

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
5	Transformation Matrices of Linear Maps	29
6	Change of Basis	31
6.1	Invertible Matrices	31
6.2	Change of Basis Formula	33
6.3	Row Rank and Column Rank	34
6.4	Elementary Matrices and Elementary Row Operations	34
7	Systems of Linear Equations	36
7.1	Geometry of the Solution Set	36
7.2	Existence and Uniqueness of Solutions	36
7.3	Gaussian Elimination	37
7.4	Representation of Affine Subspaces	37
III	Determinants and Eigenvalues	39

1	Determinants	39
1.1	Characterization of Determinants	39
1.2	Determinant of a Linear Map	42
2	Eigenvalues	43
2.1	Eigenspaces, Eigenvalues, and Eigenvectors	43
2.2	Characteristic Polynomial	44
2.3	Zeros of Polynomial Functions	47
2.4	Multiplicity of Eigenvalues	47
3	Diagonalization	48

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

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

□

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 (Basisauswahlsatz).

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 $\ker F = \{0\}$. □

Notation: Let $L(V, W)$ be the set of all linear maps from V to W .

Proposition 26. For $F \in L(V, W)$ and $G \in L(U, V)$ we have $F \circ G \in L(U, W)$.

Proof. See Exercise 5.4 a). □

5 Transformation Matrices of Linear Maps

In the sequel let V and W K -vector spaces with $\dim V = n \in \mathbb{N}$ and $\dim W = m \in \mathbb{N}$. Moreover, let $\mathcal{A} = (v_1, \dots, v_n)$ and $\mathcal{B} = (w_1, \dots, w_m)$ be bases of V and W , respectively, and let $\mathcal{E} = (e_1, \dots, e_n)$ and $\tilde{\mathcal{E}} = (\tilde{e}_1, \dots, \tilde{e}_m)$ be the standard bases of K^n and K^m , respectively, see Example 2.7.

Remark 1. For every linear map $F: V \rightarrow W$ there exists a unique matrix $A = (a_{i,j})_{i,j} \in K^{m \times n}$ such that

$$\forall 1 \leq j \leq n: F(v_j) = \sum_{i=1}^m a_{i,j} w_i, \quad (1)$$

see Proposition 2.8.(iii). For $1 \leq j \leq n$ we have

$$\begin{pmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{pmatrix} = \Phi_B^{-1}(F(v_j)).$$

Definition 2. The matrix $A \in K^{m \times n}$ from Remark 1 is denoted by $\mathcal{M}_B^A(F)$ and is called *transformation matrix* of $F \in L(V, W)$ w.r.t. the bases \mathcal{A} and \mathcal{B} .

Example 3. Consider the situation from Example 4.8.(ii). Since $e_1 = 1/2 \cdot (v_1 - v_2)$ and $e_2 = 1/2 \cdot (v_1 + v_2)$, we have

$$\begin{aligned} F(e_1) &= 1/2 \cdot (v_1 + v_2) = e_2, \\ F(e_2) &= 1/2 \cdot (v_1 - v_2) = e_1. \end{aligned}$$

Hence we get

$$\mathcal{M}_E^E(F) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

For $\mathcal{A} = (v_1, v_2)$ we have

$$\mathcal{M}_A^A(F) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Proposition 4. For $F \in L(V, W)$ and $v \in V$ we have

$$F(v) = \Phi_B(\mathcal{M}_B^A(F) \cdot \Phi_A^{-1}(v)).$$

Commutative diagram:

$$\begin{array}{ccc} V & \xrightarrow{F} & W \\ \Phi_A^{-1} \downarrow & & \uparrow \Phi_B \\ K^n & \xrightarrow{x \mapsto \mathcal{M}_B^A(F) \cdot x} & K^m \end{array}$$

Proof. The map $G: V \rightarrow W$ given by

$$G(v) = \Phi_B(\mathcal{M}_B^A(F) \cdot \Phi_A^{-1}(v))$$

is linear due to Example 4.2.(i), Remark 4.4.(ii), and Proposition 4.26. It remains to show that $F(v_j) = G(v_j)$ for all $1 \leq j \leq n$, see Lemma 4.6. Indeed we have

$$G(v_j) = \Phi_B(\mathcal{M}_B^A(F) \cdot e_j) = \Phi_B(\Phi_B^{-1}(F(v_j))) = F(v_j). \quad \square$$

Remark 5.

(i) For $A, B \in K^{m \times n}$ we have

$$(\forall x \in K^n: Ax = Bx) \Rightarrow A = B.$$

(ii) For every $F \in L(V, W)$ there exists a unique matrix $A \in K^{m \times n}$ such that $F = \Phi_B \circ \mathcal{F}_A \circ \Phi_A^{-1}$.

In the sequel let U be a finite-dimensional K -vector space with $\dim U = p \in \mathbb{N}$ and basis $\mathcal{C} = (u_1, \dots, u_p)$.

Proposition 6. For $F \in L(V, W)$ and $G \in L(W, U)$ we have

$$\mathcal{M}_{\mathcal{C}}^A(G \circ F) = \mathcal{M}_{\mathcal{C}}^B(G) \cdot \mathcal{M}_{\mathcal{B}}^A(F).$$

Proof. Let $v \in V$. Remark 3.12 and Proposition 4 show

$$\begin{aligned} G \circ F(v) &= G(\Phi_B(\mathcal{M}_{\mathcal{B}}^A(F) \cdot \Phi_A^{-1}(v))) \\ &= \Phi_C\left(\mathcal{M}_{\mathcal{C}}^B(G) \cdot \Phi_B^{-1}\left(\Phi_B(\mathcal{M}_{\mathcal{B}}^A(F) \cdot \Phi_A^{-1}(v))\right)\right) \\ &= \Phi_C\left(\mathcal{M}_{\mathcal{C}}^B(G) \cdot \mathcal{M}_{\mathcal{B}}^A(F) \cdot \Phi_A^{-1}(v)\right). \end{aligned}$$

Apply Remark 5.(ii). □

6 Change of Basis

In the sequel let $n, m \in \mathbb{N}$ and K be a field. Moreover, let $V \neq \{0\}$ and $W \neq \{0\}$ be finite-dimensional K -vector spaces.

6.1 Invertible Matrices

Definition 1. $A \in K^{n \times n}$ is *invertible* if there exists $A' \in K^{n \times n}$ such that $A'A = E_n$.

Notation: $\text{Gl}(n, K) = \{A \in K^{n \times n} : A \text{ invertible}\}.$

Lemma 2. For $A = (a_1, \dots, a_n) \in K^{n \times n}$ the following statements are equivalent:

- (i) A is invertible.
- (ii) $\mathcal{F}_A: K^n \rightarrow K^n$ is bijective.
- (iii) (a_1, \dots, a_n) is a basis of K^n .

Proof. Put $F = \mathcal{F}_A$.

“(i) \Rightarrow (ii)” : Let $x \in K^n$ such that $F(x) = 0$, and let $A' \in K^{n \times n}$ such that $A'A = E_n$. Then we have

$$x = E_n x = A'A x = A' 0 = 0.$$

Hence F is injective and thus bijective due to Corollary 4.25.

