + a_1\,X + a_2\,X^2 \cdots a_n\,X^n$$

*Using the dummy letter, we get the set $F[X]$. Its polynomial function is defined with substitution into $p(X)$.*

**Definition 7.1.2** (Operations on Polynomials)**.**

$$\begin{aligned} \big(a_n\big)_{n=0}^{\infty} + \big(b_n\big)_{n=0}^{\infty} &= \big(a_n + b_n\big)_{n=0}^{\infty} \\ t \cdot \big(a_n\big)_{n=0}^{\infty} &= \big(t \cdot a_n\big)_{n=0}^{\infty} \\ \big(a_n\big)_{n=0}^{\infty} \bullet \big(b_n\big)_{n=0}^{\infty} &= \left(\sum_{k=0}^{n} a_k \cdot b_{n-k}\right)_{n=0}^{\infty} \end{aligned}$$

*The last one being convolution, which is such that: $(p \bullet q)(X) = p(X) \cdot q(X)$.*

**Definition 7.1.3** (Degree)**.** *Let* $\deg : F[X] \to \mathbb{N}$ *as:* $\deg(p) = n \Leftrightarrow a_n \neq 0$ *and* $\forall\, i > n\,,\ a_i = 0$.

**Theorem 7.1.1** (Euclidean Division of Polynomials)**.** *Let $a, b \in F[X]$, then,* $\exists!\, q, r \in F[X] \ :\ a = b \bullet q + r$ *and* $\deg(r) < \deg(b)$

*Proof.* If $\deg(a) < \deg(b)$, it is immediate that $a = b \bullet 0 + a$. Now, if $\deg(a) = n \geq \deg(b) = k$, we work by induction on $n$. Let $a(X) = a_n\,X^n + \cdots$ and

$b(X) = b_k X^k + \cdots$. Notice that $a(X) - \dfrac{a_n}{b_k} X^{n-k} \cdot b(X)$ is a polynomial with degree less than $n$, hence, we use our induction hypothesis. □

**Corollary 7.1.1** (Divisibility of Roots). *Let $p \in F[X]$ and $\alpha \in F : p(\alpha) = 0$. Then $\exists\, q \in F[X]\ :\ p(X) = (X - \alpha)q(X)$.*

**Lemma 7.1.1** (Sequence of Surprises). *A sequence of non-zero polynomials $S = (p_1, p_2, \cdots, p_n)$ with distinct degree, that is:*

$$\forall\, i, j \in \{1, 2, \cdots, n\} : i \neq j\,,\ \deg(p_i) \neq \deg(p_j)$$

*is linearly independent.*

*Proof.* Without loss of generality, let they be in ascending order of degree: $\deg p_1 < \deg p_2 < \cdots < \deg p_n$. We write:

$$\alpha_1 \cdot p_1(X) + \alpha_2 \cdot p_2(X) + \cdots + \alpha_n \cdot p_n(X) = 0$$

If we look at $\deg p_n$, we get: $\alpha_n \cdot X^{\deg p_n} = 0 \Rightarrow \alpha_n = 0$. Apply induction, with base case that a sequence of only one non-zero element is LI. □

**Theorem 7.1.2** (Multiplicity). *Given $p \in F[X] \setminus \{0\}$ and $\lambda \in F$, there is a unique $\mu \in \mathbb{N}$ and $q \in F[X]$ where $q(\lambda) \neq 0$ such that $p(X) = (X - \lambda)^{\mu}\, q(X)$. The unique $\mu$ is called the (algebraic) multiplicity of $\lambda$ in $p$ (denote $\mathrm{am}(\lambda)$ or $\mu_p(\lambda)$).*

*Proof.* If $p(\lambda) \neq 0$, we're done. Else $p(X) = (X - \lambda)\, q_1(X)$. If $q_1(\lambda) \neq 0$, we're done. Else $p(X) = (X - \lambda)^2\, q_2(X)$. It continues at most until $p(X) = a(X - \lambda)^{\deg p}$. □

**Proposition 7.1.1** (Multiplicity with Derivatives). *$\mu_p(\lambda) = \mu \Leftrightarrow p(\lambda) = p'(\lambda) = \cdots = p^{(\mu-1)}(\lambda) = 0$ and $p^{(\mu)}(\lambda) \neq 0$.*

## 7.2 Axioms of Rings

**Definition 7.2.1** (Ring). *A ring $R$ is a set with operations $(+ : R \times R \to R, \cdot : R \times R \to R)$ iff:*

| Properties | Definition |
| --- | --- |
| Commutative | $\forall\, \alpha, \beta \in R\,,\ \alpha + \beta = \beta + \alpha$ |
| Associative | $\forall\, \alpha, \beta, \gamma \in R\,,\ (\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ |
| Neutral Element | $\exists\, 0 \in R\,:\ \forall\, \alpha \in R\,,\ \alpha + 0 = \alpha$ |
| Inverse Element | $\forall\, \alpha \in R\,,\ \exists\, \beta \in R\,:\ \alpha + \beta = 0$ |
| Associative | $\forall\, \alpha, \beta, \gamma \in R\,,\ \alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$ |
| Distributive Right | $\forall\, \alpha, \beta, \gamma \in R\,,\ \alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$ |
| Distributive Left | $\forall\, \alpha, \beta, \gamma \in R\,,\ (\alpha + \beta) \cdot \gamma = \alpha \cdot \gamma + \beta \cdot \gamma$ |
| Unital Element | $\exists\, 1 \in R\,:\ \forall\, \alpha \in R\,,\ 1 \cdot \alpha = \alpha$ |

Examples include $F[X]$ and $M_n(F)$ with addition and multiplication of polynomials and matrices, respectively. Also, notice that fields are special cases of rings.

# 8 Vector Spaces

## 8.1 Axioms of Vector Spaces

**Definition 8.1.1** (Vector Space)**.** *A vector space $V$ over a field $F$ is a set with operations* $(+ : V \times V \to V, \cdot : F \times V \to V)$ *iff:*

| Properties | Definition |
|---|---|
| Commutative | $\forall\, u, v \in V\,,\ u + v = v + u$ |
| Associative | $\forall\, u, v, w \in V\,,\ (u + v) + w = u + (v + w)$ |
| Neutral Element | $\exists\, 0 \in V\ :\ \forall\, u \in V\,,\ u + 0 = u$ |
| Inverse Element | $\forall\, u \in V\,,\ \exists\, v \in V\ :\ u + v = 0$ |
| Associative | $\forall\, u \in V\,,\ \forall\, \alpha, \beta \in F\,,\ \alpha \cdot (\beta \cdot u) = (\alpha \cdot \beta) \cdot u$ |
| Distributive Right | $\forall\, u, v \in V\,,\ \forall\, \alpha \in F\,,\ \alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$ |
| Distributive Left | $\forall\, u \in V\,,\ \forall\, \alpha, \beta \in F\,,\ (\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u$ |
| Unitary | $\forall\, u \in V\,,\ 1_F \cdot u = u$ |

*Usually, $+$ is called addition and $\cdot$ is called scalar multiplication. Also, we often denote the inverse of addition as $-u$.*

**Definition 8.1.2** (Group)**.** *The set $G$ is a group with operation $* : G \times G \to G$ iff:*

| Properties | Definition |
|---|---|
| Associative | $\forall\, a, b, c \in G\,,\ (a * b) * c = a * (b * c)$ |
| Neutral Element | $\exists\, e \in G\ :\ \forall\, a \in G\,,\ a * e = e * a = a$ |
| Inverse Element | $\forall\, a \in G\,,\ \exists\, b \in G\ :\ a * b = b * a = e$ |

Further, if the operation is commutative, the group is abelian. A field $F$ is both an abelian group $(F,+)$ and an abelian group $(F\setminus\{0\},\cdot)$. Moreover, a vector space is an abelian group $(V,+)$.

**Example 8.1.1.** *$F^n, F^{k\times n}, F[X]$ and $F^\infty$ are vector spaces we saw before.*

*The set ${}^S V$ of all functions from a set $S$ is to a vector space $V$ is again a vector space with $+$ and $\cdot$ inherited from $V$ as follows:*

$$\begin{aligned}(f+g)(s) &= f(s)+g(s)\\ (\alpha\cdot f)(s) &= \alpha\cdot f(s)\end{aligned}$$

*If $U$ and $V$ are vector spaces over $F$, then $U\times V$ is a vector space again a vector space with $+$ and $\cdot$ inherited from $U$ and $V$ as follows:*

$$\begin{aligned}(u_1,v_1)+(u_2,v_2) &= (u_1+u_2, v_1+v_2)\\ \alpha\cdot(u,v) &= (\alpha\cdot u, \alpha\cdot v)\end{aligned}$$

*The following is also a vector space $F_n[X]=\big\{p\in F[X] \mid \deg(p)\le n\big\}$.*

## 8.2 Spans and Subspaces

**Definition 8.2.1** (Finite Span)**.** *For a set $S\subseteq V$, we define:*

$$\mathrm{Span}(S)=\left\{\sum_{i=1}^{k}\alpha_i\cdot v_i \;\middle|\; k\in\mathbb{N}\,,\, v_i\in S\,,\,\alpha_i\in F\right\}$$

*If $S$ is finite, or at least countable, we may write it as a sequence.*

**Definition 8.2.2** (Vector Subspace)**.** *Given a vector space $(V,F,+,\cdot)$, a subset $U$ is a linear subspace if $(U,F,+|_{U\times U}\,,\cdot|_{F\times U})$ is a vector space, where $+|_{U\times U}$ and $\cdot|_{F\times U}$ are the operations $+$ and $\cdot$ restricted to $U$. Particularly, we just need to check:*

1. *$U$ is non-empty*
2. *$U$ is closed under addition*
3. *$U$ is closed under scalar multiplication*

**Definition 8.2.3** (Finitely Spanned Vector Spaces)**.** *Given a vector space $V$ over $F$, we say it is finitely spanned, if $\exists\, S \in V^k \,:\, \mathrm{Span}(S) = V$, for some $k$ finite.*

**Definition 8.2.4** (Spanning Sequence)**.** *Given a subspace $U$, we say $S \in V^k$, for $k$ finite, is a spanning sequence of $U$ if: $U = \mathrm{Span}(S)$. Equivalently, we write $U$ is spanned by $S$. Note, since $U$ is a linear subspace,* $\mathrm{Span}(S) \subseteq U \Leftrightarrow S \subseteq U$*.*

**Lemma 8.2.1** (Concatenation on Span)**.** *Let $S = (v_1, v_2, \cdots, v_n) \in V^n$ and $T = (v_1, \cdots, v_n, u) = S +\!\!+\, u \in V^{n+1}$. Then:*

1. $\mathrm{Span}(T) \supseteq \mathrm{Span}(S)$
2. $\mathrm{Span}(T) = \mathrm{Span}(S) \Leftrightarrow u$ *is a linear combination of $S$*
3. *$T$ is linearly independent iff $S$ is linearly independent and* $u \notin \mathrm{Span}(S)$

*Proof.* For each one:

1. $\mathrm{Span}(S) \ni \sum_{i=1}^{n} \alpha_i \cdot v_i = 0 \cdot u + \sum_{i=1}^{n} \alpha_i \cdot v_i \in \mathrm{Span}(T)$.

