

Linear Algebra

Bachelor Applied Artificial Intelligence (AAI-B2)

André Herzwurm

TH Rosenheim, SoSe 2022

Literature

S. Axler, *Linear Algebra Done Right*, Springer, 3rd Edition, 2015.

G. Fischer, *Lineare Algebra*, Springer, 18. Auflage, 2014. (German)

J. Liesen, V. Mehrmann, *Linear Algebra*, Springer, 2015.

Contents

I	Algebraic Structures	1
1	Groups	1
1.1	Permutations	3
1.2	Elementary Properties of Groups	3
1.3	Subgroups	4
2	Fields	5
2.1	Elementary Properties of Fields	6
2.2	Pointwise Addition and Multiplication of Functions	7
II	Vector Spaces and Linear Maps	9
1	Vector Spaces	9
1.1	Subspaces	11
1.2	Span and Linear Combinations	11
2	Bases and Dimension	13
2.1	Linear Independence	13
2.2	Generating Sets and Bases	14
2.3	Dimension of a Vector Space	16
3	Matrices	18
3.1	Row Space and Rank	18
3.2	Matrix Multiplication	20
4	Linear Maps	22
4.1	Construction of Linear Maps	25
4.2	Coordinate Systems	27
4.3	Kernel and Image	27

Chapter I

Algebraic Structures

In this chapter we introduce the notion of groups and fields. We also present simple examples and elementary properties of groups and fields.

1 Groups

In the sequel let G be a non-empty set.

Definition 1. A map

$$*: G \times G \rightarrow G$$

is called *operation* on G . An operation is *associative* if¹

$$\forall a, b, c \in G: (a * b) * c = a * (b * c),$$

and *commutative* if

$$\forall a, b \in G: a * b = b * a.$$

Notation: Brackets may be dropped in case of an associative operation with multiple elements. Depending on the context we may also write $a + b$, $a \cdot b$, ab etc. instead of $a * b$.

Definition 2. A set G with an operation $*$ is called *group* if $*$ is associative and there exists $e \in G$ such that

$$\forall a \in G: e * a = a \tag{1}$$

and

$$\forall a \in G \exists a' \in G: a' * a = e. \tag{2}$$

Furthermore, a group is called *commutative (abelian)* if $*$ is commutative.

Notation: We sometimes write $(G, *)$ to emphasize the operation $*$.

¹In a definition one typically uses “if” instead of “iff” (if and only if).

Example 3.

(i) Additive groups:

a) $(\mathbb{N}_0, +)$ is not a group since (2) does not hold.b) $(\mathbb{Z}, +)$ is a commutative group with $e = 0$ and

$$\forall m \in \mathbb{Z}: (-m \in \mathbb{Z} \wedge (-m) + m = 0).$$

Clearly, $+$ is associative and commutative and

$$\forall m, n \in \mathbb{Z}: m + n \in \mathbb{Z}.$$

c) $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, and $(\mathbb{C}, +)$ are commutative groups.

(ii) Multiplicative groups:

a) (\mathbb{Z}, \cdot) is not a group since (2) does not hold. In particular, we have

$$\forall m \in \mathbb{Z}: m \cdot 0 = 0 \neq 1.$$

b) $(\mathbb{Q} \setminus \{0\}, \cdot)$ is a commutative group with $e = 1$ and

$$\forall q \in \mathbb{Q} \setminus \{0\}: (1/q \in \mathbb{Q} \setminus \{0\} \wedge 1/q \cdot q = 1).$$

Clearly, \cdot is associative and commutative and

$$\forall q_1, q_2 \in \mathbb{Q} \setminus \{0\}: q_1 \cdot q_2 \in \mathbb{Q} \setminus \{0\}.$$

c) $(\mathbb{R} \setminus \{0\}, \cdot)$ and $(\mathbb{C} \setminus \{0\}, \cdot)$ are commutative groups.(iii) $G = \{0, 1\}$ together with

$*$	0	1
0	0	1
1	1	0

is a commutative group. See **Exercise 0.1**.**Definition 4.** Let $*$ be an operation on G and let $\emptyset \neq G' \subseteq G$ satisfy

$$\forall a, b \in G': a * b \in G'.$$

Then the operation $*': G' \times G' \rightarrow G'$, defined by $a *' b = a * b$, is called the *induced operation* on G' .Henceforth we do not distinguish between $*$ and $*'$.

1.1 Permutations

Notation: The set of mappings from a set X to a set Y is denoted by Y^X .

Proposition 5. Let $X \neq \emptyset$ and

$$G = \{f \in X^X : f \text{ bijective}\}.$$

G together with function composition \circ is a group. Moreover, G is not commutative if $|X| \geq 3$.

Definition 6. G from Proposition 5 is called the *symmetric group* of the set X . Its elements are called *permutations*.

1.2 Elementary Properties of Groups

In the sequel let G be a group.

Lemma 7. Let $e \in G$ satisfy (1) and (2) and let $a, a' \in G$. Then we have

$$(i) \quad a'a = e \Rightarrow aa' = e,$$

$$(ii) \quad ae = a.$$

Proof. ad (i): According to (2) there exists $a'' \in G$ such that

$$a''a' = e.$$

Using (1) we obtain

$$aa' = eaa' = a''a'aa' = a''ea' = a''a' = e.$$

ad (ii): Let $a' \in G$ such that $a'a = e$. From part (i) we get

$$ae = aa'a \stackrel{(i)}{=} ea = a.$$

□

Proposition 8. There exists a unique $e \in G$ satisfying (1) and (2).

Proof. Let (1) and (2) be satisfied for $e = e_1$ and $e = e_2$. Then we have $e_1e_2 = e_2$ and $e_2e_1 = e_1$. Lemma 7.(i) shows $e_2e_1 = e_2$ and hence $e_1 = e_2e_1 = e_2$. □

Definition 9. $e \in G$ satisfying (1) and (2) is called the *neutral element* of G .

In the sequel let e be the neutral element of G .

Proposition 10. For every $a \in G$ there exists a unique $a' \in G$ such that $a'a = e$.

Proof. Let $a', a'' \in G$ such that $a'a = a''a = e$. Lemma 7.(i) shows $aa' = e$. Using Lemma 7.(ii) we obtain

$$a'' \stackrel{(ii)}{=} a''e = a''aa' = ea' = a'.$$

Cf. Exercise 0.2.

□

Definition 11. For $a \in G$ the element $a' \in G$ satisfying $a'a = e$ is called the *inverse element* of a .

Notation: $a' = a^{-1}$, $a' = 1/a$, or $a' = -a$.

Lemma 12. For $a, b, c \in G$ we have

- (i) $ab = ac \Rightarrow b = c$,
- (ii) $ba = ca \Rightarrow b = c$,
- (iii) $(a^{-1})^{-1} = a$,
- (iv) $(ab)^{-1} = b^{-1}a^{-1}$,
- (v) $e^{-1} = e$,
- (vi) $\exists_1 x \in G: ax = b$.

Proof. ad (i): Assume $ab = ac$. Then we have $b = eb = a^{-1}ab = a^{-1}ac = ec = c$.

ad (ii): Assume $ba = ca$. By Lemma 7, we have $b = be = baa^{-1} = caa^{-1} = ce = c$.

ad (iii): By Lemma 7.(ii), we have $a = ea = (a^{-1})^{-1}a^{-1}a = (a^{-1})^{-1}e = (a^{-1})^{-1}$.

ad (iv): We have $(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}eb = b^{-1}b = e$.

ad (v): We have $ee = e$.

ad (vi): Put $x = a^{-1}b \in G$. By Lemma 7.(i), we have $ax = aa^{-1}b = eb = b$. Uniqueness follows from part (i).

Cf. **Exercise 0.3**. □

1.3 Subgroups

Definition 13. $G' \subseteq G$ is a *subgroup* of G if the following conditions hold:

- (i) $G' \neq \emptyset$,
- (ii) $\forall a, b \in G': ab \in G'$,
- (iii) $\forall a \in G': a^{-1} \in G'$.

Remark 14.

- (i) Every subgroup G' of G satisfies $e \in G'$.
(Proof: Choose $a \in G'$. Then we obtain $a^{-1} \in G'$ and $e = a^{-1}a \in G'$.)
- (ii) $\{e\}$ and G are the smallest and the largest subgroup of G , respectively, i.e., every subgroup G' of G satisfies $\{e\} \subseteq G' \subseteq G$.

Example 15. (i) $G' = \mathbb{Z}$ is a subgroup of $G = \mathbb{Q}$ w.r.t. the addition.

(ii) $G' = \{q \in \mathbb{Q}: q > 0\}$ is a subgroup of $G = \mathbb{Q} \setminus \{0\}$ w.r.t. the multiplication.

(iii) Let G be the symmetric group of X . For $X_0 \subseteq X$ put

$$G' = \{f \in G : \forall x \in X_0 : f(x) = x\}.$$

Then G' is a subgroup of G .

Proposition 16. Every subgroup of G (together with the induced operation) is a group with neutral element e .

Proof. Definition 13.(i)-(ii) ensure that the induced operation $*'$ on a subgroup G' is well-defined. Obviously, the induced operation $*'$ on G' inherits the associativity of $*$ on G . According to Remark 14.(i) we have $e \in G'$, and for $a \in G'$ we have $e *' a = e * a = a$. Moreover, Definition 13.(iii) ensures $a^{-1} \in G'$ for $a \in G'$ such that $a^{-1} *' a = a^{-1} * a = e$. \square

Lemma 17. Let G_1 and G_2 be subgroups of G . Then $G_1 \cap G_2$ is a subgroup of G .

