Mathematical Reference

Leonardo Cerasi[1]

GitHub repository: LeonardoCerasi/notes

[1] leo.cerasi@pm.me

Contents

Part I

Multilinear Algebra

Chapter 1

Vector Spaces and Applications

§1.1 Matrices

Definition 1.1.1 (Matrix)

Given a field $\mathbb{K}$ and $n, m \in \mathbb{N}$, an $n \times m$ **matrix** on $\mathbb{K}$ is the object:

$$\mathrm{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \equiv [a_{ij}]_{j=1,\dots,m}^{i=1,\dots,n} \quad : \quad a_{ij} \in \mathbb{K} \ \forall i = 1, \dots, n, \ j = 1, \dots, m$$

The set of all $n \times m$ matrices on $\mathbb{K}$ is denoted by $\mathbb{K}^{n \times m}$.

When the dimensions of the matrix A are unambiguous, we simply write $\mathrm{A} = [a_{ij}]$. We say that an $n \times n$ matrix is a **square matrix**, an $n \times 1$ matrix is a **column vector** and a $1 \times n$ matrix is a **row vector**.
It is possible to define three operations between matrices:

- sum $+ : \mathbb{K}^{n \times m} \times \mathbb{K}^{n \times m} \to \mathbb{K}^{n \times m} : [a_{ij}]_{j=1,\dots,m}^{i=1,\dots,n} + [b_{ij}]_{j=1,\dots,m}^{i=1,\dots,n} \mapsto [a_{ij} + b_{ij}]_{j=1,\dots,m}^{i=1,\dots,n}$
- product by a scalar $\cdot : \mathbb{K} \times \mathbb{K}^{n \times m} \to \mathbb{K}^{n \times m} : \alpha \cdot [a_{ij}]_{j=1,\dots,m}^{i=1,\dots,n} = [\alpha a_{ij}]_{j=1,\dots,m}^{i=1,\dots,n}$
- product $\cdot : \mathbb{K}^{n \times p} \times \mathbb{K}^{p \times m} \to \mathbb{K}^{n \times m} : [a_{ij}]_{j=1,\dots,p}^{i=1,\dots,n} \cdot [b_{ij}]_{j=1,\dots,m}^{i=1,\dots,p} = [c_{ij}]_{j=1,\dots,m}^{i=1,\dots,n}, \ c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$

Note that αa_{ij} is the $\mathbb{K}$-product.

Proposition 1.1.1

$(\mathbb{K}^{n \times m}, +)$ is an abelian group.

Proof. The matrix sum is equivalent to the $\mathbb{K}$-sum of corresponding elements, which is associative and commutative. The neutral element is the zero matrix $0_{n \times m} = [0]_{j=1,\dots,m}^{i=1,\dots,n}$, while the inverse element is $-\mathrm{A} = [-a_{ij}]_{j=1,\dots,m}^{i=1,\dots,n}$. □

Proposition 1.1.2

$(\mathbb{K}^{n\times n}, +, \cdot)$ is a non-commutative ring.

Proof. By Prop. 1.1.1, $(\mathbb{K}^{n\times n}, +)$ is an abelian group. It is trivial to show the associativity and distributivity of the matrix product, i.e.:

1. $A \cdot (B \cdot C) = (A \cdot B) \cdot C$, $\lambda(A \cdot B) = (\lambda A) \cdot B = A \cdot (\lambda B)$ $\forall A, B, C \in \mathbb{K}^{n\times n}, \lambda \in \mathbb{K}$
2. $A \cdot (B + C) = A \cdot B + A \cdot C$, $(A + B) \cdot C = A \cdot C + B \cdot C$ $\forall A, B, C \in \mathbb{K}^{n\times n}$

Finally, the neutral element of the matrix product is the identity matrix $I_n = [\delta_{ij}]_{i,j=1,\dots,n}$. □

Definition 1.1.2 (Transposed matrix)

Given a matrix $A \in \mathbb{K}^{n\times m}$, its **transpose** is defined as $A^\intercal \in \mathbb{K}^{m\times n} : [a^\intercal_{ij}]^{i=1,\dots,m}_{j=1,\dots,n} = [a_{ji}]^{j=1,\dots,n}_{i=1,\dots,m}$.

Square matrices can be further characterized: a square matrix $A \in \mathbb{K}^{n\times n}$ is said **symmetric** if $A^\intercal = A$ or **antisymmetric** if $A^\intercal = -A$, and it is **diagonal** if $a_{ij} = 0$ $\forall i \neq j \in \{1, \dots, n\}$. Moreover, we can introduce the concept of inverse matrix for square matrices.

Definition 1.1.3 (Inverse matrix)

A square matrix $A \in \mathbb{K}^{n\times n}$ is **invertible** if $\exists A^{-1} \in \mathbb{K}^{n\times n} : A^{-1} \cdot A = A \cdot A^{-1} = I_n$.

Example 1.1.1 (Non-invertible matrix)

The matrix $\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ is non-invertible, as $\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} = \begin{bmatrix} 2\alpha & 2\beta \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $\forall \alpha, \beta, \gamma, \delta \in \mathbb{R}$.

Definition 1.1.4 (General linear group)

The **general linear group** $\mathrm{GL}(n, \mathbb{K})$ is defined as the subset of $\mathbb{K}^{n\times n}$ of all invertible matrices.

Note that $\mathrm{GL}(1, \mathbb{K}) = \mathbb{K} - \{0\}$.

Proposition 1.1.3

$(\mathrm{GL}(n, \mathbb{K}), \cdot)$ is a non-abelian group.

Proof. The neutral element is I_n, as $I_n^{-1} = I_n \implies I_n \in \mathrm{GL}(n, \mathbb{K})$, while the existence of the inverse is granted by definition. We only have to show closure under matrix multiplication:

$$(AB)^{-1} = B^{-1}A^{-1} \impliedby I_n = A \cdot A^{-1} = AI_nA^{-1} = ABB^{-1}A^{-1} = (AB)(AB)^{-1}$$

Hence, $A, B \in \mathrm{GL}(n, \mathbb{K}) \implies AB \in \mathrm{GL}(n, \mathbb{K})$. □

By this result, we conclude that the inverse matrix, when it exists, is unique.

§1.1.1 Linear systems of equations

A **linear equation** with $n \in \mathbb{N}$ variables and $\mathbb{K}$-coefficients is an expression of the form:

$$a_1 x_1 + \cdots + a_n x_n = b \qquad a_i, b \in \mathbb{K}\ \forall i = 1, \dots, n$$

A **solution** of the equation is an n-tuple $(\bar{x}_1, \dots, \bar{x}_n) \in \mathbb{K}^n$ which satisfies this expression.

Definition 1.1.5 (Linear system of equations)

A linear system of equations (or simply **linear system**) is a collection of m linear equations with n variables:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \iff \qquad \mathrm{A}\mathbf{x} = \mathbf{b}$$

where we defined:

$$\mathrm{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \in \mathbb{K}^{m\times n} \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^{m\times 1} \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^{n\times 1}$$

Two linear systems with the same set of solutions are called **equivalent systems**: note that two equivalent systems must have the same number of variables, but not necessarily the same number of equations.

Based on the cardinality of its solution set, a linear system is said to be **impossible** if it has no solutions, **determined** if it has one solution and **undetermined** if it has infinitely-many solutions. Moreover, if the solution set can be parametrized by $k \in \mathbb{N}_0$ variables, the system is of kind ∞^k: a determined system is of kind ∞^0.

Linear systems can be systematically solved applying a reduction algorithm to their corresponding matrices: **Gauss algorithm**. Starting with a general composed matrix $[\mathrm{A}|\mathbf{b}] \in \mathbb{K}^{m\times(n+1)}$, first we multiply the first row by a_{11}^{-1}, so that:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & a'_{12} & \dots & a'_{1n} & b'_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right]$$

Then, at each row $\mathrm{R}_2, \dots, \mathrm{R}_m$ we apply the transformation $\mathrm{R}_k \mapsto \mathrm{R}_k - a_{k1}\mathrm{R}_1$, so that:

$$\left[\begin{array}{cccc|c} 1 & a'_{12} & \dots & a'_{1n} & b'_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & a'_{12} & \dots & a'_{1n} & b'_1 \\ 0 & a'_{22} & \dots & a'_{2n} & b'_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a'_{m2} & \dots & a'_{mn} & b'_m \end{array}\right]$$

Reiterating this process to progressively smalles submatrices, the algorithm yields the general transformation:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & a'_{12} & \dots & a'_{1n} & b'_1 \\ 0 & 1 & \dots & a'_{2n} & b'_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & b'_m \end{array}\right]$$

As these are linear transformations, the two matrices represent equivalent linear systems: the transformed linear system is substantially easier to solve, and its solution set is a solution set of the starting linear system too.

Definition 1.1.6 (Character)

Given a matrix $\mathrm{M} \in \mathbb{K}^{n\times m}$, its **character** $\mathrm{car}(\mathrm{M})$ is the number of non-zero rows remaining after Gauss reduction.

It can be proven that the character is independent of the operations performed during the reduction algorithm (see §1.3.2.2). Moreover, it is possible to prove the Rouché–Capelli theorem (Th. 1.3.6), which states that the system of equations $\mathrm{A}\mathbf{x} = \mathbf{b}$ has solutions if and only if $\mathrm{car}(\mathrm{A}) = \mathrm{car}([\mathrm{A}|\mathbf{b}])$ and, in this case, that it is of type ∞^{n-r}, with $r \equiv \mathrm{car}(\mathrm{A})$.

§1.2 Vector spaces

Definition 1.2.1 (Vector space)

Given a set $V \neq \varnothing$ and a field $\mathbb{K}$, then V is a **$\mathbb{K}$-vector space** if there exist two operations:

$$+ : V \times V \to V \,:\, (\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} + \mathbf{w} \qquad \cdot : \mathbb{K} \times V \to V \,:\, (\lambda, \mathbf{v}) \mapsto \lambda \cdot \mathbf{v}$$

such that $(V, +)$ is an abelian group and the following properties hold $\forall \lambda, \mu \in \mathbb{K}, \mathbf{v}, \mathbf{w} \in V$:

1. $(\lambda + \mu) \cdot (\mathbf{v} + \mathbf{w}) = \lambda \cdot \mathbf{v} + \mu \cdot \mathbf{v} + \lambda \cdot \mathbf{w} + \mu \cdot \mathbf{w}$
2. $(\lambda \cdot \mu) \cdot \mathbf{v} = \lambda \cdot (\mu \cdot \mathbf{v}) = \mu \cdot (\lambda \cdot \mathbf{v})$
3. $1_{\mathbb{K}} \cdot \mathbf{v} = \mathbf{v}$

Note that there are three unique neutral elements: $0_{\mathbb{K}} \equiv 0$, $1_{\mathbb{K}} \equiv 1$ and $0_V \equiv \mathbf{0}$. In the following, the multiplication symbol $\cdot$ is suppressed, as the factors clarify which multiplication is occurring ($\cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ or $\cdot : \mathbb{K} \times V \to V$, which have the same neutral element $1_{\mathbb{K}}$).

Example 1.2.1 (Complex numbers)

$V = \mathbb{C}$ is a vector space both for $\mathbb{K} = \mathbb{R}$ and $\mathbb{K} = \mathbb{C}$, although they are different objects.

Example 1.2.2 (Field as vector space)

$V = \mathbb{K}$ is a $\mathbb{K}$-vector space. Note that, in this case, $0_{\mathbb{K}} \equiv 0_V$.

Note that, by the uniqueness of 0_V, then $\forall \mathbf{v} \in V \ \exists! - \mathbf{v} \in V : \mathbf{v} + (-\mathbf{v}) = 0_V$, so the following cancellation rule holds $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$:

$$\mathbf{u} + \mathbf{v} = \mathbf{w} + \mathbf{v} \quad \Longrightarrow \quad \mathbf{u} = \mathbf{w} \tag{1.1}$$

We can now state some basic properties of vector spaces.

Lemma 1.2.1 (Basic properties of vector spaces)

Given a $\mathbb{K}$-vector space V, then $\forall \lambda \in \mathbb{K}, \mathbf{v} \in V$:

a. $0_{\mathbb{K}} \cdot \mathbf{v} = 0_V$
b. $(-\lambda) \cdot \mathbf{v} = -(\lambda \cdot \mathbf{v})$
c. $\lambda \cdot 0_V = 0_V$
d. $\lambda \cdot \mathbf{v} = 0_V \iff \lambda = 0_{\mathbb{K}} \vee \mathbf{v} = 0_V$

Proof. Respectively:

a. Consider $c \in \mathbb{K} - \{0_{\mathbb{K}}\}$; then $c\mathbf{v} + 0_V = c\mathbf{v} = (c + 0_{\mathbb{K}})\mathbf{v} = c\mathbf{v} + 0_{\mathbb{K}} \cdot \mathbf{v}$, which by Eq. 1.1 proves $0_{\mathbb{K}} \cdot \mathbf{v} = 0_V$.
b. $\lambda\mathbf{v} + (-\lambda)\mathbf{v} = (\lambda - \lambda)\mathbf{v} = 0_{\mathbb{K}} \cdot \mathbf{v} = 0_V$, which by the uniqueness of the negative element proves $(-\lambda)\mathbf{v} = -(\lambda\mathbf{v})$.
c. $\lambda \cdot 0_V = \lambda(\mathbf{v} - \mathbf{v}) = \lambda\mathbf{v} + \lambda \cdot (-1_{\mathbb{K}}) \cdot \mathbf{v} = \lambda\mathbf{v} + (-\lambda)\mathbf{v} = \lambda\mathbf{v} - (\lambda\mathbf{v}) = 0_V$
d. $\lambda = 0_{\mathbb{K}}$ is trivial, so consider $\lambda \neq 0_{\mathbb{K}}$; then $\exists! \lambda^{-1} \in \mathbb{K} : \lambda^{-1} \cdot \lambda = 1_{\mathbb{K}}$, so $0_V = \lambda^{-1} \cdot 0_V = \lambda^{-1} \cdot (\lambda\mathbf{v}) = (\lambda^{-1} \cdot \lambda)\mathbf{v} = 1_{\mathbb{K}} \cdot \mathbf{v} = \mathbf{v}$, i.e. $\mathbf{v} = 0_V$. □

§1.2.1 Subspaces

Definition 1.2.2 (Subspace)

Given a $\mathbb{K}$-vector space V and a subset $U \subseteq V : U \neq \varnothing$, then U is a **subspace** of V if it is closed under $+ : U \times U \to U$ and $\cdot : \mathbb{K} \times U \to U$.

Lemma 1.2.2

If U is a subspace of $V(\mathbb{K})$, then $0_V \in U$.

Proof. By definition $U \neq \varnothing \implies \exists \mathbf{v} \in U$. By the closure condition $\lambda\mathbf{v} \in U \ \forall \lambda \in \mathbb{K}$, hence taking $\lambda = 0_{\mathbb{K}}$ proves the thesis. □

A typical strategy to prove that U is a subspace of $V(\mathbb{K})$ is showing the closure properties, while to prove that it is *not* a subspace we usually show that $0_V \notin U$.

Example 1.2.3 (Polynomial subspaces)

Given $V = \mathbb{K}[x]$, then $U = \mathbb{K}_n[x]$ is a subspace $\forall n \in \mathbb{N}_0$.

An important concept to analyze vector spaces is that of linear combination. Given two sets $\{\lambda_k\}_{k=1,\dots,n} \subset \mathbb{K}$ and $\{\mathbf{v}_k\}_{k=1,\dots,n} \subset V$, their **linear combination** is:

$$\sum_{k=1}^{n} \lambda_k \mathbf{v}_k = \lambda_1 \mathbf{v}_1 + \dots \lambda_n \mathbf{v}_n \in V \tag{1.2}$$

Proposition 1.2.1 (Subspaces and linear combinations)

Given a $\mathbb{K}$-vector space V and $U \subset V : U \neq \varnothing$, then U is a subspace of V if and only if it is closed under linear combinations, that is:

$$\{\lambda_k\}_{k=1,\dots,n} \subset \mathbb{K}, \{\mathbf{v}_k\}_{k=1,\dots,n} \subset U \quad \Longrightarrow \quad \sum_{k=1}^{n} \lambda_k \mathbf{v}_k \in U$$

Proof. First, note that the general case of linear combinations of n vectors can be reduced to the case of 2 vectors.
($\Rightarrow$) Being U a subspace, it is closed under $+ : U \times U \to U$ and $\cdot : \mathbb{K} \times U \to U$; then, by definition $\lambda, \mu \in \mathbb{K}, \mathbf{v}, \mathbf{w} \in U \implies \lambda\mathbf{v} + \mu\mathbf{w} \in U$.
($\Leftarrow$) Given $\lambda \in \mathbb{K}$ and $\mathbf{v}, \mathbf{w} \in V$, then $\mathbf{v}+\mathbf{w} = 1_{\mathbb{K}}\mathbf{v}+1_{\mathbb{K}}\mathbf{w}$ and $\lambda\mathbf{v} = \lambda\mathbf{v}+0_{\mathbb{K}}\mathbf{w}$, hence closure under linear combinations implies closure under $+ : U \times U \to U$ and $\cdot : \mathbb{K} \times U \to U$. □

Generally, it is easier to show closure under linear combinations rather than under addition and scalar multiplication.

Lemma 1.2.3 (Intersection of subspaces)

Given two subspaces of V_1, V_2 of $V(\mathbb{K})$, then $V_1 \cap V_2$ is still a subset of $V(\mathbb{K})$.

Proof. Being V_1, V_2 subspaces, both V_1 and V_2 are closed under linear combinations, so $V_1 \cap V_2$ is too, as $\mathbf{v} \in V_1 \cap V_2 \implies \mathbf{v} \in V_1 \wedge \mathbf{v} \in V_2$. □

On the other hand, in general $V_1 \cup V_2$ is not a subspace. As a counterexample, consider e.g. $V = \mathrm{Vect}_0(\mathbb{R}^3)$, the plane $\pi : z = 0$ and the line $r : (x, y, z) = (0, 0, t), t \in \mathbb{R}$; then, consider the subspaces $V_1 = \mathrm{Vect}_0(\pi), V_2 = \mathrm{Vect}_0(r)$: their union is clearly not closed under addition, as:

$$\begin{pmatrix}1\\1\\0\end{pmatrix} \in V_1, \begin{pmatrix}0\\0\\1\end{pmatrix} \in V_2 \qquad \begin{pmatrix}1\\1\\0\end{pmatrix} + \begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}1\\1\\1\end{pmatrix} \notin V_1 \cup V_2$$

Definition 1.2.3 (Sum of subspaces)

Given a $\mathbb{K}$-vector space V and two subspaces V_1, V_2, their **sum** is defined as:

$$V_1 + V_2 \coloneqq \{\mathbf{w} \in V : \mathbf{w} = \mathbf{u} + \mathbf{v}, \mathbf{u} \in V_1, \mathbf{v} \in V_2\}$$

This is a **direct sum**, denoted by $V_1 \oplus V_2$, if every $\mathbf{w} \in V_1 + V_2$ has a unique representation as $\mathbf{w} = \mathbf{u} + \mathbf{v}, \mathbf{u} \in V_1, \mathbf{v} \in V_2$.

Trivially $V_1, V_2 \subseteq V_1 + V_2$.

Lemma 1.2.4 (Direct sum as disjoint sum)
Given two subspaces V_1, V_2 of $V(\mathbb{K})$, then $V_1 + V_2 = V_1 \oplus V_2 \iff V_1 \cap V_2 = \{\mathbf{0}\}$.

Proof. ($\Rightarrow$) Suppose $\exists \mathbf{v} \in V_1 \cap V_2 : \mathbf{v} \neq \mathbf{0}$; then $\mathbf{v} = \mathbf{v} + \mathbf{0} = \mathbf{0} + \mathbf{v}$, i.e. the expression of $\mathbf{v} \in V_1 + V_2$, but the expression of $\mathbf{v} \in V_1 \oplus V_2$ must be unique, hence $\mathbf{v} = \mathbf{0}$ $\nrightarrow\!\!\!\!\!\!\leftarrow$
($\Leftarrow$) Suppose $\exists \mathbf{w} \in V_1 + V_2 : \mathbf{w} = \mathbf{u}_1 + \mathbf{v}_1 = \mathbf{u}_2 + \mathbf{v}_2, \mathbf{u}_1 \neq \mathbf{u}_2 \in V_1, \mathbf{v}_1 \neq \mathbf{v}_2 \in V_2$; then $V_1 \ni \mathbf{u}_1 - \mathbf{u}_2 = \mathbf{v}_2 - \mathbf{v}_1 \in V_2 \implies \mathbf{v}_2 - \mathbf{v}_1 \in V_1$, so $\mathbf{v}_2 - \mathbf{v}_1 \in V_1 \cap V_2$, but $V_1 \cap V_2 = \{\mathbf{0}\}$, hence $\mathbf{v}_2 = \mathbf{v}_1$ and idem for $\mathbf{u}_1 = \mathbf{u}_2$ $\nrightarrow\!\!\!\!\!\!\leftarrow$ □

The sum of subspaces preserves the subspace structure, contrary to the simple union.

Proposition 1.2.2 (Sum as subspace)
Given a $\mathbb{K}$-vector space and two subspaces V_1, V_2, their sum $V_1 + V_2$ is still a subspace of V.

Proof. Consider $\mathbf{a}, \mathbf{b} \in V_1 + V_2$ and define $\mathbf{u}_{a,b} \in V_1, \mathbf{v}_{a,b} \in V_2 : \mathbf{a} = \mathbf{u}_a + \mathbf{v}_a \wedge \mathbf{b} = \mathbf{u}_b + \mathbf{v}_b$: as V_1, V_2 are subspaces, they are closed under linear combinations, so, given $\lambda, \mu \in \mathbb{K}$, then $\lambda\mathbf{a} + \mu\mathbf{b} = (\lambda\mathbf{u}_a + \mu\mathbf{u}_b) + (\lambda\mathbf{v}_a + \mu\mathbf{v}_b) \equiv \mathbf{u} + \mathbf{v} \in V_1 + V_2$, where $\mathbf{u} \in V_1$ and $\mathbf{v} \in V_2$, which shows that $V_1 + V_2$ too is closed under linear combinations and a subspace by Prop. 1.2.1. □

§1.2.2 Bases

To give a more explicit description of vector spaces, we have to define the concept of basis and its properties.

§1.2.2.1 Generators

Definition 1.2.4 (Linear dependence)
Given a $\mathbb{K}$-vector space V and a set $\{\mathbf{v}_j\}_{j=1,\dots,k} \equiv S \subseteq V$, then the vectors of S are:

- **linearly dependent** (LD) if $\exists \{\lambda_j\}_{j=1,\dots,k} \subset \mathbb{K} - \{0\} : \lambda_1\mathbf{v}_1 + \dots \lambda_k\mathbf{v}_k = \mathbf{0}$
- **linearly independent** (LI) if $\lambda_1\mathbf{v}_1 + \dots \lambda_k\mathbf{v}_k = \mathbf{0} \iff \lambda_j = 0 \ \forall j = 1, \dots, k$

The generalization to infinite sets is trivial: $\{\mathbf{v}_\alpha\}_{\alpha \in \mathcal{I}} \equiv S \subset V(\mathbb{K})$ is LI if every finite subset of S is LI, while it is LD if there exists at least one non-empty subset which is LD.