2. ($\Leftarrow$) We only need $\subseteq$. Let $u = \sum_{i=1}^{n} \beta_i \cdot v_i$ so that $\mathrm{Span}(T) \ni \beta \cdot u + \sum_{i=1}^{n} \alpha_i \cdot v_i = \sum_{i=1}^{n} (\beta \cdot \beta_i + \alpha_i) \cdot v_i \in \mathrm{Span}(S)$.

   ($\Rightarrow$) For $\beta \neq 0$, $\mathrm{Span}(T) \ni \beta \cdot u + \sum_{i=1}^{n} \alpha_i \cdot v_i = \sum_{i=1}^{n} \beta_i \cdot v_i \in \mathrm{Span}(S) \Rightarrow u = \sum_{i=1}^{n} \frac{\beta_i - \alpha_i}{\beta} \cdot v_i$.

3. ($\Leftarrow$) $\beta \cdot u + \sum_{i=1}^{n} \alpha_i \cdot v_i = 0$. Since $u \notin \mathrm{Span}(S)$, $\beta = 0$. Since $S$ is linearly independent $\alpha_1 = \cdots = \alpha_n = 0$.

($\Rightarrow$) By contrary, $\beta \cdot u + \sum_{i=1}^{n} \alpha_i \cdot v_i = 0$, let $u = \sum_{i=1}^{n} \beta_i \cdot v_i$, set $\beta = 1$ and $\alpha_i = -\beta_i$, so we found a linear dependency. If $u \in \mathrm{Span}(S)$, . If $S$ is linearly dependent we set $\beta = 0$ and use any $(\alpha_1, \cdots, \alpha_n) \in \mathrm{LD}(S) \setminus \{0\}$.

$\square$

## 8.3 Basis and Dimension

**Definition 8.3.1** (Hamel Basis)**.** *A set $B$ is called a (Hamel) basis of a vector space $V$ iff*

1. *Any finite subset is linearly independent*
2. $\forall\, v \in V\,,\ \exists!\, \underline{x} \in F^k, (b_1, b_2, \cdots, b_k) \subseteq B\ :\ \sum_{i=1}^{h} x_i \cdot b_i = v$

*If $B$ is finite, or at least countable, we write it as a sequence. Heretofore, we only consern ourselves with this case.*

**Lemma 8.3.1** (N&SC for Basis)**.** *$B$ is a basis of $U \subseteq V$ iff $B$ is a spanning sequence of $U$ and linearly independent.*

**Theorem 8.3.1** (Maximality of Basis)**.** *Let $V$ be a vector space over $F$, and $A \in V^k$ be a linearly independent sequence in $V$. Moreover, let $S$ be any spanning sequence of $V$ ($V = \mathrm{Span}(S)$). Then: $\#A \leq \#S$.*

*Proof.* Let $S = (v_1, v_2, \cdots, v_n) \in V^n$ such that $V = \mathrm{Span}(S)$ and $T = (u_1, u_2, \cdots, u_k) \in V^k$ with $k > n$. We want to show that $T$ is linearly dependent. Since $V = \mathrm{Span}(S)$, we can find coefficients such that:

$$u_i = \sum_{j=1}^{n} a_{i,j} \cdot v_j$$

To find LD($T$) we solve $\sum_{i=1}^{k} x_i \cdot u_i = 0$, that is,

$$\sum_{j=1}^{n} \left( \sum_{i=1}^{k} a_{i,j} \cdot x_i \right) \cdot u_j = 0$$

In particular, that solves $L = \left\{ \sum_{i=1}^{k} a_{i,j} \cdot x_i = 0 \;\middle|\; j \in \{1, 2, \cdots, n\} \right\}$ belongs to $LD(T)$. Notice this is a linear equation system of $k$ variables and $n$ equations, with $n < k$. Thereofore, there is a non-trivial solution. □

**Corollary 8.3.1** (Equality on Dimension). *Let $B_1$ and $B_2$ be two bases of $V$. Then, $\#B_1 = \#B_2$.*

**Definition 8.3.2** (Dimension). *Let $V$ be a vector space over $F$, that has a finite basis $B$. We define:*

$$\boxed{\dim V := \#B}$$

**Corollary 8.3.2** (Sequence larger than dimension). *For every $S \in V^k$, if $k > \dim V$, $S$ is linearly dependent.*

**Lemma 8.3.2** (Exact Span). *Let $S = (v_1, v_2, \cdots, v_n) \in V^n$. Then $S$ contains a basis of* Span($S$).

*Proof.* $S$ of course spans Span($S$). If $S$ is linearly independent, we are done. Conversly, if $S$ is linearly dependent, $\exists\, v_i \in S$ that is a linear combination of the rest. Take $T$ such that $T +\!\!+\, v_i = S$. Notice Span($T$) = Span($S$). Since $S$ is finite, the algorithm terminates. □

**Corollary 8.3.3** (Finite Span is very easy). *Every linear space which is finitely spanned has a basis (and therefore a dimension).*

**Definition 8.3.3** (Extension of Basis). *Let $U$ be a linear subspace of $V$ and $S = (v_1, v_2, \cdots, v_n) \in V^n$. Then $S$ can be extended to a basis of $U$ iff: $\exists\, T \supseteq S : T$ is a basis of $U$.*

**Lemma 8.3.3** (Steinitz Exchange Lemma)**.** *The necessary and sufficient criteria for $S$ to be able to be extended to basis are:*

1. *$S \subset U$*
2. *$S$ is linearly independent*

*Proof.* One direction ($\Rightarrow$) is trivial. The other: ($\Leftarrow$) By contrary, it is clear that $\text{Span}(S) \subseteq U$ and $S$ is a basis of $\text{Span}(S)$. If $\text{Span}(S) = U$, we are done. Conversely, if $\text{Span}(S) \subsetneqq U$, $\exists\, u \neq 0 \in U \setminus \text{Span}(S)$. Take $T = S \mathbin{++} \underline{u}$, since $u \notin \text{Span}(S)$. See that $T$ is also linearly independent and $T \subset U$. Since $\dim V$ is finite, the algorithm terminates (hence, $\#S \leq \dim V$). □

**Proposition 8.3.1** (Dimension of Subspace)**.** *Let $U$ be a linear subspace of $V$. Then the following statements hold:*

1. *$U$ has a basis*
2. $\dim U \leq \dim V$
3. $\dim U = \dim V \Leftrightarrow U = V$

*Proof.* We prove each one:

1. If $U = \{0\}$, we're done. Otherwise, pick $u \neq 0 \in U$ and set $S = (u)$. It fulfills the condition for extension. Hence, $U$ has a basis.
2. If $\dim U > \dim V$, there is a $\dim U$-long linearly independent sequence in $V$. Contradiction.
3. Let $B$ be a basis of $U$. If $\exists\, v \in V \setminus U$, then $B \mathbin{++} v$ is a $\dim U + 1 = \dim V + 1$-long linearly independent sequence. Contradiction.

□

**Theorem 8.3.2** (Size of Sequence)**.** *Let* $S \in V^k$

1. *If* $k > \dim V$*, then* $S$ *is linearly independent*
2. *If* $k < \dim V$*, then* $\mathrm{Span}(S) \subsetneqq V$
3. *If* $k = n$ *then either* $S$ *is a basis of* $V$ *or both* $\mathrm{Span}(S) \subsetneqq V$ *and* $S$ *is linearly dependent*

# 9 Linear Transformations

## 9.1 Linear Maps

**Definition 9.1.1** (Linear Map)**.** *A linear map between vector spaces $V$ and $W$ over the same field $F$ is a function $T : V \to W$ such that:*

| | |
|---|---|
| Additive | $\forall\, u, v \in V\ ,\ T(u + v) = T(u) + T(v)$ |
| Homogeneous | $\forall\, \alpha \in F\ ,\ \forall\, v \in V\ ,\ T(\alpha \cdot v) = \alpha \cdot T(v)$ |

*written as one:*

| | |
|---|---|
| Linearity | $\forall\, \alpha, \beta \in F\ ,\ \forall\, u, v \in V\ ,\ T(\alpha \cdot u + \beta \cdot v) = \alpha \cdot T(u) + \beta \cdot T(v)$ |

*equivalently, these diagram commute:*

$$\begin{tikzcd} V \arrow[d, "T"] & \times & V \arrow[d, "T"] \arrow[r, "+"] & V \arrow[d, "T"] \\ W & \times & W \arrow[r, "+"] & W \end{tikzcd} \quad \textit{and} \quad \begin{tikzcd} F \times V \arrow[d, "T"] \arrow[r, "\cdot"] & V \arrow[d, "T"] \\ F \times W \arrow[r, "\cdot"] & W \end{tikzcd}$$

**Definition 9.1.2** (Homomorphisms and Endomorphisms)**.** *Denote the set of all linear maps $T : V \to W$ as* $\mathrm{Hom}(V, W)$*. Also,* $\mathrm{Hom}(V, V) = \mathrm{End}(V)$*.*

**Lemma 9.1.1** (Operations on Linear Map)**.** *Let $V$ and $W$ be linear spaces over $F$.*

1. *If $T$ and $S$ are linear transformations $V \to W$, then $T \pm S$ and $\alpha \cdot T$ are linear transformations $V \to W$.*
2. *If $T : V \to W$ and $S : W \to U$, are linear transformations, then $S \circ T : V \to U$ is a linear transformation.*

3. *If $T : V \to W$ is an invertible linear transformation, then $T^{-1} : W \to V$ is a linear transformation.*

*Proof.* 1. and 2. are trivial. We prove 3. First, notice $\mathrm{Id}_V : V \to W$ is a linear transformation.