Proof. Note that $e \in G_1 \cap G_2$. Let $a, b \in G_1 \cap G_2$. Then we have $ab \in G_i$ and $a^{-1} \in G_i$ for $i = 1, 2$ since the G_i are subgroups of G . This shows $ab \in G_1 \cap G_2$ and $a^{-1} \in G_1 \cap G_2$. \square

Example 18. $G_1 = \{2k : k \in \mathbb{Z}\}$ and $G_2 = \{3k : k \in \mathbb{Z}\}$ are subgroups of $G = \mathbb{Z}$ w.r.t. the addition. We have $G_1 \cap G_2 = \{6k : k \in \mathbb{Z}\}$.

Note that $-2, 3 \in G_1 \cup G_2$, but $(-2) + 3 = 1 \notin G_1 \cup G_2$. Thus $G_1 \cup G_2$ is not a subgroup of G , cf. Exercise 0.4.

2 Fields

Definition 1. A set K together with two operations

$$\begin{aligned} + : K \times K &\rightarrow K & (\text{addition}) \\ \cdot : K \times K &\rightarrow K & (\text{multiplication}) \end{aligned}$$

is a *field* if the following conditions hold:

(i) $(K, +)$ is a commutative group.

(ii) Let 0 be the neutral element in $(K, +)$. For $K^* = K \setminus \{0\}$ we have

$$\forall a, b \in K^* : a \cdot b \in K^*,$$

and (K^*, \cdot) is a commutative group.

(iii) $\forall a, b, c \in K : (a \cdot (b + c) = a \cdot b + a \cdot c \wedge (a + b) \cdot c = a \cdot c + b \cdot c)$ (*distributivity*).

Notation: The neutral element and the inverse element of $a \in K$ in $(K, +)$ are denoted by 0 and $-a$, respectively. The neutral element in (K^*, \cdot) is denoted by 1. The inverse element of $a \in K^*$ is denoted by a^{-1} or $1/a$.

We often write ab instead of $a \cdot b$ and $a - b$ instead of $a + (-b)$ for $a, b \in K$, and a/b instead of $a \cdot b^{-1}$ for $a \in K$ and $b \in K^*$.

Convention: \cdot has precedence over $+$.

Remark 2. Per definition we have $0 \neq 1$ in every field.

Example 3. (i) \mathbb{Q} with the usual operations $+$ and \cdot is a field.

(ii) $K = \{0, 1\}$ with the addition according to Example 3.(iii) and the multiplication given by

\cdot	0	1
0	0	0
1	0	1

is a field. Note that $1 + 1 = 0$ in this case.

2.1 Elementary Properties of Fields

In the sequel let K be a field.

Lemma 4. For $a, b, c \in K$ we have

- (i) $0 \cdot a = a \cdot 0 = 0$,
- (ii) $a \cdot b = 0 \Rightarrow a = 0 \vee b = 0$,
- (iii) $a \cdot (-b) = (-a) \cdot b = -(ab)$,
- (iv) $(-a) \cdot (-b) = ab$,
- (v) $a \cdot b = a \cdot c \wedge a \neq 0 \Rightarrow b = c$,
- (vi) $a \neq 0 \Rightarrow (\exists_1 x \in K : a \cdot x = b)$.

Proof. ad (i): We have

$$0 + 0 \cdot a = 0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a.$$

Lemma 1.12.(ii) ensures $0 = 0 \cdot a$. Analogously, we obtain $0 = a \cdot 0$.

ad (ii): Note that (ii) is equivalent to (proof by contraposition)

$$a \neq 0 \wedge b \neq 0 \Rightarrow a \cdot b \neq 0,$$

which is clearly satisfied since (K^*, \cdot) is a group, see Definition 1.(ii).

ad (iii): By using (i) we obtain

$$a \cdot b + a \cdot (-b) = a \cdot (b + (-b)) = a \cdot 0 \stackrel{(i)}{=} 0.$$

Hence the unique inverse element (see Proposition 1.10) is given by $-(a \cdot b) = a \cdot (-b)$. Analogously, we obtain $-(a \cdot b) = (-a) \cdot b$.

ad (iv): Using (iii) we have

$$(-a) \cdot (-b) \stackrel{(iii)}{=} -(a \cdot (-b)) \stackrel{(iii)}{=} -(-(a \cdot b)) = a \cdot b,$$

where the last equality holds due to Lemma 1.12.(iii).

ad (v): Assume $a \cdot b = a \cdot c$. Using (iii) we thus have

$$a \cdot (b - c) = a \cdot b + a \cdot (-c) \stackrel{(iii)}{=} a \cdot b - a \cdot c = 0.$$

According to (ii) we obtain $a = 0 \vee (b - c) = 0$.

ad (vi): Let $a \neq 0$. If $b = 0$, then (ii) implies $x = 0$, and according to (i) we have

$$a \cdot x = a \cdot 0 \stackrel{(i)}{=} 0 = b,$$

i.e., $x = 0$ is the unique solution to $ax = b$. If $b \neq 0$, then $x = a^{-1}b$ is the unique solution to $ax = b$ in K^* , see Lemma 1.12.(vi). Moreover, note that $x = 0$ is not a solution to $ax = b$ in this case due to (i).

Cf. **Exercise 1.1.** □

Remark 5. Due to Lemma 4.(i) the multiplication is associative and commutative on $K = K^* \cup \{0\}$ and

$$\forall a \in K: 1 \cdot a = a.$$

Note that 0 does not have an inverse element w.r.t. the multiplication.

2.2 Pointwise Addition and Multiplication of Functions

In the sequel let $X \neq \emptyset$.

Definition 6. *Addition* and *multiplication* on K^X are defined as follows: For $f, g \in K^X$ and $x \in X$ we put

$$(f + g)(x) = f(x) + g(x)$$

and

$$(f \cdot g)(x) = f(x) \cdot g(x).$$

Terminology: pointwise addition (or sum) and pointwise multiplication (or product) of functions, respectively.

Lemma 7.

- (i) $(K^X, +)$ is a commutative group with neutral element $0 \in K^X$.
- (ii) The multiplication is associative and commutative on K^X , and for $1 \in K^X$ we have

$$\forall f \in K^X: 1 \cdot f = f.$$

- (iii) Distributivity holds.

- (iv) For all $f \in K^X$ we have

$$(\exists g \in K^X: g \cdot f = 1) \Leftrightarrow (\forall x \in X: f(x) \neq 0).$$

Proof. Exercise.

□

Notation: $f - g$ instead of $f + (-g)$ for $f, g \in K^X$ and f/g instead of $f \cdot h$ for $f \in K^X$ and $g \in K^X$ provided that $g(x) \neq 0$ and $h(x) = (g(x))^{-1}$ for all $x \in X$.

Chapter II

Vector Spaces and Linear Maps

Vector spaces and linear maps are the fundamental objects of linear algebra. In this context we study, for instance, systems of linear equations.

1 Vector Spaces

In the sequel let K be a field. We use the standard notation and conventions in the context of fields, see Definition I.2.1.

Definition 1. A set V together with two mappings

$$\begin{aligned}\oplus: V \times V &\rightarrow V & (\text{vector addition}), \\ \odot: K \times V &\rightarrow V & (\text{scalar multiplication})\end{aligned}$$

is a K -vector space (*vector space over K*) if the following conditions hold:

- (i) (V, \oplus) is a commutative group.
- (ii) For all $\lambda, \mu \in K$ and $v, w \in V$ we have

$$\begin{aligned}(\lambda + \mu) \odot v &= \lambda \odot v \oplus \mu \odot v, \\ \lambda \odot (v \oplus w) &= \lambda \odot v \oplus \lambda \odot w, \\ \lambda \odot (\mu \odot w) &= (\lambda \cdot \mu) \odot w, \\ 1 \odot v &= v.\end{aligned}$$

The elements of V are called *vectors* and the elements of K are called *scalars*.

Notation: $\mathbf{0}$ for the neutral element and $\ominus v$ for the inverse element of $v \in V$ in (V, \oplus) .

Convention: \odot has precedence over \oplus .

Example 2 (Coordinate space). For $n \in \mathbb{N}$ let

$$V = K^n = \{(x_1, \dots, x_n) : x_i \in K \text{ for } i \in \{1, \dots, n\}\}.$$

Define

$$\begin{aligned}(x_1, \dots, x_n) \oplus (y_1, \dots, y_n) &= (x_1 + y_1, \dots, x_n + y_n), \\ \lambda \odot (x_1, \dots, x_n) &= (\lambda \cdot x_1, \dots, \lambda \cdot x_n)\end{aligned}\tag{1}$$

for $\lambda \in K$ and $(x_1, \dots, x_n), (y_1, \dots, y_n) \in V$. Then V is a K -vector space, see **Exercise 1.2**. In particular, $V = K$ is a K -vector space.

Convention: If not stated otherwise we regard $V = K^n$ as a K -vector space with the vector space operations \oplus and \odot defined in (1).

Example 3 (Function spaces). Let $X \neq \emptyset$ and $V = K^X$ (set of mappings from X to K). Define

$$\begin{aligned}(f \oplus g)(x) &= f(x) + g(x), \\ (\lambda \odot f)(x) &= \lambda \cdot f(x)\end{aligned}\tag{2}$$

for $f, g \in K^X$, $\lambda \in K$ and $x \in X$. Note that the vector addition is given by the point-wise addition from Definition I.2.6. Then V is a K -vector space, see **Exercise 1.3**.