“(ii) \Rightarrow (i)” : Remark 4.4.(ii) shows $F^{-1} \in L(K^n, K^n)$. According to Exercise 5.3 a) there exists $A' \in K^{n \times n}$ such that

$$\forall x \in K^n: F^{-1}(x) = A'x.$$

This yields

$$E_n x = x = F^{-1}(F(x)) = A'Ax$$

for all $x \in K^n$. Remark 5.5.(i) shows $A'A = E_n$.

“(ii) \Leftrightarrow (iii)” : See Proposition 4.7. □

Proposition 3. Let $F: V \rightarrow W$ be linear. The following statements are equivalent:

- (i) F is a linear isomorphism.
- (ii) For all bases \mathcal{A} and \mathcal{B} of V and W , respectively, the matrix $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)$ is invertible.
- (iii) There exist bases \mathcal{A} and \mathcal{B} of V and W , respectively, such that the matrix $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)$ is invertible.

Proof. “(i) \Rightarrow (ii)” : Due to Proposition 5.4 we have

$$\mathcal{F}_{\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)} = \Phi_{\mathcal{B}}^{-1} \circ F \circ \Phi_{\mathcal{A}}.$$

In particular, $\mathcal{F}_{\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)}$ is bijective since it is the composition of bijective maps. Apply Lemma 2.

“(ii) \Rightarrow (iii)” : Obvious.

“(iii) \Rightarrow (i)” : Due to Proposition 5.4 we have

$$F = \Phi_{\mathcal{B}} \circ \mathcal{F}_{\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)} \circ \Phi_{\mathcal{A}}^{-1},$$

and $\mathcal{F}_{\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)}$ is bijective according to Lemma 2. Hence F is bijective since it is the composition of bijective maps. □

Proposition 4. $(\text{Gl}(n, K), \cdot)$ is a group with neutral element E_n .

Notation: A^{-1} is the *inverse matrix* of $A \in \text{Gl}(n, K)$ w.r.t. \cdot .

Remark 5. For $A, B \in \text{Gl}(n, K)$ we have $A \cdot A^{-1} = E_n$, $(A^{-1})^{-1} = A$, and $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$, see Lemma I.1.7 and Lemma I.1.12.

Lemma 6. Let $F: V \rightarrow W$ be a linear isomorphism, and let \mathcal{A} and \mathcal{B} be bases of V and W , respectively. Then we have

$$\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(F^{-1}) = (\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F))^{-1}.$$

Proof. Proposition 5.6 shows $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(F^{-1}) \cdot \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F) = \mathcal{M}_{\mathcal{A}}^{\mathcal{A}}(\text{id}_V) = E_n$. □

6.2 Change of Basis Formula

Definition 7. For bases \mathcal{A} and \mathcal{B} of V the transformation matrix $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(\text{id}_V)$ is called *change-of-basis matrix* from the basis \mathcal{A} to the basis \mathcal{B} .

Notation: $\mathcal{T}_{\mathcal{B}}^{\mathcal{A}} = \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(\text{id}_V)$.

Remark 8. Let $\mathcal{A} = (v_1, \dots, v_n)$ and $\mathcal{B} = (u_1, \dots, u_n)$ be bases of $V = K^n$. Put $A = (v_1, \dots, v_n) \in K^{n \times n}$ and $B = (u_1, \dots, u_n) \in K^{n \times n}$. For $x \in K^n$ we then obtain

$$\Phi_{\mathcal{A}}(x) = Ax, \quad \Phi_{\mathcal{B}}(x) = Bx.$$

Lemma 2 ensures $A, B \in \text{Gl}(n, K)$ such that $\Phi_{\mathcal{A}}^{-1}(y) = A^{-1}y$ and $\Phi_{\mathcal{B}}^{-1}(y) = B^{-1}y$ for $y \in K^n$. Hence we have

$$\mathcal{T}_{\mathcal{B}}^{\mathcal{A}} = B^{-1}A, \quad \mathcal{T}_{\mathcal{A}}^{\mathcal{B}} = A^{-1}B.$$

Example 9. For the basis $\mathcal{A} = ((1, 1)^{\top}, (-1, 1)^{\top})$ and the standard basis $\mathcal{B} = \mathcal{E} = (e_1, e_2)$ of $V = \mathbb{R}^2$ we have $B = B^{-1} = E_2$ such that

$$\mathcal{T}_{\mathcal{E}}^{\mathcal{A}} = A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Since $e_1 = 1/2 \cdot (v_1 - v_2)$ and $e_2 = 1/2 \cdot (v_1 + v_2)$, we obtain

$$\mathcal{T}_{\mathcal{A}}^{\mathcal{E}} = \frac{1}{2} \cdot \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}^{-1}.$$

Cf. Example 5.3.

Proposition 10 (Change-of-basis formula). Let $\mathcal{A}, \mathcal{A}'$ be bases of V and let $\mathcal{B}, \mathcal{B}'$ be bases of W . For $F \in L(V, W)$ we have

$$\mathcal{M}_{\mathcal{B}'}^{\mathcal{A}'}(F) = \mathcal{T}_{\mathcal{B}'}^{\mathcal{B}} \cdot \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F) \cdot (\mathcal{T}_{\mathcal{A}'}^{\mathcal{A}})^{-1}.$$

Proof. According to Proposition 5.6 we have $\mathcal{M}_{\mathcal{B}'}^{\mathcal{A}'}(F) = \mathcal{M}_{\mathcal{B}'}^{\mathcal{B}}(\text{id}_W) \cdot \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F) \cdot \mathcal{M}_{\mathcal{A}}^{\mathcal{A}'}(\text{id}_V)$. Lemma 6 shows $\mathcal{M}_{\mathcal{A}}^{\mathcal{A}'}(\text{id}_V) = (\mathcal{M}_{\mathcal{A}'}^{\mathcal{A}}(\text{id}_V))^{-1}$. \square

Example 11. Consider the situation from Example 5.3 and Example 9. Put $\mathcal{A} = \mathcal{B}$ and $\mathcal{A}' = \mathcal{B}' = \mathcal{E}$. Then we have

$$\begin{aligned} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} &= \mathcal{M}_{\mathcal{E}}^{\mathcal{E}}(F) = \mathcal{T}_{\mathcal{E}}^{\mathcal{A}} \cdot \mathcal{M}_{\mathcal{A}}^{\mathcal{A}}(F) \cdot (\mathcal{T}_{\mathcal{E}}^{\mathcal{A}})^{-1} \\ &= \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}. \end{aligned}$$

Definition 12. Two matrices $A, B \in K^{m \times n}$ are *equivalent* if there exist $S \in \text{Gl}(m, K)$ and $T \in \text{Gl}(n, K)$ such that $B = S \cdot A \cdot T^{-1}$.

Remark 13. $A, B \in K^{m \times n}$ are equivalent if and only if there exist bases $\mathcal{A}, \mathcal{A}'$ of K^n , bases $\mathcal{B}, \mathcal{B}'$ of K^m , and a linear map $F: K^n \rightarrow K^m$ with $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F) = A$ and $\mathcal{M}_{\mathcal{B}'}^{\mathcal{A}'}(F) = B$.

6.3 Row Rank and Column Rank

Definition 14. For $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \in K^{m \times n}$ the matrix

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

is called the *transposed matrix* of A .

Lemma 15.