$$\begin{aligned} T^{-1}\big(\alpha \cdot T(u) + \beta \cdot T(v)\big) &= T^{-1} \circ T(\alpha \cdot u + \beta \cdot v) \\ &= Id_V(\alpha \cdot u + \beta \cdot v) = \alpha \cdot u + \beta \cdot v = \alpha \cdot T^{-1}(T(u)) + \beta \cdot T^{-1}(T(v)) \end{aligned}$$

□

**Corollary 9.1.1** (Linear Maps are Vector Spaces)**.** $\mathrm{Hom}(V, W)$ *is a vector space over $F$ wrt to the operations we defined for ${}^S V$. Further,* $\mathrm{End}(V)$ *is a ring with composition.*

## 9.2 Kernel and Image and Dimension Theorem

**Lemma 9.2.1** (Image of Linear Map)**.** *Let $T : V \to W$ is linear transformation, then* $\mathrm{Im}(T)$ *is a linear subspace of $W$.*

*Proof.* $\mathrm{Im}(T)$ is closed under addition and scalar multiplication: $\alpha \cdot T(u) + \beta \cdot T(v) = T(\alpha \cdot u + \beta \cdot v) \in \mathrm{Im}(T)$ and is not empty since $T(0_V) = 0_W$. □

**Definition 9.2.1** (Sequence Map)**.** *Let $V, W$ be two vectors spaces over $F$. Let $f : V \to W$ be a function and $S = (v_1, v_2, \cdots, v_k) \in V^k$, we define:*

$$f(S) = \big(f(v_1), f(v_2), \cdots, f(v_k)\big) \in W^k$$

**Lemma 9.2.2** (Span and LD of Sequence Map)**.** *Let $T : V \to W$ be a linear transformation and $S \in V^k$ be a sequence in the domain.*

1. *If $V = \mathrm{Span}(S)$, then $\mathrm{Im}(T) = \mathrm{Span}(T(S))$*
2. *In particular, if $\dim V < \dim W$ then $T$ is not surjective*
3. $\mathrm{LD}(T(S)) \supseteq \mathrm{LD}(S)$
4. *If $T$ is injective, then $\mathrm{LD}(T(S)) = \mathrm{LD}(S)$.*

*Proof.* We prove each one:

1. $u \in \mathrm{Im}(T) \Leftrightarrow \exists\, v \in V = \mathrm{Span}(S) \;:\; u = T(v) = T\left(\sum_{i=1}^{k} \alpha_i \cdot v_i\right) = \sum_{i=1}^{k} \alpha_i \cdot T(v_i) \Leftrightarrow u \in \mathrm{Span}(T(S))$.

2. Take a basis $B$ of $V$: $\dim \mathrm{Im}(T) = \mathrm{Span}(T(B)) \leq \#T(B) = \dim V < \dim W \Rightarrow \mathrm{Im}(T) \subsetneqq W$.

3. $\underline{x} \in \mathrm{LD}(S) \Rightarrow S \bullet \underline{x} = \underline{0}_V \Rightarrow T(S) \bullet \underline{x} = T(S \bullet \underline{x}) = \underline{0}_W \Rightarrow \underline{x} \in \mathrm{LD}(T(S))$.

4. If $T$ is injective, there is a left inverse, so that: $\underline{x} \in \mathrm{LD}(T(S)) \Rightarrow T(S \bullet \underline{x}) = T(S) \bullet \underline{x} = 0_W \Rightarrow S \bullet \underline{x} = 0_V \Rightarrow \underline{x} \in \mathrm{LD}(S)$.

□

**Definition 9.2.2** (Kernel)**.** *Let $T : V \to W$ be a linear transformation, the kernel is the set:* $\ker(T) = \{v \in V \mid T(v) = 0_W\}$

**Lemma 9.2.3** (Kernel of linear map is Subspace)**.** *The kernel is a linear subspace of $V$.*

*Proof.* $\ker(T)$ is closed under addition and scalar multiplication: $u, v \in \ker(T) \Rightarrow 0 = \alpha \cdot T(u) + \beta \cdot T(v) = T(\alpha \cdot u + \beta \cdot v) \Rightarrow \alpha \cdot u + \beta \cdot v \in \ker(T)$ and is not empty since $T(0_V) = 0_W \Rightarrow 0_V \in \ker(T)$. □

**Lemma 9.2.4** (N&SC for Injectivity of Linear Map)**.** *Let $T : V \to W$ be a linear transformation.*

1. *$T$ is injective* $\Leftrightarrow \ker(T) = \{0_V\}$.

2. *For $V$ finetely spanned, $T$ is injective* $\Leftrightarrow \dim \mathrm{Im}(T) = \dim V$.

*Proof.* We prove each one:

1. We prove both directions:

($\Rightarrow$) $T(u) = 0_W = T(0_V) \Rightarrow u = 0_V$, that is, $\{0_V\} \supseteq \ker(T)$.

($\Leftarrow$) $T(u) = T(v) \Rightarrow T(u - v) = 0_W \Rightarrow u - v = 0_V \Rightarrow u = v$

2. We prove both directions: Let $B$ be a basis of $V$.

   ($\Rightarrow$) $\mathrm{LD}(T(B)) = \mathrm{LD}(B) = \{0\} \Rightarrow T(B)$ is linearly independent, meaning $\dim \mathrm{Im}(T) = \dim \mathrm{Span}\, T(B) = \dim V$.

   ($\Leftarrow$) $\dim \mathrm{Span}\, T(B) = \dim \mathrm{Im}(T) = \dim V$. Hence $T(B)$ is linearly independent, which implies $\ker(T) = \{0_V\}$.

□

**Theorem 9.2.1** (Dimension Theorem)**.** *If $V$ is finitely spanned, for any linear transformation $T : V \to W$:*

$$\boxed{\dim V = \dim \ker(T) + \dim \mathrm{Im}(T)}$$

*Proof.* $V$ is finitely spanned vector space ($\dim V = n$) and $\ker(T)$ is a linear subspace of $V$, therefore $\ker(T)$ is finitely spanned, hence it has a basis $A = (a_1, a_2, \cdots, a_k)$, where $k = \dim \ker(T)$.

Since $A$ is a linear independent, we can extend it to a basis of $V$, call it $A +\!\!+ B$, where $B = (b_1, b_2, \cdots, b_{n-k})$ is the remainder of the extension.

Let us show that $T(B)$ is a basis for $\mathrm{Im}(T)$: $\mathrm{Im}(T) = \mathrm{Span}(T(B))$ is easy since $\mathrm{Im}(T) = \mathrm{Span}(T(A +\!\!+ B)) = \mathrm{Span}(T(A) +\!\!+ T(B)) = \mathrm{Span}(T(B))$, we only need to prove $T(B)$ is linearly independent: $0 = \sum_{i=1}^{n-k} x_i \cdot T(b_i) = T\left(\sum_{i=1}^{n-k} x_i \cdot b_i\right) \Leftrightarrow \mathrm{Span}(B) \ni \sum_{i=1}^{n-k} x_i \cdot b_i \in \ker(T) = \mathrm{Span}(A) \Rightarrow x_1 = x_2 = \cdots = x_{n-k} = 0$ since $\mathrm{Span}(A) \cap \mathrm{Span}(B) = \varnothing$. Hence, $\dim \mathrm{Im}(T) = n - k$ □

**Corollary 9.2.1** (Character from Dimension)**.** *We have the ternary:*

1. *If $\dim V > \dim W$, then $T$ is not injective.*
2. *If $\dim V < \dim W$, then $T$ is not surjective.*

3. *If* $\dim V = \dim W$*, then* $T$ *is either bijective or it is neither surjective nor injective.*

**Corollary 9.2.2** (Sequence Dimensions)**.** *If we use for* $F_S$ *(to be defined on the next section), we get: For any sequence* $S \in V^n$*:*

$$n = \dim \mathrm{LD}(S) + \dim \mathrm{Span}(S)$$

**Corollary 9.2.3** (Dimension of Cols and Rows)**.** *For a matricial function* $T_A$*,* $A \in M_{k\times n}(F)$ *and for the matricial function* $T_{A^t}$*, we get:*

$$n = \dim \mathrm{sols}(A) + \dim \mathrm{cols}(A)$$
$$k = \dim \mathrm{sols}(A^t) + \dim \mathrm{rows}(A)$$

## 9.3 Isomorphism

**Definition 9.3.1** (Isomorphisms)**.** *A linear transformation* $T : V \to W$ *is an:*

1. *monomorphism if it is injective*

2. *epimorphism if it is surjective*

3. *isomorphism if it is bijective*

**Definition 9.3.2** (Isomorphic Vector Spaces)**.** *Two vectors spaces are* ***isomorphic****, denoted* $V \cong W$*, if there exists a* ***bijective*** *linear transformation* $T : V \to W$*, that is, an isomorphism.*

**Proposition 9.3.1** (Isomorphism Equivalence Relation)**.** *Isomorphism of vector spaces is an equivalence relation.*

*Proof.* We choose the following maps: Reflexivity: Take the identity $\mathrm{Id}_V : V \to V$; Symmetry: Take the inverse function (which exists, since it is bijective); Transitivity: Take the composition (which is linear and bijective). □

# 10 Coordinates

## 10.1 Representing Function

**Definition 10.1.1** (Sequence Function and Coordinate Map). *For $S \in V^n$, we define the sequence function:*

$$\begin{aligned} F_S : F^n &\to V \\ \underline{x} &\mapsto S \bullet \underline{x} \end{aligned}$$

*which is a linear map. Moreover:* $\ker(F_S) = \mathrm{LD}(S)$ *and* $\operatorname{Im} F_S = \operatorname{Span}(S)$.