Special cases:

- (i) For $K = \mathbb{R}$ and $X = [0, 1]$ we obtain the \mathbb{R} -vector space $\mathbb{R}^{[0,1]}$ of real-valued mappings on the unit interval $[0, 1]$.
- (ii) For $K = \mathbb{R}$ and $X = \mathbb{N}$ we obtain the \mathbb{R} -vector space $\mathbb{R}^{\mathbb{N}}$ of real-valued sequences.

Convention: We always regard $V = K^X$ as a K -vector space with the vector space operations \oplus and \odot defined in (2).

In the sequel let V be a K -vector space.

Lemma 4. For $v \in V$ and $\lambda \in K$ we have

- (i) $0 \odot v = \mathbf{0}$,
- (ii) $\lambda \odot \mathbf{0} = \mathbf{0}$,
- (iii) $\lambda \odot v = \mathbf{0} \Rightarrow \lambda = 0 \vee v = \mathbf{0}$,
- (iv) $(-1) \odot v = \ominus v$.

Proof. ad (i): We have $0 \odot v = (0 + 0) \odot v = 0 \odot v \oplus 0 \odot v$.

ad (ii): We have $\lambda \odot \mathbf{0} = \lambda \cdot (\mathbf{0} \oplus \mathbf{0}) = \lambda \odot \mathbf{0} \oplus \lambda \odot \mathbf{0}$.

ad (iii): From $\lambda \odot v = \mathbf{0}$ and $\lambda \neq 0$ we obtain using (ii) that

$$v = 1 \odot v = (\lambda^{-1} \cdot \lambda) \odot v = \lambda^{-1} \odot \mathbf{0} = \mathbf{0}.$$

ad (iv): Using (i) we obtain

$$v \oplus (-1) \odot v = 1 \odot v \oplus (-1) \odot v = (1 + (-1)) \odot v = \mathbf{0}. \quad \square$$

Notation: Henceforth we use 0 , $+$, and \cdot instead of $\mathbf{0}$, \oplus , and \odot , respectively. Moreover, we often use λv instead of $\lambda \cdot v$ for $\lambda \in K$ and $v \in V$.

1.1 Subspaces

Definition 5. $U \subseteq V$ is a *subspace* (*linear subspace*, *vector subspace*) of V if the following conditions hold:

- (i) $U \neq \emptyset$,
- (ii) $\forall v, w \in U: v + w \in U$,
- (iii) $\forall \lambda \in K \forall v \in U: \lambda \cdot v \in U$.

Remark 6. (i) Every subspace U of V satisfies $0 \in U$.

- (ii) $\{0\}$ and V are the smallest and the largest subspace of V , respectively, i.e., every subspace U of V satisfies $\{0\} \subseteq U \subseteq V$.

Cf. Remark I.1.14.

Example 7. (i) For $a \in \mathbb{R}^n$ und $b \in \mathbb{R}$ put

$$U = \left\{ x \in \mathbb{R}^n: \sum_{i=1}^n a_i \cdot x_i = b \right\}.$$

U is a subspace of \mathbb{R}^n if and only if $b = 0$, see **Exercise 1.4**.

- (ii) The set of convergent sequences in \mathbb{R} is a subspace of $\mathbb{R}^{\mathbb{N}}$.
- (iii) The set $C([0, 1])$ of continuous functions from $[0, 1]$ to \mathbb{R} is a subspace of $\mathbb{R}^{[0, 1]}$, see **Exercise 1.3**.

Proposition 8. Every subspace of V (together with the induced vector addition and the induced scalar multiplication) is a K -vector space.

Proof. Analogous to subgroups, see Proposition I.1.16. □

Lemma 9. Let U_1 and U_2 be subspaces of V . Then $U_1 \cap U_2$ is a subspace of V .

Proof. Analogous to subgroups, see Lemma I.1.17. □

Remark 10. The analogue of Lemma 9 for the union of subspaces is false, see **Exercise 1.4**.

1.2 Span and Linear Combinations

Definition 11. For $W \subseteq V$ let

$$\mathfrak{U} = \{U: U \text{ subspace of } V, W \subseteq U\}.$$

$\bigcap_{U \in \mathfrak{U}} U$ is called *span* (*linear span*, *linear hull*) of W . *Notation:*

$$\text{span}(W) = \bigcap_{U \in \mathfrak{U}} U.$$

Remark 12. $\text{span}(W)$ is the smallest subspace of V that contains W . In particular, $\text{span}(\emptyset) = \{0\}$.

Definition 13. For $n \in \mathbb{N}$ we call $v \in V$ a *linear combination* of $v_1, \dots, v_n \in V$ if

$$\exists \lambda_1, \dots, \lambda_n \in K : v = \sum_{i=1}^n \lambda_i v_i.$$

Notation: $L(v_1, \dots, v_n)$ is the set of all linear combinations of v_1, \dots, v_n .

Example 14. For $V = \mathbb{R}^2$ and $v_1 = (1, 0)$, $v_2 = (0, 1)$, and $v_3 = (3, 2)$ we have

$$L(v_1, v_2) = L(v_1, v_3) = L(v_2, v_3) = L(v_1, v_2, v_3) = \mathbb{R}^2.$$

Lemma 15. $L(v_1, \dots, v_n)$ is a subspace of V for all $n \in \mathbb{N}$ and $v_1, \dots, v_n \in V$.

Proof. We have $0 \in L(v_1, \dots, v_n)$. For $v, w \in L(v_1, \dots, v_n)$ there exist $\lambda_1, \dots, \lambda_n \in K$ and $\mu_1, \dots, \mu_n \in K$ such that

$$v = \sum_{i=1}^n \lambda_i v_i, \quad w = \sum_{i=1}^n \mu_i v_i.$$

Hence we get

$$v + w = \sum_{i=1}^n (\lambda_i + \mu_i) v_i \in L(v_1, \dots, v_n)$$

and

$$\lambda v = \sum_{i=1}^n \lambda \lambda_i v_i \in L(v_1, \dots, v_n)$$

for all $\lambda \in K$. □

Proposition 16. For $\emptyset \neq W \subseteq V$ we have

$$\text{span}(W) = \bigcup_{n \in \mathbb{N}} \bigcup_{(v_1, \dots, v_n) \in W^n} L(v_1, \dots, v_n).$$

Corollary 17. For $n \in \mathbb{N}$ and $v_1, \dots, v_n \in V$ we have

$$\text{span}(\{v_1, \dots, v_n\}) = L(v_1, \dots, v_n).$$

Proof. “ \supseteq ”: We have $\{v_1, \dots, v_n\} \subseteq \text{span}(\{v_1, \dots, v_n\})$. Since $\text{span}(\{v_1, \dots, v_n\})$ is a subspace of V , see Remark 12, we obtain $L(v_1, \dots, v_n) \subseteq \text{span}(\{v_1, \dots, v_n\})$.

“ \subseteq ”: We have $\{v_1, \dots, v_n\} \subseteq L(v_1, \dots, v_n)$. Moreover, $L(v_1, \dots, v_n)$ is a subspace according to Lemma 15, and hence $\text{span}(\{v_1, \dots, v_n\}) \subseteq L(v_1, \dots, v_n)$. □

Example 18. Let $V = \mathbb{R}^3$ and $v_1, v_2 \in V$. Corollary 17 shows

$$\text{span}(\{v_1\}) = \{\lambda v_1 : \lambda \in \mathbb{R}\}.$$

If $v_1 \neq 0$, $\text{span}(\{v_1\})$ is the straight line passing through 0 and v_1 . Furthermore, we have

$$\text{span}(\{v_1, v_2\}) = \{\lambda v_1 + \mu v_2 : \lambda, \mu \in \mathbb{R}\}.$$

If $v_1 \neq 0$ and $v_2 \notin \text{span}(\{v_1\})$, $\text{span}(\{v_1, v_2\})$ is the plane passing through 0, v_1 , and v_2 .

Example 19. Let $V = \mathbb{R}^{\mathbb{R}}$. Define $v_i \in V$ for $i \in \{0, 1, 2, 3\}$ by $v_i(x) = x^i$ for $x \in \mathbb{R}$. Then $\text{span}(\{v_i : i \in \{0, 1, 2, 3\}\})$ is the subspace of polynomial functions where the degree is at most 3.

2 Bases and Dimension

In the sequel let V be a K -vector space and $(v_i)_{i \in I}$ be a family in V for a set I . We put $V_0 = \{v_i : i \in I\}$. If $I \neq \emptyset$ is finite, we may assume w.l.o.g. that $I = \{1, \dots, n\}$.

Notation: Often (v_1, \dots, v_n) or $(v_i)_{i \in \{1, \dots, n\}}$ instead of $(v_i)_{i \in I}$.

2.1 Linear Independence

Definition 1. A family (v_1, \dots, v_n) is *linearly independent* if for all $(\lambda_i)_{i \in \{1, \dots, n\}} \in K^n$ it holds

$$\sum_{i=1}^n \lambda_i v_i = 0 \Rightarrow (\forall i \in \{1, \dots, n\} : \lambda_i = 0).$$

Otherwise the family (v_1, \dots, v_n) is *linearly dependent*. The empty family $(v_i)_{i \in \emptyset}$ is linearly independent.

Remark 2.

- (i) The family (v_1) is linearly independent if and only if $v_1 \neq 0$.
- (ii) If $v_i = 0$ for some $i \in \{1, \dots, n\}$, then (v_1, \dots, v_n) is linearly dependent.
(Proof: $1 \cdot v_i = 0$.)
- (iii) If $v_i = v_j$ for $i \neq j$, then (v_1, \dots, v_n) is linearly dependent.
(Proof: $1 \cdot v + (-1) \cdot v = 0$.)
- (iv) For $J \subseteq \{1, \dots, n\}$ it holds: $(v_i)_{i \in \{1, \dots, n\}}$ linearly independent $\Rightarrow (v_i)_{i \in J}$ linearly independent.

