LECTURE 5
LINEAR SPACES

1. Definition. Properties

Some of the most usual spaces we meet in everyday life (or science) are *Euclidean spaces*: the real line (1-dimensional), the real plane (2-dimensional), the real space (3-dimensional) and the real hyperspace (a n-dimensional space with $n \geq 4$). The notion of linear space attempts to capture the algebraic features of Euclidean spaces: addition and multiplication with a number (scalar).

Definition. Let V be a non-empty set, an operation on V, $+ : V \times V \to V$ and an external operation on V, $\cdot : \mathbb{R} \times V \to V$. We say that $(V, +, \cdot)$ is a *linear space* or a *vectorial space* if:

(i) $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z},\ \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$;
(ii) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x},\ \forall \mathbf{x}, \mathbf{y} \in V$;
(iii) $\exists \mathbf{0} \in V,\ \forall \mathbf{x} \in V : \mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$;
(iv) $\forall \mathbf{x} \in V,\ \exists (-\mathbf{x}) \in V : \mathbf{x} + (-\mathbf{x}) = (-\mathbf{x}) + \mathbf{x} = \mathbf{0}$;
(v) $\alpha \cdot (\mathbf{x} + \mathbf{y}) = \alpha \cdot \mathbf{x} + \alpha \cdot \mathbf{y}$, $\forall \alpha \in \mathbb{R}$, $\forall \mathbf{x}, \mathbf{y} \in V$;
(vi) $(\alpha + \beta) \cdot \mathbf{x} = \alpha \cdot \mathbf{x} + \beta \cdot \mathbf{x}$, $\forall \alpha, \beta \in \mathbb{R}$, $\forall \mathbf{x} \in V$;
(vii) $\alpha \cdot (\beta \cdot \mathbf{x}) = (\alpha\beta) \cdot \mathbf{x}$, $\forall \alpha, \beta \in \mathbb{R}$, $\forall \mathbf{x} \in V$;
(viii) $1 \cdot \mathbf{x} = \mathbf{x}$, $\forall \mathbf{x} \in V$.

The elements of V are usually called *vectors* and, in this context, the elements of $\mathbb{R}$ are called *scalars.* The operation $+$ is called the *addition of vectors*, while the external operation $\cdot$ is called the *multiplication with scalars.* The element $\mathbf{0}$ is called the *zero-vector*, while the vector $-\mathbf{x}$ is called the *opposite* of the vector $\mathbf{x} \in V$.

The first four properties in the above definition just assert that $(V, +)$ is a commutative group. The other four relate the multiplication with scalars $\cdot$ to the addition $+$. If $\mathbf{x}$ and $\mathbf{y}$ are two vectors, it is customary to denote $\mathbf{x} - \mathbf{y}$ instead of $\mathbf{x} + (-\mathbf{y})$ (as it is in the case of additive groups).

We often refer to V as being a linear space, instead of $(V, +, \cdot)$, especially when the operation $+$ and $\cdot$ are implied from the context.

Sometimes, the external operation $\cdot$ is defined on $K \times V$ instead of $\mathbb{R} \times V$, where K is a *field.* If we replace everywhere in the definition of linear spaces $\mathbb{R}$ by K, we obtain the notion of a K*-linear space.*

It easy to see that $\mathbb{R}$ (equipped with its canonical operations) is a linear space. A more general example (we will use it for the most part) consists in the space $\mathbb{R}^n$, the prototype of the *Euclidean space* which we told about at the beginning of the lecture.

Theorem 1.1. *Let $n \in \mathbb{N}^*$ and $\mathbb{R}^n := \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ times}}$. We define the operations $+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ and $\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ by:*

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) := (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n),\ (x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n;$$
$$\alpha \cdot (x_1, x_2, \ldots, x_n) := (\alpha x_1, a x_2, \ldots, \alpha x_n),\ \alpha \in \mathbb{R},\ (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n.$$

Then $(\mathbb{R}^n, +, \cdot)$ is a linear space, with $\mathbf{0} = (0, \ldots, 0)$.

If not specified otherwise, on $\mathbb{R}^n$ we will always consider the above two operations, named the *canonical operations* on $\mathbb{R}^n$. If $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we will call the numbers $x_1, x_2, \ldots, x_n$ the *coordinates* of $\mathbf{x}$.

Proof. Let us prove the properties which define a linear space. We will fix $\alpha, \beta \in \mathbb{R}$ and $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n)$, $\mathbf{z} = (z_1, \ldots, z_n)$ in $\mathbb{R}^n$. We also define $\mathbf{0} := (0, \ldots, 0)$ and $-\mathbf{x} := (-x_1, \ldots, -x_n)$.