*If $n = \dim V$ and $B$ is a basis of $V$, $F_B$ is bijective, so there is an (linear) inverse function $Q_B = F_B^{-1} : V \to F^{\dim V}$ which is called the coordinate function of $S$. We further denote $Q_B(v) = [v]_B$.*

*Further, for a sequence $S \in V^k$, we write $[S]_B = Q_B(S) \in M_{\dim V \times k}(F)$*

**Theorem 10.1.1** (Dimension Equality). *For $V$ and $W$ finitely spanned, we have:*

1. $V \cong F^{\dim V}$
2. $\dim_F V = \dim_F W \Leftrightarrow V \cong W$

*Proof.* Let $B$ be a basis for $V$. To show $V \cong F^{\dim V}$, take the coordinate function $Q_B$. To show the isomorphism $\dim_F V = \dim_F W \Leftrightarrow V \cong W$, take $F_C \circ Q_B$, where $C$ is a basis for $W$. $\square$

**Corollary 10.1.1** (Linear Maps are Sequence Functions). *Every linear map $T : F^n \to V$ is a sequence function $F_{T(E)}$, with $E = (\underline{e}_1, \underline{e}_2, \cdots, \underline{e}_n)$ is the standard basis.*

**Corollary 10.1.2** (Linear Maps are Matricial Functions). *Every linear map $T : F^n \to F^k$ is a matricial function. Moreover, every bijective linear transformation $T : V \to F^n$ is a coordinate map.*

**Definition 10.1.2** (Representing Function). *Let $B$ be a basis for $V$ and $C$ a basis for $W$. Let $f : V \to W$ be any function. Then, there is a function $f_C^B$ such that the following diagram commutes.*

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "Q_B"] & W \arrow[d, "Q_C"] \\ F^{\dim V} \arrow[r, dashed, "f_C^B"] & F^{\dim W} \end{tikzcd}$$

*In particular, $f_C^B = Q_C \circ f \circ Q_B^{-1} = Q_C \circ f \circ F_B$. so that $f(v) = u \Leftrightarrow f_C^B([v]_B) = [u]_C$. Moreover, $f = Q_C^{-1} \circ f_C^B \circ Q_B = F_C \circ f_C^B \circ Q_B$.*

**Proposition 10.1.1** (Character of Representing Function).

$$f \text{ is } \begin{cases} \text{injective} \\ \text{surjective} \\ \text{bijective} \\ \text{linear} \end{cases} \text{iff } f_C^B \text{ is}$$

**Definition 10.1.3** (Representing Matrix). *If $f$ is linear, then $f_C^B$ is a matricial function (by lemma). We call $[f_C^B]$ (also denoted $[f]_C^B$) the representing matrix of $f$ in $B, C$.*

**Proposition 10.1.2** (Calculation of Representing Matrix). *For $T : V \to W$ a linear map:*

$$[T]_C^B = [T(B)]_C = \begin{pmatrix} | & | & & | \\ [T(\underline{b}_1)]_C & [T(\underline{b}_2)]_C & \cdots & [T(\underline{b}_{\dim V})]_C \\ | & | & & | \end{pmatrix}$$

We can algorithmically compute $[T]_C^B$ by Gauss-Jordan Elimination: Take $E$ is a (usually standard) basis for $W$: $\left([C]_E \;\middle|\; [T(B)]_E\right) \to \left(I \;\middle|\; [T]_C^B\right)$.

**Proposition 10.1.3** (Representation of Composition).

$$\begin{tikzcd} U \arrow[r, "g"] \arrow[d, "Q_A"] & V \arrow[r, "f"] \arrow[d, "Q_B"] & W \arrow[d, "Q_C"] \\ F^{\dim U} \arrow[r, dashed, "g_B^A"] & F^{\dim V} \arrow[r, dashed, "f_C^B"] & F^{\dim W} \end{tikzcd}$$

*we have the following: $(g \circ f)_C^A = g_B^A \circ f_C^B$, for the diagram to commute.*

## 10.2 Change of Coordinates

**Definition 10.2.1** (Change-of-Coordinates Matrix)**.** *The change of coordinates in one vector space:*

$$\begin{tikzcd} & V \arrow[dl, "Q_B"'] \arrow[dr, "Q_C"] & \\ F^{\dim V} \arrow[rr, dashed, "Q_C^B"] & & F^{\dim V} \end{tikzcd}$$

*Where $Q_C^B = Q_C \circ Q_B^{-1} = F_C^{-1} \circ F_B$ is the change of coordinate function. We denote $M_C^B = [Q_C^B]$. So that:*

$$M_C^B \, [v]_B = [v]_C$$

*Equivalently, we define: $M_C^B = [\mathrm{Id}_V]_C^B$. Notice in a standard basis $E$: $M_C^B = [C]_E^{-1} \, [B]_E$*

**Proposition 10.2.1** (Changing Basis)**.** *For a linear map $T : V \to W$*

$$\begin{tikzcd} F^{\dim V} \arrow[r, dashed, "T_F^E"] & F^{\dim W} \arrow[dd, dashed, bend left, "Q_C^F"] \\ V \arrow[u, "Q_E"'] \arrow[r, "T"] \arrow[d, "Q_B"] & W \arrow[u, "Q_F"] \arrow[d, "Q_C"'] \\ F^{\dim V} \arrow[uu, dashed, bend left, "Q_E^B"] \arrow[r, dashed, "T_C^B"] & F^{\dim W} \end{tikzcd}$$

*We get:*

$$\begin{aligned} T_C^B &= Q_C \circ T \circ Q_B^{-1} = Q_C \circ \left(Q_F^{-1} \circ Q_F \circ T \circ Q_E^{-1} \circ Q_E\right) \circ Q_B^{-1} \\ &= \left(Q_C \circ Q_F^{-1}\right) \circ T_F^E \circ \left(Q_E \circ Q_B^{-1}\right) = Q_C^F \circ T_F^E \circ Q_E^B \end{aligned}$$

$\Rightarrow [T]_C^B = M_C^F \, [T]_F^E \, M_E^B$.

*For $V = W$, $B = C$ and $E = F$, we get: $[T]_B = M_B^E \, [T]_E \, M_E^B$ .*

**Proposition 10.2.2** (Calculation of Change-of-Coordinates Matrix)**.**

$$M_C^B = [B]_C = Q_C(B) = \begin{pmatrix} | & | & & | \\ [\underline{b}_1]_C & [\underline{b}_2]_C & \cdots & [\underline{b}_n]_C \\ | & | & & | \end{pmatrix}$$

**Proposition 10.2.3.** *We have a type of transitive law:* $M_C^B = M_D^B\, M_C^D$

Further, we can algorithmically compute $M_C^B$ by Gauss-Jordan Elimination: Take $E$ as standard basis and apply $\left([B]_E \,\middle|\, [C]_E\right) \to \left(I \,\middle|\, M_C^B\right)$.

**Definition 10.2.2** (Conjugation)**.** *Let* $P \in \mathrm{GL}_n(F)$*. The function:*

$$\begin{aligned} \Theta_P : M_n(F) &\to M_n(F) \\ A &\to P^{-1}\, A\, P \end{aligned}$$

*is called the conjugation (function) of* $P$*. Notice it is linear, and also:*

$$\Theta_P(AB) = \Theta_P(A) \cdot \Theta_P(B)$$

*(it is a ring homomorphism, i.e. it preserves the ring structure)*

*Further,* $\Theta_P^{-1} = \Theta_{P^{-1}}$*, so it is a bijection.*

**Definition 10.2.3** (Similar Matrices)**.** *Given two matrices* $A, B \in M_n(F)$*, we say:*

$$A \sim B \Leftrightarrow \exists\, P \in \mathrm{GL}_n(F)\ :\ \Theta_P(A) = P^{-1}\, A\, P = B$$

*the matrices are similar*

**Lemma 10.2.1** (Similarity Equivalence Relation)**.** *The similarity of matrices is an equivalence relation.*

*Proof.* We take the conjugation with the following matrices:

Reflexive: Take $I$.

Symmetric: Take $P^{-1}$.

Transitive: Take the product/composition.

$\square$

**Theorem 10.2.1** (Similar Matrices represent the same Linear Map)**.** *Let $V$ be a finitely spanned vector space over $F$.*

1. *Let $T : V \to V$ be a linear transformation and $B$ and $C$ be two basis of $V$. Then: $[T]_B = \Theta_P\big([T]_C\big)$ where $P = M_B^C$.*
2. *$\forall\, A, A' \in M_{\dim V}(F)\,,\ A \sim A' \Rightarrow \exists\, T \in \mathrm{End}(V)$ and $B, C$ basis of $V$ : $[T]_B = A$ and $[T]_C = A'$.*

*Proof.* We prove each one:

1. Simply notice: $[T]_B = M_C^B\,[T]_C\,M_B^C = \left(M_B^C\right)^{-1}[T]_C\,M_B^C$.
2. Let $P$ be such that $A = \Theta_P(A')$. Let $B$ be any basis of $V$. Simply define $T = Q_B^{-1} \circ T_A \circ Q_B$, so that $[T]_B = A$. Now, define $C$ as: $[C] = [B]\,P$ so that $P = M_B^C$. By the previous stament, $T_C = \Theta_P^{-1}([T]_B) = \Theta_{P^{-1}}(A) = A'$.

$\square$

# 11 Eigenspace

## 11.1 Eigenvectors

**Definition 11.1.1** (Eigenspace and Eigenvectors)**.** *Let $V$ be a linear space over the field $F$ and $T : V \to V$ be a linear transformation.*

*We say $v \in V \setminus \{0\}$ is an eigenvector of $T$ if $\exists\, \lambda \in F \;:\; T(v) = \lambda \cdot v$. We define:*

$$\mathrm{Eig}_\lambda(T) = \{v \in V \mid T(v) = \lambda \cdot v\} = \ker(T - \lambda \cdot \mathrm{Id}_V)$$

*notice it is a linear subspace of $V$.*

**Definition 11.1.2** (Eigenvalue)**.** *The number $\lambda$ is an eigenvalue of $T$ if* $\mathrm{Eig}_\lambda(T) \supsetneqq \{0\}$.