Example 1.2.4 (Complex numbers)
$\{1, \mathrm{i}\}$ are LD in $\mathbb{C}(\mathbb{C})$, as $1 \cdot 1 + \mathrm{i} \cdot \mathrm{i} = 0$, while they are LI in $\mathbb{C}(\mathbb{R})$.

Example 1.2.5 (Polynomials)
$\{1, x, \dots, x^n, \dots\}$ are LI in $\mathbb{K}[x]$.

We can prove some basic properties of linear dependence.

Lemma 1.2.5 (Basic properties of linear dependence)

Given a $\mathbb{K}$-vector space V and $S \subseteq V : S \neq \varnothing$, then:

a. Given $S \subseteq T \subseteq V$, then S LD $\implies$ T LD

b. $S = \{\mathbf{v}\}$ LD $\implies$ $\mathbf{v} = \mathbf{0}$;

c. $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ LD $\implies$ $\exists \lambda \in \mathbb{K} : \mathbf{v}_1 = \lambda \mathbf{v}_2$

d. If $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ LD, then at least one vector is a linear combination of the others;

e. If S LI and $S \cup \{\mathbf{w}\}$ LD, then $\mathbf{w}$ is a linear combination of the vectors of S;

f. If $\lambda_1 \mathbf{v}_1 + \dots + \lambda_n \mathbf{v}_n = \mathbf{0}$ and $\lambda_n \neq 0$, then $\mathbf{v}_n$ is a linear combination of $\{\mathbf{v}_1, \dots, \mathbf{v}_{n-1}\}$.

Proof. Respectively:

a. $S \subseteq T \implies \mathbf{v} \in T \ \forall \mathbf{v} \in S$, hence $\{\mathbf{v}_i\}_{i=1,\dots,n} \subset S$ LD $\implies$ $\{\mathbf{v}_i\}_{i=1,\dots,n} \subset T$ LD.

b. $\lambda \mathbf{v} = \mathbf{0} \iff \lambda = 0 \vee \mathbf{v} = \mathbf{0}$, so $\mathbf{v} = \mathbf{0} \implies S$ LD, while S LD $\implies \lambda \neq 0 \implies \mathbf{v} = 0$.

c. $\{\mathbf{v}_1, \mathbf{v}_2\}$ LD $\implies \exists \lambda, \mu \in \mathbb{K} - \{0\} : \lambda \mathbf{v}_1 + \mu \mathbf{v}_2 = \mathbf{0} \iff \mathbf{v}_1 = \lambda^{-1} \mu \mathbf{v}_2$

d. If $\{\mathbf{v}_j\}_{j=1,\dots,n}$ LD, then by definition $\exists \{\lambda_j\}_{j=1,\dots,n} \subset \mathbb{K} - \{0\} : \sum_{j=1}^n \lambda_j \mathbf{v}_j = \mathbf{0}$, hence WLOG $\mathbf{v}_1$ can be isolated as $\mathbf{v}_1 = -\lambda_1^{-1} \sum_{j=2}^n \lambda_j \mathbf{v}_j$.

e. $\{\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{w}\}$ LD $\implies \exists \lambda_1, \dots, \lambda_n, \alpha \in \mathbb{K} - \{0\} : \sum_{j=1}^n \lambda_j \mathbf{v}_j + \alpha \mathbf{w} = \mathbf{0}$, so $\mathbf{w}$ can be isolated as $\mathbf{w} = -\alpha^{-1} \sum_{j=1}^n \lambda_j \mathbf{v}_j$.

f. $\sum_{j=1}^n \lambda_j \mathbf{v}_j = \mathbf{0} \wedge \lambda_n \neq 0 \implies \mathbf{v}_n = -\lambda_n^{-1} \sum_{j=1}^{n-1} \lambda_j \mathbf{v}_j$ □

We can now introduce the notion of generators.

Definition 1.2.5 (Generated subset)

Given a $\mathbb{K}$-vector space V and $\{\mathbf{v}_\alpha\}_{\alpha \in \mathcal{I}} \equiv S \subseteq V$, the **subset generated by** S is the set:

$$\operatorname{span} S \coloneqq \{\mathbf{v} \in V : \exists \lambda_1, \dots, \lambda_n \in \mathbb{K}, \mathbf{v}_{\alpha_1}, \dots, \mathbf{v}_{\alpha_n} \in S : \mathbf{v} = \lambda_1 \mathbf{v}_{\alpha_1} + \dots + \lambda_n \alpha_{\mathbf{n}}\}$$

The elements of S are called **generators** of $\operatorname{span} S$.

We often denote $\operatorname{span} S \equiv \langle S \rangle$: this subset contains all vectors of V which can be expressed as linear combinations of vectors of S.

Proposition 1.2.3 (Generated subspace)

Given a $\mathbb{K}$-vector space and $S \subseteq V : S \neq \varnothing$, then $\langle S \rangle$ is a subspace of V.

Proof. Let $S = \{\mathbf{s}_\alpha\}_{\alpha\in\mathcal{I}}$ and $\mathbf{v},\mathbf{w} \in S : \mathbf{v} = \sum_{j=1}^{k} \lambda_j \mathbf{s}_{\alpha_j}, \mathbf{w} = \sum_{j=1}^{n} \mu_j \mathbf{s}_{\beta_j}$, with coefficients $\{\lambda_j\}_{j=1,\dots,k}, \{\mu_j\}_{j=1,\dots,n} \subset \mathbb{K}-\{0\}$. Adding vectors with vanishing coefficients, we can rewrite $\mathbf{v}$ and $\mathbf{w}$ in terms of the same vectors:

$$\mathbf{v} = \sum_{j=1}^{m} a_j \mathbf{s}_{\gamma_j} \qquad \mathbf{w} = \sum_{j=1}^{m} b_j \mathbf{s}_{\gamma_j} \quad \Longrightarrow \quad \zeta\mathbf{v} + \xi\mathbf{w} = \sum_{j=1}^{m} (\zeta a_j + \xi b_j)\, \mathbf{s}_{\gamma_j} \in \langle S\rangle$$

This shows that $\langle S\rangle$ is closed under linear combination, hence the thesis. □

Note that, give a subspace $U \subseteq V(\mathbb{K})$, then at most $U = \langle U\rangle$, hence every subspace admits a family of generators. If U has a finite number of generators, then it is a **finitely-generated subspace**: for example, $\mathbb{K}_n[x] = \langle 1,\dots,x^n\rangle$, $\mathbb{C}(\mathbb{C}) = \langle 1\rangle$ and $\mathbb{C}(\mathbb{R}) = \langle 1, \mathrm{i}\rangle$ are finitely-generated. We can state two trivial properties of generated subsets.

Lemma 1.2.6

Given $S \subseteq V(\mathbb{K})$ and $U = \langle S\rangle$, then:

a. Given $S \subseteq T \subseteq V$, then $U = \langle T\rangle$;

b. If $U = \langle \mathbf{s}_1,\dots,\mathbf{s}_n\rangle$ and $\mathbf{s}_n \in \langle \mathbf{s}_1, \mathbf{s}_{n-1}\rangle$, then $U = \langle \mathbf{s}_1,\dots,\mathbf{s}_{n-1}\rangle$.

Proof. Respectively:

a. If $S \subseteq T$, then each linear combination in S is a linear combination in T too, hence $\langle S\rangle = \langle T\rangle$.

b. Given $\mathbf{v} = \lambda_1\mathbf{s}_1 + \dots + \lambda_n\mathbf{s}_n \in U$ and $\mathbf{s}_n = \mu_1\mathbf{s}_1 + \dots + \mu_{n-1}\mathbf{s}_{n-1}$, then $\mathbf{v} = (\lambda_1 + \mu_1)\,\mathbf{s}_1 + \dots + (\lambda_{n-1} + \mu_{n-1})\,\mathbf{s}_{n-1}$, hence the thesis. □

§1.2.2.2 Bases of generic vector spaces

Definition 1.2.6 (Basis of a vector space)

Given a $\mathbb{K}$-vector space V, a **basis** of V is a LI subset $\mathcal{B} \subseteq V : V = \langle\mathcal{B}\rangle$.

Every non-trivial vector space (i.e. $V \neq \{\mathbf{0}\}$) admits the existence of a basis, but the proof is non-trivial as it relies on Zorn's Lemma (or equivalently to the Axiom of Choice).

Theorem 1.2.1 (Basis theorem)

Every non-trivial vector space admits a basis.

Proof. First, we prove that every LI subset of V can be extended to a basis of V. Let $A \subseteq V$ be a non-empty LI subset of V, and define S the collection of all LI supersets of A.

Lemma 1.2.7

Given a chain $\{A_\alpha\}_{\alpha\in\mathcal{I}} \subseteq S : A_1 \subseteq A_2 \subseteq \dots$, then $\bigcup_{\alpha\in\mathcal{I}} A_\alpha \in S$.

Proof. Set $\mathcal{A} \equiv \bigcup_{\alpha\in\mathcal{I}} A_\alpha$. If $A \subseteq A_\alpha \ \forall\alpha \in \mathcal{I}$, then trivially $A \subseteq \mathcal{A}$. To prove the linear independence, consider a linear combination $\lambda_1\mathbf{v}_1 + \dots + \lambda_n\mathbf{v}_n$ in $\mathcal{A}$, with $n \in \mathbb{N}$, and choose an A_{α_n} large enough so that $v_1, \dots, v_n \in A_{\alpha_n}$. Then, $\lambda_1\mathbf{v}_1 + \dots + \lambda_n\mathbf{v}_n = \mathbf{0} \implies \lambda_1, \dots, \lambda_n = 0$, as A_{α_n} is LI by definition. Since $n \in \mathbb{N}$ is generic, $\mathcal{A}$ is LI. □

It is then clear that S satisfies the hypotheses of Zorn's Lemma (Lemma A.2.1), therefore it has a maximal element $\mathcal{B}$. Now, suppose $\langle\mathcal{B}\rangle \neq V$, i.e. $\exists\mathbf{b} \in V - \langle\mathcal{B}\rangle$, and consider the linear combination $\mu\mathbf{b} + \lambda_1\mathbf{v}_1 + \dots + \lambda_n\mathbf{b}_n = \mathbf{0}$, with $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathcal{B}$ and $n \in \mathbb{N}$: then $-\mu\mathbf{b} \in \langle\mathcal{B}\rangle$, but $\mathbf{b} \notin \langle\mathcal{B}\rangle$, so $\mu = 0$ (as $\mathbf{b} \neq \mathbf{0} \in \langle\mathcal{B}\rangle$). Consequently, $\lambda_1 = \dots = \lambda_n = 0$ as $\mathcal{B}$ is LI, thus $\mathcal{B} \cup \{\mathbf{b}\}$ is LI and a superset of $\mathcal{B} \in S$, which contradicts $\mathcal{B}$ being a maximal element of S $\lightning$
Having showed that every LI subset $A \subseteq V$ can be extended to a basis $\mathcal{B}$ of V, the thesis is trivially found taking $A = \varnothing$, which is a subset of every non-trivial vector space. □

This, though trivial for finite-dimensional spaces, is quite impressive for infinite-dimensional ones (for dimensionality, see §1.2.2.4).

Proposition 1.2.4

Given a $\mathbb{K}$-vector space V, then $S \subseteq V$ is a basis of V if and only if every element of V has a unique representation as a linear combination of elements of S.

Proof. Note that two representations are equal if they differ only by vanishing coefficients.
($\Rightarrow$) As $V = \langle S\rangle$, then every $\mathbf{v} \in V$ can be written as a linear combination of elements of S. Suppose that $\mathbf{v}$ has two representations:

$$\mathbf{v} = \lambda_1\mathbf{s}_1 + \dots \lambda_n\mathbf{s}_n \qquad \mathbf{v} = \mu_1\mathbf{t}_1 + \dots + \mu_m\mathbf{t}_m$$

with $\{\mathbf{s}_j\}_{j=1,\dots,n}, \{\mathbf{t}_k\}_{k=1,\dots,m} \subseteq S$ and $\{\lambda_j\}_{j=1,\dots,n}, \{\mu_k\}_{k=1,\dots,m} \subseteq \mathbb{K}$. Now, we can extend both representations by adding vanishing coefficients, so that both include the same vectors of S:

$$\mathbf{v} = \zeta_1\mathbf{v}_1 + \dots + \zeta_r\mathbf{v}_r \qquad \mathbf{v} = \xi_1\mathbf{v}_1 + \dots + \xi_r\mathbf{v}_r$$

with $\{\mathbf{v}_j\}_{j=1,\dots,r} \subseteq S$ and $\{\zeta_j\}_{j=1,\dots,r}, \{\xi_j\}_{j=1,\dots,r} \subseteq \mathbb{K}$. Subtracting these two expressions:

$$\mathbf{0} = (\zeta_1 - \xi_1)\,\mathbf{v}_1 + \dots + (\zeta_r - \xi_r)\,\mathbf{v}_r$$

But S is LI, hence $\zeta_j = \xi_j \ \forall j = 1, \dots, r$, i.e. the two representations are equal.
($\Leftarrow$) As every $\mathbf{v} \in V$ can be written as a linear combination of elements of S, then $V = \langle S\rangle$. We only have to prove that S is LI. Consider $\mathbf{0} \in V$: by hypothesis, it has a unique representation as a linear combination of vectors in S, and a possible representation is $\mathbf{0} = 0 \cdot \mathbf{s}$ for some $\mathbf{s} \in S$, i.e. the trivial representation with all vanishing coefficients. Now, consider a linear combination in S:

$$\lambda_1\mathbf{s}_1 + \dots + \lambda_n\mathbf{s}_n = \mathbf{0}$$

with $n \in \mathbb{N}$. This too is a representation of $\mathbf{0}$, hence $\lambda_j = 0 \ \forall j = 1, \dots, n$ by the uniqueness of the representation. As $n \in \mathbb{N}$ is generic, this is the definition of S being LI. □

§1.2.2.3 Bases of finitely-generated vector spaces

We now turn our attention to finitely-generated vector spaces, i.e. $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ with $n \in \mathbb{N}$.

Proposition 1.2.5

Given a $\mathbb{K}$-vector space $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$, then $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ contains a basis of V.

Proof. If $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is LI, then it is a basis of V, so consider $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ LD, i.e. $\exists \mathbf{v} \in \langle \{\mathbf{v}_1, \dots, \mathbf{v}_n\} - \{\mathbf{v}\} \rangle$. WLOG, consider $\mathbf{v} = \mathbf{v}_n$, so that $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_{n-1} \rangle$: reiterating this procedure, all LD vectors are eliminated, leaving a basis of V, as at most only a single vector $\mathbf{v}_1$ remains ($\mathbf{v}_1 \neq \mathbf{0}$ as it is LI). □

A direct corollary is that every finitely-generated vector space admits a finite basis, found by the elimination algorithm highlighted in the previous proof.

Definition 1.2.7 (MSLIV)

Given a $\mathbb{K}$-vector space V, then a LI subset $\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ is a **maximal set of linearly-independent vectors** (MSLIV) if $\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \cup \{\mathbf{v}\}$ is LD $\forall \mathbf{v} \in V$.

We extend this notion considering $V = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$: then, a LI subset $\{\mathbf{v}_{j_1}, \dots, \mathbf{v}_{j_r}\} \subseteq \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, with $r \leq n$, is a **maximal subset of linearly-independent vectors** (MSLIV) if $\{\mathbf{v}_{j_1}, \dots, \mathbf{v}_{j_r}\} \cup \{\mathbf{v}_j\}$ is LD $\forall j \in \{1, \dots, n\} - \{j_1, \dots, j_r\}$. Trivially, a maximal subset of LI vectors is also a maximal set of LI vectors in V, so the redundant acronym MSLIV is justified. We can now prove that bases and MSLIVs are equivalent notions.

Theorem 1.2.2 (Bases as MSLIVs)

Given a non-trivial $\mathbb{K}$-vector space $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$, then $\mathcal{B} \subseteq V$ is a basis if and only if it is a MSLIV.

Proof. ($\Leftarrow$) WLOG let $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$, with $r \leq n$, be a MSLIV of $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$: then WTS $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_r \rangle$. If $r = n$ the proof is complete, so consider $r < n$ and $\mathbf{v}_j : r < j \leq n$: by definition $\{\mathbf{v}_1, \dots, \mathbf{v}_r\} \cup \{\mathbf{v}_j\}$ is LD, i.e. $\exists \{\lambda_{j_k}\}_{k=1,\dots,r} \subseteq \mathbb{K} : \mathbf{v}_i = \lambda_{j_1}\mathbf{v}_1 + \dots + \lambda_{j_r}\mathbf{v}_r$, which means that $\mathbf{v}_i \in \langle \mathbf{v}_1, \dots, \mathbf{v}_r \rangle \implies V = \langle \{\mathbf{v}_1, \dots, \mathbf{v}_n\} - \{\mathbf{v}_i\} \rangle$. This holds $\forall i \in [r+1, n] \subseteq \mathbb{N}$, hence $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_r \rangle$.
($\Rightarrow$) Let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of $V$and $\{\mathbf{w}_1, \dots, \mathbf{w}_m\} \subseteq V : m > n$, and suppose this is LI. By definition $\exists \lambda_1, \dots, \lambda_n \in \mathbb{K} : \mathbf{w}_1 = \lambda_1\mathbf{v}_1 + \dots + \lambda_n\mathbf{v}_n$, but $\mathbf{w}_1$ is LI, therefore $\exists j \in [1, \dots, n] \subseteq \mathbb{N} : \lambda_j \neq 0$. WLOG $j = 1$, hence $\mathbf{v}_1 \in \langle \mathbf{w}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \rangle$. Iterating, we can substitute $\mathbf{v}_1, \dots, \mathbf{v}_n$ with $\mathbf{w}_1, \dots, \mathbf{w}_n$: indeed, supposing that $\{v_1, \dots, \mathbf{v}_r\}$ have been substituted with $\mathbf{w}_1, \dots, \mathbf{w}_r$, with $1 \leq r < n$, then $\mathbf{v}_{r+1}$ can be substituted with $\mathbf{w}_{r+1}$ as $V = \langle \mathbf{w}_1, \dots, \mathbf{w}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n \rangle \implies \exists \alpha_1, \dots, \alpha_r, \beta_{r+1}, \dots, \beta_n \in \mathbb{K} : \mathbf{w}_{r+1} = \alpha_1\mathbf{w}_1 + \dots + \alpha_r\mathbf{w}_r + \beta_{r+1}\mathbf{v}_{r+1} + \dots + \beta_n\mathbf{v}_n$, but $\{\mathbf{w}_1, \dots, \mathbf{w}_{r+1}\}$ are LI, thus $\exists j \in [r+1, n] \subseteq \mathbb{N} : \beta_j \neq 0$,

and WLOG $j = r+1$ by reordering indices. Performing the reiteration $V = \langle \mathbf{w}_1, \dots, \mathbf{w}_n \rangle$, so $\mathbf{w}_{n+1}$ is a linear combination of $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ $\lightning$ □

There is still another equivalent concept to introduce.

Definition 1.2.8 (MSG)

Given a $\mathbb{K}$-vector space V, then $\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ is a **minimal set of generators** (MSG) if $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ and $\{\mathbf{v}_1, \dots, \mathbf{v}_n\} - \{\mathbf{v}_j\}$ does not generate V $\forall j = 1, \dots, n$.

Theorem 1.2.3 (Bases ad MSGs)

Given a non-trivial $\mathbb{K}$-vector space V, then $\mathcal{B} \subseteq V$ is a basis of V if and only if it is a MSG.

Proof. ($\Leftarrow$) Let $\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ be a MSG: then WTS $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is LI. Consider a linear combination $\lambda_1 \mathbf{v}_1 + \dots + \lambda_n \mathbf{v}_n = \mathbf{0}$ and suppose $\lambda_1 \neq 0$: this allows to express $\mathbf{v}_1$ as a linear combination of $\{\mathbf{v}_2, \dots, \mathbf{v}_n\}$, but then $V = \langle \mathbf{v}_2, \dots, \mathbf{v}_n \rangle$ $\lightning$
($\Rightarrow$) Suppose $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is not a MSG, and WLOG $V = \langle \mathbf{v}_2, \dots, \mathbf{v}_n \rangle$: then $\mathbf{v}_1$ can be expressed as linear combination of $\{\mathbf{v}_2, \dots, \mathbf{v}_n\}$, i.e. $\mathcal{B}$ is LD $\lightning$ □

This shows that bases, MSLIVs and MSGs are all equivalent notions.

§1.2.2.4 Dimensionality

To properly define the concept of dimensionality of a vector space, we first have to prove that all bases are equivalent.

Theorem 1.2.4 (Equicardinality of bases)

Given a non-trivial $\mathbb{K}$-vector space V and two bases $\mathcal{B}_1 = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}, \mathcal{B}_2 = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$, then $n = m$.

Proof. As $\mathcal{B}_1$ is a MSLIV by Th. 1.2.2, then every subset of $n+1$ vectors in V is LD, hence $m \leq n$ as $\mathcal{B}_2$ must be LI. The vice versa applies too, hence $n = m$. □

By this theorem, all bases of finitely-generated spaces are equivalent, since the equicardinality ensures that we can define a bijection $f : \mathcal{B}_1 \leftrightarrow \mathcal{B}_2$ $\forall \mathcal{B}_1, \mathcal{B}_2$ bases of V.
Moreover, this result hints to the fact that the cardinality of the bases of V is a fundamental property of the vector space, linked to hits dimensionality, so we give a proper definition of this quantity.

Definition 1.2.9 (Dimension)

Given a $\mathbb{K}$-vector space V, then we define its **dimension** as:

$$\dim_{\mathbb{K}} V := \begin{cases} 0 & V = \{\mathbf{0}\} \\ n & |\mathcal{B}| = n \ \forall \mathcal{B} \text{ basis of } V \\ \infty & V \text{ not finitely-generated} \end{cases}$$

The dimension of a vector space is a well-defined quantity by Th. 1.2.1 and Th. 1.2.4.

Example 1.2.6 (Various spaces)
Trivially, $\dim_{\mathbb{K}} \mathbb{K}^n = n$, so $\dim_{\mathbb{C}} \mathbb{C}^n = n$ and $\dim_{\mathbb{R}} \mathbb{C}^n = 2n$, while $\dim_{\mathbb{R}} \mathbb{R}^{\mathbb{R}} = \infty$.

We can now give some trivial properties of dimensionality.