- (i) For $A \in K^{m \times n}$ we have $(A^\top)^\top = A$.
- (ii) For $A \in K^{m \times n}$ and $B \in K^{n \times p}$ we have $(AB)^\top = B^\top A^\top$.
- (iii) For $A \in \text{Gl}(n, K)$ we have $A^\top \in \text{Gl}(n, K)$ and $(A^\top)^{-1} = (A^{-1})^\top$.

Proof. ad (i): Obvious.

ad (ii): This is an immediate consequence of Definition 3.11 and Definition 14.

ad (iii): Using (ii) we get $(A^{-1})^\top A^\top = (AA^{-1})^\top = E_n^\top = E_n$. □

Proposition 16. For $A \in K^{m \times n}$ we have⁴

$$\text{rank } A = \text{rank } A^\top.$$

Moreover, $A \in K^{n \times n}$ is invertible if and only if $\text{rank } A = n$.

Corollary 17. Let \mathcal{A} and \mathcal{B} be bases of V and W , respectively. For $F \in L(V, W)$ we have

$$\text{rank } F = \text{rank } \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F).$$

Corollary 18. Let $A, B \in K^{m \times n}$. The following statements are equivalent:

- (i) A and B are equivalent.
- (ii) $\text{rank } A = \text{rank } B$.

6.4 Elementary Matrices and Elementary Row Operations

Definition 19. For $i, j \in \{1, \dots, m\}$ with $i \neq j$ and $\lambda \in K \setminus \{0\}$ the *elementary matrices*

$$S_i(\lambda) = (s_{k,\ell})_{k,\ell} \in K^{m \times m}, \quad Q_i^j(\lambda) = (q_{k,\ell}) \in K^{m \times m}, \quad P_i^j = (p_{k,\ell}) \in K^{m \times m}$$

⁴ $\text{rank } A^\top$ is called column rank of A .

are defined by

$$s_{k,\ell} = \begin{cases} \lambda, & \text{if } k = \ell = i, \\ 1, & \text{if } k = \ell \neq i, \\ 0, & \text{else,} \end{cases}$$

$$q_{k,\ell} = \begin{cases} \lambda, & \text{if } k = i \text{ and } \ell = j, \\ 1, & \text{if } k = \ell, \\ 0, & \text{else,} \end{cases}$$

$$p_{k,\ell} = \begin{cases} 1, & \text{if } k = \ell \in \{1, \dots, m\} \setminus \{i, j\} \text{ or } \{k, \ell\} = \{i, j\}, \\ 0, & \text{else.} \end{cases}$$

Remark 20. Let $A \in K^{m \times n}$. Left multiplication of A by an elementary matrix represents an elementary row operation.

- (i) $S_i(\lambda) \cdot A$ results from A by multiplying the i -th row by λ (type I),
- (ii) $Q_i^j(\lambda) \cdot A$ results from A by adding λ times the j -th row to the i -th row (type II),
- (iii) $P_i^j \cdot A$ results from A by switching the i -th and the j -th row (type III).

See **Exercise 4.3.a**).

Moreover, we have $S_i(\lambda), Q_i^j(\lambda), P_i^j \in \text{Gl}(m, K)$ with

$$(S_i(\lambda))^{-1} = S_i(1/\lambda), \quad (Q_i^j(\lambda))^{-1} = Q_i^j(-\lambda), \quad (P_i^j)^{-1} = P_i^j.$$

Proposition 21. Every invertible matrix is the product of finitely many elementary matrices.

Proof. Let $A \in K^{n \times n}$. According to Proposition 3.8 and Remark 20 there exist $k \in \mathbb{N}$ and elementary matrices $C_1, \dots, C_k \in K^{n \times n}$ such that

$$\tilde{A} = C_k \cdot \dots \cdot C_1 \cdot A$$

is in row echolon form. Lemma 3.7 shows $\text{rank } \tilde{A} = \text{rank } A$ such that $A \in \text{Gl}(n, K)$ is equivalent to $\text{rank } \tilde{A} = n$, see Proposition 16. In case of $\text{rank } A = n$ there exist $\ell \in \mathbb{N}$ and elementary matrices $C_{k+1}, \dots, C_{k+\ell} \in K^{n \times n}$ such that

$$C_{k+\ell} \cdot \dots \cdot C_{k+1} \cdot C_k \cdot \dots \cdot C_1 \cdot A = E_n$$

and hence $A^{-1} = C_{k+\ell} \cdot \dots \cdot C_1$ as well as $A = C_1^{-1} \cdot \dots \cdot C_{k+\ell}^{-1}$. \square

Remark 22. The proof of Proposition 21 yields an algorithm to check whether $A \in K^{n \times n}$ is invertible and, if so, to the computation of A^{-1} . Let $A = (a_1, \dots, a_n)$ and $E_n = (e_1, \dots, e_n)$. Put

$$(A | E_n) = (a_1, \dots, a_n, e_1, \dots, e_n) \in K^{n \times 2n}.$$

Using the notation from the proof of Proposition 21 we have in case of $\text{rank } \tilde{A} = n$ that

$$C_{k+\ell} \cdot \dots \cdot C_1 \cdot (A | E_n) = (E_n | A^{-1}),$$

see **Exercise 4.3.b**).

7 Systems of Linear Equations

In the sequel let K be a field, $m, n \in \mathbb{N}$, $A \in K^{m \times n}$, and $b \in K^m$. Moreover, let $r = \text{rank } A$ and $F = \mathcal{F}_A$.

Definition 1. The set

$$\mathfrak{L}(A, b) = \{x \in K^n : Ax = b\}$$

is called *solution set* of the *system of linear equations* $Ax = b$. Such a system of linear equations is *homogeneous* if $b = 0$, and *inhomogeneous* if $b \neq 0$.

Notation: The matrix A is called *coefficient matrix*⁵.

7.1 Geometry of the Solution Set

Proposition 2.

- (i) $\mathfrak{L}(A, 0)$ is a subspace of K^n with dimension $n - r$.
- (ii) $\mathfrak{L}(A, b)$ is either empty or an affine subspace of K^n with dimension $n - r$. If $x \in \mathfrak{L}(A, b)$, then

$$\mathfrak{L}(A, b) = x + \mathfrak{L}(A, 0).$$

Proof. Remark 4.15.(i) shows that $\ker F = \mathfrak{L}(A, 0)$ is a subspace of K^n . Moreover, Corollary 6.17 shows $\text{rank } A = \text{rank } F$.

ad (i): Corollary 4.24 (Rank-Nullity Theorem) yields $\dim \mathfrak{L}(A, 0) = n - r$.

ad (ii): We have $\mathfrak{L}(A, b) = F^{-1}(\{b\})$. Apply Proposition 4.20. □

7.2 Existence and Uniqueness of Solutions

Notation: Augmented matrix $(A, b) = (a_1, \dots, a_n, b) \in K^{m \times (n+1)}$ if $A = (a_1, \dots, a_n)$.

Proposition 3. We have $\mathfrak{L}(A, b) \neq \emptyset$ if and only if

$$\text{rank}(A, b) = \text{rank } A. \tag{1}$$

Proof. Due to Corollary 2.15 and Proposition 6.16 we have (1) if and only if

$$\text{span}(\{a_1, \dots, a_n, b\}) = \text{span}(\{a_1, \dots, a_n\}).$$

The latter is equivalent to $b \in \text{span}(\{a_1, \dots, a_n\})$. □

Proposition 4.