**Lemma 11.1.1** (Eigenspaces are disjoint)**.**

$$\lambda \neq \mu \Rightarrow \mathrm{Eig}_\lambda(T) \cap \mathrm{Eig}_\mu(T) = \{0\}$$

*Proof.* Let $v \in \mathrm{Eig}_\lambda(T) \cap \mathrm{Eig}_\mu(T) \Rightarrow T(v) = \lambda \cdot v = \mu \cdot v \Rightarrow (\lambda - \mu) \cdot v = 0 \Rightarrow v = 0$. So, $\mathrm{Eig}_\lambda(T) \cap \mathrm{Eig}_\mu(T) \subseteq \{0\}$. Moreover, $\mathrm{Eig}_\lambda(T) \cap \mathrm{Eig}_\mu(T) \supseteq \{0\}$ is clear. □

**Lemma 11.1.2** (Power of Linear Maps)**.** *If $\lambda$ is an eigenvalue of $T : V \to V$, then $\forall\, k \in \mathbb{N}\,,\; \lambda^k$ is an eigenvalue of $T^k = \underbrace{T \circ T \circ \cdots \circ T}^{k \text{ times}}$. Moreover,* $\mathrm{Eig}_\lambda(T) \subseteq \mathrm{Eig}_{\lambda^k}(T^k)$

*Proof.* If $\lambda$ is an eigenvalue, there is at least one eigenvector $u \neq 0$. Therefore, $T^k(u) = T^{k-1}(\lambda \cdot u) = \lambda \cdot T^{k-1}(u) = \cdots = \lambda^k \cdot u$ , hence $\lambda^k$ is an eigenvalue of $T^k$. Further, we showed $u \in \mathrm{Eig}_\lambda(T) \Rightarrow u \in \mathrm{Eig}_{\lambda^k}(T^k)$. □

**Lemma 11.1.3** (Reciprocal of Eigenvalue)**.** *If $T$ is invertible, $\lambda^{-1}$ is an eigenvalue of $T^{-1}$.*

*Proof.* If $T$ is invertible: $u = T^{-1} \circ T(u) = T^{-1}(\lambda \cdot u) = \lambda \cdot T^{-1}(u) \Rightarrow T^{-1}(u) = \lambda^{-1} \cdot u$ , hence $\lambda^{-1}$ is an eigenvalue of $T^{-1}$. Further, observe that $T$ invertible $\Rightarrow$ $T$ injective $\Rightarrow \ker(T) = \{0\} \Rightarrow \lambda = 0$ is not an eigenvalue. $\square$

**Corollary 11.1.1** (N&SC for Invertibility)**.** *If $\lambda = 0$ is an eigenvalue of $T$, then $T$ is not invertible.*

## 11.2 Characteristic Polynomial

**Proposition 11.2.1** (Determinant in Finite Case)**.** *Let $V = F^n$ and $T = T_A : F^n \to F^n$, for a matrix $A$. We have:*

$$\{0\} \subsetneqq \mathrm{Eig}_\lambda(T) = \ker\big(T - \lambda\, \mathrm{Id}_V\big) = \mathrm{sols}(A - \lambda\, I) \Leftrightarrow \det\big(A - \lambda\, I\big) = 0$$

**Definition 11.2.1** (Characteristic Polynomial)**.** *We define*

$$p_A(\lambda) = \det\big(\lambda\, I - A\big)$$

*the characteristic polynomial of $A$ (which is a monic polynomial). Notice the eigenvalues of $T_A$ are exactly the roots of $p_A$.*

**Theorem 11.2.1** (Cayley-Hamilton)**.**

$$p_A(A) = 0$$

*Proof.* Let $B = \mathrm{adj}(\lambda\, I - A)$. First, we must have $(\lambda\, I - A)\, B = \det(\lambda\, I - A)\, I = p_A(\lambda)\, I$. Now, we can expand $B$ as $B = \sum_{k=0}^{n-1} \lambda^k\, B_i$. Now,

$$p_A(\lambda)\, I = (\lambda\, I - A)\, B = (\lambda\, I - A) \sum_{k=0}^{n-1} \lambda^k\, B_i$$

$$= \sum_{k=0}^{n-1} \lambda^{k+1}\, B_i - \sum_{k=0}^{n-1} \lambda^k\, A\, B_i = \lambda^n\, B_{n-1} + \sum_{k=1}^{n-1} \lambda^k \big(B_{k-1} - A\, B_k\big) - A\, B_0$$

$$p_A(\lambda)\, I = \sum_{k=0}^{n} c_k\, \lambda^k\, I \Rightarrow \begin{cases} B_{n-1} = c_n\, I = I \\ B_{k-1} - A\, B_k = c_k\, I \\ -AB_0 = c_0\, I \end{cases}$$

Therefore,

$$
\begin{aligned}
p_A(A) = \sum_{k=0}^{n} A^k \left(c_k\, I\right) &= A^n\, B_{n-1} + \sum_{k=1}^{n-1} A^k \big(B_{k-1} - A\, B_k\big) - A\, B_0 \\
&= A^n\, B_{n-1} + \sum_{k=1}^{n-1} A^k\, B_{k-1} - \sum_{k=1}^{n-1} A^{k+1}\, B_k - A\, B_0 \\
&= A^n\, B_{n-1} + A\, B_0 - A^n\, B_{n-1} - A\, B_0 = 0
\end{aligned}
$$

□

## 11.3 Diagonalizing

**Definition 11.3.1** (Diagonalizable Matrix)**.** *A matrix is diagonalizable if is similar to a diagonal matrix* $\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$*, or, equivalently* $\exists\, P \in \mathrm{GL}_n(F)\ :\ A = P\,\Lambda\,P^{-1}$.

**Theorem 11.3.1** (EigenBasis)**.** *If* $P^{-1}\,A\,P$ *is a diagonal matrix* $\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$*, then* $\forall\, \underline{p}_i \in S_c(P)\,,\ A\,\underline{p}_i = \lambda \cdot \underline{p}_i$.

*Proof.* $P^{-1}\,A\,P = \Lambda \Leftrightarrow A\,P = P\,\Lambda$:

$$
\begin{aligned}
A\,P = A \begin{pmatrix} | & | & & | \\ \underline{p}_1 & \underline{p}_2 & \cdots & \underline{p}_n \\ | & | & & | \end{pmatrix} &= \begin{pmatrix} | & | & & | \\ A(\underline{p}_1) & A(\underline{p}_2) & \cdots & A(\underline{p}_n) \\ | & | & & | \end{pmatrix} \\
= P\,\Lambda = P \begin{pmatrix} | & & | \\ \lambda_1 \cdot \underline{e}_1 & \cdots & \lambda_n \cdot \underline{e}_n \\ | & & | \end{pmatrix} &= \begin{pmatrix} | & & | \\ \lambda_1 \cdot P(\underline{e}_1) & \cdots & \lambda_n \cdot P(\underline{e}_n) \\ | & & | \end{pmatrix} \\
= \begin{pmatrix} | & | & & | \\ \lambda_1 \cdot \underline{p}_1 & \lambda_2 \cdot \underline{p}_2 & \cdots & \lambda_n \cdot \underline{p}_n \\ | & | & & | \end{pmatrix}
\end{aligned}
$$

□

**Corollary 11.3.1** (N&SC for Diagonalizability). *$A$ is diagonalizable iff there is a sequence on $n$ linearly independent eigentuples $\underline{p}_i$ (called an eigenbasis).*

**Theorem 11.3.2** (Sylvester's law of Inertia). *Every symmetric matrix is diagonalizable*

**Definition 11.3.2** (Geometric Multiplicity).

$$\mathrm{gm}_A(\lambda) = \dim \mathrm{Eig}_\lambda(T_A) = \dim \mathrm{sols}(A - \lambda\, I)$$

**Definition 11.3.3** (Algebraic Multiplicity). $\mathrm{am}_A(\lambda)$ *is the multiplicity of $\lambda \in F$ in the polynomial $p_A$.*

**Lemma 11.3.1** (AM-GM Inequality). $\forall\, \lambda \in F\,,\ \mathrm{am}_A(\lambda) \geq \mathrm{gm}_A(\lambda)$.

**Theorem 11.3.3** (N&SC for Diagonalizability). *$A$ is a diagonalizable iff $p_A$ can be split into linear factors and $\forall\, \lambda \in F\,,\ \mathrm{am}_A(\lambda) = \mathrm{gm}_A(\lambda)$.*

*Proof.* We use the previous theorem to show that $p_A$ can be split into linear factors and $\forall\, \lambda \in F\,,\ \mathrm{am}_A(\lambda) = \mathrm{gm}_A(\lambda) \Leftrightarrow$ there is an eigenbasis for $A$. Notice, from the lemma above:

$$n = \sum_{\lambda \in F} \mathrm{am}_A(\lambda) \geq \sum_{\lambda \in F} \mathrm{gm}_A(\lambda)$$

($\Rightarrow$) Each $\mathrm{Eig}_\lambda(T_A)$ has a basis $B_\lambda$, which are linearly independent from each other so $B = B_{\lambda_1} +\!\!+\, B_{\lambda_2} +\!\!+ \cdots +\!\!+\, B_{\lambda_N}$ is a linearly independent set. Further, $n = \sum_{\lambda \in F} \mathrm{am}_A(\lambda) = \sum_{\lambda \in F} \mathrm{gm}_A(\lambda)$, so the length of $B$ is $n$. Hence, it is an eigenbasis.

($\Leftarrow$) By contrary, if $\exists\, \lambda \in F\ :\ \mathrm{am}_A(\lambda) > \mathrm{gm}_A(\lambda) \Rightarrow n > \sum_{\lambda \in F} \mathrm{gm}_A(\lambda)$. If there is an eigenbasis $B$, $\dim \mathrm{Span}(B) = \sum_{\lambda \in F} \mathrm{gm}_A(\lambda) < n$, contradiction.

$\square$

**Definition 11.3.4** (Diagonalizable Linear Map). *Let $V$ be a finetely spanned vector space over $F$ and $T : V \to V$ be a linear transformation. $T$ is called diagonalizable if there is a basis $B$ of $V$ such that $[T]_B$ is diagonal.*

**Lemma 11.3.2** (EigenBasis of Linear Map)**.** *If $B = (b_1, b_2, \cdots, b_n)$ satisfies that $[T]_B$ is diagonal, then each $b_i$ must be an eigenvector of $T$.*

*Proof.* Let $[T]_B = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$, then:

$$\underline{e}_i = Q_B(b_i) \Rightarrow T_B(\underline{e}_i) = \lambda_i \cdot \underline{e}_i \Rightarrow T(b_i) = \lambda_i \cdot b_i$$

$\square$

*Moreover,* $T(v) = \lambda \cdot v \Leftrightarrow [T]_B\,[v]_B = \lambda \cdot [v]_B$.

**Example 11.3.1.** $V = \mathbb{Q}_1[X]$,

$$\begin{aligned} T : \mathbb{Q}_1[X] &\to \mathbb{Q}_1[X] \\ a + bX &\mapsto (a + 2b) + (2a + b)X \end{aligned}$$

*If we pick* $E = (1, X)$*, we find* $A = [T]_E = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$*, calculating:* $p_A(\lambda) = \lambda^2 - 2\lambda - 3 = (\lambda - 3)(\lambda + 1)$.

$$\begin{aligned} \operatorname{sols}(A - 3I) &= \operatorname{sols}\begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix} = \operatorname{Span}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) \\ \operatorname{sols}(A + I) &= \operatorname{sols}\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} = \operatorname{Span}\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right) \end{aligned}$$

*Pulling it back to* $V$ *:* $\operatorname{Eig}_3(T) = \operatorname{Span}(1 + X)$ *and* $\operatorname{Eig}_{-1}(T) = \operatorname{Span}(1 - X)$.