Lemma 3. Let $n \geq 2$. The following statements are equivalent:

- (i) (v_1, \dots, v_n) linearly dependent,
- (ii) $\exists i \in \{1, \dots, n\} : v_i \in L((v_j)_{j \in \{1, \dots, n\} \setminus \{i\}})$.

Proof. “(i) \Rightarrow (ii)” : There exist $(\lambda_j)_{j \in \{1, \dots, n\}} \in K^n$ and $i \in \{1, \dots, n\}$ such that $\sum_{j=1}^n \lambda_j v_j = 0$ and $\lambda_i \neq 0$. This shows

$$v_i = \sum_{j \in \{1, \dots, n\} \setminus \{i\}} (-\lambda_j / \lambda_i) \cdot v_j \in L((v_j)_{j \in \{1, \dots, n\} \setminus \{i\}}).$$

“(ii) \Rightarrow (i)” : There exist $i \in \{1, \dots, n\}$ and $(\lambda_j)_{j \in \{1, \dots, n\} \setminus \{i\}} \in K^{n-1}$ such that $v_i = \sum_{j \in \{1, \dots, n\} \setminus \{i\}} \lambda_j v_j$. Put $\lambda_i = -1$ to obtain

$$\sum_{j=1}^n \lambda_j v_j = \lambda_i v_i + \sum_{j \in \{1, \dots, n\} \setminus \{i\}} \lambda_j v_j = -v_i + v_i = 0. \quad \square$$

Example 4. (i) Consider the situation from Example 1.14. Then (v_1, v_2) , (v_1, v_3) , and (v_2, v_3) are linearly independent. E.g., for (v_1, v_3) we have

$$\lambda_1 v_1 + \lambda_3 v_3 = 0 \Leftrightarrow \lambda_1 \cdot (1, 0) + \lambda_3 \cdot (3, 2) = 0 \Leftrightarrow \begin{cases} \lambda_1 + 3\lambda_3 = 0 \\ 2\lambda_3 = 0 \end{cases}$$

for $\lambda_1, \lambda_3 \in \mathbb{R}$, and hence $\lambda_1 = \lambda_3 = 0$. Furthermore, (v_1, v_2, v_3) is linearly dependent since $v_3 = 3v_1 + 2v_2$.

(ii) Consider the situation from Example 1.19 with $v_0(x) = 1$ and $v_1(x) = x$ for $x \in \mathbb{R}$. Then (v_0, v_1) is linearly independent. To see this, let $\lambda_0 v_0 + \lambda_1 v_1 = 0$ for $\lambda_0, \lambda_1 \in \mathbb{R}$. At the evaluation sites $x = 0$ and $x = 1$ we then have

$$\begin{aligned} 0 &= (\lambda_0 v_0 + \lambda_1 v_1)(0) = \lambda_0 \cdot v_0(0) + \lambda_1 \cdot v_1(0) = \lambda_0 \cdot 1 + \lambda_1 \cdot 0 = \lambda_0, \\ 0 &= (\lambda_0 v_0 + \lambda_1 v_1)(1) = \lambda_0 \cdot v_0(1) + \lambda_1 \cdot v_1(1) = \lambda_0 \cdot 1 + \lambda_1 \cdot 1 = \lambda_0 + \lambda_1, \end{aligned}$$

and hence $\lambda_0 = \lambda_1 = 0$.

Cf. Exercise 2.2 and Exercise 2.4.

Proposition 5. The following statements are equivalent:

- (i) (v_1, \dots, v_n) is linearly independent,
- (ii) for all $v \in \text{span}(\{v_1, \dots, v_n\})$ there exists a unique $(\lambda_i)_{i \in \{1, \dots, n\}} \in K^n$ such that

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

Proof. “(i) \Rightarrow (ii)”: Existence follows from Corollary 1.17. Moreover, the fact

$$\sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n \mu_i v_i \Leftrightarrow \sum_{i=1}^n (\lambda_i - \mu_i) v_i = 0$$

for $(\lambda_i)_{i \in \{1, \dots, n\}}, (\mu_i)_{i \in \{1, \dots, n\}} \in K^n$ and the linear independence of (v_1, \dots, v_n) shows uniqueness.

“(ii) \Rightarrow (i)”: The unique representation of $v = 0$ corresponds to $\lambda_1 = \dots = \lambda_n = 0$. \square

2.2 Generating Sets and Bases

Terminology: For a finite index set I we call $(v_i)_{i \in I}$ finite and $|I|$ the length of $(v_i)_{i \in I}$. For an infinite index set I we call $(v_i)_{i \in I}$ infinite.

Definition 6.

- (i) $(v_i)_{i \in I}$ is a *generating set* (*spanning set*) of V if $V = \text{span}(\{v_i : i \in I\})$.
- (ii) A linearly independent generating set of V is called *basis* of V .

(iii) V is *finitely generated* if there exists a finite generating set of V .

Example 7. Let $V = \mathbb{R}^n$ and $e_1, \dots, e_n \in \mathbb{R}^n$ be given by

$$e_i(j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{else,} \end{cases}$$

where $e_i(j)$ denotes the j -th component of e_i for $i, j \in \{1, \dots, n\}$. Clearly, (e_1, \dots, e_n) is linearly independent and every $x \in \mathbb{R}^n$ can be expressed by

$$x = (x_1, \dots, x_n) = \sum_{i=1}^n x_i \cdot e_i.$$

Hence (e_1, \dots, e_n) is a basis of \mathbb{R}^n . It is called the *standard basis* or the *canonical basis* of the coordinate space \mathbb{R}^n .

Proposition 8. Let $n \in \mathbb{N}$. The following statements are equivalent:

- (i) (v_1, \dots, v_n) is a basis of V .
- (ii) (v_1, \dots, v_n) is a generating set of V , and $(v_j)_{j \in \{1, \dots, n\} \setminus \{i\}}$ is not a generating set of V for every $i \in \{1, \dots, n\}$.
- (iii) For all $v \in V$ there exists a unique $(\lambda_1, \dots, \lambda_n) \in K^n$ such that $\sum_{i=1}^n \lambda_i v_i = v$.
- (iv) (v_1, \dots, v_n) is linearly independent and every family (v_1, \dots, v_n, v) in V with $v \in V$ is linearly dependent.

Proof. We consider the non-trivial case $n \geq 2$.

“(i) \Rightarrow (ii)” : Proof by contradiction: Let $i \in \{1, \dots, n\}$ such that $(v_j)_{j \in \{1, \dots, n\} \setminus \{i\}}$ is a generating set of V . Then we have $v_i \in L((v_j)_{j \in \{1, \dots, n\} \setminus \{i\}})$, see Corollary 1.17. Lemma 3 shows that (v_1, \dots, v_n) is linearly dependent, which is a contradiction to (i).

“(ii) \Rightarrow (iii)” : Existence is obvious. Proof by contradiction to show uniqueness: Let $(\lambda_1, \dots, \lambda_n), (\mu_1, \dots, \mu_n) \in K^n$ and $i \in \{1, \dots, n\}$ such that $\lambda_i \neq \mu_i$ and

$$\sum_{j=1}^n \lambda_j v_j = \sum_{j=1}^n \mu_j v_j.$$

Then we have $v_i \in L((v_j)_{j \in \{1, \dots, n\} \setminus \{i\}})$, see Proposition 5 and Lemma 3, and hence $L(v_1, \dots, v_n) = L((v_j)_{j \in \{1, \dots, n\} \setminus \{i\}})$, which is a contradiction to (ii).

“(iii) \Rightarrow (iv)” : The linear independence of (v_1, \dots, v_n) follows from Proposition 5, the linear dependence of (v_1, \dots, v_n, v) follows from Lemma 3 since $v \in L(v_1, \dots, v_n)$.

“(iv) \Rightarrow (i)” : It remains to show $V \subseteq \text{span}(\{v_1, \dots, v_n\})$. For $v \in V$ there exist $(\lambda_1, \dots, \lambda_n) \in K^n$ and $\lambda \in K$ such that $\sum_{i=1}^n \lambda_i v_i + \lambda v = 0$ and additionally $(\lambda_1, \dots, \lambda_n, \lambda) \neq 0 \in K^{n+1}$. Since (v_1, \dots, v_n) is linearly independent, we have $\lambda \neq 0$ and hence

$$v = -1/\lambda \cdot \sum_{i=1}^n \lambda_i v_i \in \text{span}(\{v_1, \dots, v_n\}). \quad \square$$

Corollary 9 (“Basisauswahlsatz”). Let $(v_i)_{i \in I}$ be a finite generating set of V . Then there exists $I_0 \subseteq I$ such that $(v_i)_{i \in I_0}$ is a basis of V . In particular, every finitely generated vector space has a basis.

Proof. We consider the non-trivial case $V \neq \{0\}$. Let $V = L((v_i)_{i \in I})$ with a finite set $I \neq \emptyset$.