- (i) $\text{rank } A = m \Leftrightarrow \forall b \in K^m : \mathfrak{L}(A, b) \neq \emptyset$.
- (ii) $\text{rank } A = n \Leftrightarrow \forall b \in K^m : |\mathfrak{L}(A, b)| \leq 1$.

Proof. F is surjective if and only if $\text{rank } A = m$, and F is injective if and only if $\text{rank } A = n$. □

⁵Common notation: $\ker A = \mathfrak{L}(A, 0)$.

7.3 Gaussian Elimination

Remark 5. For an application of Proposition 3 one can transform (A, b) into row echelon form using elementary row operations, see Lemma 3.7 and Proposition 3.8.

Lemma 6. Let (\tilde{A}, \tilde{b}) result from finitely many elementary row operations applied to (A, b) . Then we have $\mathfrak{L}(\tilde{A}, \tilde{b}) = \mathfrak{L}(A, b)$.

Proof. Apply Remark 6.20. □

Proposition 7 (Back substitution). Let (A, b) be in row echelon form. If $\text{rank}(A, b) = \text{rank } A = n$, we have $\mathfrak{L}(A, b) = \{x\}$ for $x \in K^n$ given by

$$x_n = b_n / a_{n,n}$$

and

$$x_i = \left(b_i - \sum_{j=i+1}^n a_{i,j} \cdot x_j \right) / a_{i,i}$$

for $i = n - 1, \dots, 1$.

Example 8. Let $c, d \in \mathbb{R}$ and let $A \in \mathbb{R}^{3 \times 3}$ and $b \in \mathbb{R}^3$ be given by

$$A = \begin{pmatrix} 2 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} -1 \\ 0 \\ d \end{pmatrix}.$$

Then we have

$$\text{rank } A = 3 \Leftrightarrow c \neq 0.$$

If $c \neq 0$, Proposition 4 shows $|\mathfrak{L}(A, b)| = 1$ for all $b \in \mathbb{R}^3$, and we obtain

$$\mathfrak{L}(A, b) = \{(-1/2, 0, d/c)^\top\},$$

see Proposition 7. If $c = 0$, Proposition 3 shows that $\mathfrak{L}(A, b) \neq \emptyset$ if and only if $d = 0$. If $c = d = 0$, we have $\text{rank } A = \text{rank}(A, b) = 2$, $\mathfrak{L}(A, 0) = \text{span}(\{e_3\})$, and

$$\mathfrak{L}(A, b) = (-1/2, 0, 0)^\top + \text{span}(\{e_3\})$$

see Proposition 2.

7.4 Representation of Affine Subspaces

Proposition 9. Let $n \in \mathbb{N}$ and $p \in \mathbb{N}_0$ with $p \leq n - 1$, and let $X \subseteq K^n$. Then the following statements are equivalent:

- (i) X is a p -dimensional affine subspace of K^n .
- (ii) There exist $A \in K^{(n-p) \times n}$ and $b \in K^{n-p}$ such that

$$\text{rank } A = \text{rank}(A, b) = n - p \wedge X = \mathfrak{L}(A, b).$$

Proof. “(ii) \Rightarrow (i)”: Proposition 3 ensures $\mathfrak{L}(A, b) \neq \emptyset$ and $\mathfrak{L}(A, b)$ is a p -dimensional affine subspace of K^n according to Proposition 2.(ii).

“(i) \Rightarrow (ii)”: If $p = 0$, choose $A = E_n$ and $b = x$ for $X = \{x\}$.

Let $p \geq 1$. At first, we consider the case of a p -dimensional subspace $U \subseteq K^n$ and choose a basis (v_1, \dots, v_p) of U . Let $A \in K^{r \times n}$ with $r = n - p \in \{1, \dots, n - 1\}$. Then, $\mathfrak{L}(A, 0) = U$ is equivalent to

$$\forall 1 \leq i \leq p: Av_i = 0 \quad (2)$$

and

$$p = n - \text{rank } A, \quad (3)$$

see Corollary 2.15 and Proposition 2.(i). Clearly, (3) is equivalent to $\text{rank } A = r$. Put $B = (v_1, \dots, v_p) \in K^{n \times p}$. Then, (2) is equivalent to

$$B^\top A^\top = 0,$$

and we have $\dim \mathfrak{L}(B^\top, 0) = r$. In sum, conditions (2) and (3) hold if and only if the columns of A^\top form a basis of $\mathfrak{L}(B^\top, 0)$.

Consider the general case of $X = x + U$ with $x \in K^n$ and U as above. For A with (2) and (3) we define b by $Ax = b$ to get $\mathfrak{L}(A, b) = X$. \square

Example 10. Consider the situation from Proposition 9 with $K = \mathbb{R}$, $n = 5$ and

$$U = \text{span}(\{(1, 0, 1, 1, 1)^\top, (1, 2, 0, 2, 0)^\top, (3, 1, 1, 1, 1)^\top\}).$$

We obtain

$$B^\top = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 2 & 0 \\ 3 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

By using elementary row operations we can transform B^\top into

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 2 & -1 & 1 & -1 \\ 0 & 0 & -3 & -5 & -3 \end{pmatrix},$$

which yields $\dim U = \text{rank } B^\top = 3$ and a basis $((0, 0, -1, 0, 1)^\top, (2, -4, -5, 3, 0)^\top)$ of $\mathfrak{L}(B^\top, 0)$. Hence we obtain

$$A = \begin{pmatrix} 0 & 0 & -1 & 0 & 1 \\ 2 & -4 & -5 & 3 & 0 \end{pmatrix}$$

and $b = 0$ for the case $X = U$.

Chapter III

Determinants and Eigenvalues

Given a finite-dimensional vector space V and a linear map $F: V \rightarrow V$ we try to find bases \mathcal{A} such that the transformation matrix $\mathcal{M}_{\mathcal{A}}^{\mathcal{A}}(F)$ has a simple form. To this end we first discuss the concept of determinants.

1 Determinants

1.1 Characterization of Determinants

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

Definition 1 (Weierstrass axioms). A map¹ $\det: K^{n \times n} \rightarrow K$ is called *determinant* if the following conditions hold:

- (i) For all $1 \leq i \leq n$, $a_1, \dots, a_n, b_i \in K^{1 \times n}$, and $\lambda \in K$ we have

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_i + b_i \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_i \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ b_i \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix}$$

and

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ \lambda \cdot a_i \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ a_i \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix}.$$

¹Notation: $\det A$ or $\det(A)$.

(ii) For all $A \in K^{n \times n}$ with two identical rows we have $\det A = 0$.

(iii) We have $\det E_n = 1$.

Proposition 2. Every determinant $\det: K^{n \times n} \rightarrow K$ has the following properties:

(i) For all $A \in K^{n \times n}$ and $\lambda \in K$ we have

$$\det(\lambda \cdot A) = \lambda^n \cdot \det A.$$

(ii) For all $A \in K^{n \times n}$ with a zero row we have $\det A = 0$.

(iii) Let $A \in K^{n \times n}$ and $\lambda \in K$. If $B \in K^{n \times n}$ results from an elementary row operation of type II (adding λ times a row to another row) applied to A , then we have $\det B = \det A$.