Lemma 1.2.8 (Basic property of dimension)
Given an n-dimensional $\mathbb{K}$-vector space V, then:

a. $\{\mathbf{v}_1, \dots, \mathbf{v}_m\} \subseteq V$ is LD $\forall m > n$;

b. $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ LI is a basis of V;

c. $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ set of generators of V is a basis of V.

Proof. These results are corollaries of Th. 1.2.2 and Th. 1.2.3. □

Proposition 1.2.6 (Dimension of subspaces)
Given $\dim_{\mathbb{K}} V = n$ and a subspace $U \subseteq V$, then $\dim_{\mathbb{K}} U \equiv k \leq n$ and $k = n \iff U = V$.

Proof. The case $U = \{\mathbf{0}\}$ is trivial, so consider $U \neq \{\mathbf{0}\}$. Let $\mathbf{u}_1 \in U$ LI and add $\mathbf{u}_2, \mathbf{u}_3, \dots \in U$ to get $\{\mathbf{u}_1, \mathbf{u}_2\}, \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}, \dots$: a LD subset is reached in at most n steps. Let WLOG $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ the MSLIV of $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$, with $k \leq n$: by Th. 1.2.2, this is a basis of U, hence $k = \dim_{\mathbb{K}} U \leq n$.
$U = V \implies k = n$ is trivial, while $k = n \implies \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is a MSLIV of V, hence a basis of V, so $V = \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle = V$. □

A consequence of this theorem is the fact that LI subset $\{\mathbf{v}_1, \dots, \mathbf{v}_r\} \subseteq V$, with $r < n$, can always be completed to a basis, i.e. $\exists \mathbf{w}_{r+1}, \dots, \mathbf{w}_n \in V : \{\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{w}_{r+1}, \mathbf{w}_n\}$ is a basis of V.

Theorem 1.2.5 (Grassmann's Theorem)
Given a $\mathbb{K}$-vector space V and finitely-generated subspaces $X, Y \subseteq V$, then:

$$\dim_{\mathbb{K}} X + \dim_{\mathbb{K}} Y = \dim_{\mathbb{K}} (X + Y) + \dim_{\mathbb{K}} (X \cap Y) \tag{1.3}$$

Proof. Let $\mathcal{B}_X = \{\mathbf{x}_1, \dots, \mathbf{x}_r\}, \mathcal{B}_Y = \{\mathbf{y}_1, \dots, \mathbf{y}_s\}$ be bases of X, Y and $m \equiv \dim_{\mathbb{K}} (X \cap Y)$. If $m = 0$, then $X \cap Y = \{\mathbf{0}\}$, while if $m \geq 1$ let $\mathcal{B}_{XY} = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ be a basis of $X \cap Y$, which is a finitely-generated subspace by Lemma 1.2.3. Then, completing the bases, $\exists \mathbf{x}_{m+1}, \dots, \mathbf{x}_r \in X : \{\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{x}_{m+1}, \dots, \mathbf{x}_r\}$ is a basis of X and $\exists \mathbf{y}_{m+1}, \dots, \mathbf{y}_s \in Y : \{\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{y}_{m+1}, \dots, \mathbf{y}_s\}$ is a basis of Y (WLOG same vectors as in $\mathcal{B}_X$ and $\mathcal{B}_Y$). Now, WTS $\dim_{\mathbb{K}} (X + Y) = r + s - m$, so consider $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{x}_{m+1}, \dots, \mathbf{x}_r, \mathbf{y}_{m+1}, \dots, \mathbf{y}_s\}$:

- $X + Y := \{\mathbf{v} = \mathbf{x} + \mathbf{y} : \mathbf{x} \in X, \mathbf{y} \in Y\}$, but $\mathbf{x} \in \langle \mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{x}_{m+1}, \dots, \mathbf{x}_r \rangle$ and $\mathbf{y} \in \langle \mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{y}_{m+1}, \dots, \mathbf{y}_s \rangle$, so $\mathbf{x}+\mathbf{y} \in \langle \mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{x}_{m+1}, \dots, \mathbf{x}_r, \mathbf{y}_{m+1}, \dots, \mathbf{y}_s \rangle$, i.e.

$X + Y = \langle \mathcal{B} \rangle$;

- consider the following linear combination:

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_m \mathbf{v}_m + \beta_{m+1} \mathbf{x}_{m+1} + \cdots + \beta_r \mathbf{x}_r + \gamma_{m+1} \mathbf{y}_{m+1} + \cdots + \gamma_s \mathbf{y}_s = \mathbf{0}$$

and rearrange it as:

$$\underbrace{\alpha_1 \mathbf{v}_1 + \cdots + \alpha_m \mathbf{v}_m + \beta_{m+1} \mathbf{x}_{m+1} + \cdots + \beta_r \mathbf{x}_r}_{\in X} = \underbrace{-\gamma_{m+1} \mathbf{y}_{m+1} - \cdots - \gamma_s \mathbf{y}_s}_{\in Y}$$

Therefore, both expressions are in $X \cap Y = \langle \mathbf{v}_1, \ldots, \mathbf{v}_m \rangle$, hence $\exists \delta_1, \ldots, \delta_m \in \mathbb{K}$ such that:

$$\delta_1 \mathbf{v}_1 + \cdots + \delta_m \mathbf{v}_m + \gamma_{m+1} \mathbf{y}_{m+1} + \cdots + \gamma_s \mathbf{y}_s = \mathbf{0}$$

But $\mathcal{B}_Y$ is a basis of Y, i.e. LI, so $\delta_1 = \cdots = \delta_m = \gamma_{m+1} = \cdots = \gamma_s = 0$, thus:

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_m \mathbf{v}_m + \beta_{m+1} \mathbf{x}_{m+1} + \cdots + \beta_r \mathbf{x}_r = \mathbf{0}$$

But $\mathcal{B}_X$ is a basis of X, i.e. LI, so $\alpha_1 = \cdots = \alpha_m = \beta_{m+1} = \cdots = \beta_r = 0$. This shows that $\mathcal{B}$ is LI.

By Def. 1.2.6, $\mathcal{B}$ is a basis of $X + Y$, i.e. $\dim_{\mathbb{K}}(X + Y) = r + s - m$. □

Example 1.2.7 (Euclidean geometry)

Consider $V = \mathrm{Vect}_0(\mathbb{R}^3)$ and α, β planes such that $\mathbf{0} \in \alpha, \beta$: they then determine a line $r \equiv \alpha \cap \beta \ni \{\mathbf{0}\}$. Setting $X = \mathrm{Vect}_0(\alpha)$, $Y = \mathrm{Vect}_0(\beta)$ and $X \cap Y = \mathrm{Vect}_0(r)$, we correctly have $2 + 2 = 3 + 1$.

§1.3 Linear applications

A fundamental tool in mathematics are linear applications, which arise in every one of its fields of study.

Definition 1.3.1 (Linear application)

Given $\mathbb{K}$-vector spaces V, W, an application $f : V \to W$ is **$\mathbb{K}$-linear** if:

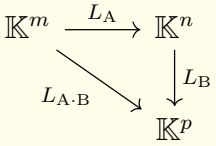
$$f(\lambda \mathbf{v} + \mu \mathbf{w}) = \lambda f(\mathbf{v}) + \mu f(\mathbf{w}) \quad \forall \lambda, \mu \in \mathbb{K}, \mathbf{v}, \mathbf{w} \in V$$

This condition means that $\mathbb{K}$-linear applications preserve linear combinations.

Example 1.3.1 (Matrices)

Given $\mathrm{A} \in \mathbb{K}^{m \times n}$, we can associate to it an application $L_{\mathrm{A}} : \mathbb{K}^n \to \mathbb{K}^m : \mathbf{v} \mapsto \mathrm{A}\mathbf{v}$, which is $\mathbb{K}$-linear by the linearity of the matrix product. Note that $L_{\mathrm{I}_n} = \mathrm{id}_{\mathbb{K}^n}$.
Moreover, given $\mathrm{A} \in \mathbb{K}^{m \times n}$ and $\mathrm{B} \in \mathbb{K}^{n \times p}$, then $L_{\mathrm{A}} \circ L_{\mathrm{B}} = L_{\mathrm{A} \cdot \mathrm{B}} : \mathbb{K}^p \to \mathbb{K}^m$ by the following

commutative diagram:



We can now state some properties of linear applications.

Lemma 1.3.1 (Basic properties of linear applications)

Given $\mathbb{K}$-vector spaces V, W, Z and $\mathbb{K}$-linear applications $f : V \to W, g : W \to Z$, then:

a. $f(\mathbf{0}_V) = \mathbf{0}_W$

b. $g \circ f : V \to Z$ is $\mathbb{K}$-linear;

c. If f is bijective, then $f^{-1} : W \to V$ is $\mathbb{K}$-linear.

Proof. Respectively:

a. $f(\mathbf{0}_V) = f(0_\mathbb{K} \cdot \mathbf{v}) = 0_\mathbb{K} \cdot f(\mathbf{v}) = \mathbf{0}_W$

b. $g \circ f(\lambda\mathbf{u} + \mu\mathbf{v}) = g(\lambda f(\mathbf{u}) + \mu f(\mathbf{v})) = \lambda g(f(\mathbf{u})) + \mu g(f(\mathbf{v}))$

c. $f(\lambda f^{-1}(\mathbf{u}) + \mu f^{-1}(\mathbf{v})) = \lambda\mathbf{u} + \mu\mathbf{v} \implies f^{-1}(\lambda\mathbf{u} + \mu\mathbf{v}) = \lambda f^{-1}(\mathbf{u}) + \mu f^{-1}(\mathbf{v})$ □

We can also prove an existence-uniqueness theorem for linear applications.

Theorem 1.3.1 (Existence and uniqueness)

Let V, W be $\mathbb{K}$-vector spaces, $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ a basis of V and $\{\mathbf{w}_1, \dots, \mathbf{w}_n\} \subseteq W$ an ordered set of vectors. Then $\exists! \varphi : V \to W : \varphi(\mathbf{b}_j) = \mathbf{w}_j \ \forall j = 1, \dots, n$ which is $\mathbb{K}$-linear.

Proof. Let $\mathbf{v} \in V$; then $\exists \alpha_1, \dots, \alpha_n \in \mathbb{K} : \mathbf{v} = \alpha_1\mathbf{b}_1 + \dots + \alpha_n\mathbf{b}_n$ fixed since $\mathcal{B}$ is a basis. Now, define $\varphi : V \to W : \varphi(\mathbf{v}) = \alpha_1\mathbf{w}_1 + \dots \alpha_n\mathbf{w}_n$: clearly $\varphi(\mathbf{b}_j) = \mathbf{w}_j \ \forall j = 1, \dots, n$, and also φ is unique since both $\{\alpha_j\}_{j=1,\dots,n} \subseteq \mathbb{K}$ and $\{\mathbf{w}_j\}_{j=1,\dots,n} \subseteq W$ are fixed. Finally, φ is $\mathbb{K}$-linear, since $f(\lambda\mathbf{v}_1 + \mu\mathbf{v}_2) = (\lambda\alpha_1 + \mu\beta_1)\,\mathbf{w}_1 + \dots + (\lambda\alpha_n + \mu\beta_n)\,\mathbf{w}_n = \lambda f(\mathbf{v}_1) + \mu f(\mathbf{v}_2)$. □

In general, fixed $\dim_\mathbb{K} V = n$, then given two sets $\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subseteq V$ and $\{\mathbf{w}_1, \dots, \mathbf{w}_k\} \subseteq W$, with $k \in \mathbb{N}$, then the existence of $\varphi : V \to W : \varphi(\mathbf{v}_j) = \mathbf{w}_j \ \forall j = 1, \dots, k$ is only granted if $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is LI: in this case, if $n = k$ then φ is unique too, by the previous theorem, while if $k < n$ in general we can define multiple φ with such property, as we can complete $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ to a basis of V, which can then be mapped to arbitrary vectors in W. On the other hand, if $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is LD, then φ can be defined only if $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ satisfies the same linear-dependence relations, otherwise linearity cannot be satisfied.

Given two $\mathbb{K}$-vector space V and W, we denote the set of all $\mathbb{K}$-linear applications $f : V \to W$ as $\mathrm{Hom}_\mathbb{K}(V, W)$: this has a natural structure of $\mathbb{K}$-vector space with $(f + g)(\mathbf{v}) \equiv f(\mathbf{v}) + g(\mathbf{v})$ and $(\lambda \cdot f)(\mathbf{v}) = \lambda \cdot f(\mathbf{v})$.

Definition 1.3.2 (Kernel and image)

Given $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$, its **kernel** is defined as $\ker f \coloneqq \{\mathbf{v} \in V : f(\mathbf{v}) = \mathbf{0}_W\} \subseteq V$, while its **image** (or range) is defined as $\mathrm{ran}\, f \coloneqq \{\mathbf{w} \in W : \exists \mathbf{v} \in V : \mathbf{w} = f(\mathbf{v})\} \subseteq W$.

Lemma 1.3.2 (Kernel and image as subspaces)

Given $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$, then $\ker f$ is a subspace of V and $\mathrm{ran}\, f$ is a subspace of W.

Proof. By the linearity of f, given $\mathbf{v}, \mathbf{v}' \in V$ and $\mathbf{w}, \mathbf{w}' \in W$:

$$f(\lambda\mathbf{v} + \mu\mathbf{v}') = \lambda f(\mathbf{v}) + \mu f(\mathbf{v}') = \lambda \cdot \mathbf{0}_W + \mu \mathbf{0}_W = \mathbf{0}_{\mathbf{w}}$$

$$\mathrm{ran}\, f \ni \lambda\mathbf{w} + \mu\mathbf{w}' = \lambda f(\mathbf{v}) + \mu f(\mathbf{v}') = f(\lambda\mathbf{v} + \mu\mathbf{v}')$$

Thus, both $\ker f$ and $\mathrm{ran}\, f$ are closed under linear combinations, i.e. vector spaces. □

We can further carachterize the kernel and the image of a linear application.

Proposition 1.3.1 (Kernel and injections)

Let V, W be finitely-generated $\mathbb{K}$-vector spaces and $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. Then the following conditions are equivalent:

a. f is injective;

b. $\ker f = \{\mathbf{0}_V\}$

c. $\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subseteq V$ LI $\implies$ $\{f(\mathbf{v}_1, \dots, f(\mathbf{v}_k)\} \subseteq W$ LI

Proof. Consider the following implications:
(a $\Rightarrow$ b) Suppose $\exists \mathbf{v} \in \ker f : \mathbf{v} \neq \mathbf{0}_V$; then $f(\mathbf{v}) = \mathbf{0}_W = f(\mathbf{0}_V)$, but f is injective $\lightning$
(b $\Rightarrow$ c) Let $\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subseteq V$ LI and consider $\mathbf{0}_W = \lambda_1 f(\mathbf{v}_1) + \cdots + \lambda_k f(\mathbf{v}_k) = f(\lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k)$, hence $\lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k = \mathbf{0}_V$ as $\ker f = \{\mathbf{0}_V\}$, therefore $\lambda_1 = \cdots = \lambda_k = 0$ as $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ LI.
(c $\Rightarrow$ a) Given $\mathbf{v}_1, \mathbf{v}_2 \in V$, by linearity $f(\mathbf{v}_1) = f(\mathbf{v}_2) \implies f(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0}_W$, so suppose $\mathbf{v}_1 \neq \mathbf{v}_2$: then $\mathbf{v} \equiv \mathbf{v}_1 - \mathbf{v}_2 \neq \mathbf{0}_V$, i.e. LI, but $f(\mathbf{v}) = \mathbf{0}_W$, i.e. LD $\lightning$ □

Proposition 1.3.2 (Image and surjections)

Let V, W be finitely-generated $\mathbb{K}$-vector spaces and $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. Then the following conditions are equivalent:

a. f is surjective;

b. $\mathrm{ran}\, f = W$

c. $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_k \rangle \implies W = \langle f(\mathbf{v}_1), \dots, f(\mathbf{v}_k) \rangle$

Proof. Consider the following implications:
(a $\Rightarrow$ b) Suppose $\operatorname{ran} f \subsetneq W \implies \exists \mathbf{w} \in W : \nexists \mathbf{v} \in V : \mathbf{w} = f(\mathbf{v}) \implies f$ not surjective $\lightning$
(b $\Rightarrow$ c) $\operatorname{ran} f = W \implies \forall \mathbf{w} \in W\, \exists \mathbf{v} \in V : \mathbf{w} = f(\mathbf{v})$; moreover, $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_k \rangle \implies \forall \mathbf{v} \in V\, \exists \lambda_1, \dots, \lambda_k \in \mathbb{K} : \mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k$. Then $\forall \mathbf{w} \in W\, \exists \lambda_1, \dots, \lambda_k \in \mathbb{K} : \mathbf{w} = \lambda_1 f(\mathbf{v}_1) + \cdots + \lambda_k f(\mathbf{v}_k)$, i.e. $W = \langle f(\mathbf{v}_1), \dots, f(\mathbf{v}_k) \rangle$.
(c $\Rightarrow$ a) $W = \langle f(\mathbf{v}_1), \dots, f(\mathbf{v}_k) \rangle \implies \forall \mathbf{w} \in W\, \exists \lambda_1, \dots, \lambda_k \in \mathbb{K} : \mathbf{w} = \lambda_1 f(\mathbf{v}_1) + \cdots + \lambda_k f(\mathbf{v}_k) = f(\lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k)$, but $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_k \rangle$, so $\forall \mathbf{w} \in W\, \exists \mathbf{v} \in V : \mathbf{w} = f(\mathbf{v})$. □

In general, even for non-surjective $f \in \operatorname{Hom}_{\mathbb{K}}$, it is true that $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_k \rangle \implies \operatorname{ran} f = \langle f(\mathbf{v}_1), \dots, f(\mathbf{v}_k) \rangle$ with a reasoning analogous to the previous proof.
As injections map LI vectors to LI vectors and surjections map generators to generators, we see that bijections map bases to bases.

Theorem 1.3.2 (Rank–nullity theorem)

Let V, W be finitely-generated $\mathbb{K}$-vector spaces and $f \in \operatorname{Hom}_{\mathbb{K}}(V, W)$. Then:

$$\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} \ker f + \dim_{\mathbb{K}} \operatorname{ran} f \tag{1.4}$$

Proof. As $\ker f \subseteq V$ and $\operatorname{ran} f \subseteq W$, they are both finitely-generated. If $\operatorname{ran} f = \{\mathbf{0}_W\}$ (trivial map), then $\ker f = V$ and the thesis is verified.
Consider $\operatorname{ran} f \neq \{\mathbf{0}_W\}$ and choose a basis $\{\mathbf{c}_1, \dots, \mathbf{c}_k\}$ of $\operatorname{ran} f$: this means that $\exists \mathbf{b}_1, \dots, \mathbf{b}_k \in V : f(\mathbf{b}_j) = \mathbf{c}_j\ \forall j = 1, \dots, k$. Now, if $\ker f \neq \{\mathbf{0}_V\}$ choose a basis $\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ of $\ker f$, otherwise consider no other vectors, and set $\mathcal{B} \equiv \{\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{b}_1, \dots, \mathbf{b}_k\} \subseteq V$. WTS $\mathcal{B}$ is a basis of V:

- consider the following linear combination:

  $$\alpha_1 \mathbf{a}_1 + \cdots + \alpha_r \mathbf{a}_r + \beta_1 \mathbf{b}_1 + \cdots + \beta_k \mathbf{b}_k = \mathbf{0}_V$$

  Then, by the linearity of f:

  $$\begin{aligned} \mathbf{0}_W = f(\mathbf{0}_V) &= f(\alpha_1 \mathbf{a}_1 + \cdots + \alpha_r \mathbf{a}_r + \beta_1 \mathbf{b}_1 + \cdots + \beta_k \mathbf{b}_k) \\ &= \alpha_1 \cdot \mathbf{0}_W + \cdots + \alpha_r \cdot \mathbf{0}_W + \beta_1 \mathbf{c}_1 + \cdots + \beta_k \mathbf{c}_k = \beta_1 \mathbf{c}_1 + \cdots + \beta_k \mathbf{c}_k \end{aligned}$$

  But $\{\mathbf{c}_1, \dots, \mathbf{c}_k\}$ is a basis of $\operatorname{ran} f$, hence $\beta_1 = \cdots = \beta_k$ due to linear independence. Then $\alpha_1 \mathbf{a}_1 + \cdots + \alpha_r \mathbf{a}_r = \mathbf{0}_V$, but $\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ is a basis of $\ker f$, so $\alpha_1 = \cdots = \alpha_r = 0$;

- $\mathbf{v} \in V \implies f(\mathbf{v}) \in \operatorname{ran} f = \langle f(\mathbf{b}_1), \dots, f(\mathbf{b}_k) \rangle$, so $\exists \gamma_1, \dots, \gamma_k \in \mathbb{K} : f(\mathbf{v}) = \gamma_1 f(\mathbf{b}_1) + \cdots + \gamma_k f(\mathbf{b}_k)$, which rearranging and using the linearity of f becomes $f(\mathbf{v} - \gamma_1 \mathbf{v}_1 - \cdots - \gamma_k \mathbf{b}_k) = \mathbf{0}_W$, i.e. $\mathbf{v} - \gamma_1 \mathbf{v}_1 - \cdots - \gamma_k \mathbf{v}_k \in \ker f = \langle \mathbf{a}_1, \dots, \mathbf{a}_r \rangle$. Then, $\exists \delta_1, \dots, \delta_r \in \mathbb{K} : \mathbf{v} = \gamma_1 \mathbf{v}_1 + \cdots + \gamma_k \mathbf{v}_k + \delta_1 \mathbf{a}_1 + \cdots + \delta_r \mathbf{a}_r$, which shows that $V = \langle \mathcal{B} \rangle$.

By Def. 1.2.6, $\mathcal{B}$ is a basis of V, i.e. $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} \ker f + \dim_{\mathbb{K}} \operatorname{ran} f$. □

The name of this theorem will be clear in §1.3.2.2.

Corollary 1.3.2.1 (Equidimensionality and bijections)

Let V, W be finitely-generated $\mathbb{K}$-vector spaces and $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. Then:

$$\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W \quad \Longrightarrow \quad f \text{ injective} \iff f \text{ surjective} \iff f \text{ bijective}$$

Proof. By Prop. 1.3.1, f injective $\iff \ker f = \{\mathbf{0}_V\} \iff \dim_{\mathbb{K}} \ker f = 0$. By Th. 1.3.2 $\dim_{\mathbb{K}} \ker f = 0 \iff \dim_{\mathbb{K}} \mathrm{ran}\, f = \dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W \iff \mathrm{ran}\, f = W$, and by Prop. 1.3.2 $\mathrm{ran}\, f = W \iff f$ surjective. Hence, f is both injective and surjective, i.e. a bijection. □

We can further classify applications:

- $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ is a **homomorphism**;
- $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ bijective is an **isomorphism**;
- $f \in \mathrm{Hom}_{\mathbb{K}}(V, V) \equiv \mathrm{End}(V)$ is an **endomorphism**;
- $f \in \mathrm{End}(V)$ bijective is an **automorphism**.