If $|I| = 1$, then $(v_i)_{i \in I}$ is a basis, see Remark 2.(i). If $|I| \geq 2$, we either have

$$\forall i \in I: L((v_j)_{j \in I \setminus \{i\}}) \neq V,$$

which implies that $(v_i)_{i \in I}$ is a basis according to Proposition 8, or there exists $i \in I$ such that

$$L((v_j)_{j \in I \setminus \{i\}}) = V.$$

In this case we consider $(v_j)_{j \in I \setminus \{i\}}$ and start over. □

2.3 Dimension of a Vector Space

Lemma 10 (“Basisaustauschlemma”). Let $n \in \mathbb{N}$ and let (v_1, \dots, v_n) be a basis of V . Moreover, let $w = \sum_{j=1}^n \lambda_j v_j$ with $(\lambda_1, \dots, \lambda_n) \in K^n$ such that $\lambda_i \neq 0$ for some $i \in \{1, \dots, n\}$. Then $(\tilde{v}_1, \dots, \tilde{v}_n)$ given by

$$\tilde{v}_j = \begin{cases} v_j, & \text{if } j \neq i, \\ w, & \text{if } j = i, \end{cases}$$

is also a basis of V .

Proposition 11. If V has a finite basis, then every basis of V is finite and the lengths of all bases coincide.

Definition 12. If V has a finite basis, the length of a basis of V is called *dimension* of V and V is called *finite-dimensional*. Otherwise, V is called *infinite-dimensional*.

Notation: $\dim V$ for the dimension of V .

Example 13. For $n \in \mathbb{N}$ we have $\dim \mathbb{R}^n = n$, see Example 7.

Definition 14. A subspace U of \mathbb{R}^n is a (*straight*) *line* passing through zero if $\dim U = 1$. A subspace U of \mathbb{R}^n is a *plane* passing through zero if $\dim U = 2$.

Corollary 15. Let V be finitely generated and let U be a subspace of V . Then we have

- (i) $\dim U \leq \dim V$,
- (ii) $\dim U = \dim V \Rightarrow U = V$.

Corollary 16 (“Basisergänzungssatz”). Let V be finite-dimensional with $\dim V = n \geq 2$. Moreover, let (w_1, \dots, w_r) be linearly independent with $r \in \{1, \dots, n-1\}$. Then there exist $w_{r+1}, \dots, w_n \in V$ such that (w_1, \dots, w_n) is a basis of V .

Proposition 17. Let U_1, U_2 be finite-dimensional subspaces of V . Then we have

$$\dim(U_1 \cap U_2) + \dim(U_1 + U_2) = \dim(U_1) + \dim(U_2).$$

Proof. According to Corollary 15 the subspace $U_1 \cap U_2$ of U_1 is finitely generated and hence there exists a basis (u_1, \dots, u_n) of $U_1 \cap U_2$ due to Corollary 9. Using Corollary 16 we may extend the basis (u_1, \dots, u_n) of $U_1 \cap U_2$ to a basis $(u_1, \dots, u_n, u'_1, \dots, u'_m)$ of U_1 and a basis $(u_1, \dots, u_n, u''_1, \dots, u''_p)$ of U_2 for $m, p \in \mathbb{N}_0$.

We show that $(u_1, \dots, u_n, u'_1, \dots, u'_m, u''_1, \dots, u''_p)$ is a basis of $U_1 + U_2$: Since, by construction, $(u_1, \dots, u_n, u'_1, \dots, u'_m, u''_1, \dots, u''_p)$ is a generating set of $U_1 + U_2$, it remains to show the linear independence. Let $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_m, \nu_1, \dots, \nu_p \in K$ such that

$$\sum_{i=1}^n \lambda_i u_i + \sum_{i=1}^m \mu_i u'_i + \sum_{i=1}^p \nu_i u''_i = 0$$

Then we have

$$\underbrace{\sum_{i=1}^n \lambda_i u_i + \sum_{i=1}^m \mu_i u'_i}_{\in U_1} = - \underbrace{\sum_{i=1}^p \nu_i u''_i}_{\in U_2} \in U_1 \cap U_2.$$

Hence there exist $\lambda'_1, \dots, \lambda'_n \in K$ such that

$$\sum_{i=1}^n \lambda_i u_i + \sum_{i=1}^m \mu_i u'_i = \sum_{i=1}^n \lambda'_i u_i.$$

Since $(u_1, \dots, u_n, u'_1, \dots, u'_m)$ is a basis of U_1 , Proposition 8.(iii) shows

$$\forall i \in \{1, \dots, n\}: \lambda_i = \lambda'_i \quad \wedge \quad \forall i \in \{1, \dots, m\}: \mu_i = 0.$$

This yields

$$\sum_{i=1}^n \lambda_i u_i + \sum_{i=1}^p \nu_i u''_i = 0.$$

Since $(u_1, \dots, u_n, u''_1, \dots, u''_p)$ is a basis of U_2 , we obtain

$$\forall i \in \{1, \dots, n\}: \lambda_i = 0 \quad \wedge \quad \forall i \in \{1, \dots, p\}: \nu_i = 0,$$

i.e., $(u_1, \dots, u_n, u'_1, \dots, u'_m, u''_1, \dots, u''_p)$ is linearly independent and hence a basis of $U_1 + U_2$. Finally, we obtain

$$\begin{aligned} \dim(U_1 \cap U_2) + \dim(U_1 + U_2) &= n + (n + m + p) \\ &= (n + m) + (n + p) = \dim(U_1) + \dim(U_2). \quad \square \end{aligned}$$

Outlook 18. Based on Zorn's lemma (axiom of choice) one can show that every vector space has a basis.

3 Matrices

In the sequel let K be a field and $m, n, p \in \mathbb{N}$.

Given: $v_1, \dots, v_m \in K^n$.

Aim: Find a basis of $\text{span}(\{v_1, \dots, v_m\})$, cf. Corollary 2.9.

Definition 1. An $m \times n$ -matrix $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n}$ (over K) is a rectangular array of elements of K . Its elements $a_{i,j} \in K$ are called *entries* of A .

Notation: $K^{m \times n}$ denotes the set of all $m \times n$ -matrices over K . We write

$$A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} = (a_{i,j})_{i,j} \in K^{m \times n}.$$

Henceforth we identify¹ K^m with $K^{m \times 1}$. We put (“*transposed*”)

$$(x_1, \dots, x_m)^\top = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in K^m$$

and

$$\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}^\top = (x_1, \dots, x_m) \in K^{1 \times m}$$

for $x_1, \dots, x_m \in K$.

Definition 2.

- (i) For $a_1, \dots, a_n \in K^m$, $A = (a_1, \dots, a_n) \in K^{m \times n}$, and $1 \leq j \leq n$ we call a_j the *j-th column* of A .
- (ii) For $a_1, \dots, a_m \in K^n$,

$$A = \begin{pmatrix} a_1^\top \\ \vdots \\ a_m^\top \end{pmatrix} \in K^{m \times n}, \tag{1}$$

and $1 \leq i \leq m$ we call $a_i^\top \in K^{1 \times n}$ the *i-th row* of A .

3.1 Row Space and Rank

Definition 3.

- (i) For A according to (1) the *row space* of A is given by

$$\text{RS}(A) = \text{span}(\{a_1, \dots, a_m\}) \subseteq K^n.$$

The *rank* of A is given by $\text{rank } A = \dim \text{RS}(A)$.

¹Elements of K^m are “column vectors”.

(ii) $A \in K^{m \times n}$ is in *row echelon form* if there exists $r \in \{0, \dots, \min(m, n)\}$ such that:

- (a) $\forall 1 \leq i \leq r \exists 1 \leq j \leq n: a_{i,j} \neq 0$,
- (b) $\forall r < i \leq m \forall 1 \leq j \leq n: a_{i,j} = 0$,
- (c) if $r \geq 2$, we have $j_1 < \dots < j_r$ for $j_i = \min\{j: a_{i,j} \neq 0\}$.

The leading non-zero entries a_{i,j_i} for $i \in \{1, \dots, r\}$ are called *pivots*.

Example 4. Consider the matrix $A \in \mathbb{R}^{4 \times 6}$ given by

$$A = \begin{pmatrix} 0 & 2 & -3 & 0 & 1 & 0 \\ 0 & 0 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then, A is in row echelon form with $m = 4$, $n = 6$, $r = 3$, $j_1 = 2$, $j_2 = 3$, and $j_3 = 6$, see Definition 3.(ii). Moreover, the pivots are $a_{1,2} = 2$, $a_{2,3} = -1$, and $a_{3,6} = 3$.

Remark 5. Let A be given by (1) and in row echelon form, and let r be according to Definition 3.(ii). Then we have $r = \text{rank } A$. Moreover, if $r > 0$, we have $\text{RS}(A) = \text{span}(\{a_1, \dots, a_r\})$, and (a_1, \dots, a_r) is a basis of the subspace $\text{RS}(A)$.

Definition 6. There are three types of *elementary row operations*, which can be applied to $A \in K^{m \times n}$ with $\lambda \in K \setminus \{0\}$:

- (i) multiplication of a row of A by λ (type I),
- (ii) addition of the λ -fold of a row of A to another row of A (type II),
- (iii) switching two rows within A (type III).

Lemma 7. Let $\tilde{A} \in K^{m \times n}$ result from finitely many elementary row operations applied to $A \in K^{m \times n}$. Then we have $\text{RS}(\tilde{A}) = \text{RS}(A)$.

Proof. For a single elementary row operation we clearly have $\text{RS}(\tilde{A}) \subseteq \text{RS}(A)$. Since every elementary row operation is reversible using an elementary row operation of the same type, we also obtain $\text{RS}(A) \subseteq \text{RS}(\tilde{A})$. The general case follows by induction. \square

Proposition 8. Every matrix $A \in K^{m \times n}$ can be transformed into row echelon form by using finitely many elementary row operations.