(iv) If $B \in K^{n \times n}$ results from an elementary row operation of type III (switching two rows) applied to A , then we have $\det B = -\det A$.

(v) Let $A \in K^{n \times n}$ with² $a_{i,j} = 0$ for all $1 \leq j < i \leq n$. Then we have

$$\det A = \prod_{i=1}^n a_{i,i}.$$

(vi) For $A \in K^{n \times n}$ we have $\det A = 0$ if and only if $\text{rank } A < n$.

(vii) For $A, B \in K^{n \times n}$ we have

$$\det(A \cdot B) = \det A \cdot \det B.$$

(viii) For $A \in \text{Gl}(n, K)$ we have

$$\det(A^{-1}) = (\det A)^{-1}.$$

Proof. ad (i), (ii): This follows immediately from Definition 1.(i).

ad (iii): For $i \neq j$ we have

$$\det \begin{pmatrix} \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \lambda \cdot \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix}.$$

ad (iv): For $i \neq j$ we have

$$\det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_i + a_j \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} = 0.$$

²Such matrices are called *upper triangular*.

ad (v): If $\prod_{i=1}^n a_{i,i} \neq 0$, elementary row operations of type II and part (iii) show

$$\det A = \det \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix} = \prod_{i=1}^n a_{i,i} \cdot \det E_n = \prod_{i=1}^n a_{i,i}.$$

If $\prod_{i=1}^n a_{i,i} = 0$, elementary row operations of type II lead to matrix with a zero row such that $\det A = 0$ according to part (ii).

ad (vi): Use elementary row operations of type II and type III as well as parts (iii), (iv), and (v).

ad (vii): If $\text{rank } A < n$, we have $\text{rank}(A \cdot B) < n$ such that $\det A \cdot \det B = 0$ and $\det(A \cdot B) = 0$ due to part (vi).

Let $\text{rank } A = n$. According to Proposition II.6.21 there exist elementary matrices $C_1, \dots, C_k \in K^{n \times n}$ such that $A = C_k \cdots C_1$. Definition 1.(i) shows for elementary matrices C_I of type I (multiplying a row by $\lambda \in K \setminus \{0\}$)

$$\det(C_I \cdot B) = \lambda \cdot \det B.$$

For elementary matrices C_{II} and C_{III} of types II and III, respectively, we have

$$\det(C_{II} \cdot B) = \det B,$$

see (iii), and

$$\det(C_{III} \cdot B) = -\det B$$

see (iv). By Definition 1.(iii) we have $\det C_I = \lambda$, $\det C_{II} = 1$, and $\det C_{III} = -1$. Hence we get $\det(C \cdot B) = \det C \cdot \det B$ for all elementary matrices C . The general case follows by induction.

ad (viii): We have

$$\det(A^{-1}) \cdot \det A = \det(A^{-1} \cdot A) = \det E_n = 1,$$

according to (vii) and Definition 1.(iii). □

Remark 3. Determinants can be calculated by using elementary row operations, see Definition 1 and Proposition 2

Example 4.

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

Proposition 5. There exists a unique map $\det: K^{n \times n} \rightarrow K$.

Proof sketch. Uniqueness follows from the proof of Proposition 2. Existence can be shown, e.g., via the Leibniz formula. □

Notation: $\det A$ is called *the determinant of $A \in K^{n \times n}$* .

Example 6. For $a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2} \in K$ we have

$$\det \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = a_{1,1} \cdot a_{2,2} - a_{1,2} \cdot a_{2,1}.$$

For

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \in K^{3 \times 3}$$

we have

$$\begin{aligned} \det A &= a_{1,1} \cdot a_{2,2} \cdot a_{3,3} + a_{1,2} \cdot a_{2,3} \cdot a_{3,1} + a_{1,3} \cdot a_{2,1} \cdot a_{3,2} \\ &\quad - a_{1,3} \cdot a_{2,2} \cdot a_{3,1} - a_{1,2} \cdot a_{2,1} \cdot a_{3,3} - a_{1,1} \cdot a_{2,3} \cdot a_{3,2}. \end{aligned}$$

Lemma 7. For $a \in K$, $b \in K^{1 \times n}$, and $A \in K^{n \times n}$ we have

$$\det \begin{pmatrix} a & b \\ 0 & A \end{pmatrix} = a \cdot \det(A).$$

Proof. Use elementary row operations and Proposition 2. □

Proposition 8. For $A \in K^{n \times n}$ we have

$$\det A = \det(A^\top).$$

Remark 9. According to Proposition 8 we may use “elementary column operations” to calculate the determinant.

1.2 Determinant of a Linear Map

In the sequel let V be a finite-dimensional K -vector space with $n = \dim V \in \mathbb{N}$ and let $F: V \rightarrow V$ be linear.

Notation: $\mathcal{M}_{\mathcal{A}}(F)$ instead of $\mathcal{M}_{\mathcal{A}}^A(F)$ for bases \mathcal{A} of V .

Remark 10. For bases \mathcal{A} and \mathcal{B} of V we have

$$\mathcal{M}_{\mathcal{B}}(F) = \mathcal{T}_{\mathcal{B}}^{\mathcal{A}} \cdot \mathcal{M}_{\mathcal{A}}(F) \cdot (\mathcal{T}_{\mathcal{B}}^{\mathcal{A}})^{-1},$$

see Proposition II.6.10 (change-of-basis formula).

Definition 11. $A, B \in K^{n \times n}$ are *similar* if there exists $S \in \text{Gl}(n, K)$ such that

$$B = S \cdot A \cdot S^{-1}.$$

Remark 12. Two matrices $A, B \in K^{n \times n}$ are similar if and only if there exist bases \mathcal{A} and \mathcal{B} of K^n and a linear map $F: K^n \rightarrow K^n$ with $\mathcal{M}_{\mathcal{A}}(F) = A$ and $\mathcal{M}_{\mathcal{B}}(F) = B$, cf. Remark II.6.13.

Proof of “ \Rightarrow ”: Choose the standard basis \mathcal{E} of K^n and \mathcal{B} given by the columns of S^{-1} . Then we have $\mathcal{M}_{\mathcal{E}}(F) = A$ und $\mathcal{T}_{\mathcal{E}}^{\mathcal{B}} = S^{-1}$. Apply Remark 10.

Lemma 13. For similar matrices $A, B \in K^{n \times n}$ we have $\det A = \det B$.

Proof. According to Proposition 2.(vii) and (viii) we have

$$\det(S \cdot A \cdot S^{-1}) = \det S \cdot \det A \cdot \det(S^{-1}) = \det S \cdot \det A \cdot (\det S)^{-1} = \det A$$

for every $S \in \text{Gl}(n, K)$. □

According to Lemma 13 the following definition is well-defined.

Definition 14. Let \mathcal{A} be a basis of V . The *determinant* of a linear map F is defined by $\det \mathcal{M}_{\mathcal{A}}(F)$. *Notation:* $\det F$.

Corollary 15. The following statements are equivalent:

- (i) F is injective,
- (ii) F is surjective,
- (iii) $\det F \neq 0$.

Proof. “(i) \Leftrightarrow (ii)” : See Corollary II.4.25.