Isomorphisms are particularly interesting.

Lemma 1.3.3 (Basic properties of isomorphisms)

Given three $\mathbb{K}$-vector spaces V, W, Z and $f \in \mathrm{Hom}_{\mathbb{K}}(V, W), g \in \mathrm{Hom}_{\mathbb{K}}(W, Z)$, then:

a. f is an isomorphism if and only if it is invertible;

b. If f and g are isomorphisms, then $g \circ f \in \mathrm{Hom}_{\mathbb{K}}(V, Z)$ is an isomorphism.

Proof. Trivial by the fact that invertibility is equivalent to bijectivity and that the composition of bijections is a bijection. □

Example 1.3.2 (Matrices as endomorphisms)

Given $\mathrm{A} \in \mathbb{K}^{n\times n}$, then $L_{\mathrm{A}} \in \mathrm{End}(\mathbb{K}^n)$. Moreover, if $\mathrm{A} \in \mathrm{GL}(n, \mathbb{K})$, then L_{A} is an automorphism.

Isomorphism induce an equivalence relation between vector spaces.

Definition 1.3.3 (Isomorphism relation)

Two $\mathbb{K}$-vector spaces V, W are **isomorphic** $V \cong W$ if $\exists f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ isomorphism.

This is an equivalence relation since, if f is an isomorphism, then f^{-1} is an isomorphism too.

Theorem 1.3.3 (Equidimensionality and isomorphisms)

Let V, W be finitely-generated $\mathbb{K}$-vector spaces. Then:

$$V \cong W \iff \dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W$$

Proof. ($\Rightarrow$) $V \cong W \implies \exists f \in \mathrm{Hom}_{\mathbb{K}}(V,W)$ isomorphism, which maps bases to bases, hence $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W$
($\Leftarrow$) Consider $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\} \subseteq V$ basis of V, so that $\forall \mathbf{v} \in V \, \exists! \alpha_1, \dots, \alpha_n \in \mathbb{K} : \mathbf{v} = \alpha_1 \mathbf{b}_1 + \cdots + \alpha_n \mathbf{b}_n$. Then, define $\varphi : V \to \mathbb{K}^n : \varphi(\mathbf{v}) = (\alpha_1, \dots, \alpha_n)$, which is clearly linear, so $\varphi \in \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K}^n)$. Moreover, $\forall \boldsymbol{\alpha} \in \mathbb{K}^n \, \exists \mathbf{v} \in V : \mathbf{v} = \sum_{j=1}^n \alpha_j \mathbf{b}_j$, as $V = \langle \mathcal{B} \rangle$, hence $\forall \boldsymbol{\alpha} \in \mathbb{K}^n \, \exists \mathbf{v} \in V : \varphi(\mathbf{v}) = \boldsymbol{\alpha}$, i.e. φ is a surjection. Since $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} \mathbb{K}^n$, by Cor. 1.3.2.1 φ is a bijection too, thus $V \cong \mathbb{K}^n$.
Analogously, given a basis $\mathcal{C} \subseteq W$ of W, we can construct an equivalent isomorphism $\psi : W \to \mathbb{K}^n$, so $W \cong \mathbb{K}^n$. By the transitivity of the isomorphism relation, $V \cong W$. □

The isomorphism relation then partitions the set of all finitely-generated vector spaces into equivalence classes composed of equidimensional spaces: for example, $\mathbb{C}^n(\mathbb{R}) \cong \mathbb{R}^{2n}$ and $\mathbb{C}_n[x] \cong \mathbb{C}^{n+1}$.

§1.3.1 Representative matrices

Recalling that we can associate to each matrix $\mathrm{A} \in \mathbb{K}^{m \times n}$ an application $L_{\mathrm{A}} \in \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m) : \mathbf{v} \mapsto \mathrm{A}\mathbf{v}$, it is clear that $\ker L_{\mathrm{A}}$ is the solution space of the homogeneus linear system determined by $\mathrm{A}\mathbf{x} = \mathbf{0}$, while $\operatorname{ran} L_{\mathrm{A}}$ is the space of all constant terms $\mathbf{b}$ which make the system $\mathrm{A}\mathbf{x} = \mathbf{b}$ solvable. Moreover, the generators of $\operatorname{ran} L_{\mathrm{A}}$ are the images of the generators of $\mathbb{K}^n$: taking the Euclidean base $\{\mathbf{e}_j\}_{j=1,\dots,n}$, then:

$$L_{\mathrm{A}}(\mathbf{e}_j) = \begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

We see, then, that the n columns of A are the n column vectors which generate $\operatorname{ran} L_{\mathrm{A}}$.
Now, the converse is possible too, i.e. to associate a matrix to a linear application. Consider two $\mathbb{K}$-vector spaces V, W with respective bases $\mathcal{A} = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}, \mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$, and take $f \in \mathrm{Hom}_{\mathbb{K}}(V,W)$. By linearity, f is determined by its values on $\mathcal{A}$, so suppose that:

$$\begin{aligned} f(\mathbf{a}_1) &= \alpha_{11}\mathbf{b}_1 + \cdots + \alpha_{1m}\mathbf{b}_m \\ &\vdots \\ f(\mathbf{a}_n) &= \alpha_{n1}\mathbf{b}_1 + \cdots + \alpha_{nm}\mathbf{b}_m \end{aligned} \qquad \Longrightarrow \qquad \mathrm{A} \equiv \begin{bmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{bmatrix} \tag{1.5}$$

We want to show that f and L_{A} are the "same" application, i.e. we want to show that the following diagram commutes:

$$\begin{array}{ccc} V & \xrightarrow{\ f\ } & W \\ {\scriptstyle \varphi_{\mathcal{A}}}\downarrow & & \downarrow{\scriptstyle \varphi_{\mathcal{B}}} \\ \mathbb{K}^n & \xrightarrow[L_{\mathrm{A}}]{} & \mathbb{K}^m \end{array}$$

where $\varphi_{\mathcal{A}} : V \to \mathbb{K}^n$ and $\varphi_{\mathcal{B}} : W \to \mathbb{K}^m$ are the representations of V and W on $\mathbb{K}^n$ and $\mathbb{K}^m$ in

the respective bases, defined as:

$$V \ni \lambda_1\mathbf{a}_1 + \cdots + \lambda_n\mathbf{a}_n = \mathbf{v} \mapsto \begin{pmatrix}\lambda_1\\ \vdots \\ \lambda_n\end{pmatrix} \in \mathbb{K}^n \qquad W \ni \mu_1\mathbf{b}_1 + \cdots + \mu_m\mathbf{b}_m = \mathbf{w} \mapsto \begin{pmatrix}\mu_1\\ \vdots \\ \mu_m\end{pmatrix} \in \mathbb{K}^m$$

Now, we can directly verify that $L_\mathrm{A} \circ \varphi_\mathcal{A} = \varphi_\mathcal{B} \circ f$, and in particular it is sufficient to show it on a basis:

$$L_\mathrm{A} \circ \varphi_\mathcal{A}(\mathbf{a}_j) = L_\mathrm{A}(\mathbf{e}_j) = \begin{pmatrix}\alpha_{1j}\\ \vdots \\ \alpha_{mj}\end{pmatrix} = \varphi_\mathcal{B}(\alpha_{1j}\mathbf{b}_1 + \cdots + \alpha_{mj}\mathbf{b}_m) = \varphi_\mathcal{B} \circ f(\mathbf{a}_j)$$

Hence, the association between matrices and linear applications is bidirectional and well-defined, and in fact it defines an isomorphism $\mathrm{Hom}_\mathbb{K}(V,W) \cong \mathbb{K}^{m\times n}$.

Definition 1.3.4 (Representative matrix)

Let V, W be finitely-generated $\mathbb{K}$-vector spaces with respective bases $\mathcal{A}, \mathcal{B}$. Then, the **representative matrix** of $f \in \mathrm{Hom}_\mathbb{K}(V,W)$ is the matrix $\mathrm{M}_\mathcal{B}^\mathcal{A}(f)$ determined by the isomorphism $\mathrm{Hom}_\mathbb{K}(V,W) \leftrightarrow \mathbb{K}^{m\times n} : f \leftrightarrow \mathrm{M}_\mathcal{B}^\mathcal{A}(f)$ defined by Eq. 1.5.

Lemma 1.3.4 (Basic properties of representative matrices)

Given three finitely-generated $\mathbb{K}$-vector spaces X, Y, Z with respective bases $\mathcal{A}, \mathcal{B}, \mathcal{C}$ and $f \in \mathrm{Hom}_\mathbb{K}(V,W), g \in \mathrm{Hom}_\mathbb{K}(W,Z)$, then:

a. $\mathrm{M}_\mathcal{C}^\mathcal{A}(g \circ f) = \mathrm{M}_\mathcal{C}^\mathcal{B}(g) \cdot \mathrm{M}_\mathcal{B}^\mathcal{A}(f)$

b. $V = W \wedge \mathcal{A} = \mathcal{B} \implies \mathrm{M}_\mathcal{A}^\mathcal{A}(\mathrm{id}_V) = \mathrm{I}_{\dim_\mathbb{K} V}$

c. f isomorphism $\implies \mathrm{M}_\mathcal{B}^\mathcal{A}(f)$ invertible $\wedge\ [\mathrm{M}_\mathcal{B}^\mathcal{A}(f)]^{-1} = \mathrm{M}_\mathcal{A}^\mathcal{B}(f^{-1})$

Proof. The first two propositions are true by the linearity of f and g, while the last one is proved solving $f^{-1} \circ f = \mathrm{id}_V \implies \mathrm{M}_\mathcal{A}^\mathcal{B}(f^{-1}) \cdot \mathrm{M}_\mathcal{B}^\mathcal{A}(f) = \mathrm{id}_{\dim_\mathbb{K} V}$, where the first two properties where applied. □

When the bases $\mathcal{A}$ and $\mathcal{B}$ are both the respective canonical bases of V and W, then we denote the representative matrix of f simply as M_f.

§1.3.1.1 Change of bases

To discuss how to perform a change of basis in a vector space, first we have to introduce two equivalence relations.

Definition 1.3.5 (Equivalent matrices)

Two matrices $\mathrm{A}, \mathrm{B} \in \mathbb{K}^{m\times n}$ are **equivalent** if $\exists \mathrm{E} \in \mathrm{GL}(m,\mathbb{K}), \mathrm{F} \in \mathrm{GL}(n,\mathbb{K}) : \mathrm{B} = \mathrm{EAF}$.

Definition 1.3.6 (Similar matrices)

Two square matrices $\mathrm{A}, \mathrm{B} \in \mathbb{K}^{n\times n}$ are **similar** if $\exists \mathrm{N} \in \mathrm{GL}(n,\mathbb{K}) : \mathrm{B} = \mathrm{N}^{-1}\mathrm{AN}$.

To illustrate how representation matrices change under a change of basis, consider a $\mathbb{K}$-vector space V and two bases $\mathcal{A}, \mathcal{B} \subseteq V$ (we denote $V_\mathcal{A}, V_\mathcal{B}$ the space with basis $\mathcal{A}$ and $\mathcal{B}$ respectively), and take $f \in \operatorname{End} V$. Then, consider the following commutative diagram:

$$\begin{array}{ccc} V_\mathcal{A} & \xrightarrow{f} & V_\mathcal{A} \\ {\scriptstyle \mathrm{id}_V}\downarrow & & \downarrow{\scriptstyle \mathrm{id}_V} \\ V_\mathcal{B} & \xrightarrow{f} & V_\mathcal{B} \end{array} \quad \Longrightarrow \quad \underbrace{\mathrm{M}_\mathcal{B}^\mathcal{B}(f)}_{\in \mathbb{K}^{n\times n}} \cdot \underbrace{\mathrm{M}_\mathcal{B}^\mathcal{A}(\mathrm{id}_V)}_{\in \mathrm{GL}(n,\mathbb{K})} = \underbrace{\mathrm{M}_\mathcal{B}^\mathcal{A}(\mathrm{id}_V)}_{\in \mathrm{GL}(n,\mathbb{K})} \cdot \underbrace{\mathrm{M}_\mathcal{A}^\mathcal{A}(f)}_{\mathbb{K}^{n\times n}}$$

Hence, we see that representative matrices of the same endomorphism are similar. Moreover, we can define the change-of-basis matrix $\mathrm{N}_\mathcal{B}^\mathcal{A} \equiv \mathrm{M}_\mathcal{B}^\mathcal{A}(\mathrm{id}_V)$, whose columns are the coefficients of the representations on $\mathcal{B}$ of the vectors of $\mathcal{A}$. Note that, in the particular case $f = \mathrm{id}_V$, the above equation proves that $[\mathrm{N}_\mathcal{B}^\mathcal{A}]^{-1} = \mathrm{N}_\mathcal{A}^\mathcal{B}$.
A similar diagram can be drawn for the generalized case of $f \in \mathrm{Hom}_\mathbb{K}(V, W)$:

$$\begin{array}{ccc} V_\mathcal{A} & \xrightarrow{f} & W_\mathcal{B} \\ {\scriptstyle \mathrm{id}_V}\downarrow & & \downarrow{\scriptstyle \mathrm{id}_W} \\ V_{\mathcal{A}'} & \xrightarrow{f} & W_{\mathcal{B}'} \end{array} \quad \Longrightarrow \quad \underbrace{\mathrm{M}_{\mathcal{B}'}^{\mathcal{A}'}(f)}_{\in \mathbb{K}^{m\times n}} \cdot \underbrace{\mathrm{N}_\mathcal{A}^{\mathcal{A}'}}_{\in \mathrm{GL}(n,\mathbb{K})} = \underbrace{\mathrm{N}_{\mathcal{B}'}^\mathcal{B}}_{\in \mathrm{GL}(m,\mathbb{K})} \cdot \underbrace{\mathrm{M}_\mathcal{B}^\mathcal{A}(f)}_{\mathbb{K}^{m\times n}}$$

Therefore, representative matrices of the same linear application are equivalent.

§1.3.2 Determinant and rank

In order to continue our analysis of linear applications, we need to introduce two important notions: the determinant and the rank of a matrix.

§1.3.2.1 Determinant

Definition 1.3.7 (Determinant)

Given a square matrix $\mathrm{A} \in \mathbb{K}^{n\times n} : \mathrm{A} = [a_{ij}]_{i,j=1,\dots,n}$, its **determinant** is defined as:

$$\det \mathrm{A} := \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} \tag{1.6}$$

Note that the determinant has $n!$ terms, each containing one and only one element from each row and each column of A; moreover, it can be interpreted as an application $\det : \mathbb{K}^{n\times n} \to \mathbb{K}$. We can now prove some trivial properties.

Lemma 1.3.5 (Basic properties of determinants)

Let $\mathrm{A} \in \mathbb{K}^{n\times n}$. Then:

a. $\det \mathrm{A}^\intercal = \det \mathrm{A}$

b. Swapping two rows or two columns, the determinant changes sign;

c. If two rows or two columns are equal, then $\det \mathrm{A} = 0$;

d. Keeping $n-1$ columns (or rows) fixed, the determinant is a $\mathbb{K}$-linear application with respect to the other column (or row);

e. $\det(\lambda \mathrm{A}) = \lambda^n \det \mathrm{A}\ \forall \lambda \in \mathbb{K}$

f. $\det \mathrm{I}_n = 1$

Proof. Respectively:

a. Transposition exchanges columns and rows without altering their structure, but each term of the determinant has one and only one element from each row and each column, hence it is unchanged.

b. Swapping two rows or two columns is achieved through a permutation $\rho \in S_n : \operatorname{sgn} \rho = -1$, thus, as $\operatorname{sgn}(\sigma \circ \rho) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\rho) = -\operatorname{sgn} \sigma$, the determinant changes sign.

c. By the previous property, exchanging two equal rows or columns $\det \mathrm{A} = -\det \mathrm{A}$, hence $\det \mathrm{A} = 0$.

d. Linearity follows from the fact that each term of the determinant contains one and only one element of each row and each column.

e. Follows from the previous property, recalling that $\lambda \mathrm{A}$ means multiplying each row (or column) of A by λ, and each term of the determinant has n elements of A.

f. Trivial by direct computation. □

In particular, property (d) shows that the determinant is a $\mathbb{K}$-multilinear applications, as it is linear with respect to each row (or column) of A, while property (b) is easily generalized to:

$$\det(\mathbf{a}_{\sigma(1)}, \dots, \mathbf{a}_{\sigma(n)}) = \operatorname{sgn}(\sigma) \det \mathrm{A} \tag{1.7}$$

where $\mathbf{a}_1, \dots, \mathbf{a}_n$ can denote either the rows ($\in \mathbb{K}^{1\times n}$) or the columns ($\in \mathbb{K}^{n\times 1}$) of A. Furthermore, we can prove a powerful theorem for computing determinants.

Theorem 1.3.4 (Binet theorem)

Given $\mathrm{A}, \mathrm{B} \in \mathbb{K}^{n\times n}$, then:

$$\det(\mathrm{AB}) = \det(\mathrm{A}) \det(\mathrm{B}) \tag{1.8}$$

Proof. Denote the rows of A and B as $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}, \{\mathbf{b}_1, \dots, \mathbf{b}_n\} \in \mathbb{K}^{n\times 1}$ respectively, and set the Euclidean base of $\mathbb{K}^{n\times 1}$ as $\mathbf{e}_j = (0, \dots, 1, \dots, 0)$, so that the rows of AB are:

$$\mathbf{r}_i = \mathbf{a}_i \mathrm{B} = \sum_{j=1}^{n} a_{ij} \mathbf{e}_j \mathrm{B} = \sum_{j=1}^{n} a_{ij} \mathbf{b}_j$$

Then, by the multilinearity of the determinant:

$$\det(\mathrm{AB}) = \det\left(\sum_{j_1=1}^{n} a_{1j_1}\mathbf{b}_{j_1}, \dots, \sum_{j_n=1}^{n} a_{nj_n}\mathbf{b}_{j_n}\right) = \sum_{j_1=1}^{n}\cdots\sum_{j_n=1}^{n} a_{1j_1}\dots a_{nj_n}\det(\mathbf{b}_{j_1}, \dots, \mathbf{b}_{j_n})$$

By Lemma 1.3.5, if $j_i = j_k$ for some $i \neq k$, then the determinant vanishes, so the summation is restricted to $(j_1, \dots, j_n) = (\sigma(1), \dots, \sigma(n))$, with $\sigma \in S_n$, i.e.:

$$\det(\mathrm{AB}) = \sum_{\sigma\in S_n} a_{1\sigma(1)}\dots a_{n\sigma(n)}\det(\mathbf{b}_{\sigma(1)}, \dots, \mathbf{b}_{\sigma(n)}) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)}\dots a_{n\sigma(n)}\det \mathrm{B}$$

where Eq. 1.7 was used. By Def. 1.3.7, the proof is complete. □

The determinant can also be used to establish the linear (in)dependence of a set of vectors.

Proposition 1.3.3

Let $\mathrm{A} \in \mathbb{K}^{n\times n}$ and $\{\mathbf{a}_1, \dots, \mathbf{a}_n\} \in \mathbb{K}^{n\times 1}$ be its columns. Then $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ is LD if and only if $\det \mathrm{A} = 0$.

Proof. ($\Rightarrow$) WLOG $\exists \lambda_2, \dots, \lambda_n \in \mathbb{K} : \mathbf{a}_1 = \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n$, so, by the multilinearity of the determinant:

$$\begin{aligned}\det \mathrm{A} &= \det(\lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n, \mathbf{a}_2, \dots, \mathbf{a}_n)\\ &= \lambda_2\det(\mathbf{a}_2, \mathbf{a}_2, \dots, \mathbf{a}_n) + \cdots + \lambda_n\det(\mathbf{a}_n, \mathbf{a}_2, \dots, \mathbf{a}_n) = \sum_{i=1}^{n}\lambda_i\det(\mathbf{a}_i, \dots, \mathbf{a}_i, \dots) = 0\end{aligned}$$

($\Leftarrow$) Suppose $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ LI, i.e. they form a basis of $\mathbb{K}^{n\times 1}$. Let $\mathrm{B} \equiv [\beta_{ij}] \in \mathrm{GL}(n, \mathbb{K})$ be the matrix representing the change of basis to the Euclidean basis, i.e.:

$$\mathbf{e}_i = \beta_{1i}\mathbf{a}_1 + \cdots + \beta_{ni}\mathbf{a}_n$$

and consider $\mathrm{C} \equiv \mathrm{AB} \in \mathbb{K}^{n\times n}$, whose columns are:

$$\mathbf{c}_i = \sum_{j=1}^{n}\mathbf{a}_j\beta_{ji} = \beta_{1i}\mathbf{a}_1 + \cdots + \beta_{ni}\mathbf{a}_n = \mathbf{e}_i$$

Hence, $\mathrm{C} = \mathrm{I}_n$, but by Binet's theorem $1 = \det \mathrm{C} = \det \mathrm{A}\det \mathrm{B} = 0 \cdot \det \mathrm{B} = 0 \rightarrow\!\!\!\!\leftarrow$ □

Since the determinant is invariant under transposition, this proposition holds considering rows too. We can now prove an alternative way to compute determinants.

Definition 1.3.8 (Submatrices and minors)

Let $\mathrm{M} \in \mathbb{K}^{m\times n}$. Then any matrix obtained by eliminating any number of rows and/or columns from M is a **submatrix** of M, and the determinant of any square submatrix is a **minor** of M.

Given a square matrix $\mathbf{A} = [a_{ij}] \in \mathbb{K}^{n\times n}$, then we can associate to each element a_{ij} the square submatrix $\mathrm{M}_{ij} \in \mathbb{K}^{(n-1)\times(n-1)}$ obtained eliminating the i^{th} row and the j^{th} columns from A: the quantity $\tilde{a}_{ij} \equiv (-1)^{i+j} \det \mathrm{M}_{ij}$ is denoted as the **cofactor** of a_{ij}, and we define the cofactor matrix $\operatorname{cof} \mathrm{A} \equiv [\tilde{a}_{ij}] \in \mathbb{K}^{n\times n}$.

Example 1.3.3 (2×2 matrices)

Consider $\mathrm{A} \in \mathbb{K}^{2\times 2}$. The cofactors then are:

$$\tilde{a}_{11} = (-1)^{1+1} \det[a_{22}] = a_{22} \qquad \tilde{a}_{12} = (-1)^{1+2} \det[a_{21}] = -a_{21}$$
$$\tilde{a}_{21} = (-1)^{2+1} \det[a_{12}] = -a_{12} \qquad \tilde{a}_2 2 = (-1)^{2+2} \det[a_{11}] = a_{11}$$

So, in general:

$$\mathrm{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Longrightarrow \quad \operatorname{cof} \mathrm{A} = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$$

Cofactors allow us to compute determinants without explicitly dealing with permutations.