Proof. *Gaussian elimination* for $A = (a_1, \dots, a_n) = (a_{i,j})_{i,j} \in K^{m \times n}$:

- 1.) If $A = 0$ or $m = 1$: STOP.
- 2.) Let $j^* = \min\{j \in \{1, \dots, n\}: a_j \neq 0\}$.
- 3.) Choose $i \in \{1, \dots, m\}$ such that $a_{i,j^*} \neq 0$.
- 4.) If $i \neq 1$: Switch rows 1 and i .

- 5.) For $i = 2, \dots, m$: Subtract the $a_{i,j^*}/a_{1,j^*}$ -fold of row 1 from row i .
- 6.) If $j^* = n$: STOP.
- 7.) Remove row 1 and columns $1, \dots, j^*$. Take the resulting submatrix and continue with step 1.).

By induction, this algorithm transforms every $A \in K^{m \times n}$ into row echelon form. \square

Example 9. For $m = n = 3$ we consider $v_1, v_2, v_3 \in \mathbb{R}^3$ given by

$$v_1 = (0, 0, 1)^\top, \quad v_2 = (0, 1, 0)^\top, \quad v_3 = (0, 1, 1)^\top.$$

Gaussian elimination for A consisting of the rows $v_1^\top, v_2^\top, v_3^\top$ yields²

$$\begin{aligned} & \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{(I) \leftrightarrow (II)} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{(III) = (III) - (I)} \begin{pmatrix} 0 & \boxed{1} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ & \rightarrow \begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 \\ \cdot & \cdot & 1 \end{pmatrix} \xrightarrow{(II) = (II) - (I)} \begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \boxed{1} \\ \cdot & \cdot & 0 \end{pmatrix} \\ & \rightarrow \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Hence we get $\text{rank } A = 2$ and

$$\text{span}(\{v_1, v_2, v_3\}) = \text{span}(\{(0, 1, 0)^\top, (0, 0, 1)^\top\}) = \{x \in \mathbb{R}^3 : x_1 = 0\}.$$

Moreover, $((0, 1, 0)^\top, (0, 0, 1)^\top)$ is a basis of this subspace.

3.2 Matrix Multiplication

Remark 10. $K^{m \times n}$ is a K -vector space with the coordinatewise operations from Example 1.2. We have $\dim K^{m \times n} = \dim K^{m \cdot n} = m \cdot n$, see Example 2.13.

Definition 11. The *matrix product* $C = (c_{i,j})_{i,j} \in K^{m \times p}$ of two matrices $A = (a_{i,\ell})_{i,\ell} \in K^{m \times n}$ and $B = (b_{\ell,j})_{\ell,j} \in K^{n \times p}$ is defined by

$$c_{i,j} = \sum_{\ell=1}^n a_{i,\ell} \cdot b_{\ell,j}, \quad 1 \leq i \leq m, \quad 1 \leq j \leq p.$$

²The elementary row operation, e.g., of type II where the “new” third row results from the sum of the previous third row and the λ -fold of the first row is indicated by “(III)=(III)+ λ ·(I)”.

The matrix product can be illustrated by the following scheme:

$$\underbrace{\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{i,1} & \cdots & a_{i,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}}_n \underbrace{\begin{pmatrix} b_{1,1} & \cdots & b_{1,j} & \cdots & b_{1,p} \\ \vdots & & \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,j} & \cdots & b_{n,p} \end{pmatrix}}_p \left. \begin{matrix} \left. \begin{pmatrix} c_{1,1} & \cdots & c_{1,j} & \cdots & c_{1,p} \\ \vdots & & \vdots & & \vdots \\ c_{i,1} & \cdots & c_{i,j} & \cdots & c_{i,p} \\ \vdots & & \vdots & & \vdots \\ c_{m,1} & \cdots & c_{m,j} & \cdots & c_{m,p} \end{pmatrix} \right\} m \\ \left. \begin{pmatrix} b_{1,1} & \cdots & b_{1,j} & \cdots & b_{1,p} \\ \vdots & & \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,j} & \cdots & b_{n,p} \end{pmatrix} \right\} n \end{matrix} \right\}$$

Notation: $C = A \cdot B$ or $C = AB$.

Convention: \cdot has precedence over $+$.

Remark 12. The matrix multiplication is associative, see **Exercise 4.2**. However, it is neither commutative nor does $AB = 0$ imply $A = 0$ or $B = 0$, see, e.g.,

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Remark 13. Let

$$A = \begin{pmatrix} a_1^\top \\ \vdots \\ a_m^\top \end{pmatrix} \in K^{m \times n}, \quad B = (b_1, \dots, b_p) \in K^{n \times p}.$$

Then we have

$$A \cdot B = (A \cdot b_1, \dots, A \cdot b_p) = \begin{pmatrix} a_1^\top \cdot B \\ \vdots \\ a_m^\top \cdot B \end{pmatrix}.$$

Special cases:

(i) For $A = (a_1, \dots, a_n) = (a_{i,j})_{i,j} \in K^{m \times n}$ and $x = (x_1, \dots, x_n)^\top \in K^n$ we have

$$Ax = \begin{pmatrix} \sum_{\ell=1}^n a_{1,\ell} \cdot x_\ell \\ \vdots \\ \sum_{\ell=1}^n a_{m,\ell} \cdot x_\ell \end{pmatrix} = \sum_{\ell=1}^n x_\ell \cdot a_\ell \in K^m.$$

(ii) For $y \in K^n$ and

$$B = \begin{pmatrix} b_1^\top \\ \vdots \\ b_n^\top \end{pmatrix} = (b_{i,j})_{i,j} \in K^{n \times p}$$

we have

$$y^\top \cdot B = \left(\sum_{\ell=1}^n y_\ell \cdot b_{\ell,1}, \dots, \sum_{\ell=1}^n y_\ell \cdot b_{\ell,p} \right) = \sum_{\ell=1}^n y_\ell \cdot b_\ell^\top \in K^{1 \times p}.$$

Definition 14. $E_n = (e_{i,j})_{i,j} \in K^{n \times n}$ given by

$$e_{i,j} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{else,} \end{cases}$$

is called $n \times n$ -identity matrix.

Lemma 15. For $A, A_1, A_2 \in K^{m \times n}$, $B, B_1, B_2 \in K^{n \times p}$, and $\lambda \in K$ we have³

- (i) $(A_1 + A_2) \cdot B = A_1 \cdot B + A_2 \cdot B$,
- (ii) $A \cdot (B_1 + B_2) = A \cdot B_1 + A \cdot B_2$,
- (iii) $(\lambda \cdot A) \cdot B = A \cdot (\lambda \cdot B) = \lambda \cdot (A \cdot B)$,
- (iv) $A \cdot E_n = A$ and $E_n \cdot B = B$.

Proof. Exercise. □

Notation: Due to Remark 12 and Lemma 15.(iii) we may omit brackets in the context of scalar and matrix multiplication.

4 Linear Maps

In the sequel let U , V , and W be K -vector spaces.

Definition 1. A map $F: V \rightarrow W$ is *linear* if

- (i) $\forall v, w \in V: F(v + w) = F(v) + F(w)$,
- (ii) $\forall v \in V \forall \lambda \in K: F(\lambda v) = \lambda F(v)$.

F is a *linear isomorphism* if F is linear and bijective. Two vector spaces V and W are *isomorphic* if there exists a linear isomorphism from V to W .

Example 2.

- (i) For $A \in K^{m \times n}$ let $\mathcal{F}_A: K^n \rightarrow K^m$ be given by

$$\mathcal{F}_A(v) = Av, \quad v \in K^n.$$

According to Lemma 3.15.(ii) and (iii) we have

$$\begin{aligned} \mathcal{F}_A(v + w) &= A(v + w) = Av + Aw = \mathcal{F}_A(v) + \mathcal{F}_A(w), \\ \mathcal{F}_A(\lambda v) &= A(\lambda v) = \lambda Av = \lambda \mathcal{F}_A(v) \end{aligned}$$

for all $v, w \in K^n$ and $\lambda \in K$. Hence \mathcal{F}_A is a linear map from K^n to K^m .

³ $(K^{n \times n}, +, \cdot)$ is a ring with unity E_n .

(ii) For $a, b \in \mathbb{R}$ with $a < b$ let $F: C([a, b]) \rightarrow \mathbb{R}$ be given by

$$F(v) = \int_a^b v(x) dx, \quad v \in C([a, b]).$$

Note that

$$\begin{aligned} F(v + w) &= \int_a^b v(x) + w(x) dx = \int_a^b v(x) dx + \int_a^b w(x) dx = F(v) + F(w), \\ F(\lambda v) &= \int_a^b \lambda \cdot v(x) dx = \lambda \cdot \int_a^b v(x) dx = \lambda F(v) \end{aligned}$$

for all $v, w \in C([a, b])$ and $\lambda \in \mathbb{R}$, see Analysis I. Hence F is a linear map from $C([a, b])$ to \mathbb{R} .

(iii) For $x \in [0, 1]$ let $F: \mathbb{R}^{[0,1]} \rightarrow \mathbb{R}$ be given by

$$F(\varphi) = \varphi(x), \quad \varphi \in \mathbb{R}^{[0,1]}.$$

Note that

$$\begin{aligned} F(\varphi + \psi) &= (\varphi + \psi)(x) = \varphi(x) + \psi(x) = F(\varphi) + F(\psi), \\ F(\lambda \varphi) &= (\lambda \varphi)(x) = \lambda \cdot \varphi(x) = \lambda F(\varphi) \end{aligned}$$

for all $\varphi, \psi \in \mathbb{R}^{[0,1]}$ and $\lambda \in \mathbb{R}$, see Definition I.2.6. Hence F is a linear map from $\mathbb{R}^{[0,1]}$ to \mathbb{R} . It is called function *evaluation* at x .