“(i)/(ii) \Leftrightarrow (iii)” : Proposition II.6.3, Proposition II.6.16, and Proposition 2.(vi) show that F is a linear isomorphism if and only if $\det F \neq 0$. □

2 Eigenvalues

In the sequel let V be a K -vector space, let $F: V \rightarrow V$ be linear, and let $\lambda \in K$.

2.1 Eigenspaces, Eigenvalues, and Eigenvectors

Definition 1.

$$\text{Eig}(F, \lambda) = \{v \in V : F(v) = \lambda v\}$$

is called *eigenspace* of F w.r.t. λ .

Remark 2.

- (i) $\text{Eig}(F, \lambda)$ is a subspace of V .
- (ii) For $\lambda, \mu \in K$ with $\lambda \neq \mu$ we have $\text{Eig}(F, \lambda) \cap \text{Eig}(F, \mu) = \{0\}$.

Definition 3.

- (i) λ is an *eigenvalue* of F if $\text{Eig}(F, \lambda) \neq \{0\}$.

Notation: $\sigma(F)$ denotes the set of all eigenvalues of F .

- (ii) For $\lambda \in \sigma(F)$ the elements of $\text{Eig}(F, \lambda) \setminus \{0\}$ are called *eigenvectors* of F corresponding to λ .

Definition 4. For $A \in K^{n \times n}$ with $n \in \mathbb{N}$ the set

$$\text{Eig}(A, \lambda) = \text{Eig}(\mathcal{F}_A, \lambda)$$

is called *eigenspace* of A w.r.t. λ , and λ is an *eigenvalue* of A if $\text{Eig}(A, \lambda) \neq \{0\}$.

Notation: $\sigma(A)$ denotes the set of all eigenvalues of A .

For $\lambda \in \sigma(A)$ the elements of $\text{Eig}(A, \lambda) \setminus \{0\}$ are called *eigenvectors* of A corresponding to λ .

Example 5. Let $K = \mathbb{R}$, $V = \mathbb{R}^2$, and

$$A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

with $\alpha \in [0, 2\pi[$. Recall that \mathcal{F}_A represents a rotation of α , see Example II.4.8.(i). If $\alpha = 0$, we have $\sigma(A) = \{1\}$ and $\text{Eig}(A, 1) = \mathbb{R}^2$. If $\alpha = \pi$, we have $\sigma(A) = \{-1\}$ and $\text{Eig}(A, -1) = \mathbb{R}^2$. If $\alpha \in [0, 2\pi[\setminus \{0, \pi\}$, we have $\sigma(A) = \emptyset$, see Example 12 below.

Lemma 6. Let $k \in \mathbb{N}$, and let $v_1, \dots, v_k \in V$ be eigenvectors of F corresponding to pairwise distinct eigenvalues. Then (v_1, \dots, v_k) is linearly independent.

Proof. Induction: For $k = 1$ the statement is obviously true. Let $k \geq 2$, and let $F(v_\ell) = \lambda_\ell v_\ell$ for $\ell = 1, \dots, k$ with pairwise distinct eigenvalues $\lambda_1, \dots, \lambda_k \in K$. Moreover, let $\sum_{\ell=1}^k \mu_\ell v_\ell = 0$ for $\mu_1, \dots, \mu_k \in K$. Applying F and multiplying this equation by λ_k yields

$$\sum_{\ell=1}^{k-1} \lambda_\ell \mu_\ell v_\ell + \lambda_k \mu_k v_k = 0, \quad \sum_{\ell=1}^{k-1} \lambda_k \mu_\ell v_\ell + \lambda_k \mu_k v_k = 0,$$

respectively, and hence

$$\sum_{\ell=1}^{k-1} (\lambda_\ell - \lambda_k) \mu_\ell v_\ell = 0.$$

Since $\lambda_\ell - \lambda_k \neq 0$ for all $\ell = 1, \dots, k-1$, the induction hypothesis shows $\mu_\ell = 0$ for $\ell = 1, \dots, k-1$, and hence $\mu_k = 0$. \square

2.2 Characteristic Polynomial

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

We define $P_A: K \rightarrow K$ for $A \in K^{n \times n}$ by

$$P_A(\lambda) = \det(A - \lambda E_n).$$

Proposition 7.

(i) For $A \in K^{n \times n}$ we have

$$\sigma(A) = \{\lambda \in K : P_A(\lambda) = 0\}.$$

(ii) For similar matrices $A, B \in K^{n \times n}$ we have $P_A = P_B$.

Proof. ad (i): For $x \in K^n$ we have

$$Ax = \lambda x \Leftrightarrow (A - \lambda E_n)x = 0.$$

This shows $\lambda \in \sigma(A)$ if and only if $\text{rank}(A - \lambda E_n) < n$, see Proposition II.7.2.(i). Moreover, Proposition 1.2.(vi) shows

$$\text{rank}(A - \lambda E_n) < n \Leftrightarrow P_A(\lambda) = \det(A - \lambda E_n) = 0.$$

ad (ii): Let $B = SAS^{-1}$ for $S \in \text{Gl}(n, K)$. Then we have

$$B - \lambda E_n = SAS^{-1} - S\lambda E_n S^{-1} = S \cdot (A - \lambda E_n) \cdot S^{-1},$$

see Lemma II.3.15. Thus $B - \lambda E_n$ and $A - \lambda E_n$ are also similar, and similar matrices have the same determinant, see Lemma 1.13. \square

In the sequel let $K = \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Proposition 8. For all $A = (a_{i,j})_{i,j} \in \mathbb{K}^{n \times n}$ there exists $Q_A \in \Pi_{n-2}$ such that³

$$P_A(\lambda) = (-1)^n \cdot \lambda^n + (-1)^{n-1} \left(\sum_{i=1}^n a_{i,i} \right) \cdot \lambda^{n-1} + Q_A(\lambda)$$

for all $\lambda \in \mathbb{K}$. In particular, $P_A: \mathbb{K} \rightarrow \mathbb{K}$ is a polynomial function of degree n .

Corollary 9. For $A \in \mathbb{K}^{n \times n}$ we have $|\sigma(A)| \leq n$.

Proof. Apply Proposition 7 and Proposition 8, and note that a (non-zero) polynomial function of degree n has at most n roots. \square

Definition 10. P_A is called *characteristic polynomial* of $A \in \mathbb{K}^{n \times n}$.

According to Proposition 7.(ii) the following definition is well-defined.

Definition 11. Let $V \neq \{0\}$ be finite-dimensional with basis \mathcal{A} , and let $F: V \rightarrow V$ be linear. The *characteristic polynomial* of F is defined by $P_{\mathcal{M}_{\mathcal{A}}(F)}$. *Notation:* P_F .

Example 12. For A according to Example 5 with $\alpha \in [0, 2\pi[$ and $\lambda \in \mathbb{R}$ we have

$$P_A(\lambda) = (\cos(\alpha) - \lambda)^2 + \sin^2(\alpha).$$

Proposition 7.(i) shows

$$\sigma(A) = \begin{cases} \{1\} & \text{if } \alpha = 0, \\ \{-1\} & \text{if } \alpha = \pi, \\ \emptyset & \text{else.} \end{cases}$$

³ $\sum_{i=1}^n a_{i,i}$ is called *trace* of A .