(i)
$$\begin{aligned}\mathbf{x} + (\mathbf{y} + \mathbf{z}) &= (x_1, \ldots, x_n) + ((y_1, \ldots, y_n) + (z_1, \ldots, z_n)) = (x_1, \ldots, x_n) + (y_1 + z_1, \ldots, y_n + z_n)\\ &= (x_1 + y_1 + z_1, \ldots, x_n + y_n + z_n) = (x_1 + y_1, \ldots, x_n + y_n) + (z_1, \ldots, z_n)\\ &= ((x_1, \ldots, x_n) + (y_1, \ldots, y_n)) + (z_1, \ldots, z_n) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}.\end{aligned}$$

(ii)
$$\begin{aligned}\mathbf{x} + \mathbf{y} &= (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n) = (y_1 + x_1, \ldots, y_n + x_n)\\ &= (y_1, \ldots, y_n) + (x_1, \ldots, x_n) = \mathbf{y} + \mathbf{x}.\end{aligned}$$

(iii) $\mathbf{x} + \mathbf{0} = (x_1, \ldots, x_n) + (0, \ldots, 0) = (x_1, \ldots, x_n) = \mathbf{x}$.

(iv) $(x_1, \ldots, x_n) + (-x_1, \ldots, -x_n) = (x_1 - x_1, \ldots, x_n - x_n) = (0, \ldots, 0) = \mathbf{0}$.

(v) $\alpha \cdot (\mathbf{x}+\mathbf{y}) = \alpha \cdot (x_1+y_1,\ldots,x_n+y_n) = (\alpha(x_1+y_1),\ldots,\alpha(x_n+y_n)) = (\alpha x_1+\alpha y_1,\ldots,\alpha x_n+\alpha y_n)$
$= (\alpha x_1,\ldots,\alpha x_n)+(\alpha y_1,\ldots,\alpha y_n) = \alpha\cdot(x_1,\ldots,x_n)+\alpha\cdot(y_1,\ldots,y_n) = \alpha\cdot\mathbf{x}+\alpha\cdot\mathbf{y}$..

(vi) $(\alpha+\beta)\cdot\mathbf{x} = (\alpha+\beta)\cdot(x_1,\ldots,x_n) = ((\alpha+\beta)x_1,\ldots,(\alpha+\beta)x_n) = (\alpha x_1+\beta x_1,\ldots,\alpha x_n+\beta x_n)$
$= (\alpha x_1,\ldots,\alpha x_n)+(\beta x_1,\ldots,\beta x_n) = \alpha\cdot(x_1,\ldots,x_n)+\beta\cdot(x_1,\ldots,x_n) = \alpha\cdot\mathbf{x}+\beta\cdot\mathbf{x}$.

(vii) $\alpha\cdot(\beta\cdot\mathbf{x}) = \alpha\cdot(\beta x_1,\ldots,\beta x_n) = (\alpha\beta x_1,\ldots,\alpha\beta x_n) = (\alpha\beta)\cdot(x_1,\ldots,x_n) = (\alpha\beta)\cdot\mathbf{x}$.

(viii) $1\cdot\mathbf{x} = 1\cdot(x_1,\ldots,x_n) = (1\cdot x_1,\ldots,1\cdot x_n) = (x_1,\ldots,x_n) = \mathbf{x}$.

□

Further examples.
1. Let, for $m,n\in\mathbb{N}^*$, $\mathscr{M}_{m,n}$ be the set of all $m\times n$-matrices. If $+$ denotes the usual addition between matrices, while $\cdot$ is the multiplication of matrices with reals, then $(\mathscr{M}_{m,n},+,\cdot)$ is a linear space.
2. Let $\mathbb{R}[X]$ be the set of all *polynomials* with real coefficients. If $+$ denotes the usual addition between polynomials, while $\cdot$ is the multiplication of polynomials with reals, then $(\mathbb{R}[X],+,\cdot)$ is a linear space.
3. Let E be a set, $(V,+,\cdot)$ a linear space and $\mathscr{F}(E;V)$ the collection of all functions $f:E\to V$. If we define $+:\mathscr{F}(E;V)\times\mathscr{F}(E;V)\to\mathscr{F}(E;V)$ and $\cdot:\mathbb{R}\times\mathscr{F}(E;V)\to\mathscr{F}(E;V)$ by:

$(f+g)(x) := f(x)+g(x)$, $f,g\in\mathscr{F}(E;V)$, $x\in E$;
$(\alpha\cdot f)(x) := \alpha\cdot f(x)$, $\alpha\in\mathbb{R}$, $f\in\mathscr{F}(E;V)$, $x\in E$,

then $(\mathscr{F}(E;V),+,\cdot)$ is a linear space.
Particularizing E and $(V,+,\cdot)$ we get other or already known examples.

- For instance, if we take $E:=\{1,\ldots,m\}\times\{1,\ldots,n\}$ and $V:=\mathbb{R}$, we obtain once again the linear space $(\mathscr{M}_{m,n},+,\cdot)$, since $\mathscr{M}_{m,n}$ is precisely $\mathscr{F}(\{1,\ldots,m\}\times\{1,\ldots,n\};\mathbb{R})$.
- If $m,n\in\mathbb{N}^*$, $E\subseteq\mathbb{R}^n$ and $V:=\mathbb{R}^m$, then $(\mathscr{F}(E;\mathbb{R}^m),+,\cdot)$ is a vectorial space of functions of n variables with values in $\mathbb{R}^m$.
- If $E:=\mathbb{N}$ and $V:=\mathbb{R}$, then $\mathscr{F}(E;V)$ is the space of real sequences $(x_n)_{n\in\mathbb{N}}$.