Theorem 1.3.5 (Laplace's Theorem)

Let $\mathrm{A} = [a_{ij}] \in \mathbb{K}^{n\times n}$. Then:

$$\sum_{k=1}^{n} a_{ik}\tilde{a}_{jk} = \sum_{k=1}^{n} a_{ki}\tilde{a}_{kj} = \delta_{ij} \det \mathrm{A} \quad \forall i, j \in \{1, \dots, n\} \tag{1.9}$$

Proof. The proofs of both equalities are analogous, so WLOG we prove the second one.
$(i = j)$ Consider $k, i \in \{1, \dots, n\}$ fixed and denote as $\mathrm{M}_{ki} = [m_{p,q}]_{p,q=1,\dots,n-1}$ the submatrix obtained by eliminating the k^{th} row and the i^{th} column. Then, each term in the expansion of $\det \mathrm{A}$ which contains a_{ki} can be rewritten as:

$$\operatorname{sgn}(\sigma) a_{1\sigma(1)} \dots a_{ki} \dots a_{n\sigma(n)} = \operatorname{sgn}(\tau) a_{ki} m_{1,\tau(1)} \dots m_{n-1,\tau(n-1)}$$

WTS that $\{\sigma \in S_n : \sigma(k) = i\} \leftrightarrow S_{n-1}$ is a bijection. Firs, the map $\sigma \mapsto \tau$ is:

$$\tau = \begin{pmatrix} 1 & \dots & k-1 & k & \dots & n-1 \\ \gamma_i \circ \sigma(1) & \dots & \gamma_i \circ \sigma(k-1) & \gamma_i \circ \sigma(k+1) & \dots & \gamma_i \circ \sigma(n-1) \end{pmatrix}$$

where $\gamma_i \equiv (n, n-1, \dots, i+1, i)$ is a cycle which decrements all indices larger than i. Then, to construct the map $\tau \mapsto \sigma$, consider $\tau' \in S_n : \tau'(p) = \tau(p)\ \forall p \in \{1, \dots, n-1\} \wedge \tau'(n) = n$, so:

$$\tau' = \begin{pmatrix} 1 & \dots & k-1 & k & \dots & n-1 & n \\ \gamma_i \circ \sigma(1) & \dots & \gamma_i \circ \sigma(k-1) & \gamma_i \circ \sigma(k+1) & \dots & \gamma_i \circ \sigma(n-1) & n \end{pmatrix}$$

Now, consider the following compositions:

$$\tau' \circ \gamma_k = \begin{pmatrix} 1 & \dots & k-1 & k & k+1 & \dots & n-1 & n \\ \gamma_i \circ \sigma(1) & \dots & \gamma_i \circ \sigma(k-1) & n & \gamma_i \circ \sigma(k+1) & \dots & \gamma_i \circ \sigma(n-1) & \gamma_i \circ \sigma(n) \end{pmatrix}$$

$$\gamma_i \circ \sigma = \begin{pmatrix} 1 & \dots & k-1 & k & k+1 & \dots & n-1 & n \\ \gamma_i \circ \sigma(1) & \dots & \gamma_i \circ \sigma(k-1) & n & \gamma_i \circ \sigma(k+1) & \dots & \gamma_i \circ \sigma(n-1) & \gamma_i \circ \sigma(n) \end{pmatrix}$$

Therefore, these are the same permutation, i.e. $\sigma = \gamma_i^{-1} \circ \tau' \circ \gamma_k$, proving the bijection $\sigma \leftrightarrow \tau$. As the cycle γ_k can be written as $n-k$ transpositions, $\operatorname{sgn}\gamma_k = (-1)^{n-k}$, hence:

$$\operatorname{sgn}\sigma = (-1)^{2n-k-i}\operatorname{sgn}\tau' = (-1)^{k+i}\operatorname{sgn}\tau$$

Putting everything together:

$$\sigma(\sigma)a_{1\sigma(1)}\dots a_{ki}\dots a_{n\sigma(n)} = a_{ki}(-1)^{k+i}\operatorname{sgn}(\tau)m_{1,\tau(1)}\dots m_{n-1,\tau(n-1)}$$

Then, the expansion of $\det \mathrm{A}$ can be rewritten as:

$$\begin{aligned}\det \mathrm{A} &= \sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)a_{1\sigma(1)}\dots a_{n\sigma(n)} = \sum_{k=1}^{n}\sum_{\sigma\in S_n:\,\sigma(k)=i}\operatorname{sgn}(\sigma)a_{1\sigma(1)}\dots a_{ki}\dots a_{n\sigma(n)} \\ &= \sum_{k=1}^{n} a_{ki}(-1)^{k+i}\sum_{\tau\in S_{n-1}}\operatorname{sgn}(\tau)m_{1,\tau(1)}\dots m_{n-1,\tau(n-1)} =: \sum_{k=1}^{n} a_{ki}(-1)^{k+i}\det \mathrm{M}_{ki}\end{aligned}$$

which is valid $\forall i \in \{1,\dots,n\}$. Using the definition of cofactor $\tilde{a}_{ki} \equiv (-1)^{k+i}\det \mathrm{M}_{ki}$ concludes the proof.
$(i \neq j)$ Consider $i \neq j \in \{1,\dots,n\}$ and define a matrix B by replacing the j^{th} column of A with the i^{th} one: then $\det \mathrm{B} = 0$. Applying the just-proved expansion to B yields:

$$0 = \det \mathrm{B} = \sum_{k=1}^{n} b_{kj}\tilde{b}_{kj}$$

Now, note that $b_{kj} = a_{ki}$ by definition, while $\tilde{b}_{kj} = \tilde{a}_{kj}$, as A and B only differ by the j^{th} column, which does not affect the definition of the cofactors of the j^{th} column's elements. Hence, the proof is complete. □

Moreover, cofactors also allow us to compute inverse matrices in a straightforward way.

Proposition 1.3.4 (Inverse matrix from cofactors)

Let $\mathrm{A} \in \mathbb{K}^{n\times n}$. Then, A is invertible if and only if $\det \mathrm{A} \neq 0$, and:

$$\mathrm{A}^{-1} = \frac{1}{\det \mathrm{A}}(\operatorname{cof}\mathrm{A})^{\intercal} \tag{1.10}$$

Proof. $(\Rightarrow)$ Invertibility means that $\exists \mathrm{A}^{-1} \in \mathbb{K}^{n\times n} : \mathrm{A}^{-1}\mathrm{A} = \mathrm{A}\mathrm{A}^{-1} = \mathrm{I}_n$, so $\det(\mathrm{A}^{-1}\mathrm{A}) = 1$, but $\det(\mathrm{A}^{-1}\mathrm{A}) = \det(\mathrm{A}^{-1})\det(\mathrm{A})$ by Binet's theorem, hence $\det \mathrm{A} \neq 0$.
$(\Leftarrow)$ To prove this implication, consider $\mathrm{A} = [a_{ij}]_{i,j=1,\dots,n}$. Then, we prove the following lemma.

Lemma 1.3.6

$$\mathrm{A}(\operatorname{cof}\mathrm{A})^{\intercal} = (\operatorname{cof}\mathrm{A})^{\intercal}\mathrm{A} = \det(\mathrm{A})\mathrm{I}_n$$

Proof. The general elements of $\mathrm{A}\,(\operatorname{cof}\mathrm{A})^\intercal$ and $(\operatorname{cof}\mathrm{A})^\intercal\,\mathrm{A}$ are:

$$[\mathrm{A}\,(\operatorname{cof}\mathrm{A})^\intercal]_{ij} = \sum_{k=1}^{n} a_{ik}\tilde{a}_{jk} = \delta_{ij}\det\mathrm{A} \qquad [(\operatorname{cof}\mathrm{A})^\intercal\,\mathrm{A}]_{ij} = \sum_{k=1}^{n} a_{kj}\tilde{a}_{ki} = \delta_{ij}\det\mathrm{A}$$

where Eq. 1.9 was used. This completes the proof. □

Using this lemma, if $\det\mathrm{A} \neq 0$, then it is clear that $\mathrm{A}^{-1} = (\det\mathrm{A})^{-1}\,(\operatorname{cof}\mathrm{A})^\intercal$. □

Corollary 1.3.4.1 (Cramer's Theorem)

Consider a linear system $\mathrm{A}\mathbf{x} = \mathbf{b}$, with $\mathrm{A} \in \mathbb{K}^{n\times n}$ and $\mathbf{b} \in \mathbb{K}^{n\times 1}$. Then:

$$\det\mathrm{A} \neq 0 \implies \exists!\mathbf{x} \in \mathbb{K}^{n\times 1} \text{ solution} : x_i = \frac{1}{\det\mathrm{A}}\det(\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_n)$$

where $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{K}^{n\times 1}$ are the columns of A.

Proof. $\det\mathrm{A} \neq 0 \iff \exists\mathrm{A}^{-1} \in \mathbb{K}^{n\times n}$ by the previous proposition. Then:

$$\mathbf{x} = \mathrm{A}^{-1}\mathbf{b} = \frac{1}{\det\mathrm{A}}(\operatorname{cof}\mathrm{A})^\intercal\,\mathbf{b}$$

Its elements are:

$$x_i = \frac{1}{\det\mathrm{A}}\sum_{k=1}^{n}\tilde{a}_{ij}b_j$$

which, by Eq. 1.9, is the Laplace expansion of the matrix obtained by substituting the i^{th} column of A with $\mathbf{b}$, as the cofactors relative to this column are not changed by this substitution. □

These results clearly show the importance of cofactors and determinants in Linear Algebra.

§1.3.2.2 Rank

Definition 1.3.9 (Rank)

Given $\mathrm{A} \in \mathbb{K}^{m\times n}$ with rows $\mathbf{r}_1, \ldots, \mathbf{r}_m \in \mathbb{K}^{1\times n}$ and columns $\mathbf{c}_1, \ldots, \mathbf{c}_n \in \mathbb{K}^{m\times 1}$, the **rank by rows** of A is defined as $\mathrm{rk_r}(\mathrm{A}) \coloneqq \dim_{\mathbb{K}}\langle\mathbf{r}_1, \ldots, \mathbf{r}_m\rangle$, and the **rank by columns** as $\mathrm{rk_c}(\mathrm{A}) \coloneqq \dim_{\mathbb{K}}\langle\mathbf{c}_1, \ldots, \mathbf{c}_n\rangle$.

The rank by rows (or columns) is just the number of LI rows (or columns). We can show that the two ranks are the same.

Proposition 1.3.5

Given $\mathrm{A} \in \mathbb{K}^{m\times n}$, then $\mathrm{rk_r}(\mathrm{A}) = \mathrm{rk_c}(\mathrm{A})$.

Proof. Set $r \equiv \mathrm{rk_r}(\mathrm{A})$ and $c \equiv \mathrm{rk_c}(\mathrm{A})$. If $r = 0$, then $\mathrm{A} \equiv 0_{m\times n}$ and $c = 0$ too, so consider $r > 0$. A LD relation between $\mathbf{c}_1, \dots, \mathbf{c}_n$ can be written as:

$$x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n = \mathbf{0} \quad \Longrightarrow \quad \exists \mathbf{x} \equiv \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \neq \mathbf{0} \ : \ \mathrm{A}\mathbf{x} = \mathbf{0}$$

Now, consider the associated linear application $L_\mathrm{A} \in \mathrm{Hom}_\mathbb{K}(\mathbb{K}^n, \mathbb{K}^m)$: then $\dim_\mathbb{K} \operatorname{ran} L_\mathrm{A} = n - \dim_\mathbb{K} \ker L_\mathrm{A}$ by Eq. 1.4. But $\dim_\mathbb{K} \operatorname{ran} L_\mathrm{A}$ is the number of LI columns of A (as the columns of A are the images of the Euclidean basis of $\mathbb{K}^n$), so $\dim_\mathbb{K} \operatorname{ran} L_\mathrm{A} = c$, while $\dim_\mathbb{K} \ker L_\mathrm{A} = \dim_\mathbb{K}\{\mathbf{x} \in \mathbb{K}^{n\times 1} : \mathrm{A}\mathbf{x} = \mathbf{0}\}$ by the above equation.
WLOG let $\mathbf{r}_1, \dots, \mathbf{r}_r$ the r LI rows of A, so $\mathbf{r}_{r+1}, \dots, \mathbf{r}_m$ are linear combinations of $\{\mathbf{r}_1, \dots, \mathbf{r}_r\}$. Then, the linear system reduces to $\mathrm{A}'\mathbf{x} = \mathbf{0}$, where $\mathrm{A}' \in \mathbb{K}^{r\times n}$ is only composed of $\mathbf{r}_1, \dots, \mathbf{r}_r$, hence $\mathrm{rk_c}(\mathrm{A}') = \mathrm{rk_c}(\mathrm{A}) = c$ since they represent equivalent systems. But the columns of A′ are vectors in $\mathbb{K}^{r\times 1}$, thus $c \leq r$.
The same reasoning can be applied to $\mathrm{A}^\intercal$, finding $r \leq c$, therefore $r = c$. □

We then set the **rank** of A to be $\mathrm{rk}(\mathrm{A}) \equiv \mathrm{rk_r}(\mathrm{A}) = \mathrm{rk_c}(\mathrm{A})$, and also $\mathrm{rk}(\mathrm{A}) = \mathrm{car}(\mathrm{A})$, that is the number of the so-called "pivots" of A. We can now prove some trivial properties of the rank.

Lemma 1.3.7 (Basic properties of rank)

a. If $\mathrm{A} \in \mathbb{K}^{m\times n}$ and B is a submatrix of A, then $\operatorname{rk} \mathrm{B} \leq \operatorname{rk} \mathrm{A}$;

b. If $\mathrm{A} \in \mathbb{K}^{n\times n}$, then A is invertible if and only if $\operatorname{rk} \mathrm{A} = n$.

Proof. Respectively:

a. Let $\mathrm{A} = [a_{ij}] \in \mathbb{K}^{m\times n}$ and B be formed by rows $i_1, \dots, i_p$ and columns $j_1, \dots, j_q$ of A. Moreover, define C as the submatrix of A formed by the same rows of B and all the columns of A: then, obviously $\mathrm{rk_c}(\mathrm{B}) \leq \mathrm{rk_c}(\mathrm{C})$ and $\mathrm{rk_r}(\mathrm{C}) \leq \mathrm{rk_r}(\mathrm{A})$, hence $\operatorname{rk} \mathrm{B} \leq \operatorname{rk} \mathrm{A}$.

b. A is invertible if and only if $\det \mathrm{A} \neq 0$, which, by Prop. 1.3.3, is equivalent to all the columns of A being LI, i.e. $\operatorname{rk} \mathrm{A} = n$. □

With the notion of rank defined, we can prove the Rouché–Capelli theorem.

Theorem 1.3.6 (Rouché–Capelli Theorem)

Let $\mathrm{A} \in \mathbb{K}^{m\times n}$ and $\mathbf{b} \in \mathbb{K}^{m\times 1}$. Then, the lineas system $\mathrm{A}\mathbf{x} = \mathbf{b}$ has solutions if and only if $\operatorname{rk} \mathrm{A} = \operatorname{rk} [\mathrm{A}|\mathbf{b}]$ and, in this case, it is of kind ∞^{n-r}, with $r \equiv \operatorname{rk} \mathrm{A}$.

Proof. Consider the associated linear application $L_\mathrm{A} \in \mathrm{Hom}_\mathbb{K}(\mathbb{K}^n, \mathbb{K}^m)$: then, the system has solutions if and only if $\mathbf{b} \in \operatorname{ran} L_\mathrm{A} = \langle \mathbf{c}_1, \dots, \mathbf{c}_n \rangle$, where $\mathbf{c}_1, \dots, \mathbf{c}_n \in \mathbb{K}^{m\times 1}$ are the columns of A. But $\mathbf{b} \in \langle \mathbf{c}_1, \dots, \mathbf{c}_n \rangle \iff \langle \mathbf{c}_1, \dots, \mathbf{c}_n \rangle = \langle \mathbf{c}_1, \dots, \mathbf{c}_n, \mathbf{b} \rangle$, which is equivalent to $\dim_\mathbb{K} \langle \mathbf{c}_1, \dots, \mathbf{c}_n \rangle = \dim_\mathbb{K} \langle \mathbf{c}_1, \dots, \mathbf{c}_n, \mathbf{b} \rangle$, i.e. $\operatorname{rk} \mathrm{A} = \operatorname{rk} [\mathrm{A}|\mathbf{b}]$.

Now, assume the system has solutions and $\operatorname{rk} A = \operatorname{rk}[A|\mathbf{b}] \equiv r$. WLOG, let $\mathbf{r}_1, \dots, \mathbf{r}_r$ be the r LI rows of A, so that the equations from the $(r+1)^{\text{th}}$ to the m^{th} can be eliminated from the linear system. Then:

$$A\mathbf{x} = \mathbf{b} \iff \begin{cases} a_{11}x_1 + \cdots + a_{1r}x_r + a_{1(r+1)}x_{r+1} + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{r1}x_1 + \cdots + a_{rr}x_r + a_{r(r+1)}x_{r+1} + \cdots + a_{rn}x_n = b_r \end{cases}$$

$x_{r+1}, \dots, x_n$ can be WLOG interpreted as free parameters, which can then be absorbed into the constant terms: in this way, the system reduces to $A'\mathbf{x}' = \mathbf{b}'$, with $A \in \mathbb{K}^{r\times r}$ and $\mathbf{x}', \mathbf{b} \in \mathbb{K}^{r\times 1}$. By hypothesis $\operatorname{rk} A = r$, hence it is invertible, and the system has a unique solution $\mathbf{x}' = A^{-1}\mathbf{b}$ which depends on $n - r$ parameters, i.e. of kind ∞^{n-r}. □

Finally, rank can be used to prove some properties of linear applications too.

Proposition 1.3.6

Given two finite-dimensional $\mathbb{K}$-vector spaces V, W and an application $f \in \operatorname{Hom}_{\mathbb{K}}(V, W)$ with representative matrix $M_f \in \mathbb{K}^{m\times n}$, then:

a. $\dim_{\mathbb{K}} \operatorname{ran} f = \operatorname{rk} M_f$

b. $\dim_{\mathbb{K}} \ker f = n - \operatorname{rk} M_f$

c. If $n = m$ and $\operatorname{rk} M_f = n$, then f is an isomorphism.

Proof. As $\dim_{\mathbb{K}} \operatorname{ran} f = \dim_{\mathbb{K}} \langle f(\mathbf{e}_1), \dots, f(\mathbf{e}_n)\rangle$, with $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ the canonical basis of V, it is clear that $\dim_{\mathbb{K}} \operatorname{ran} f = \operatorname{rk} M_f$ by Def. 1.3.9. Then, $\dim_{\mathbb{K}} \ker f = n - \operatorname{rk} M_f$ by the rank–nullity theorem (Th. 1.3.2).
Now, consider $n = m$, i.e. $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W$. If $\operatorname{rk} M_f = n$, then $\operatorname{rk} \operatorname{ran} f = \dim_{\mathbb{K}} W$ and $\dim_{\mathbb{K}} \ker f = 0$, so f is both a surjection and an injection, i.e. a bijection. □

§1.3.3 Eigenvalues and eigenvectors

Consider two $\mathbb{K}$-vector spaces V, W of dimensions $\dim_{\mathbb{K}} V = n, \dim_{\mathbb{K}} W = m$ and an application $f \in \operatorname{Hom}_{\mathbb{K}}(V, W)$. Then, by Prop. 1.3.6, is possible to chose two bases $\mathcal{B} \subseteq V, \mathcal{C} \subseteq W$ such that:

$$M_{\mathcal{C}}^{\mathcal{B}}(f) = \begin{bmatrix} I_k & 0_{n-k} \\ 0_{m-k} & 0_{(m-k)\times(n-k)} \end{bmatrix}$$

with $k \equiv \dim_{\mathbb{K}} \operatorname{ran} f$. To show that this is possible, consider $k < n$ (i.e. $\ker f \neq \{\mathbf{0}\}$), so that, by Th. 1.3.2, $\dim_{\mathbb{K}} \ker f = n - k$, and let $\{\mathbf{v}_{k+1}, \dots, \mathbf{v}_n\} \subseteq V$ be a basis of $\ker f$ and WLOG $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ its extension to a basis of V: then, $\{f(\mathbf{v}_1), \dots, f(\mathbf{v}_k)\} \subseteq W$ is a basis of $\operatorname{ran} f$, and WLOG $\mathcal{C} = \{f(\mathbf{v}_1), \dots, f(\mathbf{v}_k), \mathbf{w}_{k+1}, \dots, \mathbf{w}_m\}$ its extension to a basis of W. It is trivial to see that $M_{\mathcal{C}}^{\mathcal{B}}(f)$ has the desired form, for the so-defined bases.
A matrix of this form resembles a diagonal matrix, and it is in fact diagonal if $n = m$. A particular such case is $W = V$, i.e. that of endomorphisms that can be diagonalized, which

we now analyze. Note that our discussion in this section is limited to finite-dimensional vector spaces, for which the concepts used are well-defined.

Definition 1.3.10 (Diagonalizable endomorphism)

Let V be a $\mathbb{K}$-vector space. Then, $f \in \operatorname{End} V$ is **diagonalizable** if $\exists \mathcal{B} \subseteq V$ basis of V, called **diagonalizing basis**, such that:

$$\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

for some $\lambda_1, \dots, \lambda_n \in \mathbb{K}$.

Lemma 1.3.8 (Diagonalizable matrix)

Given $\mathrm{A} \in \mathbb{K}^{n \times n}$, then A is diagonalizable if and only if it is similar to a diagonal matrix.

Proof. ($\Rightarrow$) $\mathrm{A} \in \mathbb{K}^{n \times n}$ diagonalizable is equivalent to $L_{\mathrm{A}} \in \operatorname{End} \mathbb{K}^n$ diagonalizable, which means that $\exists \mathcal{B} \subseteq \mathbb{K}^n : \mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(L_{\mathrm{A}})$ is diagonal. Denoting the canonical basis of $\mathbb{K}^n$ as $\mathcal{E}$, then $\mathrm{A} \equiv \mathrm{M}_{\mathcal{E}}^{\mathcal{E}}(L_{\mathrm{A}})$ and $\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(L_{\mathrm{A}}) = [\mathrm{N}_{\mathcal{E}}^{\mathcal{B}}]^{-1} \mathrm{M}_{\mathcal{E}}^{\mathcal{E}}(L_{\mathrm{A}}) \mathrm{N}_{\mathcal{E}}^{\mathcal{B}}$, which shows that A is similar to a diagonal matrix.
($\Leftarrow$) Let $\mathrm{S} \in \mathrm{GL}(n, \mathbb{K}) : \mathrm{S}^{-1}\mathrm{AS} = \mathrm{D}$, with $\mathrm{D} \in \mathbb{K}^{n \times n}$ diagonal. Consider $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{K}^{n \times 1}$ the columns of S: as $\det \mathrm{S} \neq 0$, they form a basis $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq \mathbb{K}^n$ since they are LI, and $\mathrm{S} = \mathrm{N}_{\mathcal{E}}^{\mathcal{B}}$ (with $\mathcal{E}$ canonical basis of $\mathbb{K}^n$). Then $\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(L_{\mathrm{A}}) = [\mathrm{N}_{\mathcal{E}}^{\mathcal{B}}]^{-1} \mathrm{M}_{\mathcal{E}}^{\mathcal{E}}(L_{\mathrm{A}}) \mathrm{N}_{\mathcal{E}}^{\mathcal{B}} = \mathrm{S}^{-1}\mathrm{AS} = \mathrm{D}$, i.e. L_{A} is diagonalizable with diagonalizing basis $\mathcal{B}$. □

It is possible to characterize the diagonalizing basis in order to get an algorithm to establish whether an endomorphism is diagonalizable.