*Now, if we pick the basis* $B = (1 + X, 1 - X)$*,* $[T]_B = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$.

# 12 Normed and Scalar Product Spaces

## 12.1 Euclidean Product

**Definition 12.1.1** (Norm)**.** *Let $V$ be a linear space over $\mathbb{R}$. A norm is a function $\|\cdot\| : V \to \mathbb{R}$ such that:*

| | |
|---|---|
| Positive-Definite | $\forall\, u \in V\,,\ \|u\| = 0 \Leftrightarrow u = 0$ |
| Homogeneous | $\forall\, \alpha \in K\,,\ \forall\, u \in V\,,\ \|\alpha \cdot u\| = \lvert\alpha\rvert \cdot \|u\|$ |
| Triangle Inequality | $\forall\, u, v \in V\,,\ \|u + v\| \leq \|u\| + \|v\|$ |

**Definition 12.1.2** (Normalizing)**.** *We say $u$ is a* ***unit vector*** *if $\|u\| = 1$. If $u \neq 0$, its normalized vector is $\widehat{u} = \dfrac{1}{\|u\|} \cdot u$.*

**Definition 12.1.3** (Scalar/Inner Product)**.** *An scalar product is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ such that:*

| | |
|---|---|
| Homogeneous | $\forall\, \alpha \in K\,,\ \forall\, u, v \in V\,,\ \langle \alpha \cdot u, v\rangle = \langle u, \alpha \cdot v\rangle = \alpha \cdot \langle u, v\rangle$ |
| Distributivity | $\forall\, u, v \in V\,,\quad \begin{matrix} \langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle \\ \langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle \end{matrix}$ |
| Symmetric | $\forall\, u, v \in V\,,\ \langle u, v\rangle = \langle v, u\rangle$ |

*If we have:*

| | |
|---|---|
| Positivity | $\forall\, u \in V\,,\ \langle u, u\rangle \geq 0$ and $\langle u, u\rangle \Leftrightarrow u = 0$ |

*the scalar product is said to be Euclidean (positive). A vector space with a positive scalar product is called a Euclidean space.*

**Example 12.1.1.** *We have these examples:*

$\mathbb{R}^n$ , *Sumprod:* $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \bullet \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^n x_i \cdot y_i = x_1\, y_1 + x_2\, y_2 + \cdots + x_n\, y_n.$

$\mathbb{R}^n$ , *Sumprod with (positive-definite) symmetric matrix A:*

$$x \underset{A}{\bullet} y = x \bullet (A\, y) = (Ax) \bullet y = x^t\, A\, y = \sum_{i=1}^n \sum_{j=1}^n x_i \cdot y_j \cdot a_{ij}$$

$\mathbb{R}[X]$ , *Integration:* $\langle p, q \rangle = \int_0^1 p(x) \cdot q(x)\, dx$

$M_n(F)$ , *Trace:* $\langle A, B \rangle = \operatorname{tr}(A^t\, B)$

**Definition 12.1.4** (Induced Norm)**.** *With a Euclidean scalar product, we can induce a norm, that is, we define*

$$\|u\| := \sqrt{\langle u, u \rangle}$$

*which we can check it obeys all the axioms.*

**Theorem 12.1.1** (Cauchy-Schwarz)**.**

$$\forall\, u, v \in V\ ,\ |\, \langle u, v \rangle\, | \leq \|u\| \cdot \|v\|$$

*Proof.* Let $t = \dfrac{\langle u, v \rangle}{\|v\|^2}$

$$\begin{aligned} \langle u - t \cdot v, u - t \cdot v \rangle &\geq 0 \\ \Leftrightarrow \langle u, u \rangle - 2t\ \langle u, v \rangle + t^2\ \langle v, v \rangle &\geq 0 \\ \Leftrightarrow \|u\|^2 - \frac{2|\, \langle u, v \rangle\, |^2}{\|v\|^2} + \frac{|\, \langle u, v \rangle\, |^2}{\|v\|^2} &\geq 0 \\ \|u\|^2\, \|v\|^2 - |\, \langle u, v \rangle\, |^2 &\geq 0 \end{aligned}$$

We have the result, with equality if, and only if:

$$u = \frac{\langle u, v \rangle}{\|v\|^2} v$$

That is, the vectors are parallel. □

**Theorem 12.1.2** (Polarization). *A norm $\|\cdot\|$ is induced by a scalar product iff*

$$\forall\, u, v \in V\,,\ \|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$$

*Further, the scalar product is defined by:* $\langle u, v\rangle = \dfrac{\|u+v\|^2 - \|u-v\|^2}{4}$

*Proof.* We prove each direction:

($\Rightarrow$) We have: $\|u+v\|^2 = \|u\|^2 + 2\langle u, v\rangle + \|v\|^2$ and $\|u-v\|^2 = \|u\|^2 - 2\langle u, v\rangle + \|v\|^2 \Rightarrow \|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$

($\Leftarrow$) We prove that $\langle u, v\rangle = \dfrac{\|u+v\|^2 - \|u-v\|^2}{4}$ satisfies the definition of scalar product. We need to use the Cauchy functional equation.

$\square$

**Definition 12.1.5** (Gram Matrix). *From a basis $B = (b_1, b_2, \cdots, b_n)$ of $V$, an Euclidean space, we define:*

$$\mathrm{Gram}(B) = \begin{pmatrix} \langle b_1, b_1\rangle & \langle b_1, b_2\rangle & \cdots & \langle b_1, b_n\rangle \\ \langle b_2, b_1\rangle & \langle b_2, b_2\rangle & \cdots & \langle b_2, b_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle b_n, b_1\rangle & \langle b_n, b_2\rangle & \cdots & \langle b_n, b_n\rangle \end{pmatrix} = [\langle b_i, b_j\rangle]_{ij}$$

**Lemma 12.1.1** (Calculating Scalar Products). *Let $x, y \in V$, then:*

$$\langle x, y\rangle = [x]_B^t\ \mathrm{Gram}(B)\,[y]_B = [x]_B \underset{\mathrm{Gram}(B)}{\bullet} [y]_B$$

*Proof.* Let $[x]_B = (x_1, x_2, \cdots, x_n)$ and $[y]_B = (y_1, y_2, \cdots, y_n)$, In every scalar product, by linearity:

$$\left\langle \sum_{i=1}^{n} x_i \cdot b_i\ ,\ \sum_{j=1}^{n} y_j \cdot b_j \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i \cdot y_j \cdot \langle b_i, b_j\rangle$$

$\square$

## 12.2 Orthogonality

**Definition 12.2.1** (Orthogonal)**.** *We say* $u, v \in V : u \perp v$ *iff* $\langle u, v\rangle = 0$

**Example 12.2.1.** *Let* $E = (\underline{e}_1, \underline{e}_2, \cdots, \underline{e}_n)$ *be the standard basis of* $\mathbb{R}^n$*, then* $\langle \underline{e}_i, \underline{e}_j \rangle = \underline{e}_i \bullet \underline{e}_{,j} = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$*, that is, they are both orthogonal and normalized.*

**Definition 12.2.2** (Orthogonal Complement)**.** *Given a subset of* $U \subseteq V$*, we define:*

$$U^{\perp} = \{v \in V \mid \forall\, u \in U\,,\ u \perp v\} = \{v \in V \mid \forall\, u \in U\,,\ \langle u, v\rangle = 0\}$$

**Lemma 12.2.1** (Complement is Subspace)**.** *For any subset* $U \subseteq V$*,* $U^{\perp}$ *is a linear subspace of* $V$*.*

*Proof.* We check each condition:

1. $0 \in U^{\perp}$, since $\forall\, v \in V\,,\ \langle 0, v\rangle = \langle v, 0\rangle = 0$
2. $v, w \in U^{\perp} \Rightarrow \forall\, u \in U\,,\ \langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle = 0 + 0 + = 0 \Rightarrow v + w \in U^{\perp}$
3. $v \in U^{\perp} \Rightarrow \forall\, u \in U\,,\ \langle u, \alpha \cdot v\rangle = \alpha \cdot v \in U^{\perp}$

□

**Proposition 12.2.1** (Double Perp)**.** *For any* $U \subseteq V$*,* $U \subseteq U^{\perp\perp} = \left(U^{\perp}\right)^{\perp}$

**Definition 12.2.3** (Orthogonal Sets)**.** *We write* $v \perp U \Leftrightarrow \forall\, u \in U\,,\ v \perp u \Leftrightarrow v \in U^{\perp}$ *and* $U \perp W \Leftrightarrow \forall\, u \in U\,, w \in W\,,\ u \perp w$

**Lemma 12.2.2** (N&SC for Orthogonality)**.**

$$U \perp W \Leftrightarrow U \subseteq W^{\perp} \text{ and } W \subseteq U^{\perp}$$

**Proposition 12.2.2** (Spans don't change the Complement)**.** $S \perp T \Leftrightarrow \operatorname{Span}(S) \perp \operatorname{Span}(T)$.