Lemma 3. Let $F: V \rightarrow W$ be linear. Then we have:

- (i) $F(0) = 0$.
- (ii) F is injective if and only if $F(v) \neq 0$ for all $v \in V \setminus \{0\}$.
- (iii) $F(\sum_{i=1}^n \lambda_i v_i) = \sum_{i=1}^n \lambda_i F(v_i)$ for $\lambda_1, \dots, \lambda_n \in K$ and $v_1, \dots, v_n \in V$.
- (iv) If (v_1, \dots, v_n) is linearly dependent in V , then $(F(v_1), \dots, F(v_n))$ is also linearly dependent in W .
- (v) If V' is a subspace of V , then $F(V')$ is a subspace of W .
- (vi) If W' is a subspace of W , then $F^{-1}(W')$ is a subspace of V .
- (vii) If V is finite-dimensional, then $F(V)$ is also finite-dimensional and $\dim F(V) \leq \dim V$.

Proof. ad (i): $F(0) = F(0 \cdot 0) = 0 \cdot F(0) = 0$.

ad (ii): For $v, w \in V$ we have $F(v - w) = F(v) - F(w)$ and hence

$$F(v - w) = 0 \Leftrightarrow F(v) = F(w).$$

Thus F is injective if and only if

$$F(v - w) = 0 \Leftrightarrow F(v) = F(w) \Leftrightarrow v - w = 0.$$

ad (iii): Induction.

ad (iv): There exists $i \in \{1, \dots, n\}$ such that

$$v_i = \sum_{j \in \{1, \dots, n\} \setminus \{i\}} \lambda_j v_j$$

for $\lambda_j \in K$, see Lemma 2.3. According to part (iii) we obtain

$$F(v_i) = \sum_{j \in \{1, \dots, n\} \setminus \{i\}} \lambda_j F(v_j).$$

According to Lemma 2.3 the family $(F(v_i), \dots, F(v_n))$ is linearly dependent in W . Cf. **Exercise 5.1**.

ad (v): We have $0 = F(0) \in F(V')$. Moreover, for $v, w \in V'$ and $\lambda \in K$ we have $v + w \in V'$ and $\lambda v \in V'$. Hence we get $F(v) + F(w) = F(v + w) \in F(V')$ and $\lambda F(v) = F(\lambda v) \in F(V')$. Cf. **Exercise 5.1**.

ad (vi): We have $F(0) = 0 \in W'$. Moreover, for $v, w \in V$ with $F(v), F(w) \in W'$ and $\lambda \in K$ we have $F(v) + F(w) \in W'$ and $\lambda F(v) \in W'$. Hence we get $F(v + w) = F(v) + F(w) \in W'$ and $F(\lambda v) = \lambda F(v) \in W'$.

ad (vii): Let (v_1, \dots, v_n) be a basis of V . Then $(F(v_1), \dots, F(v_n))$ is a generating system of $F(V)$ according to part (iii). Apply Corollary 2.9 (Basisauswahlsatz). \square

Remark 4. Let $F: V \rightarrow W$.

- (i) F is linear if and only if $F(\lambda v + w) = \lambda F(v) + F(w)$ for all $v, w \in V$ and $\lambda \in K$, cf. **Exercise 5.2**.

Proof: “ \Rightarrow ”: Let F be linear. We have

$$F(\lambda v + w) = F(\lambda v) + F(w) = \lambda F(v) + F(w)$$

for $v, w \in V$ and $\lambda \in K$.

“ \Leftarrow ”: For $w \in V$ we have

$$F(w) = F(1 \cdot 0 + w) = 1 \cdot F(0) + F(w) = F(0) + F(w),$$

which shows $F(0) = 0$. Hence we have

$$F(v + w) = F(1 \cdot v + w) = 1 \cdot F(v) + F(w) = F(v) + F(w)$$

and

$$F(\lambda v) = F(\lambda v + 0) = \lambda F(v) + F(0) = \lambda F(v),$$

respectively, for $v, w \in V$ and $\lambda \in K$.

(ii) If F is a linear isomorphism, then F^{-1} is also linear.

Proof: Let $w, w' \in W$ and $\lambda \in K$. For $v = F^{-1}(w)$ and $v' = F^{-1}(w')$ we have

$$F(\lambda v + v') = \lambda w + w'$$

and thus

$$\lambda F^{-1}(w) + F^{-1}(w') = \lambda v + v' = F^{-1}(\lambda w + w').$$

Proposition 5. Let $F: V \rightarrow W$ be linear and $n \in \mathbb{N}$.

- (i) If F is injective and (v_1, \dots, v_n) is a linearly independent family in V , then $(F(v_1), \dots, F(v_n))$ is a linear independent family in W .
- (ii) If F is surjective and (v_1, \dots, v_n) is a generating set of V , then $(F(v_1), \dots, F(v_n))$ is a generating set of W .
- (iii) If F is a linear isomorphism and V is finite-dimensional, then $\dim V = \dim W$.

Proof. ad (i): Let $\sum_{i=1}^n \lambda_i F(v_i) = 0$ for $\lambda_1, \dots, \lambda_n \in K$. Then we have $F(\sum_{i=1}^n \lambda_i v_i) = 0$ and thus $\sum_{i=1}^n \lambda_i v_i = 0$ since F is injective, see Lemma 3.(ii). The linear independence of (v_1, \dots, v_n) implies $\lambda_i = 0$ for all $i \in \{1, \dots, n\}$.

ad (ii): Let $w \in W$. By assumption, there exist $v \in V$ and $\lambda_1, \dots, \lambda_n \in K$ such that $F(v) = w$ and $v = \sum_{i=1}^n \lambda_i v_i$. Then we have $w = \sum_{i=1}^n \lambda_i F(v_i)$.

ad (iii): Let (v_1, \dots, v_n) be a basis of V . According to (i) and (ii) the family $(F(v_1), \dots, F(v_n))$ is a basis of W . \square

4.1 Construction of Linear Maps

In the sequel let $n, m \in \mathbb{N}$ and $\mathcal{A} = (v_1, \dots, v_n)$ be a basis of V .

Lemma 6. For linear maps $F, G: V \rightarrow W$ we have

$$F = G \Leftrightarrow \forall i \in \{1, \dots, n\}: F(v_i) = G(v_i).$$

Proof. “ \Rightarrow ”: Obvious.

“ \Leftarrow ”: Let $v \in V$. Since \mathcal{A} is a basis of V and thus a generating set of V , there exist $\lambda_1, \dots, \lambda_n \in K$ such that $v = \sum_{i=1}^n \lambda_i v_i$. By using Lemma 3.(iii) we obtain

$$F(v) = F\left(\sum_{i=1}^n \lambda_i v_i\right) = \sum_{i=1}^n \lambda_i F(v_i) = \sum_{i=1}^n \lambda_i G(v_i) = G\left(\sum_{i=1}^n \lambda_i v_i\right) = G(v). \quad \square$$

Proposition 7. Let (w_1, \dots, w_n) be a family in W . Then there exists a unique linear map $F: V \rightarrow W$ such that

$$\forall i \in \{1, \dots, n\}: F(v_i) = w_i.$$

Moreover, F is injective if and only if (w_1, \dots, w_n) is linearly independent, and F is surjective if and only if (w_1, \dots, w_n) is a generating set of W .

Proof. “Uniqueness”: See Lemma 6.

“Existence”: For $v \in V$ there exist unique $\lambda_1, \dots, \lambda_n \in K$ such that $v = \sum_{i=1}^n \lambda_i v_i$. We define $F: V \rightarrow W$ by

$$F(v) = \sum_{i=1}^n \lambda_i w_i.$$

For $w \in V$ there exist unique $\mu_1, \dots, \mu_n \in K$ such that $w = \sum_{i=1}^n \mu_i v_i$ and hence $v + w = \sum_{i=1}^n (\lambda_i + \mu_i) v_i$. This shows

$$F(v + w) = \sum_{i=1}^n (\lambda_i + \mu_i) w_i = \sum_{i=1}^n \lambda_i w_i + \sum_{i=1}^n \mu_i w_i = F(v) + F(w).$$

Moreover, for $\lambda \in K$ we have $\lambda v = \sum_{i=1}^n (\lambda \cdot \lambda_i) v_i$ and hence

$$F(\lambda v) = \sum_{i=1}^n (\lambda \cdot \lambda_i) w_i = \lambda \cdot \sum_{i=1}^n \lambda_i w_i = \lambda F(v).$$

This shows that F is linear.

“Injectivity”: By Lemma 3.(ii), F is injective if and only if

$$F(v) = 0 \Rightarrow v = 0.$$

Finally, note that $F(v) = \sum_{i=1}^n \lambda_i w_i$, $v = \sum_{i=1}^n \lambda_i v_i$, and

$$v = 0 \Leftrightarrow (\forall i \in \{1, \dots, n\}: \lambda_i = 0),$$

since \mathcal{A} is a basis and thus linearly independent.

“Surjectivity”: Obvious. □

Example 8. Let $V = W = \mathbb{R}^2$.

(i) For $\alpha \in \mathbb{R}$ let

$$\begin{aligned} v_1 &= (\cos(0), \sin(0))^\top = (1, 0)^\top, \\ v_2 &= (\cos(\pi/2), \sin(\pi/2))^\top = (0, 1)^\top, \\ w_1 &= (\cos(\alpha), \sin(\alpha))^\top, \\ w_2 &= (\cos(\pi/2 + \alpha), \sin(\pi/2 + \alpha))^\top = (-\sin(\alpha), \cos(\alpha))^\top. \end{aligned}$$

The unique linear map $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ satisfying $F(v_1) = w_1$ and $F(v_2) = w_2$ represents a rotation of α (counterclockwise).