Definition 1.3.11 (Eigenvectors and eigenvalues)

Given a $\mathbb{K}$-vector space V and $f \in \operatorname{End} V$, then $\mathbf{v} \in V : \mathbf{v} \neq \mathbf{0}$ is an **eigenvector** of f if $\exists \lambda \in \mathbb{K} : f(\mathbf{v}) = \lambda \mathbf{v}$, called **eigenvalue** relative to $\mathbf{v}$.

The null vector $\mathbf{0}$ is excluded from the formal definition of eigevector since it can be regarded as an eigenvector with $\mathbb{K}$-finitely-many eigenvalues.

Proposition 1.3.7 (Diagonalizing basis)

Given a $\mathbb{K}$-vector space V and $f \in \operatorname{End} V$, then f is diagonalizable if and only if $\exists \mathcal{B} \subseteq V$ basis of V composed of eigenvectors of f.

Proof. ($\Rightarrow$) If $\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, then, given $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, clearly $f(\mathbf{v}_i) = 0 \cdot \mathbf{v}_1 + \dots + \lambda_i \mathbf{v}_i + \dots + 0 \cdot \mathbf{v}_n = \lambda_i \mathbf{v}_i \ \forall i \in \{1, \dots, n\}$.
($\Leftarrow$) Given $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} : f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i \ \forall i \in \{1, \dots, n\}$, then by Def. 1.3.4 $\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) =$

$\operatorname{diag}(\lambda_1, \dots, \lambda_n)$. □

Moreover, the eigenvalues of $f \in \operatorname{End} V$ determine subspaces of V.

Proposition 1.3.8 (Eigenspaces)

Let V be a $\mathbb{K}$-vector space and $f \in \operatorname{End} V$. Then, given an eigenvalue $\lambda \in \mathbb{K}$ of f, the relative **eigenspace** $V_\lambda(f) \coloneqq \{\mathbf{v} \in V : f(\mathbf{v}) = \lambda \mathbf{v}\}$ is a subspace of V.

Proof. Trivially $\{\mathbf{0}\} \in V_\lambda(\mathbf{v})\ \forall \lambda \in \mathbb{K}$. Then, given $\mathbf{v}_1, \mathbf{v}_2 \in V_\lambda(f)$:

$$f(\mu_1 \mathbf{v}_1 + \mu_2 \mathbf{v}_2) = \mu_1 f(\mathbf{v}_1) + \mu_2 f(\mathbf{v}_2) = \mu_1 \lambda \mathbf{v}_1 + \mu_2 \lambda \mathbf{v}_2 = \lambda(\mu_1 \mathbf{v}_1 + \mu_2 \mathbf{v}_2)$$

which shows that $V_\lambda(f) \subseteq V$ is closed under linear combinations, i.e. a subspace of V. □

Example 1.3.4 (Euclidean geometry)

Set $V = \operatorname{Vect}_0(\mathbb{R}^2)$. Then, if f is a reflection with respect to the line r, then it has two eigenspaces: $V_1(f) = \{\mathbf{v} \in V : \mathbf{v} \parallel r\}$ and $V_{-1}(f) = \{\mathbf{v} \in V : \mathbf{v} \perp r\}$. On the other hand, if g is a rotation by an angle $\alpha \in [0, 2\pi)$, then there are three possible cases:

- $\alpha = 0$, i.e. $g = \operatorname{id}_V$ and $V_1(g) = V$;
- $\alpha = \pi$, i.e. $g = -\operatorname{id}_V$ and $V_{-1}(g) = V$;
- $\alpha \in (0, \pi) \cup (\pi, 2\pi)$, and it has no eigenvalues.

We can prove that these eigenspaces are "linearly-independent" from one another.

Theorem 1.3.7 (LI eigenvectors)

Consider a $\mathbb{K}$-vector space V and $f \in \operatorname{End} V$. If $\lambda_1 \neq \ldots \neq \lambda_k$ are distinct eigenvalues of f, then their relative eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ are LI.

Proof. We use induction on k:

- $k = 1$ is true as $\mathbf{v}_1 \neq \mathbf{0}$ is LI.
- Assume that $\mathbf{v}_1, \dots, \mathbf{v}_{k-1}$ are LI.
- Consider $a_1 \mathbf{v}_1 + \cdots + a_k \mathbf{v}_k = \mathbf{0}$, with $a_1, \dots, a_k \in \mathbb{K}$. Then $f(a_1 \mathbf{v}_1 + \cdots + a_k \mathbf{v}_k) = \mathbf{0}$, i.e. $a_1 \lambda_1 \mathbf{v}_1 + \ldots a_k \lambda_k \mathbf{v}_k = \mathbf{0}$, but $\mathbf{0} = \lambda_k \mathbf{0} = a_1 \lambda_k \mathbf{v}_1 + \cdots + a_k \lambda_k \mathbf{v}_k$, so the linear combination reduces to $a_1(\lambda_1 - \lambda_k)\mathbf{v}_1 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k)\mathbf{v}_k = \mathbf{0}$. The eigenvalues are distinct by hypothesis and the eigenvectors are LI by the inductive step, hence $a_1 = \cdots = a_{k-1} = 0$, and $a_k = 0$ as $\mathbf{v}_k \neq \mathbf{0}$, i.e. $\mathbf{v}_1, \dots, \mathbf{v}_k$ are LI. □

By Prop. 1.3.7, if f has $n = \dim_{\mathbb{K}} V$ distinct eigenvalues, then it is diagonalizable, as they form a diagonalizing basis of V by Th. 1.3.7.

§1.3.3.1 Computation of eigenvalues

In order to systematically compute the eigenvalues of $f \in \operatorname{End} V$, note that the condition for $\lambda \in \mathbb{K}$ to be an eigenvalue is $\exists \mathbf{v} \in V : f(\mathbf{v}) = \lambda \mathbf{v} \iff (f - \lambda \operatorname{id}_V)(\mathbf{v}) = \mathbf{0}$, hence the relative eigenspace can be written as $V_\lambda(f) = \ker(f - \lambda \operatorname{id}_V)$.
Translating this in matrix terms, given a basis $\mathcal{B} \subseteq V$ and $\mathrm{A} \equiv \mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f)$, then λ is an eigenvalue of f if $V_\lambda(f) \neq \{\mathbf{0}\}$, which means that $f - \lambda \operatorname{id}_V$ must not be an injection (by Prop. 1.3.1): by Prop. 1.3.6 this is equivalent to $\operatorname{rk}(\mathrm{A} - \lambda \mathrm{I}_n) < n$, i.e. $\det(\mathrm{A} - \lambda \mathrm{I}_n) = 0$ by Prop. 1.3.3.

Definition 1.3.12 (Characteristic polynomial)

Given $\mathrm{A} \in \mathbb{K}^{n \times n}$, its **characteristic polynomial** is $p_{\mathrm{A}}(t) \coloneqq \det(\mathrm{A} - t\mathrm{I}_n) \in \mathbb{K}_n[t]$.

We see then that the eigenvalues of f are the roots of $p_{\mathrm{A}}(\lambda)$, with $\mathrm{A} \equiv \mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f)$. Note that the characteristic polynomial is basis-independent, and so are the eigenvalues of f.

Lemma 1.3.9 (Basis-independence of eigenvalues)

Given V a $\mathbb{K}$-vector space, $\mathcal{B}, \mathcal{C} \subseteq V$ two bases and $f \in \operatorname{End} V$, setting $\mathrm{B} \equiv \mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ and $\mathrm{C} \equiv \mathrm{M}_{\mathcal{C}}^{\mathcal{C}}(f)$, then $p_{\mathrm{B}}(\lambda) = p_{\mathrm{C}}(\lambda)$.

Proof. Let $\mathrm{N} \equiv \mathrm{N}_{\mathcal{C}}^{\mathcal{B}} \in \mathrm{GL}(n, \mathbb{K})$, so that $\mathrm{B} = \mathrm{N}^{-1}\mathrm{C}\mathrm{N}$. Then:

$$\det(\mathrm{B} - \lambda \mathrm{I}_n) = \det(\mathrm{N}^{-1}\mathrm{C}\mathrm{N} - \lambda \mathrm{N}^{-1}\mathrm{I}_n\mathrm{N}) = \det(\mathrm{N}^{-1}) \det(\mathrm{C} - \lambda \mathrm{I}_n) \det(\mathrm{N}) = \det(\mathrm{C} - \lambda \mathrm{I}_n)$$

where we used Binet's theorem. This shows that $p_{\mathrm{B}}(\lambda) = p_{\mathrm{C}}(\lambda)$. □

A particular corollary, obtained setting $\lambda = 0$, is that similar matrices have the same determinant.
The basis-independence of the characteristic polynomial allows us to define the notions of determinant $\det f$ and characteristic polynomial $p_f(\lambda)$ directly for the endomorphism f, since they are well-defined by Lemma 1.3.9. In particular, since $p_f(\lambda) \in \mathbb{K}_n[\lambda]$ with $n = \dim_{\mathbb{K}} V$, if V is finite-dimensional, then f as finitely-many eigenvalues.

Example 1.3.5 (Complex and real endomorphisms)

If $\mathbb{K} = \mathbb{C}$, then $p_f(\lambda)$ has n roots by the fundamental theorem of algebra. On the other hand, if $\mathbb{K} = \mathbb{R}$, then only the real roots of $p_f(\lambda)$ are eigenvalues of f.

In general, then, the eigevalues λ of f are found by solving $p_f(\lambda) = 0$, and then the eigenspaces $V_\lambda(f)$ are found by solving the systems $(\mathrm{A} - \lambda \mathrm{I}_n)\mathbf{x} = \mathbf{0}$.

§1.3.3.2 Multiplicity of eigevalues

From Ruffini's theorem, if $\alpha \in \mathbb{K}$ is a root of $p(t) \in \mathbb{K}[t]$, then $\exists q(t) \in \mathbb{K}[t] : p(t) = (t - \alpha)q(t)$, so in general we define the **multiplicity** of α as $m_p(\alpha) \coloneqq \max_{\mathbb{N}_0}\{k \in \mathbb{N}_0 : \exists q(t) \in \mathbb{K}[t] : p(t) = (t - \alpha)^k q(t)\}$, i.e. the exponent of the highest power of $(t - \alpha)$ that divides $p(t)$. This concept can be extended to eigenvalues in a two-fold way.

Definition 1.3.13 (Multiplicity)

Given a $\mathbb{K}$-vector space V and $f \in \operatorname{End} V$, for an eigenvalue $\lambda \in \mathbb{K}$ of f the **algebraic multiplicity** is $m_{\mathrm{a}}(\lambda) \coloneqq m_{p_f}(\lambda)$ and the **geometric multiplicity** is $m_{\mathrm{g}}(\lambda) \coloneqq \dim_{\mathbb{K}} V_\lambda(f)$.

Though different in nature, we can prove an important relation between the two multiplicities.

Lemma 1.3.10

Given a $\mathbb{K}$-vector space V, $f \in \operatorname{End} V$ and an eigenvalue $\lambda \in \mathbb{K}$, then $1 \leq m_{\mathrm{g}}(\lambda) \leq m_{\mathrm{a}}(\lambda)$.

Proof. As λ is an eigenvalue of f, $V_\lambda(f) \neq \{\mathbf{0}\}$, hence $m_{\mathrm{g}}(\lambda) \geq 1$. Now, let $s \equiv m_{\mathrm{g}}(\lambda)$ and $\{\mathbf{v}_1, \dots, \mathbf{v}_s\} \subseteq V_\lambda(f)$ a basis of $V_\lambda(f)$, and extend it to a basis $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_s, \mathbf{w}_{s+1}, \dots, \mathbf{w}_n\}$ of V. By definition $f(\mathbf{v}_i) = \lambda \mathbf{v}_i \ \forall i = 1, \dots, s$, hence:

$$\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{bmatrix} \lambda \mathrm{I}_s & * \\ 0_{s\times(n-s)} & * \end{bmatrix}$$

Clearly, then, $p_f(t) = (t-\lambda)^s q(t)$, with $q(t) \in \mathbb{K}_{n-s}[t]$ determined by $\{\mathbf{w}_{s+1}, \dots, \mathbf{w}_n\}$, hence $m_{\mathrm{a}}(\lambda) \geq s$, which is the thesis. □

The diagonalizability of an endomorphism is linked to the reducibility of its characteristic polynomial and the multiplicity of its eigenvalues.

Theorem 1.3.8 (Diagonalizability and multiplicity)

Let V be a $\mathbb{K}$-vector space and $f \in \operatorname{End} V$. Then, f is diagonalizable if and only if $p_f(\lambda)$ is fully reducible in $\mathbb{K}$ and $m_{\mathrm{a}}(\lambda) = m_{\mathrm{g}}(\lambda) \ \forall \lambda \in \mathbb{K}$ eigenvalue of f.

Proof. ($\Rightarrow$) Let $\mathcal{B} \subseteq V$ be the diagonalizing basis of f: then, given $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ distinct eigenvalues of f, $\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \operatorname{diag}(\lambda_1, \dots, \lambda_1, \dots, \lambda_r, \dots \lambda_r)$ with $m_i \equiv m_{\mathrm{a}}(\lambda_i) \ \forall i \in \{1, \dots, r\}$, and $p_f(t) = (\lambda_1 - t)^{m_1} \dots (\lambda_r - t)^{m_r}$ with $m_1 + \dots + m_r = n \equiv \dim_{\mathbb{K}} V$. This shows that $p_f(t)$ is fully reducible in $\mathbb{K}$.
Now, consider $V_{\lambda_i}(f) = \ker(f - \lambda_i \operatorname{id}_V)$: by Prop. 1.3.6 $\dim_{\mathbb{K}} V_{\lambda_i}(f) = n - \operatorname{rk}(\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) - \lambda_i \mathrm{I}_n)$, but:

$$\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) - \lambda_i \mathrm{I}_n = \operatorname{diag}(\underbrace{\lambda_1 - \lambda_i, \dots, \lambda_1 - \lambda_i}_{m_1}, \dots, \underbrace{0, \dots, 0}_{m_i}, \dots, \underbrace{\lambda_r - \lambda_i, \dots, \lambda_r - \lambda_i}_{m_r})$$

hence $\operatorname{rk}(\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) - \lambda_i \mathrm{I}_n) = \sum_{j=1}^{n} m_j - m_i = n - m_i$, i.e. $m_{\mathrm{g}}(\lambda_i) = m_i \equiv m_{\mathrm{a}}(\lambda_i)$.
($\Leftarrow$) Let $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ be the distinct eigenvalues of f, with $m_i \equiv m_{\mathrm{g}}(\lambda_i) \ \forall i \in \{1, \dots, r\}$, and be $\mathcal{B}_i = \{\mathbf{v}_{i,1}, \dots, \mathbf{v}_{i,m_i}\} \subseteq V_{\lambda_i}(f)$ a basis of $V_{\lambda_i}(f)$. By hypothesis, the total number of basis-eigenvectors is $m_1 + \dots + m_r = m_{\mathrm{a}}(\lambda_1) + \dots + m_{\mathrm{a}}(\lambda_r) = n$, as $p_f(t)$ is fully reducible in $\mathbb{K}$, so consider the following linear combination:

$$\underbrace{\mu_{1,1}\mathbf{v}_{1,1} + \dots + \mu_{1,m_1}\mathbf{v}_{1,m_1}}_{\mathbf{w}_1} + \dots + \underbrace{\mu_{r,1}\mathbf{v}_{r,1} + \dots + \mu_{r,m_r}\mathbf{v}_{r,m_r}}_{\mathbf{w}_r} = \mathbf{0}$$

Since $\mathbf{w}_i \in V_{\lambda_i}(f)\ \forall i \in \{1,\dots,r\}$, they are LI by Th. 1.3.7, so $\mathbf{w}_1 = \dots = \mathbf{w}_r = \mathbf{0}$. Then, $\mu_{i,1}\mathbf{v}_{i,1} + \dots + \mu_{i,m_i}\mathbf{v}_{i,m_i} = \mathbf{0}\ \forall i \in \{1,\dots,r\}$, but $\mathcal{B}_i$ is a basis of $V_{\lambda_i}(f)$, thus $\mu_{i,1} = \dots = \mu_{i,m_i} = 0\ \forall i \in \{1,\dots,r\}$: this means that $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_r$ is a basis of V, hence f is diagonalizable with diagonalizing basis $\mathcal{B}$. □

§1.4 Inner-product spaces

A particular class of vector spaces is of great interest in mathematics: these are inner-product spaces, in which we can introduce the notion of "distance".

§1.4.1 Dual spaces

In order to define inner-product spaces, we first have to analyze dual spaces, which we will generalize in the next chapter.

Definition 1.4.1 (Dual space)

Given a $\mathbb{K}$-vector space V, its **dual space** is $V^* := \mathrm{Hom}_{\mathbb{K}}(V,\mathbb{K})$, whose elements are **linear forms** (or functionals) on V.

Clearly, V^* is a $\mathbb{K}$-vector space too, since $\mathrm{Hom}_{\mathbb{K}}(V,\mathbb{K})$ has a natural structure of $\mathbb{K}$-vector space due to linearity.

Lemma 1.4.1 (Dual basis)

Given a finite-dimensional $\mathbb{K}$-vector space V with a basis $\mathcal{B} = \{\mathbf{b}_1,\dots,\mathbf{b}_n\} \subseteq V$, then $\dim_{\mathbb{K}} V^* = \dim_{\mathbb{K}} V$ and $\mathcal{B}^* \equiv \{b_1^*,\dots,b_n^*\} : b_i^*(\mathbf{b}_j) = \delta_{ij}\ \forall i,j \in \{1,\dots,n\}$ is a basis of V^*, called **dual basis** relative to $\mathcal{B}$.

Proof. Fix the canonical basis $\mathcal{E} = \{1_{\mathbb{K}}\} \subseteq \mathbb{K}$ of $\mathbb{K}$: then, the application $\varphi : V^* \to \mathbb{K}^{1\times n} : \varphi(\omega) = \mathrm{M}_{\mathcal{E}}^{\mathcal{B}}(\omega)$ is an isomorphism, so $\dim_{\mathbb{K}} V = n$ by Th. 1.3.3.
To prove that $\mathcal{B}^* \subseteq V^*$ is a basis, WTS $\mathcal{B}^*$ is LI. Take $\mathbf{v} \in V : \mathbf{v} = \sum_{i=1}^n v_i\mathbf{b}_i$, so that:

$$b_i^*(\mathbf{v}) = \sum_{j=1}^n v_j b_i^*(\mathbf{b}_j) = \sum_{j=1}^n v_j\delta_{ij} = v_i \quad \Longrightarrow \quad \mathbf{v} = \sum_{i=1}^n b_i^*(\mathbf{v})\mathbf{b}_i$$

Now, consider a linear combination $\lambda_1 b_1^* + \dots + \lambda_n b_n^* = 0 \in V^*$, and rewrite:

$$0_{\mathbb{K}} = 0(\mathbf{b}_i) = \sum_{j=1}^n \lambda_j b_j^*(\mathbf{b}_i) = \sum_{j=1}^n \lambda_j\delta_{ij} = \lambda_i \quad \forall i \in \{1,\dots,n\}$$

This concludes the proof. □

Given the isomorphisms $V \cong \mathbb{K}^n \cong V^*$, it is possible to represent both V and V^* on $\mathbb{K}^n$: in particular, V is usually represente on $\mathbb{K}^{n\times 1}$, while V^* on $\mathbb{K}^{1\times n}$, so that $\omega(\mathbf{v}) \in \mathbb{K}\ \forall \mathbf{v} \in V, \omega \in V^*$.

Definition 1.4.2 (Transposed homomosphism)

Given two $\mathbb{K}$-vector spaces V, W and $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$, its **transposed homomorphism** is the application $f^\intercal : W^* \to V^*$ define as $W^* \ni \omega \mapsto f^\intercal(\omega) \coloneqq \omega \circ f \in V^*$.

As composition preserves linearity, $f^\intercal \in \mathrm{Hom}_{\mathbb{K}}(W^*, V^*)$, making the following diagrams commutative:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ & \searrow^{f^\intercal(\omega)} & \downarrow{\scriptstyle \omega} \\ & & \mathbb{K} \end{array} \iff \begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle \varphi_V}\uparrow\downarrow & & \uparrow\downarrow{\scriptstyle \varphi_W} \\ V^* & \xleftarrow[f^\intercal]{} & W^* \end{array}$$

Proposition 1.4.1

Consider two finite-dimensional $\mathbb{K}$-vector spaces V, W with bases $\mathcal{B} \subseteq V, \mathcal{C} \subseteq W$, and $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. Then:

$$\mathrm{M}^{\mathcal{C}^*}_{\mathcal{B}^*}(f^\intercal) = [\mathrm{M}^{\mathcal{B}}_{\mathcal{C}}(f)]^\intercal \tag{1.11}$$

Proof. Set $n \equiv \dim_{\mathbb{K}} V$ and $m \equiv \dim_{\mathbb{K}} W$. Clearly $\mathrm{M}^{\mathcal{B}}_{\mathcal{C}}(f) \in \mathbb{K}^{m\times n}$ and $\mathrm{M}^{\mathcal{C}^*}_{\mathcal{B}^*}(f^\intercal) \in \mathbb{K}^{n\times m}$, so, fixing the canonical basis $\mathcal{E} = \{1_{\mathbb{K}}\} \subseteq \mathbb{K}$ of $\mathbb{K}$ and given $\omega \in W^*$, $\mathrm{M}^{\mathcal{C}}_{\mathcal{E}}(\omega) = [\beta_1, \dots, \beta_m] \in \mathbb{K}^{1\times m}$, but $\omega = \omega_1 c_1^* + \cdots + \omega_m c_m^*$ on the basis $\mathcal{C}^*$, thus, given $\mathbf{w} \in W$:

$$\begin{aligned} \omega(\mathbf{w}) = \mathrm{M}^{\mathcal{C}}_{\mathcal{E}}(\omega)\mathbf{w} &= \begin{bmatrix} \beta_1 & \dots & \beta_m \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix} = \sum_{i=1}^{m} \beta_i w_i \\ &= \sum_{i=1}^{m} \omega_i c_i^*(\mathbf{w}) = \sum_{i,j=1}^{m} \omega_i w_j c_i^*(\mathbf{c}_j) = \sum_{i,j=1}^{m} \omega_i w_j \delta_{ij} = \sum_{i=1}^{m} \omega_i w_i \end{aligned}$$

Hence, $[\mathrm{M}^{\mathcal{C}}_{\mathcal{E}}(\omega)]^\intercal$ is precisely the representation of $\omega \in W^*$ on $\mathbb{K}^{m\times 1}$ with basis $\mathcal{C}^*$.
Now, since $f^\intercal(\omega) = \omega \circ f \in V^*$, by Lemma 1.3.4 $\mathrm{M}^{\mathcal{B}}_{\mathcal{E}}(f^\intercal(\omega)) = \mathrm{M}^{\mathcal{C}}_{\mathcal{E}}(\omega)\mathrm{M}^{\mathcal{B}}_{\mathcal{C}}(f)$, therefore, by the same reasoning, the representation of $f^\intercal(\omega) \in V^*$ on $\mathbb{K}^{n\times 1}$ with basis $\mathcal{B}^*$ is $[\mathrm{M}^{\mathcal{B}}_{\mathcal{E}}(f^\intercal(\omega))]^\intercal = [\mathrm{M}^{\mathcal{B}}_{\mathcal{C}}(f)]^\intercal[\mathrm{M}^{\mathcal{C}}_{\mathcal{E}}(\omega)]^\intercal$, but $f^\intercal \in \mathrm{Hom}_{\mathbb{K}}(W^*, V^*)$ has representative matrix $\mathrm{M}^{\mathcal{C}^*}_{\mathcal{B}^*}(f^\intercal)$, thus:

$$f^\intercal(\omega) = \mathrm{M}^{\mathcal{B}^*}_{\mathcal{C}^*}(f^\intercal)\boldsymbol{\omega} = \mathrm{M}^{\mathcal{B}^*}_{\mathcal{C}^*}(f^\intercal)[\mathrm{M}^{\mathcal{B}}_{\mathcal{E}}(\omega)]^\intercal$$

Comparing the two expressions for $f^\intercal(\omega)$, the proof is complete. □

§1.4.1.1 Bilinear forms

We can generalize linear forms to multilinear forms, i.e. maps which are linear with respect to each of their arguments. Here, we consider maps with two arguments.