**Lemma 12.2.3.** *If $A, B \subseteq V$ and $A \perp B$, then $A \cap B \subseteq \{0\}$.*

*Proof.* Let $v \in A \cap B \Rightarrow v \in A$ and $v \in B$, $A \perp B \Rightarrow \langle v, v\rangle = 0 \Rightarrow v = 0$. $\square$

**Lemma 12.2.4** (Rows Perp Sols)**.** *In $\mathbb{R}^n$, $S = (\underline{a}_1, \underline{a}_2, \cdots, \underline{a}_k)$:*

$$S^{\perp} = \text{sols}\begin{pmatrix} — & \underline{a}_1 & — \\ — & \underline{a}_2 & — \\ & \vdots & \\ — & \underline{a}_k & — \end{pmatrix} = \text{sols}[S]^t$$

*Then, for every matrix $A \in M_{n\times k}(F)$ $\text{rows}(A) \perp \text{sols}(A)$, in the standard scalar product.*

*Proof.* Let $S = S_r(A) \Rightarrow \text{rows}(A)^{\perp} = S_r(A)^{\perp} = \text{sols}(A) \Rightarrow \text{rows}(A) \subseteq \text{rows}(A)^{\perp\perp} = \text{sols}(A)^{\perp} \Rightarrow \text{rows}(A) \perp \text{sols}(A)$ $\square$

**Definition 12.2.4** (Projection onto one vector)**.** *Let $u \in V \setminus \{0\}$, we define $\text{proj}_u = \dfrac{\langle u, \cdot\rangle}{\|u\|^2} \cdot u$, that is $\text{proj}_u : V \to V$ so that $\text{proj}_u : v \mapsto \dfrac{\langle u, v\rangle}{\|u\|^2} \cdot u$, which is linear due the linearity of the scalar product.*

**Lemma 12.2.5** (Calculations on Projection)**.** *We have:*

1. $\ker(\text{proj}_u) = \{u\}^{\perp}$
2. $\text{Im}(\text{proj}_u) = \text{Span}(u)$
3. $\text{proj}_u^2 = \text{proj}_u$
4. $u \perp w \Rightarrow \text{proj}_u \circ \text{proj}_w = \text{proj}_w \circ \text{proj}_u = 0$, *the zero map.*

*Proof.* We prove each one:

1. By definition: $\ker(\text{proj}_u) = \left\{ v \in V \;\middle|\; \text{proj}_u(v) = \dfrac{\langle u, v\rangle}{\|u\|^2} \cdot u = 0 \right\}$ since $u \neq 0$, we get: $\ker(\text{proj}_u) = \{v \in V \mid \langle u, v\rangle = 0\} = \{u\}^{\perp}$

2. Notice $\mathrm{proj}_u(\lambda \cdot u) = \lambda \cdot u \Rightarrow \mathrm{Span}(u) \subseteq \mathrm{Im}(\mathrm{proj}_u)$. Also, $\mathrm{Im}(\mathrm{proj}_u) \subseteq \mathrm{Span}(u)$ is trivially given by the definition of $\mathrm{proj}_u$.

3. $\mathrm{proj}_u^2(v) = \mathrm{proj}_u \left( \dfrac{\langle u, v\rangle}{\|u\|^2} \cdot u \right) = \dfrac{\langle u, v\rangle}{\|u\|^2} \cdot \mathrm{proj}_u(u) = \dfrac{\langle u, v\rangle}{\|u\|^2} \cdot u = \mathrm{proj}_u(v)$

4. $u \perp w \Rightarrow \mathrm{Im}(\mathrm{proj}_w) = \mathrm{Span}(w) \subseteq \{u\}^\perp = \ker(\mathrm{proj}_u)$ then, we must have $\forall\, v \in V\,,\ \mathrm{proj}_w(\mathrm{proj}_u(v)) = 0 \Rightarrow \mathrm{proj}_u \circ \mathrm{proj}_w = 0$.

□

## 12.3 Orthogonal Sequences

**Definition 12.3.1** (Orthogonal/Orthonormal Sequences)**.** *Let $V$ be an Euclidean space, $K = (e_1, e_2, \cdots, e_k) \in V^k$ is called **orthogonal** if: $0 \notin K$ and $\forall\, i, j \in \{1, 2, \cdots, n\} : i \neq j\,,\ e_i \perp e_j$.*

*It is called **orthonormal** if it is orthogonal and $\forall\, i \in \{1, 2, \cdots, k\}\,,\ \|e_i\| = 1$.*

**Proposition 12.3.1** (Kronecker Delta)**.** *A sequence $K = (e_1, e_2, \cdots, e_k) \in V^k$ is orthonormal iff $\langle e_i, e_j\rangle = \delta_{ij}$. Moreover, it is orthogonal iff $\langle e_i, e_j\rangle = \|e_i\|^2\delta_{ij}$.*

**Theorem 12.3.1** (Orthogonal Sequences are LI)**.** *Let $K = (e_1, e_2, \cdots, e_n)$ be orthogonal, then $K$ is LI.*

*Proof.* Let $K = (e_1, e_2, \cdots, e_k) \in V^k$, let $(\alpha_1, \alpha_2, \cdots, \alpha_k) \in F^k$ such that: $\displaystyle\sum_{i=1}^{k} \alpha_i \cdot e_i = 0 \Rightarrow 0 = \left\langle e_i\,, \sum_{j=1}^{k} \alpha_j \cdot e_j \right\rangle = \sum_{j=1}^{k} \alpha_j \cdot \langle e_i, e_j\rangle = \alpha_i \Rightarrow$ $(\alpha_1, \alpha_2, \cdots, \alpha_k) = \underline{0}$ □

**Lemma 12.3.1** (Coordinates in Orthogonal Basis)**.** *Let $K = (e_1, e_2, \cdots, e_n)$ is an orthogonal basis of $V$, then, for any $v \in V$:*

$$[v]_K = \begin{pmatrix} \dfrac{\langle e_1, v\rangle}{\|e_1\|^2} \\ \vdots \\ \dfrac{\langle e_n, v\rangle}{\|e_n\|^2} \end{pmatrix} \qquad \textit{that is,} \qquad v = \sum_{i=1}^{n} \mathrm{proj}_{e_i}(v) = \sum_{i=1}^{n} \frac{\langle e_i, v\rangle}{\|e_i\|^2} \cdot e_i$$

*Proof.* Let $v = \sum_{i=1}^{n} x_i \cdot e_i$

$$\Rightarrow \langle e_i, v\rangle = \left\langle e_i\,, \sum_{j=1}^{n} x_j \cdot e_j \right\rangle = \sum_{j=1}^{n} x_j \cdot \langle e_i, e_j\rangle = \sum_{j=1}^{n} x_j \cdot \|e_i\|^2\,\delta_{ij} = x_i \cdot \|e_i\|^2$$

□

**Theorem 12.3.2** (Parseval's Identity)**.** *Let $K = (e_1, e_2, \cdots, e_n)$ is an orthonormal basis of $V$, then, for any $v \in V$:*

$$\|v\|^2 = \sum_{i=1}^{n} |\,\langle e_i, v\rangle\,|^2$$

*Proof.* From the lemma above:

$$v = \sum_{i=1}^{n} \langle e_i, v\rangle \cdot e_i \Rightarrow \langle v, v\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \langle e_i, v\rangle \cdot \langle e_j, v\rangle \cdot \overbrace{\langle e_i, e_j\rangle}^{\delta_{ij}} = \sum_{i=1}^{n} |\,\langle e_i, v\rangle\,|^2$$

□

**Theorem 12.3.3** (Gram-Schmidt Process)**.** *Let $S = (v_1, v_2, \cdots, v_n) \in V^n$ be a linearly independent sequence, then there is a orthonormal sequence $K = (e_1, e_2, \cdots, e_n)$ such that* $\mathrm{Span}(K) = \mathrm{Span}(S)$*. In particular:*

$$\begin{aligned}
u_1 &= v_1 & e_1 &= \frac{u_1}{\|u_1\|} \\
u_2 &= v_2 - \mathrm{proj}_{u_1}(v_2) & e_2 &= \frac{u_2}{\|u_2\|} \\
u_3 &= v_3 - \mathrm{proj}_{u_1}(v_3) - \mathrm{proj}_{u_2}(v_3) & e_3 &= \frac{u_3}{\|u_3\|} \\
&\ \vdots & &\ \vdots \\
u_k &= v_k - \sum_{i=1}^{k-1} \mathrm{proj}_{u_i}(v_k) & e_k &= \frac{u_k}{\|u_k\|} \\
&\ \vdots & &\ \vdots \\
u_n &= v_n - \sum_{i=1}^{n-1} \mathrm{proj}_{u_i}(v_n) & e_n &= \frac{u_n}{\|u_n\|}
\end{aligned}$$

*Proof.* To prove $(u_1, u_2, \cdots, u_n)$ is orthogonal, we use induction:

Base $\text{proj}_{u_1}(u_2) = \text{proj}_{u_1}(v_2) - \text{proj}^2_{u_1}(v_2) = \text{proj}_{u_1}(v_2) - \text{proj}_{u_1}(v_2) = 0 \Rightarrow u_2 \in \ker(\text{proj}_{u_1}) = \{u_1\}^\perp \Rightarrow u_2 \perp u_1$

Step $j \in \{1, \cdots, k-1\}$ : $\text{proj}_{u_j}(u_k) = \text{proj}_{u_j}(v_k) - \sum_{i=1}^{k-1} \text{proj}_{u_j}\left(\text{proj}_{u_i}(v_k)\right) = \text{proj}_{u_j}(v_k) - \text{proj}^2_{u_j}(v_k) = 0 \Rightarrow u_k \in \ker(\text{proj}_{u_j}) = \{u_j\}^\perp \Rightarrow u_k \perp u_j$

To prove the span is the same, notice $\dim \text{Span}(K) = n = \dim \text{Span}(S)$ and $K \subset \text{Span}(S)$. □

**Corollary 12.3.1** (Finite Span is easy). *Every finitely spanned Euclidean vector space has an orthonormal basis.*

## 12.4 Orthogonal Maps

**Definition 12.4.1** (Orthogonal Map). *For two Euclidean spaces $(V, \langle\cdot,\cdot\rangle_V)$ and $(W, \langle\cdot,\cdot\rangle_W)$. A linear map $Q : V \to W$ is orthogonal if preserves the scalar product, that is:*

$$\forall\, u, v \in V\,,\ \langle Q(u), Q(v)\rangle_W = \langle u, v\rangle_V$$

**Lemma 12.4.1** (Geometrical Properties). *Let $Q$ be an orthogonal map. We have:*

1. $\forall\, u, v \in V\,,\ u \perp v \Leftrightarrow Q(u) \perp Q(v)$
2. $\forall\, u \in V\,,\ \|Q(u)\| = \|u\|$
3. *$Q$ is injective.*
4. *If $\lambda$ is an eigenvalue of $Q$, then $|\lambda| = 1$*

*Proof.* We prove each one:

1. $\forall\, u, v \in V\,,\ u \perp v \Leftrightarrow 0 = \langle u, v\rangle_V = \langle Q(u), Q(v)\rangle_W \Leftrightarrow Q(u) \perp Q(v)$
2. $\forall\, u \in V\,,\ \|Q(u)\|^2 = \langle Q(u), Q(u)\rangle = \langle u, u\rangle = \|u\|^2$
3. $u \in \ker(Q) \Rightarrow 0 = \|Q(u)\| = \|u\| \Rightarrow u = 0 \Rightarrow \ker(Q) = \{0\}$
4. If $\lambda$ is an eigenvalue, there is a $u \neq 0$ such that $Q(u) = \lambda \cdot u \Rightarrow \|u\| = \|Q(u)\| = \|\lambda \cdot u\| = |\lambda| \cdot \|u\| \Rightarrow |\lambda| = 1$