Example 13. For

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

and $\lambda \in \mathbb{K}$ we obtain using Lemma 1.7

$$\begin{aligned} P_A(\lambda) &= \det \begin{pmatrix} -\lambda & -1 & 1 \\ -3 & -2-\lambda & 3 \\ -2 & -2 & 3-\lambda \end{pmatrix} = \det \begin{pmatrix} 2 & 2 & \lambda-3 \\ -3 & -2-\lambda & 3 \\ -\lambda & -1 & 1 \end{pmatrix} \\ &= \det \begin{pmatrix} 2 & 2 & \lambda-3 \\ 0 & 1-\lambda & 3/2(\lambda-1) \\ 0 & \lambda-1 & 1+\lambda/2(\lambda-3) \end{pmatrix} \\ &= 2 \cdot (1-\lambda) \cdot \det \begin{pmatrix} 1 & -3/2 \\ \lambda-1 & 1+\lambda/2(\lambda-3) \end{pmatrix} \\ &= (1-\lambda) \cdot (\lambda^2 - 1) \\ &= -(\lambda-1)^2 \cdot (\lambda+1), \end{aligned}$$

and hence $\sigma(A) = \{1, -1\}$.

Let $x \in \mathbb{K}^3$. For $\lambda_1 = 1$ elementary row operations show that $(A - \lambda_1 E_3)x = 0$ is equivalent to

$$\begin{pmatrix} -1 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x = 0.$$

We obtain

$$\text{Eig}(A, \lambda_1) = \text{span}(\{(1, 0, 1)^\top, (-1, 1, 0)^\top\}).$$

For $\lambda_2 = -1$ elementary row operations show that $(A - \lambda_2 E_3)x = 0$ is equivalent to

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

We obtain

$$\text{Eig}(A, \lambda_2) = \text{span}(\{(1, 3, 2)^\top\}).$$

Example 14. For

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

and $\lambda \in \mathbb{K}$ we obtain

$$P_A(\lambda) = \det \begin{pmatrix} -1-\lambda & 0 & 0 \\ 0 & -\lambda & -1 \\ 0 & 1 & -\lambda \end{pmatrix} = -(\lambda+1) \cdot (\lambda^2 + 1).$$

Case $\mathbb{K} = \mathbb{R}$: We have $\sigma(A) = \{-1\}$ and $\text{Eig}(A, -1) = \text{span}(\{(1, 0, 0)^\top\})$.

Case $\mathbb{K} = \mathbb{C}$: We have $\sigma(A) = \{-1, i, -i\}$ and $\text{Eig}(A, -1) = \text{span}(\{(1, 0, 0)^\top\})$, $\text{Eig}(A, i) = \text{span}(\{(0, i, 1)^\top\})$, and $\text{Eig}(A, -i) = \text{span}(\{(0, -i, 1)^\top\})$.

2.3 Zeros of Polynomial Functions

Remark 15. Let $P: \mathbb{K} \rightarrow \mathbb{K}$ be a polynomial function with $P \neq 0$. For $\lambda \in \mathbb{K}$ let $P(\lambda) = 0$. Then there exists $r \in \mathbb{N}$ and a polynomial function $Q: \mathbb{K} \rightarrow \mathbb{K}$ such that

- (i) $P(t) = (t - \lambda)^r \cdot Q(t)$ for all $t \in \mathbb{K}$,
- (ii) $Q(\lambda) \neq 0$.

Moreover, r and Q are uniquely determined by (i) and (ii).

Definition 16. Consider the situation of Remark 15. $\lambda \in \mathbb{K}$ is *root* of P with *multiplicity* $r \in \mathbb{N}$. *Notation:* $\mu(P, \lambda) = r$.

Proposition 17. Every polynomial function $P \in \Pi_n \setminus \{0\}$ with pairwise distinct roots $\lambda_1, \dots, \lambda_k$ satisfies

$$\sum_{\ell=1}^k \mu(P, \lambda_\ell) \leq n.$$

Definition 18. A polynomial function P *decomposes into linear factors* if $P \in \Pi_0$ or if $P \in \Pi_n \setminus \Pi_{n-1}$ and there exist $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ and $a \in \mathbb{K}$ such that

$$\forall t \in \mathbb{K}: P(t) = a \cdot \prod_{i=1}^n (t - \lambda_i).$$

Proposition 19 (Fundamental theorem of algebra). Every polynomial function $P: \mathbb{C} \rightarrow \mathbb{C}$ decomposes into linear factors.

Corollary 20. For $A \in \mathbb{C}^{n \times n}$ we have $\sigma(A) \neq \emptyset$ and

$$\sum_{\lambda \in \sigma(A)} \mu(P_A, \lambda) = n.$$

Proof. Apply Proposition 7, Proposition 8, and Proposition 19. □

2.4 Multiplicity of Eigenvalues

In the sequel let $V \neq \{0\}$ be finite-dimensional.

Lemma 21. For $\lambda \in \sigma(F)$ we have⁴

$$1 \leq \dim \text{Eig}(F, \lambda) \leq \mu(P_F, \lambda) \leq n.$$

Proof. By definition, we have $1 \leq \dim \text{Eig}(F, \lambda) \leq n$ for all $\lambda \in \sigma(F)$. Moreover, Proposition 17 implies $\mu(P_F, \lambda) \leq n$.

⁴ $\dim \text{Eig}(F, \lambda)$ and $\mu(P_F, \lambda)$ are called *geometric multiplicity* and *algebraic multiplicity* of the eigenvalue $\lambda \in \mathbb{K}$, respectively.

Put $m = \dim \operatorname{Eig}(F, \lambda)$ such that $1 \leq m \leq n$. Choose a basis $\mathcal{A} = (v_1, \dots, v_n)$ of V such that $F(v_i) = \lambda v_i$ for $i = 1, \dots, m$, see Corollary II.2.16 (Basisergänzungssatz). Then we have

$$\mathcal{M}_{\mathcal{A}}(F) = \begin{pmatrix} D & B \\ 0 & C \end{pmatrix}$$

for some $B \in \mathbb{K}^{m \times (n-m)}$, $C \in \mathbb{K}^{(n-m) \times (n-m)}$, and

$$D = \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{pmatrix}.$$

Iterative application of Lemma 1.7 shows

$$P_F(t) = P_{\mathcal{M}_{\mathcal{A}}(F)}(t) = \det(\mathcal{M}_{\mathcal{A}}(F) - tE_n) = (\lambda - t)^m \cdot \det(C - tE_{n-m}),$$

for $t \in \mathbb{K}$, and hence $\mu(P_F, \lambda) \geq m$. □

Example 22. Consider the situation from Example 13. We have $\dim \operatorname{Eig}(A, \lambda) = \mu(P_A, \lambda)$ for all $\lambda \in \sigma(A) = \{1, -1\}$. For

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

we have $P_B(\lambda) = (1 - \lambda)^2$ and $\dim \operatorname{Eig}(B, 1) = 1 \leq 2 = \mu(P_B, 1)$.

3 Diagonalization

In the sequel let V be a finite-dimensional K -vector space with $n = \dim V \in \mathbb{N}$, and let $F: V \rightarrow V$ be linear.

Definition 1. F is *diagonalizable* if there exists a basis of V consisting of eigenvectors of F .

Remark 2. If $|\sigma(F)| = n$, then F is diagonalizable, see Lemma 2.6.