Definition 1.4.3 (Bilinear forms)

Given a $\mathbb{K}$-vector space V, a **bilinear form** is an application $b : V \times V \to \mathbb{K}$ which is linear with respect to both its arguments.

By linearity, it is trivial to see that $b(\mathbf{0},\mathbf{v}) = b(\mathbf{v},\mathbf{0}) = 0_{\mathbb{K}}$. Moreover, we can associate to every $\mathrm{A} \in \mathbb{K}^{n\times n}$ a bilinear form on $b_{\mathrm{A}} : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ defined as $b_{\mathrm{A}}(\mathbf{x},\mathbf{y}) = \mathbf{x}^\intercal \mathrm{A}\mathbf{y}$, i.e., setting $\mathrm{A} = [a_{ij}]$, $b_{\mathrm{A}}(\mathbf{x},\mathbf{y}) = \sum_{i,j=1}^n a_{ij}x_i y_j$. Analogously, we can in general define the representative matrix of $b : V \times V \to \mathbb{K}$ on a basis $\mathcal{B}\{\mathbf{v}_1,\dots,\mathbf{v}_n\} \subseteq V$ as $\mathrm{M}_{\mathcal{B}}(b) \equiv [b(\mathbf{v}_i,\mathbf{v}_j)]_{i,j=1,\dots,n}$, which defines the actions of b on V as:

$$b(\mathbf{v},\mathbf{w}) = \sum_{i,j=1}^{n} v_i w_j b(\mathbf{v}_i,\mathbf{w}_j) = \mathbf{v}^\intercal \mathrm{M}_{\mathcal{B}}(b)\mathbf{w}$$

where the abuse of notation $\mathbf{v},\mathbf{w} \in V$ and $\mathbf{v},\mathbf{w} \in \mathbb{K}^n$ is justified by the isomorphism $V \cong \mathbb{K}^n$. To see how the representative matrices in different bases are related, we need to introduce another equivalence relation for square matrices.

Definition 1.4.4 (Congruent matrices)

Given $\mathrm{A},\mathrm{B} \in \mathbb{K}^{n\times n}$, they are **congruent** if $\exists \mathrm{C} \in \mathrm{GL}(n,\mathbb{K}) : \mathrm{B} = \mathrm{C}^\intercal \mathrm{A}\mathrm{C}$.

Note that congruent matrices have the same rank. Indeed, $\operatorname{rk}\mathrm{A} = \dim_{\mathbb{K}} \operatorname{ran} L_{\mathrm{A}}$, but $\mathrm{C} \in \mathrm{GL}(n,\mathbb{K})$ is invertible, hence it defines an automorphism of $\mathbb{K}^{n\times n}$: then, $\operatorname{rk}(\mathrm{C}\mathrm{A}) = \operatorname{rk}\mathrm{A}$ by Th. 1.3.3, and $\operatorname{rk}(\mathrm{A}\mathrm{C}) = \operatorname{rk}\mathrm{A}$ since $\operatorname{rk}\mathrm{A}^\intercal = \operatorname{rk}\mathrm{A}$.

Theorem 1.4.1 (Congruence and bilinear forms)

Given a $\mathbb{K}$-vector space V, then two matrices represent the same bilinear form $b : V \times V \to \mathbb{K}$ if and only if they are congruent.

Proof. ($\Rightarrow$) Consider $\mathcal{B},\mathcal{B}' \subseteq V$ bases of V, and let $\mathrm{C} \equiv \mathrm{N}_{\mathcal{B}'}^{\mathcal{B}} \in \mathrm{GL}(n,\mathbb{K})$ be the basis-change matrix, so that $\mathbf{v}' = \mathrm{C}\mathbf{v}\ \forall \mathbf{v} \in V$ (with an abuse of notation since $V \cong \mathbb{K}^n$). Then:

$$b(\mathbf{v},\mathbf{w}) = \mathbf{v}^\intercal \mathrm{M}_{\mathcal{B}}(b)\mathbf{w} = (\mathrm{C}\mathbf{v})^\intercal \mathrm{M}_{\mathcal{B}'}(b)(\mathrm{C}\mathbf{w}) = \mathbf{v}^\intercal(\mathrm{C}^\intercal \mathrm{M}_{\mathcal{B}'}(b)\mathrm{C})\mathbf{w} \quad\Longrightarrow\quad \mathrm{M}_{\mathcal{B}}(b) = \mathrm{C}^\intercal \mathrm{M}_{\mathcal{B}'}(b)\mathrm{C}$$

($\Leftarrow$) Consider two congruent matrices $\mathrm{A},\mathrm{B} \in \mathbb{K}^{n\times n} : \exists \mathrm{C} \in \mathrm{GL}(n,\mathbb{K}) : \mathrm{B} = \mathrm{C}^\intercal \mathrm{A}\mathrm{C}$. Then, C determines a change of basis on $\mathbb{K}^{n\times n}$, so that, setting $b_{\mathrm{A}}(\mathbf{v}',\mathbf{w}') = \mathbf{v}'^\intercal \mathrm{A}\mathbf{w}'$:

$$b_{\mathrm{B}}(\mathbf{v},\mathbf{w}) = \mathbf{v}^\intercal \mathrm{B}\mathbf{w} = \mathbf{v}^\intercal \mathrm{C}^\intercal \mathrm{A}\mathrm{C}\mathrm{w} = (\mathrm{C}\mathbf{v})^\intercal \mathrm{A}(\mathrm{C}\mathbf{w}) = \mathbf{v}'^\intercal \mathrm{A}\mathbf{w} = b_{\mathrm{A}}(\mathbf{v}',\mathbf{w}')$$

Hence, A and B represent the same bilinear form in different bases of V. □

By this theorem, it is clear that symmetry is a well-defined property for bilinear forms, since it is preserved by congruence.

Lemma 1.4.2 (Symmetric bilinear form)

Let V be a $\mathbb{K}$-vector space and $b : V \times V \to \mathbb{K}$ a bilinear form on $\mathbb{K}$. Then b is **symmetric**, i.e. $b(\mathbf{v},\mathbf{w}) = b(\mathbf{w},\mathbf{v})\ \forall \mathbf{v},\mathbf{w} \in V$, if and only if its representative matrix is symmetric.

Proof. Set $\mathrm{M}_{\mathcal{B}}(b) = [b_{ij}]$. Then $b(\mathbf{v},\mathbf{w}) = b(\mathbf{w},\mathbf{v}) \iff \sum_{i,j=1}^n v_i w_j b_{ij} = \sum_{i,j=1}^n v_i w_j b_{ji} \iff b_{ij} = b_{ji}\ \forall i,j \in \{1,\dots,n\}$. □

§1.4.2 Euclidean vector spaces

In this subsection, we always consider V to be a $\mathbb{R}$-vector space of dimension $\dim_{\mathbb{R}} V = n \in \mathbb{N}$.

Definition 1.4.5 (Positive-definite bilinear form)

A symmetric bilinear form $b : V \times V \to \mathbb{R}$ is **positive-definite** if $b(\mathbf{v}, \mathbf{v}) \geq 0 \; \forall \mathbf{v} \in V$ and $b(\mathbf{v}, \mathbf{v}) = 0 \iff \mathbf{v} = \mathbf{0}$.

We refer to symmetric positive-definite bilinear forms as **inner products**, and we set the notation $b(\mathbf{v}, \mathbf{w}) \equiv \langle \mathbf{v}, \mathbf{w} \rangle$.

Example 1.4.1 (Canonical inner product)

Consider $V = \mathbb{R}^n$. Then, the **canonical inner product** is $\langle \mathbf{v}, \mathbf{w} \rangle \equiv b_{\mathrm{I}_n}(\mathbf{v}, \mathbf{w}) = \mathbf{v}^\intercal \mathbf{w}$.

Definition 1.4.6 (Euclidean vector space)

V is a **Euclidean vector space** (EVS) if it is possible to define an inner product on it, and it is denoted as $(V, \langle \cdot, \cdot \rangle)$.

In an EVS it is possible to define two important notions (with $\mathbf{v}, \mathbf{w} \in V$):

- **norm**: $\|\mathbf{v}\| := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$
- **distance**: $\mathrm{d}(\mathbf{v}, \mathbf{w}) := \|\mathbf{v} - \mathbf{w}\|$

For the resto of this subsection, we assume that V has the structure of EVS.

Lemma 1.4.3 (Cauchy–Schwarz inequality)

Given $\mathbf{v}, \mathbf{w} \in V$, then:

$$|\langle \mathbf{v}, \mathbf{w} \rangle|^2 \leq \|\mathbf{v}\|^2 \|\mathbf{w}\|^2 \quad \wedge \quad |\langle \mathbf{v}, \mathbf{w} \rangle|^2 = \|\mathbf{v}\|^2 \|\mathbf{w}\|^2 \iff \mathbf{v}, \mathbf{w} \text{ LI} \tag{1.12}$$

Proof. The case $\mathbf{v} = \mathbf{0} \vee \mathbf{w} = \mathbf{0}$ is trivial, so WLOG $\mathbf{v}, \mathbf{w} \neq \mathbf{0}$.
Given $\mathbf{v}, \mathbf{w} \in V$, consider $p(t) \equiv \langle t\mathbf{v} + \mathbf{w}, t\mathbf{v} + \mathbf{w} \rangle \in \mathbb{R}_2[t]$: by definition $p(t) \geq 0 \; \forall t \in \mathbb{R}$, hence, by linearity and symmetry:

$$p(t) = t^2 \|\mathbf{v}\|^2 + 2t \langle \mathbf{v}, \mathbf{w} \rangle + \|\mathbf{w}\|^2 \geq 0 \quad \forall t \in \mathbb{R}$$

To be true, this requires that the discriminant of the polynomial is non-positive, i.e.:

$$\Delta_p \equiv 4 \left(\langle \mathbf{v}, \mathbf{w} \rangle^2 - \|\mathbf{v}\|^2 \|\mathbf{w}\|^2 \right) \leq 0$$

which is the thesis. Moreover, assume $\Delta_p = 0$: then $p(t)$ has a single root λ with multiplicity 2, but $p(t) = \|t\mathbf{v} + \mathbf{w}\|^2$, hence $p(\lambda) = 0 \iff \lambda\mathbf{v} + \mathbf{w} = \mathbf{0}$, i.e. $\mathbf{v}, \mathbf{w}$ LI.
On the other hand, if $\mathbf{v}$ and $\mathbf{w}$ are LI, then $\exists \lambda \in \mathbb{R} : \mathbf{v} = \lambda \mathbf{w}$, so by linearity $\langle \mathbf{v}, \mathbf{w} \rangle^2 = \lambda^2 \langle \mathbf{v}, \mathbf{v} \rangle^2 = \lambda^2 \|\mathbf{v}\|^4 = \|\mathbf{v}\|^2 \|\mathbf{w}\|^2$. □

We can now state some properties of the norm.

Lemma 1.4.4 (Basic properties of norm)

Given $\mathbf{v}, \mathbf{w} \in V$ and $\lambda \in \mathbb{R}$, then:

a. $|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \|\mathbf{v}\| \, \|\mathbf{w}\|$

b. $\|\mathbf{v}\| \geq 0 \wedge \|\mathbf{v}\| = 0 \iff \mathbf{v} = \mathbf{0}$

c. $\|\lambda \mathbf{v}\| = |\lambda| \, \|\mathbf{v}\|$

d. $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$

Proof. Property (a) is a direct consequence of the Cauchy–Schwarz inequality, while (b) and (c) are trivial by the definition of inner product. Now, consider $\mathbf{v}, \mathbf{w} \in V$:

$$\begin{aligned} \|\mathbf{v} + \mathbf{w}\|^2 = \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w} \rangle &= \|\mathbf{v}\|^2 + 2 \langle \mathbf{v}, \mathbf{w} \rangle + \|\mathbf{w}\|^2 \\ &\leq \|\mathbf{v}\|^2 + 2 \|\mathbf{v}\| \, \|\mathbf{w}\| + \|\mathbf{w}\|^2 = (\|\mathbf{v}\| + \|\mathbf{w}\|)^2 \end{aligned}$$

where property (a) was used. This concludes the proof. □

We now introduce some more notation: if $\mathbf{v} \in V : \|\mathbf{v}\| = 1$, then $\mathbf{v}$ is a **versor** (or unit vector). Moreover, given $\mathbf{v}, \mathbf{w} \in V - \{\mathbf{0}\}$, then the **angle** $\vartheta \in [0, \pi]$ between them is defined as:

$$\cos \theta = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\| \, \|\mathbf{w}\|} \tag{1.13}$$

which is well-defined by the Cauchy–Schwarz inequality. Two vectors $\mathbf{v}, \mathbf{w} \in V$ are **orthogonal** if $\langle \mathbf{v}, \mathbf{w} \rangle = 0$, i.e. if their angle is $\vartheta = \frac{\pi}{2}$, and we denote this by $\mathbf{v} \perp \mathbf{w}$.

Definition 1.4.7 (Orthogonal subspace)

Given a subspace $U \subseteq V$, its orthogonal space is $U^\perp \coloneqq \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{u} \rangle = 0 \ \forall \mathbf{u} \in U\}$.

Lemma 1.4.5

Given a subspace $U \subseteq V$ and a basis $\{\mathbf{s}_1, \ldots, \mathbf{s}_k\} \subseteq U$, then $U^\perp = \{\mathbf{v} \in V : \mathbf{v} \perp \mathbf{s}_i \ \forall i = 1, \ldots, k\}$ and it is a subspace of V.

Proof. Consider $\mathbf{v}, \mathbf{w} \in U^\perp$ and $\mathbf{u} \in U$: then $\langle \mathbf{u}, \lambda \mathbf{v} + \mu \mathbf{w} \rangle = \lambda \langle \mathbf{u}, \mathbf{v} \rangle + \mu \langle \mathbf{u}, \mathbf{w} \rangle = 0$, hence $U^\perp$ is closed under linear combinations, i.e. a subspace of V.
Now, since $U = \langle \mathbf{s}_1, \ldots, \mathbf{s}_k \rangle$, $\forall \mathbf{u} \in U \ \exists! u_1, \ldots, u_k \in \mathbb{R} : \mathbf{u} = u_1 \mathbf{s}_1 + \cdots + u_k \mathbf{s}_k$, so:

$$0 = \langle \mathbf{v}, \mathbf{u} \rangle = \sum_{i=1}^{k} u_i \langle \mathbf{v}, \mathbf{s}_i \rangle \quad \forall \mathbf{v} \in U^\perp \quad \iff \quad \langle \mathbf{v}, \mathbf{s}_i \rangle = 0 \quad \forall i \in \{1, \ldots, k\}$$

which is the thesis. □

§1.4.2.1 Orthonormal bases

It is possible to link orthogonality and linear independence in a straightforward way.

Lemma 1.4.6 (Orthogonality and LI)

Given $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subseteq V - \{\mathbf{0}\}$, if $\mathbf{v}_i \perp \mathbf{v}_j \ \forall i \neq j \in \{1, \dots, k\}$, then S is LI.

Proof. Consider $\lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k = \mathbf{0}$, so that:

$$0 = \langle \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k, \mathbf{v}_i \rangle = \lambda_i \|\mathbf{v}_i\|^2 \quad \forall i \in \{1, \dots, k\}$$

Since $\mathbf{v}_i \neq \mathbf{0}$, then $\lambda_i = 0 \ \forall i \in \{1, \dots, k\}$, which is the thesis. □

An important consequence of this lemma is that a set of mutually-orthogonal vectors cannot have more than $n = \dim_{\mathbb{R}} V$ vectors. Now, we can characterize a particular classes of bases of Euclidean spaces: orthonormal bases.

Definition 1.4.8 (Orthonormal basis)

A basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\} \subseteq V$ is **orthonormal** if it is composed by versors such that $\mathbf{b}_i \perp \mathbf{b}_j \ \forall i \neq j \in \{1, \dots, n\}$.

Theorem 1.4.2 (Gram–Schmidt Theorem)

Every non-trivial Euclidea vector space admits an orthonormal basis.

Proof. To prove this theorem we use the following lemma.

Lemma 1.4.7

Given orthonormal versors $\mathbf{a}_1, \dots, \mathbf{a}_k \in V$ and $\mathbf{w} \in V - \langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle$, then:

$$\mathbf{a} \equiv \mathbf{w} - \sum_{i=1}^{k} \langle \mathbf{w}, \mathbf{a}_i \rangle \, \mathbf{a}_i \quad : \quad \mathbf{a} \perp \mathbf{a}_i \ \forall i \in \{1, \dots, k\}$$

Proof. Suppose $\mathbf{a} = \mathbf{0}$: then $\mathbf{w} = \sum_{i=1}^{k} \langle \mathbf{w}, \mathbf{a}_i \rangle \, \mathbf{a}_i \in \langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle \ \not\rightarrow$
Then:

$$\langle \mathbf{a}, \mathbf{a}_i \rangle = \langle \mathbf{w}, \mathbf{a}_i \rangle - \sum_{j=1}^{k} \langle \mathbf{w}, \mathbf{a}_j \rangle \langle \mathbf{a}_j, \mathbf{a}_i \rangle = \langle \mathbf{w}, \mathbf{a}_i \rangle - \sum_{j=1}^{k} \langle \mathbf{w}, \mathbf{a}_j \rangle \, \delta_{ij} = \langle \mathbf{w}, \mathbf{a}_i \rangle - \langle \mathbf{w}, \mathbf{a}_i \rangle = 0$$

which is the thesis. □

Now, consider $\mathbf{v} \in V - \{\mathbf{0}\}$ and set $\mathbf{a}_1 \equiv \frac{\mathbf{v}}{\|\mathbf{v}\|}$: if $V = \langle \mathbf{a}_1 \rangle$ the proof is complete, otherwise $\exists \mathbf{w} \in V - \langle \mathbf{a}_1 \rangle$ and, by Lemma 1.4.7, $\exists \mathbf{a} \in V : \mathbf{a} \perp \mathbf{a}_1$. Set $\mathbf{a}_2 \equiv \frac{\mathbf{a}}{\|\mathbf{a}\|}$: if $V = \langle \mathbf{a}_1, \mathbf{a}_2 \rangle$ the proof is complete, otherwise we iterate this process, which finds an orthonormal basis of V in $n = \dim_{\mathbb{R}} V$ steps. □

This result allows us to define additional structure on V.

Proposition 1.4.2 (Orthogonal decomposition)
Given a subspace $W \subseteq V$, then $V = W \oplus W^\perp$.

Proof. If $W = \{\mathbf{0}\}$ then trivially $W^\perp = V$, so consider $W \neq \{\mathbf{0}\}$: WTS $W \cap W^\perp = \{\mathbf{0}\}$ and $W + W^\perp = V$. First, $\mathbf{w} \neq \mathbf{0} \implies \langle \mathbf{w}, \mathbf{w} \rangle \neq 0 \implies W \cap W^\perp = \{\mathbf{0}\}$. Then, take an orthonormal basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_k\} \subseteq W$: thus $\forall \mathbf{v} \in V \; \exists \mathbf{u}_1 \in W, \mathbf{u}_2 \in W^\perp : \mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2$ by Lemma 1.4.7:

$$\mathbf{u}_1 = \sum_{i=1}^{k} \langle \mathbf{v}, \mathbf{b}_i \rangle \, \mathbf{b}_i \qquad \mathbf{u}_2 = \mathbf{v} - \sum_{i=1}^{k} \langle \mathbf{v}, \mathbf{b}_i \rangle \, \mathbf{b}_i$$

Hence, $W + W^\perp$, which completes the proof. □

Given a subspace $W \subseteq V$ with orthonormal basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_k\}$, we denote the **orthogonal projection** on W as the linear application $\pi_W : V \to W$ defined as:

$$\pi_W(\mathbf{v}) := \sum_{i=1}^{k} \langle \mathbf{v}, \mathbf{b}_i \rangle \, \mathbf{b}_i \tag{1.14}$$

§1.4.2.2 Symmetric endomorphisms

Definition 1.4.9 (Symmetric endomorphisms)
Given $f \in \operatorname{End} V$, it is **symmetric** (or self-adjoint) if $\langle f(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, f(\mathbf{w}) \rangle \;\; \forall \mathbf{v}, \mathbf{w} \in V$.

From the linearity of the inner product, it is clear that the symmetry condition can be checked on the vectors of a particular basis, instead that on the whole vector space.

Lemma 1.4.8 (Symmetric endomorphisms and representative matrices)
Given an orthonormal basis $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ of V, then $f \in \operatorname{End} V$ is symmetric if and only if $\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ is symmetric.