(ii) Let $v_1 = (1, 1)^\top$ and $v_2 = (-1, 1)^\top$. The unique linear map $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ satisfying $F(v_1) = v_1$ and $F(v_2) = -v_2$ represents a reflection through the axis $\text{span}(\{v_1\})$.

Corollary 9. Let V and W be finite-dimensional. V and W are isomorphic if and only if $\dim V = \dim W$.

Proof. “ \Rightarrow ”: See Proposition 5.(iii).

“ \Leftarrow ”: Choose bases (v_1, \dots, v_n) and (w_1, \dots, w_n) of V and W , respectively, and apply Proposition 7. □

4.2 Coordinate Systems

Example 10. Consider the standard basis $\mathcal{E} = (e_1, \dots, e_n)$ of \mathbb{R}^n , see Example 2.7, and let $\mathcal{A} = (v_1, \dots, v_n)$ be a basis of the \mathbb{R} -vector space V . According to Proposition 7 we define a linear isomorphism $\Phi_{\mathcal{A}}: \mathbb{R}^n \rightarrow V$ by

$$\Phi_{\mathcal{A}}(e_i) = v_i.$$

For $(\lambda_1, \dots, \lambda_n)^\top \in \mathbb{R}^n$ we have

$$\Phi_{\mathcal{A}}((\lambda_1, \dots, \lambda_n)^\top) = \Phi_{\mathcal{A}}\left(\sum_{i=1}^n \lambda_i e_i\right) = \sum_{i=1}^n \lambda_i v_i.$$

Definition 11. $\Phi_{\mathcal{A}}$ from Example 10 is the *coordinate system* in V that is induced by \mathcal{A} , and $\Phi_{\mathcal{A}}^{-1}(v)$ is the *family of coordinates* of $v \in V$ w.r.t. \mathcal{A} .

Corollary 12. Every finite-dimensional K -vector space $V \neq \{0\}$ is isomorphic to K^n for $n = \dim V$.

Example 13. Consider the \mathbb{R} -vector space V of polynomial functions where the degree is at most 3, i.e.,

$$V = \text{span}(\{v_0, \dots, v_3\}) \subseteq \mathbb{R}^{\mathbb{R}}$$

with $v_i(x) = x^i$ for $x \in \mathbb{R}$ and $i \in \{0, 1, 2, 3\}$, see Example 1.19. Moreover, the family $\mathcal{A} = (v_0, \dots, v_3)$ is linearly independent, see Example 2.4.(ii) for (v_0, v_1) . Hence \mathcal{A} is a basis and $\dim V = 4$. By Corollary 12, V is isomorphic to \mathbb{R}^4 . The polynomial function $v \in V$ given by

$$v(x) = 4x^3 - 2x^2 + x - 7, \quad x \in \mathbb{R},$$

satisfies $v = -7v_0 + 1v_1 - 2v_2 + 4v_3$ such that the family of coordinates of v w.r.t. the basis \mathcal{A} is given by $\Phi_{\mathcal{A}}^{-1}(v) = (-7, 1, -2, 4)^\top \in \mathbb{R}^4$.

4.3 Kernel and Image

In the sequel let $F: V \rightarrow W$ be linear.

Definition 14. The *kernel* of F is defined by

$$\ker F = F^{-1}(\{0\}) = \{v \in V : F(v) = 0\}.$$

The *image* of F is defined by

$$\text{im } F = F(V) = \{F(v) : v \in V\}.$$

Remark 15.

- (i) $\ker F$ and $\text{im } F$ are subspaces of V and W , respectively, see Lemma 3.(v)-(vi).

(ii) F is injective if and only if $\ker F = \{0\}$, see Lemma 3.(ii).

Definition 16. If $\operatorname{im} F$ is finite-dimensional, the *rank* of F is given by $\operatorname{rank} F = \dim \operatorname{im} F$.

Example 17. For $V = W = \mathbb{R}^2$ and $A = (a_1, a_2) \in \mathbb{R}^{2 \times 2}$ with $a_1 = (2, 1)^\top$ and $a_2 = (-2, -1)^\top$ let $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be given by $F(x) = Ax$ for $x = (x_1, x_2)^\top \in \mathbb{R}^2$. Then we have

$$Ax = x_1 a_1 + x_2 a_2 = (x_1 - x_2) a_1,$$

see Remark 3.13, such that $\operatorname{im} F = \operatorname{span}(\{a_1\})$, which implies $\operatorname{rank} F = 1$, and $\ker F = \operatorname{span}(\{(1, 1)^\top\})$.

Notation: For $u \in V$ and $V_0 \subseteq V$ let

$$u + V_0 = \{u + v_0 : v_0 \in V_0\}.$$

It describes a *translation* of V_0 in direction u .

Definition 18. $X \subseteq V$ is an *affine subspace* of V if $X = \emptyset$ or if there exists a subspace V_0 of V and $u \in V$ such that $X = u + V_0$. If V_0 is finite-dimensional, the *dimension* of X is defined by $\dim X = \dim V_0$ and X is called *finite-dimensional*. Otherwise X is called *infinite-dimensional*.

Definition 19. An affine subspace X of \mathbb{R}^n is a (*straight*) *line* if $\dim X = 1$. An affine subspace X of \mathbb{R}^n is a *plane* if $\dim X = 2$.

Proposition 20. Let $w \in W$. Then $F^{-1}(\{w\})$ is an affine subspace of V . If $w \in \operatorname{im} F$ we have

$$F^{-1}(\{w\}) = u + \ker F$$

for every $u \in F^{-1}(\{w\})$.

Proof. If $w \notin \operatorname{im} F$, we have $F^{-1}(\{w\}) = \emptyset$. Let $w \in \operatorname{im} F$ and $u \in F^{-1}(\{w\})$, i.e., $F(u) = w$. For $v \in V$ we have $v = u + (v - u)$ and $F(v) = F(u) + F(v - u)$. Hence we get

$$v \in F^{-1}(\{w\}) \Leftrightarrow u - v \in \ker F. \quad \square$$

Example 21. Consider the situation from Example 17. For every $\lambda \in \mathbb{R}$ we have $\lambda a_1 \in \operatorname{im} F$ and

$$\begin{aligned} F^{-1}(\{\lambda a_1\}) &= \{v \in V : F(v) = \lambda a_1\} = \lambda \cdot (1, 0)^\top + \{\mu \cdot (1, 1)^\top : \mu \in \mathbb{R}\} \\ &= \{\lambda \cdot (1, 0)^\top + \mu \cdot (1, 1)^\top : \mu \in \mathbb{R}\}. \end{aligned}$$

Proposition 22. Let $k, p \in \mathbb{N}_0$, let (v_1, \dots, v_k) be a basis of $\ker F$, and let (w_1, \dots, w_p) be a basis of $\operatorname{im} F$. Moreover, let $v_{k+j} \in F^{-1}(\{w_j\})$ for $j \in \{1, \dots, p\}$. Then (v_1, \dots, v_{k+p}) is a basis of V .

Proof. If $p = 0$, we have $\text{im } F = \{0\}$. This shows $F = 0$ and $\ker F = V$. In the following let $p \in \mathbb{N}$.

We show that the family (v_1, \dots, v_{k+p}) is linearly independent: Let $\sum_{i=1}^{k+p} \lambda_i v_i = 0$ for $\lambda_1, \dots, \lambda_{k+p} \in K$. Applying F shows

$$0 = F(0) = F\left(\sum_{i=1}^{k+p} \lambda_i v_i\right) = \sum_{i=1}^{k+p} \lambda_i F(v_i) = \sum_{i=1}^p \lambda_{k+i} w_i$$

and thus $\lambda_{k+1} = \dots = \lambda_p = 0$ since (w_1, \dots, w_p) is a basis of $\text{im } F$. If $k \in \mathbb{N}$, we have $\sum_{i=1}^k \lambda_i v_i = 0$, which implies $\lambda_1 = \dots = \lambda_k = 0$ since (v_1, \dots, v_k) is a basis of $\ker F$.

Let $v \in V$. We show $v \in \text{span}(\{v_1, \dots, v_{k+p}\})$: There exists $w = \sum_{i=1}^p \lambda_i w_i$ such that $F(v) = w$. Proposition 20 shows

$$v \in F^{-1}(\{w\}) = u + \text{span}(\{v_1, \dots, v_k\})$$

with $u = \sum_{i=1}^p \lambda_i v_{k+i} \in \text{span}(\{v_{k+1}, \dots, v_{k+p}\})$. □

Example 23. Consider the situation of Example 17. Then $((1, 1)^\top, (1, 0)^\top)$ is a basis of \mathbb{R}^2 according to Proposition 22.

Corollary 24 (Rank-Nullity Theorem). Let V be finite-dimensional. Then we have

$$\dim V = \dim \ker F + \dim \text{im } F.$$

Corollary 25. Let V and W be finite-dimensional with $\dim V = \dim W$. Then, F is injective if and only if F is surjective.

Proof. “ \Rightarrow ”: Let F be injective. Corollary 24 shows $\dim W = \dim V = 0 + \dim \text{im } F$ and thus $W = \text{im } F$ by Corollary 2.15.

“ \Leftarrow ”: Let F be surjective. Corollary 24 shows $\dim \ker F = 0$ and thus $\ker F = \{0\}$. □