Definition 3. $A = (a_{i,j})_{i,j} \in K^{n \times n}$ is a *diagonal matrix* if $a_{i,j} = 0$ for all $i, j \in \{1, \dots, n\}$ with $i \neq j$. *Notation:* For $\lambda_1, \dots, \lambda_n \in K$ let

$$\operatorname{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in K^{n \times n}.$$

Remark 4. F is diagonalizable if and only if there exists a basis \mathcal{A} of V such that $\mathcal{M}_{\mathcal{A}}(F)$ is a diagonal matrix.

In the sequel let $A \in K^{n \times n}$.

Definition 5. A is *diagonalizable* if \mathcal{F}_A is diagonalizable.

Lemma 6. Let $\mathcal{A} = (v_1, \dots, v_n)$ be a basis of K^n , $S = (v_1, \dots, v_n) \in K^{n \times n}$, and $\lambda_1, \dots, \lambda_n \in K$. Then the following statements are equivalent:

- (i) $\mathcal{M}_{\mathcal{A}}(\mathcal{F}_A) = \text{diag}(\lambda_1, \dots, \lambda_n)$,
- (ii) $\forall 1 \leq i \leq n: Av_i = \lambda_i v_i$,
- (iii) $\text{diag}(\lambda_1, \dots, \lambda_n) = S^{-1} \cdot A \cdot S$.

Proof. “(i) \Leftrightarrow (ii)”: Obvious. “(ii) \Leftrightarrow (iii)”: Note that

$$S \cdot \text{diag}(\lambda_1, \dots, \lambda_n) = (\lambda_1 v_1, \dots, \lambda_n v_n) \wedge AS = (Av_1, \dots, Av_n). \quad \square$$

Remark 7. $A \in K^{n \times n}$ is diagonalizable if and only if A is similar to a diagonal matrix, see Lemma 6.

Example 8. Consider the situation from Example 2.14.

Case $\mathbb{K} = \mathbb{R}$: A is not diagonalizable since every eigenvector lies in $\text{span}(\{(1, 0, 0)^\top\})$.
Case $\mathbb{K} = \mathbb{C}$: A is diagonalizable, and the family $((1, 0, 0)^\top, (0, i, 1)^\top, (0, -i, 1)^\top)$ of eigenvectors is a basis of \mathbb{C}^3 , see Remark 2.

Example 9. Consider the matrix $A \in \mathbb{K}^{3 \times 3}$ from Example 2.13, and put

$$S = (v_1, v_2, v_3) = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 3 \\ 1 & 0 & 2 \end{pmatrix}$$

consisting of the corresponding eigenvectors v_1, v_2, v_3 of A . It can be readily checked that (v_1, v_2, v_3) is linearly independent and thus a basis of \mathbb{K}^3 . Lemma 6 shows

$$\mathcal{M}_{\mathcal{A}}(\mathcal{F}_A) = \text{diag}(1, 1, -1) = S^{-1} \cdot A \cdot S.$$

In the sequel let V be a \mathbb{K} -vector space, i.e., $K = \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Proposition 10. The following statements are equivalent:

- (i) F is diagonalizable,
- (ii) P_F decomposes into linear factors (over \mathbb{K}) and

$$\forall \lambda \in \sigma(F): \dim \text{Eig}(F, \lambda) = \mu(P_F, \lambda). \quad (1)$$

Proof. “(i) \Rightarrow (ii)”: Let (v_1, \dots, v_n) be a basis of V consisting of eigenvectors of F , and let $F(v_i) = \lambda_i v_i$ with $\lambda_i \in \mathbb{K}$ for $i = 1, \dots, n$. Then we have

$$\sum_{\lambda \in \sigma(F)} \dim \text{Eig}(F, \lambda) \geq \sum_{\lambda \in \sigma(F)} |\{i \in \{1, \dots, n\}: v_i \in \text{Eig}(F, \lambda)\}| \geq n.$$

Combining Proposition 2.17 and Lemma 2.21 shows (1). Moreover, for $\mathcal{A} = (v_1, \dots, v_n)$ we have

$$\mathcal{M}_{\mathcal{A}}(F) = \text{diag}(\lambda_1, \dots, \lambda_n),$$

and hence

$$\forall t \in \mathbb{K}: P_F(t) = \prod_{i=1}^n (\lambda_i - t).$$

“(ii) \Rightarrow (i)”: From (ii) we obtain

$$\sum_{\lambda \in \sigma(F)} \dim \operatorname{Eig}(F, \lambda) = \sum_{\lambda \in \sigma(F)} \mu(F, \lambda) = n.$$

Hence there exist $k \in \{1, \dots, n\}$, pairwise distinct eigenvalues $\lambda_1, \dots, \lambda_k \in \mathbb{K}$, and bases $(v_{i,1}, \dots, v_{i,j_i})$ of $\operatorname{Eig}(F, \lambda_i)$ for $1 \leq i \leq k$ such that $\sum_{i=1}^k j_i = n$. We will show that the family $(v_{1,1}, \dots, v_{1,j_1}, \dots, v_{k,1}, \dots, v_{k,j_k})$ consisting of n eigenvectors of F is linearly independent and hence a basis of V .

Let $\sum_{i=1}^k \sum_{j=1}^{j_i} \mu_{i,j} v_{i,j} = 0$ for $\mu_{1,1}, \dots, \mu_{k,j_k} \in \mathbb{K}$. Since $\sum_{j=1}^{j_i} \mu_{i,j} v_{i,j} \in \operatorname{Eig}(F, \lambda_i)$ for all $1 \leq i \leq k$, Lemma 2.6 implies

$$\forall i \in \{1, \dots, k\}: \sum_{j=1}^{j_i} \mu_{i,j} v_{i,j} = 0.$$

Moreover, since $(v_{i,1}, \dots, v_{i,j_i})$ is a basis of $\operatorname{Eig}(F, \lambda_i)$ and thus linearly independent for all $1 \leq i \leq k$, we obtain $\mu_{i,j} = 0$ for all $i \in \{1, \dots, k\}$ and all $j \in \{1, \dots, j_i\}$. \square

Remark 11. Algorithm to diagonalize a matrix A and to find a basis of \mathbb{K}^n consisting of eigenvectors of A :

1. Compute P_A and its roots.
2. If $\sigma(A) = \emptyset$ or $\sum_{\lambda \in \sigma(A)} \mu(P_A, \lambda) < n$: STOP – A is not diagonalizable.
3. Determine (by using elementary row operations) $\dim \operatorname{Eig}(A, \lambda)$ for $\lambda \in \sigma(A)$. If $\dim \operatorname{Eig}(A, \lambda) < \mu(P_A, \lambda)$ for some $\lambda \in \sigma(A)$: STOP – A is not diagonalizable.
4. Determine (by solving the corresponding homogeneous systems of linear equations) bases of $\operatorname{Eig}(A, \lambda)$ for all $\lambda \in \sigma(A)$ and build up the change-of-basis matrix S such that $S^{-1} \cdot A \cdot S = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

Example 12. In Example 9 we have $n = 3$, P_A decomposes into linear factors, and we have

$$\dim \operatorname{Eig}(A, -1) = 1 = \mu(P_A, -1), \quad \dim \operatorname{Eig}(A, 1) = 2 = \mu(P_A, 1).$$

As seen before, A is diagonalizable.