Proof. Set $\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) = [m_{ij}]_{i,j=1,\dots,n}$, so that $f(\mathbf{v}_i) = \sum_{k=1}^{n} m_{ki} \mathbf{v}_k$. Then:

$$\langle f(\mathbf{v}_i), \mathbf{v}_j \rangle = \sum_{k=1}^{n} m_{ki} \delta_{kj} = m_{ji} \qquad \langle \mathbf{v}_i, f(\mathbf{v}_j) \rangle = \sum_{k=1}^{n} m_{kj} \delta_{ik} = m_{ij}$$

Hence, f is symmetric if and only if $m_{ij} = m_{ji}$, i.e. if $\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ is symmetric. □

It is possible to state two necessary conditions for an endomorphism to be symmetric.

Proposition 1.4.3
Given a symmetric $f \in \operatorname{End} V$, then:

a. given two distinct eigenvalues λ, μ of f and $\mathbf{v} \in V_\lambda(f), \mathbf{w} \in V_\mu(f)$, then $\mathbf{v} \perp \mathbf{w}$;

b. the roots of $p_f(t) \in \mathbb{R}_n[t]$ are all real.

Proof. Respectively:

a. As $f(\mathbf{v}) = \lambda\mathbf{v}$ and $f(\mathbf{w}) = \mu\mathbf{w}$, then $\langle f(\mathbf{v}), \mathbf{w}\rangle = \lambda\langle \mathbf{v}, \mathbf{w}\rangle$ and $\langle \mathbf{v}, f(\mathbf{w})\rangle = \mu\langle \mathbf{v}, \mathbf{w}\rangle$, but f is symmetric, hence $(\lambda - \mu)\langle \mathbf{v}, \mathbf{w}\rangle$: since λ and μ are distinct, $\langle \mathbf{v}, \mathbf{w}\rangle = 0$.

b. Consider an orthonormal basis $\mathcal{B} \subseteq V$ and $\mathrm{A} \equiv \mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f)$, so that $\mathrm{A}^\intercal = \mathrm{A} \in \mathbb{R}^{n\times n} \subseteq \mathbb{C}^{n\times n}$ by Lemma 1.4.8. Then, let $\lambda \in \mathbb{C} : p_f(\lambda) = 0$, which means that $\exists \mathbf{v} \in \mathbb{C}^{n\times 1} - \{\mathbf{0}\} : \mathrm{A}\mathbf{v} = \lambda\mathbf{v}$: thus, $\mathbf{v}^\intercal\mathrm{A}^\intercal = \lambda\mathbf{v}^\intercal$ and $\overline{\mathrm{A}\mathbf{v}} = \overline{\lambda\mathbf{v}}$, but A is real and symmetric, so $\mathbf{v}^\intercal\mathrm{A} = \lambda\mathbf{v}^\intercal$ and $\mathrm{A}\overline{\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}$. With further manipulation:

$$\lambda\mathbf{v}^\intercal\overline{\mathbf{v}} = \mathbf{v}^\intercal\mathrm{A}\overline{\mathbf{v}} = \overline{\lambda}\mathbf{v}^\intercal\overline{\mathbf{v}} \quad \Longrightarrow \quad (\lambda - \overline{\lambda})\mathbf{v}^\intercal\overline{\mathbf{v}} = 0$$

However, $\mathbf{v}^\intercal\overline{\mathbf{v}} = \sum_{i=1}^{n} |v_i|^2 > 0$, hence $\lambda = \overline{\lambda}$, i.e. $\lambda \in \mathbb{R}$. □

We can now prove the most important result for symmetric endomorphism: the real spectral theorem.

Theorem 1.4.3 (Real spectral theorem)

Given a non-trivial Euclidean vector space V, then $f \in \operatorname{End} V$ is symmetric if and only if $\exists\mathcal{B} \subseteq V$ orthonormal basis of V composed of eigenvectors of f, i.e. if f is diagonalizable.

Proof. ($\Leftarrow$) Trivially, if f is diagonalizable, then in the diagonalizing basis its representative matrix is diagonal, i.e. symmetric, and so is f too.
($\Rightarrow$) We use induction on $n \equiv \dim_{\mathbb{R}} V$:

- $n = 1$ is true as $V = \langle\mathbf{v}\rangle$, with $\mathbf{v} \neq \mathbf{0}$, so $\mathcal{B} = \{\mathbf{a}\}$, with $\mathbf{a} \equiv \frac{\mathbf{v}}{\|\mathbf{v}\|}$, is an orthonormal basis and, since $f(\mathbf{a}) \in V = \langle\mathbf{a}\rangle$, it is composed of eigenvectors of f as $\exists\lambda \in \mathbb{R} : f(\mathbf{a}) = \lambda\mathbf{a}$.
- Assume that the implication is true for $\dim_{\mathbb{R}} V = n - 1$.
- If $f \in \operatorname{End} V$ is symmetric, then all its eigenvalues are real by Prop. 1.4.3. Let $\lambda \in \mathbb{R}$ be an eigenvalue of f and $\mathbf{v} \in V_\lambda(f) - \{\mathbf{0}\}$, and define $S \equiv \langle\mathbf{v}\rangle$: then, f is $S^\perp$-invariant, i.e. $f(S^\perp) \subseteq S^\perp$, since:

  $$\mathbf{w} \in S^\perp \quad \Longrightarrow \quad \langle \mathbf{v}, f(\mathbf{w})\rangle = \langle f(\mathbf{v}), \mathbf{w}\rangle = \lambda\langle \mathbf{v}, \mathbf{w}\rangle = 0 \quad \Longrightarrow \quad f(\mathbf{w}) \in S^\perp$$

  This means that $f|_{S^\perp} \in \operatorname{End} S^\perp$, but $\dim_{\mathbb{R}} S^\perp = n-1$ since $V = S \oplus S^\perp$ (by Prop. 1.4.2), hence, by the inductive step, $\exists\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_{n-1}\} \subseteq S^\perp$ orthonormal basis of $S^\perp$ such that $\mathbf{b}_1, \dots, \mathbf{b}_{n-1}$ are eigenvectors of f.
  Now, set $\mathbf{b}_n \equiv \frac{\mathbf{v}}{\|\mathbf{v}\|}$: then, $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is an orthonormal basis of V composed of eigenvectors of f, since $\mathbf{b}_n \perp \mathbf{b}_i \ \forall i \in \{1, \dots, n-1\} \implies \mathcal{B}$ is LI by Lemma 1.4.6, and $f(\mathbf{b}_n) = \lambda\mathbf{b}_n$ by definition of $\mathbf{v}$. □

A computational approach to check if $f \in \operatorname{End} V$ is symmetric is to consider a basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\} \subseteq V$ and $\mathrm{P} \equiv [\langle \mathbf{b}_i, \mathbf{b}_j\rangle]_{i,j=1,\dots,n}$: then, given $\mathbf{u}, \mathbf{v} \in V \cong \mathbb{R}^n$, we write $\langle \mathbf{u}, f(\mathbf{v})\rangle =$

$\mathbf{u}^\intercal \mathrm{PA}\mathbf{v}$ and $\langle f(\mathbf{u}), \mathbf{v}\rangle = \mathbf{u}^\intercal \mathrm{A}^\intercal \mathrm{P}\mathbf{v}$, with $\mathrm{A} \equiv \mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f)$, so the symmetry condition becomes $\mathrm{A}^\intercal \mathrm{P} = \mathrm{PA}$.

§1.4.3 Hermitian vector spaces

We can generalize the results for Euclidean vector spaces to the case $\mathbb{K} = \mathbb{C}$. In this subsection, we always consider V to be a $\mathbb{C}$-vector space of dimension $\dim_{\mathbb{C}} V = n \in \mathbb{N}$.

Definition 1.4.10 (Hermitian product)

An application $b : V \times V \to \mathbb{C}$ is a **Hermitian product** if it has the following properties:

1. $b(\mathbf{u}, \mathbf{v}) = \overline{b(\mathbf{v}, \mathbf{u})}\ \forall \mathbf{u}, \mathbf{v} \in V$
2. $b(\mathbf{u} + \mathbf{v}, \mathbf{w}) = b(\mathbf{u}, \mathbf{w}) + b(\mathbf{v}, \mathbf{w})\ \wedge\ b(\mathbf{u}, \mathbf{v} + \mathbf{w}) = b(\mathbf{u}, \mathbf{v}) + b(\mathbf{u}, \mathbf{w})\ \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$
3. $b(\lambda\mathbf{u}, \mathbf{v}) = \overline{\lambda} b(\mathbf{u}, \mathbf{v})\ \wedge\ b(\mathbf{u}, \lambda\mathbf{v}) = \lambda b(\mathbf{u}, \mathbf{v})\ \forall \lambda \in \mathbb{C}, \mathbf{u}, \mathbf{v} \in V$
4. $b(\mathbf{u}, \mathbf{u}) \geq 0\ \forall \mathbf{u} \in V\ \wedge\ b(\mathbf{u}, \mathbf{u}) = 0 \iff \mathbf{u} = \mathbf{0}$

A vector space with a Hermitian product is a **Hermitian space**, in which the definitions of norm and orthogonality are analogous to those in the Euclidean case.

Definition 1.4.11 (Hermitian endomorphism)

Given $f \in \operatorname{End} V$, it is **Hermitian** if $\langle f(\mathbf{v}), \mathbf{w}\rangle = \langle \mathbf{v}, f(\mathbf{w})\rangle\ \ \forall \mathbf{v}, \mathbf{w} \in V$.

To extend this notation to matrices, we define the **Hermitian conjugate** of a matrix A as $\mathrm{A}^\dagger := \overline{\mathrm{A}}^\intercal$: if $\mathrm{A} = \mathrm{A}^\dagger$, we say that A is a Hermitian matrix.
The main results for Euclidean vector spaces hold for Hermitian spaces too: in particular, Lemma 1.4.8 (with the Hermitianity condition, in place of the symmetry one) and Prop. 1.4.3 remain as-is, while for Th. 1.4.3 only the forward implication holds, since diagonalizability does not necessarily imply Hermitianity.

Example 1.4.2 (Counterexample for the complex spectral theorem)

Consider $V = \mathbb{C}^2$ with canonical basis $\mathcal{B} = \{\binom{1}{0}, \binom{0}{1}\}$ and the endomorphism $f \in \operatorname{End} \mathbb{C}^2$ represented by:

$$\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{bmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{bmatrix}$$

Clearly this is not a Hermitian endomorphism, even though it is diagonalizable.

Proposition 1.4.4

Given $f \in \operatorname{End} V$, then f is Hermitian if and only if $\langle f(\mathbf{v}), \mathbf{v}\rangle \in \mathbb{R}\ \forall \mathbf{v} \in V$.

Proof. $(\Rightarrow)$ $\langle f(\mathbf{v}), \mathbf{v}\rangle = \langle \mathbf{v}, f(\mathbf{v})\rangle = \overline{\langle f(\mathbf{v}), \mathbf{v}\rangle} \implies \langle f(\mathbf{v}), \mathbf{v}\rangle \in \mathbb{R}$
$(\Leftarrow)$ $\langle f(\mathbf{v}), \mathbf{v}\rangle \in \mathbb{R} \implies \langle f(\mathbf{v}), \mathbf{v}\rangle = \overline{\langle f(\mathbf{v}), \mathbf{v}\rangle} = \langle \mathbf{v}, f(\mathbf{v})\rangle$ □

§1.4.4 Unitary endomorphisms

An interesting class of endomorphisms are unitary endomorphisms.

Definition 1.4.12 (Unitary endomorphism)

Given a Euclidean or Hermitian vector space V, then $f \in \operatorname{End} V$ is **unitary** if $\langle f(\mathbf{v}), f(\mathbf{w})\rangle = \langle \mathbf{v}, \mathbf{w}\rangle \ \ \forall \mathbf{v}, \mathbf{w} \in V$.

Lemma 1.4.9 (Unitary endomorphisms as automorphisms)

Every unitary endomorphisms is an automorphism.

Proof. Let $\mathbf{v} \in \ker f$. Then $\langle \mathbf{v}, \mathbf{v}\rangle = \langle f(\mathbf{v}), f(\mathbf{v})\rangle = \langle \mathbf{0}, \mathbf{0}\rangle = 0 \iff \mathbf{v} = \mathbf{0}$, hence f is injective by Prop. 1.3.1, and so a bijection by Cor. 1.3.2.1. □

It is easy to see from the definition that an endomorphism is an isometry if and only if it maps orthonormal bases to orthonormal bases.
In the case of a Euclidean vector space, unitary endomorphisms are called isometries.

Proposition 1.4.5 (Isometries and representative matrices)

Given an n-dimensional Euclidean vector space V and an orthonormal basis $\mathcal{B} \subseteq V$, then $f \in \operatorname{End} V$ is an isometry if and only if $\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) \in \mathrm{O}(n)$.

Proof. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) = [m_{ij}]_{i,j=1,\dots,n}$. WTS $\mathcal{C} \equiv \{f(\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is an orthonormal basis of V:

$$\langle f(\mathbf{b}_i), f(\mathbf{b}_j)\rangle = \sum_{k,\ell=1}^{n} m_{ki} m_{\ell j} \langle \mathbf{b}_k, \mathbf{b}_\ell\rangle = \sum_{k=1}^{n} m_{ki} m_{kj} = [(\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f))^{\intercal} \mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f)]_{ij}$$

It is then clear that f is an isometry if and only if $(\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f))^{\intercal} \mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \mathrm{I}_n$. □

In a generic basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\} \subseteq V$, to check if $f \in \operatorname{End} V$ is an isometry, consider $\mathrm{P} \equiv [\langle \mathbf{b}_i, \mathbf{b}_j\rangle]_{i,j=1,\dots,n}$ and set $\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) \equiv \mathrm{A}$: then, $\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{u}^{\intercal}\mathrm{P}\mathbf{v}$ and $\langle f(\mathbf{u}), f(\mathbf{v})\rangle = \mathbf{u}^{\intercal}\mathrm{A}^{\intercal}\mathrm{P}\mathrm{A}\mathbf{v}$, so that the unitarity condition becomes $\mathrm{A}^{\intercal}\mathrm{P}\mathrm{A} = \mathrm{P}$.

Proposition 1.4.6 (Unitary endomorphisms and representative matrices)

Given an n-dimensional Hermitian vector space V and an orthonormal basis $\mathcal{B} \subseteq V$, then $f \in \operatorname{End} V$ is unitary if and only if $\mathrm{M}_{\mathcal{B}}^{\mathcal{B}}(f) \in \mathrm{U}(n)$.

Proof. The proof is analogous to that of Prop. 1.4.5, recalling Def. 1.4.10. □

Appendices

Appendix A

Logic

§A.1 Binary relations

Definition A.1.1 (Binary relation)

Given two sets $\mathcal{A}$, $\mathcal{B}$ and their cartesian product $\mathcal{A} \times \mathcal{B} \coloneqq \{(a, b) : a \in \mathcal{A} \land b \in \mathcal{B}\}$, a **binary relation** $\mathfrak{R}$ is a subset of $\mathcal{A} \times \mathcal{B}$. Two elements $a \in \mathcal{A}$, $b \in \mathcal{B}$ are related, and we write $a\mathfrak{R}b$, if $(a, b) \in \mathfrak{R} \subseteq \mathcal{A} \times \mathcal{B}$.

If $\mathcal{B} = \mathcal{A}$, we say that $\mathfrak{R}$ is a relation "on" $\mathcal{A}$.

Definition A.1.2 (Function)

A **function** between two sets $\mathcal{A}$, $\mathcal{B}$ is a relation $\mathfrak{R}_f$ such that, given an element $a \in \mathcal{A}$, then there exists at most one element $b \in \mathcal{B} : a\mathfrak{R}_f b$.

We usually write $b = f(a)$ in place of $a\mathfrak{R}_f b$.

Definition A.1.3 (Equivalence relation)

Given a set $\mathcal{A}$, a relation $\mathfrak{R}$ on $\mathcal{A}$ is an **equivalence relation** if it has the following properties:

1. reflexivity: $a\mathfrak{R}a \ \forall a \in \mathcal{A}$
2. symmetry: $a\mathfrak{R}b \iff b\mathfrak{R}a \ \forall a, b \in \mathcal{A}$
3. transitivity: $a\mathfrak{R}b \land b\mathfrak{R}c \implies a\mathfrak{R}c \ \forall a, b, c \in \mathcal{A}$

Example A.1.1

Take $\mathcal{A} = \mathbb{Z}$. Then, the relation $a\mathfrak{R}b \iff \exists k \in \mathbb{Z} : a - b = 2k$ is an equivalence relation: $a - a = 2k$ with $k = 0$ (reflexivity), $a - b = 2k \iff b - a = 2h$ with $h = -k$ (symmetry) and $a - b = 2k, b - c = 2h \implies a - c = 2l$ with $l = k + h$ (transitivity.

Definition A.1.4 (Equivalence class)

Given a set $\mathcal{A}$ and an equivalence relation $\mathfrak{R}$ on $\mathcal{A}$, then the **equivalence relation** of $a \in \mathcal{A}$ is defined as $[a]_{\mathfrak{R}} \coloneqq \{b \in \mathcal{A} : a\mathfrak{R}b\}$.

In absence of ambiguity, the subscript $\mathfrak{R}$ is dropped, and the equivalence class $a \in \mathcal{A}$ is simply denoted by $[a]$.

Theorem A.1.1

Given a set $\mathcal{A}$, an **equivalence** relation $\mathfrak{R}$ on $\mathcal{A}$ and two elements $a, b \in \mathcal{A}$, then:

1. $a \in [a]_{\mathfrak{R}}$
2. $a \mathfrak{R} b \implies [a]_{\mathfrak{R}} = [b]_{\mathfrak{R}}$
3. $a \not\mathfrak{R} b \implies [a]_{\mathfrak{R}} \cap [b]_{\mathfrak{R}} = \varnothing$

Proof. The first proposition is true by reflexivity. To prove the second proposition, let $x \in [a]_{\mathfrak{R}}$: then, $x \mathfrak{R} a$, but also $x \mathfrak{R} b$ by transitivity, hence $x \in [b]_{\mathfrak{R}}$. This proves $[b]_{\mathfrak{R}} \subseteq [a]_{\mathfrak{R}}$, and the vice versa is equivalently proven, hence $[a]_{\mathfrak{R}} = [b]_{\mathfrak{R}}$. To prove the third proposition, suppose $\exists x \in [b]_{\mathfrak{R}} \cap [a]_{\mathfrak{R}}$: then, $x \mathfrak{R} a \wedge x \mathfrak{R} b \implies a \mathfrak{R} b$ by transitivity, which is absurd. □

This theorem shows that an equivalence relation splits the set in separated equivalence classes.

Definition A.1.5 (Partition)

Given a set $\mathcal{X} \neq \varnothing$ and its power set $\wp(\mathcal{X}) \coloneqq \{\mathcal{A} : \mathcal{A} \subseteq \mathcal{X}\}$, a **partition** of $\mathcal{X}$ is a collection of subsets $\{\mathcal{A}_i\}_{i \in \mathcal{I}} \subseteq \wp(\mathcal{X})$ which satisfies the following propeties:

1. $\mathcal{A}_i \neq \varnothing \ \forall i \in \mathcal{I}$
2. $\mathcal{A}_i \cap \mathcal{A}_j = \varnothing \ \forall i \neq j \in \mathcal{I}$
3. $\mathcal{X} = \bigcup_{i \in \mathcal{I}} \mathcal{A}_i$

The equivalence classes determined by an equivalence relation form a partition of the set it is defined on.

Definition A.1.6 (Quotient set)

Given a set $\mathcal{A}$ and an equivalence relation $\mathfrak{R}$ on $\mathcal{A}$, the **quotient set** $\mathcal{A}/\mathfrak{R}$ is defined as the set of all equivalence classes of $\mathcal{A}$ determined by $\mathfrak{R}$.

Example A.1.2 ($\mathbb{Z}$ as a quotient set)

The set $\mathbb{Z}$ can be seen as a quotient set $\mathbb{Z} = (\mathbb{N} \times \mathbb{N})/\mathfrak{R}$ with $(n, m) \mathfrak{R} (n', m') \iff n - m = n' - m'$. Indeed, there are three kinds of equivalence classes: $[(n, 0)] \equiv n$, $[(0, n)] \equiv -n$ and $[(0, 0)] \equiv 0$.

Example A.1.3 (Modular equivalence)

Given $n \in \mathbb{N}$, the **congruence modulo** n relation is an equivalence relation on $\mathbb{Z}$ defined as $a \equiv_n b \iff \exists k \in \mathbb{Z} : a - b = kn$. This relation defines the quotient set $\mathbb{Z}_n \equiv \mathbb{Z}/(\mathrm{mod}\, n)$, which in general is $\mathbb{Z}_n = \{[0]_n, [1]_n, \ldots, [n-1]_n\}$.

§A.2 Zorn's Lemma

Zorn's Lemma is an equivalent expression of the Axiom of Choice.

Definition A.2.1 (Order relation)

Given a set $\mathcal{X}$, an **order relation** is a relation $\leq$ with the following properties:

1. reflexivity: $x \leq x \ \forall x \in \mathcal{X}$
2. antisymmetry: $x \leq y \wedge y \leq x \iff x = y$
3. transitivity: $x \leq y \wedge y \leq z \implies x \leq z$

Then, $(\mathcal{X}, \leq)$ is an **ordered set**.

Note that we define $x < y$ as $x \leq y \wedge x \neq y$. Moreover, trivially, every subset of an ordered set is an ordered set too, with the induced order relation.

Example A.2.1 (Inclusion)

Let $\mathcal{X}$ be a set. Then the **inclusion** $\subseteq$ is an order relation on $\wp(\mathcal{X})$.

An order relation on $\mathcal{X}$ is a **total ordering** is $x \leq y \vee y \leq x \ \forall x, y \in \mathcal{X}$, and $\mathcal{X}$ is a **totally-ordered set**[1].

Definition A.2.2 (Chains)

Given an ordered set $(\mathcal{X}, \leq)$, then:

1. a subset $\mathcal{C} \subseteq \mathcal{X}$ is a **chain** if $(\mathcal{C}, \leq)$ is a totally-ordered set
2. given $\mathcal{C} \subseteq \mathcal{X}$ and $x \in \mathcal{X}$, then x is an **upper bound** of $\mathcal{C}$ if $y \leq x \ \forall y \in \mathcal{C}$
3. an element $m \in \mathcal{X}$ is a **maximal element** of $\mathcal{X}$ if $\{x \in \mathcal{X} : m \leq x\} \equiv \{m\}$

Lemma A.2.1 (Zorn's Lemma)

Let $(\mathcal{X}, \leq)$ be a non-empty ordered set. If every chain in $\mathcal{X}$ has at least one upper bound, then $\mathcal{X}$ has at least one maximal element.

[1] Not a universal convention: some refer to ordered set as "partially-ordered sets" and to totally-ordered sets as "ordered sets". We use the convention of e.g. [1]

Index

Bibliography

[1] M. Manetti. *Topologia*. Springer Milano, 2014. DOI: 10.1007/978-88-470-5662-6.