□

**Definition 12.4.2** (Gram Matrix)**.** *For $A \in M_{k\times n}(F)$, we define:*

$$\mathrm{Gram}(A) = A^t\, A$$

**Lemma 12.4.2** (N&SC of Orthonormality)**.** *For $A \in M_{k\times n}(F)$, $S_c(A)$ is orthonormal* $\Leftrightarrow \mathrm{Gram}(A) = I$

**Corollary 12.4.1** (Rotation Equation)**.** *$Q : \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal map* $\Leftrightarrow \mathrm{Gram}([Q]) = [Q]^t\,[Q] = I$

**Proposition 12.4.1.** *This coincides with the definition we had before by taking the sumprod as the inner product and writing the sequence as a matrix.*

**Proposition 12.4.2** (Orthogonal Coordinate Maps)**.** *$K$ is an orthonormal sequence iff $Q_K$ (coordinate map) is an orthogonal map.*

*Proof.* By definition, $K$ is orthonormal iff $\mathrm{Gram}(K) = I$ so, by a previous lemma:

$$\langle u, v\rangle = [u]_K^t\ \mathrm{Gram}(K)\,[v]_K = Q_K(u) \bullet \Big(\mathrm{Gram}(K)\, Q_K(v)\Big)$$

Hence, $\mathrm{Gram}(K) = I \Leftrightarrow Q_K(u) \bullet Q_K(v) = \langle u, v\rangle$ □

# 13 Direct Sum

## 13.1 Sum of Subspaces

**Definition 13.1.1** (Sum of Subspaces)**.** *Let $V$ be a linear space and $U$ and $W$ subspaces of $V$.*

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

**Lemma 13.1.1** (Analog of "Union" of Subspaces)**.** *We have:*

1. *$U + W$ is a linear subspace of $V$.*
2. *If $U = \mathrm{Span}(S)$ and $W = \mathrm{Span}(T)$, then $U + W = \mathrm{Span}(S +\!\!+ T)$*

*Proof.* We prove each one:

1. $0+0 = 0 \in U+W$ , $(u_1+w_1)+(u_2+w_2) = (u_1+u_2)+(w_1+w_2) \in U+W$ , $\alpha \cdot \big(u + w\big) = (\alpha \cdot u) + (\alpha \cdot w) \in U + W$ .
2. $U+W = \left\{ \sum_{i=1}^{n} a_i \cdot s_i + \sum_{j=1}^{k} b_j \cdot t_j \;\middle|\; a_1, a_2, \cdots, a_n, b_1, b_2, \cdots, b_k \in \mathbb{R} \right\} = \mathrm{Span}(S +\!\!+ T)$

$\square$

**Corollary 13.1.1** (Inequality of Dimensions)**.** *For finitely spanned subspaces $U, W$:*

$$\dim U, \dim W \le \dim(U + W) \le \dim U + \dim W$$

**Lemma 13.1.2** (Intersection of Subspaces is Subspace)**.** *Let $U, W \subseteq V$ be subspaces. Then, $U \cap W$ is also a subspace.*

*Proof.* $0 \in U$ and $0 \in W \Rightarrow 0 \in U \cap W$. $u_1, u_2 \in U \cap W \Rightarrow u, v \in U$ and $u, v \in W \Rightarrow u + v, \alpha \cdot u \in U$ and $u + v, \alpha \cdot u \in W$ since they are subspaces, $\Rightarrow u + v, \alpha \cdot u \in U \cap W$. $\square$

**Definition 13.1.2** (Direct Sum)**.** *If $U \cap W = \{0\}$, we write $U \oplus W = U + W$.*

**Definition 13.1.3** (Sum Function)**.** *We define the sum function as:*

$$\Sigma : U \times W \to U + W$$
$$(u, w) \mapsto u + w$$

**Lemma 13.1.3** (Kernel and Image of Sum)**.** *We have:*

1. $\ker(\Sigma) \cong U \cap W$
2. $\mathrm{Im}(\Sigma) = U + W$

*Proof.* We prove each one:

1. $\ker(\Sigma) = \{(u, w) \in U \times W \mid u + w = 0\} = \{(u, -u) \in U \times W\}$, so we need $u \in U$ and $u \in W$, so: $\ker(\Sigma) = \{(u, -u) \mid u \in U \cap W\} \cong U \cap W$, with the map $u \mapsto (u, -u)$.
2. $\mathrm{Im}(\Sigma) = \{\Sigma(u, w) = u + w \mid u \in U \ , \ w \in W\} = U + W$

□

**Theorem 13.1.1** (Grassman's Formula)**.** *For $V$ finetely spanned:*

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$$

*Proof.* $\dim(U \times W) = \dim U + \dim W$, use the dimension theorem with $\Sigma$. □

**Theorem 13.1.2** (Equivalence of Direct Sum)**.** *The following are equivalent:*

1. $U \cap W = \{0\}$
2. *Every $v \in U + W$ has an* ***unique*** *representation as a sum $v = u + w$ where $u \in U$ and $w \in W$*
3. *$\Sigma$ is an isomorphism.*

*Proof.* We prove the directions:

$(1 \Leftrightarrow 3)$ $(\Leftarrow)$ $\Sigma$ is an isomorphism $\Rightarrow U \cap W \cong \ker\Sigma = \{\underline{0}\} \Rightarrow U \cap W = \{0\}$.

$(\Rightarrow)$ If $U \cap W = \{0\} \Rightarrow \ker\Sigma = \{\underline{0}\}$ .

$(2 \Leftrightarrow 3)$ Trivial, since we can find an inverse for $\Sigma$, and shows $\Sigma^{-1}$ is a function.

□

**Definition 13.1.4** (Decomposition)**.** *If $U \oplus W = V$, we say that $W$ is the complement of $U$ (or $U, W$ are complement subspaces) in this case, every $v \in V$ has an* ***unique*** *representation as a sum $v = u + w$ where $u \in U$ and $w \in W$, called the $U$-$W$ decomposition of $V$. We can write the representation as $\Sigma^{-1}(v)$.*

## 13.2 Orthogonal Decomposition

**Lemma 13.2.1** (Properties of Orthogonal Complement)**.** *For any subspaces $U, W \subseteq V$*

*1.* $U \cap U^{\perp} = \{0\}$

*2.* $(U + W)^{\perp} = U^{\perp} \cap W^{\perp}$

*Proof.* We prove each one:

1. Double Inclusion:

   $(\supseteq)$ Trivially given.

   $(\subseteq)$ $v \in U \cap U^{\perp} \Rightarrow v \in U$ and $\forall\, u \in U$ , $\langle u, v\rangle = 0 \overset{u=v}{\Longrightarrow} \langle v, v\rangle = 0 \Rightarrow v = 0$

2. Double Inclusion:

($\supseteq$) $v \in U^\perp \cap W^\perp \Rightarrow \begin{array}{l} \forall\, u \in U\,,\ \langle u, v\rangle = 0 \\ \forall\, w \in W\,,\ \langle w, v\rangle = 0 \end{array} \Rightarrow \forall\, u \in U + W\,,\ \langle u, v\rangle = 0 \Rightarrow v \in (U + W)^\perp$

($\subseteq$) By contrary: $v \notin U^\perp \cap W^\perp \Rightarrow$ either:

(a) $v \notin U^\perp \Rightarrow \exists\, u \in U\ :\ \langle u, v\rangle \neq 0$

(b) $v \notin W^\perp \Rightarrow \exists\, w \in W\ :\ \langle w, v\rangle \neq 0$

$\Rightarrow \exists\, u \in U + W\ :\ \langle u, v\rangle \neq 0 \Rightarrow v \notin (U + W)^\perp$

□

**Theorem 13.2.1** (Orthogonal Decomposition)**.** *Let $V$ be a finitely spanned Euclidean space and $U$ a linear subspace. Then:*

$$V = U \oplus U^\perp$$

*Proof.* Let $K$ be an orthonormal basis for $U$. Let $\sum_{i=1}^{\dim U} x_i \cdot e_i = u \in U$. We want to prove that $\forall\, v \in V\,,\ \exists\, u \in U : (v - u) \in U^\perp$. It is necessary and sufficient to check $\forall\, e_i \in K\,,\ (v - u) \perp e_i$, which has unique solution: $x_i = \langle v, e_i\rangle$ (by orthonormality). □

**Corollary 13.2.1** (Dimension of the Complement)**.**

$$\dim U^\perp = \dim V - \dim U$$

**Lemma 13.2.2** (Double Perp)**.** *Let $V$ finite-dimensional vector space. For any subspace $U \subseteq V$: $U^{\perp\perp} = U$.*

*Proof.* We use the orthogonal decomposition:

$$\dim U^{\perp\perp} = \dim V - \dim U^\perp = \dim V - (\dim V - \dim U) = \dim U$$

Since $U \subseteq U^{\perp\perp}$ is given, by Dimension Equality, $U^{\perp\perp} = U$. □

**Corollary 13.2.2.** $U^\perp + W^\perp = (U \cap W)^\perp$

**Definition 13.2.1** (Projections). *Let the following classes of functions:*

$$\pi_j : (a_1, a_2, \cdots, a_n) \mapsto a_j$$

*we define:*

*Projection:* $P_U = \pi_1 \circ \Sigma^{-1}_{U \times W}$

*Complement/Rejection:* $P_W = \pi_2 \circ \Sigma^{-1}_{U \times W} = Id_V - P_U$

**Lemma 13.2.3** (Orthogonal Projection in a Basis). *Let* $P_U = \pi_1 \circ \Sigma^{-1}_{U \times U^\perp}$ *be the projection map and* $K = (e_1, e_2, \cdots, e_{\dim U})$ *an ON basis of* $U$*. Then:*

$$P_U : v \mapsto \sum_{i=1}^{\dim U} \langle v, e_i \rangle \cdot e_i$$

*Proof.* As before, with the proof of the Orthogonal Decomposition Theorem, let $u = \sum_{i=1}^{\dim U} \langle v, e_i \rangle \cdot e_i$, then $\forall\, e_i \in K$, $(v - u) \perp e_i \Rightarrow v - u \in U^\perp$. And by the uniqueness of the orthogonal projection, we have the desired result. $\square$