Let us give some properties of the operations on a linear space.

Theorem 1.2. *Let $(V,+,\cdot)$ be a linear space. Then, for any $\alpha\in\mathbb{R}$ and $\mathbf{x}\in V$ we have:*

i) $\alpha\cdot\mathbf{0} = 0\cdot\mathbf{x} = \mathbf{0}$;
ii) $(-\alpha)\cdot\mathbf{x} = \alpha\cdot(-\mathbf{x}) = -(\alpha\cdot\mathbf{x})$;
iii) $(-\alpha)\cdot(-\mathbf{x}) = \alpha\cdot\mathbf{x}$;
iv) $\alpha\cdot\mathbf{x}=\mathbf{0}\Rightarrow\alpha=0$ *or* $\mathbf{x}=\mathbf{0}$.

Proof. Let $\alpha\in\mathbb{R}$ and $\mathbf{x}\in V$.

i) We have
$$\alpha\cdot\mathbf{0} = \alpha\cdot(\mathbf{0}+\mathbf{0}) = (\alpha\cdot\mathbf{0})+(\alpha\cdot\mathbf{0}),$$
hence $\alpha\cdot\mathbf{0}=\mathbf{0}$ (we can substract $\alpha\cdot\mathbf{0}$). Also,
$$0\cdot\mathbf{x} = (0+0)\cdot\mathbf{x} = 0\cdot\mathbf{x}+0\cdot\mathbf{x},$$
hence $0\cdot\mathbf{x}=\mathbf{0}$.

ii) We have
$$(-\alpha)\cdot\mathbf{x}+\alpha\cdot\mathbf{x} = ((-\alpha)+\alpha)\cdot\mathbf{x} = 0\cdot\mathbf{x} = \mathbf{0},$$
therefore $(-\alpha)\cdot\mathbf{x} = -(\alpha\cdot\mathbf{x})$. Also,
$$\alpha\cdot(-\mathbf{x})+\alpha\cdot\mathbf{x} = \alpha\cdot((-\mathbf{x})+\mathbf{x}) = \alpha\cdot\mathbf{0} = \mathbf{0},$$
so $\alpha\cdot(-\mathbf{x}) = -(\alpha\cdot\mathbf{x})$.

iii) By the previous point,
$$(-\alpha)\cdot(-\mathbf{x}) = -(\alpha\cdot(-\mathbf{x})) = -(-(\alpha\cdot\mathbf{x})) = \alpha\cdot\mathbf{x}.$$

iv) Suppose that $\alpha\neq 0$. Then
$$\mathbf{x} = 1\cdot\mathbf{x} = (\alpha^{-1}\alpha)\cdot\mathbf{x} = \alpha^{-1}\cdot(\alpha\cdot\mathbf{x}) = \alpha^{-1}\cdot\mathbf{0} = \mathbf{0}.$$

□

Definition. Let $(V,+,\cdot)$ be a linear space and W a non-empty subset of V. We say that W is a *linear subspace* of V if for any $\alpha\in\mathbb{R}$ and $\mathbf{x},\mathbf{y}\in W$ we have that $\mathbf{x}+\mathbf{y}\in W$ and $\alpha\cdot\mathbf{x}\in W$.

Examples.
1. If $m,n\in\mathbb{N}^*$ and $m\le n$, the set
$$W_m := \{(x_1,\ldots,x_m,0,\ldots,0)\in\mathbb{R}^n \mid (x_1,\ldots,x_m)\in\mathbb{R}^m\}$$

is a linear subspace of $\mathbb{R}^n$. Since we can identify W_m with $\mathbb{R}^m$, we often consider $\mathbb{R}^m$ as a subset of $\mathbb{R}^n$ (as we consider $\mathbb{R}$ a subset of $\mathbb{C}$).

2. Let $n \in \mathbb{N}^*$ and $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ such that not all $\alpha_1, \dots, \alpha_n$ are 0 (*i.e.*, $(\alpha_1, \dots, \alpha_n) \neq \mathbf{0}$). The set

$$H := \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid \alpha_1 x_1 + \dots + \alpha_n x_n = 0\}$$

is a linear subspace of $\mathbb{R}^n$, called a *hyperplane.*

3. The set of *even* real functions,

$$\{f \in \mathscr{F}(\mathbb{R};\mathbb{R}) \mid f(-x) = f(x),\ \forall x \in \mathbb{R}\}$$

is a linear subspace of $\mathscr{F}(\mathbb{R};\mathbb{R})$.

Proposition 1.3. *Let W_1 and W_2 be two linear subspaces of a linear space $(V, +, \cdot)$. Then $W_1 \cap W_2$ is again a linear subspace of V.*

We let the proof of this result as an exercice.

In contrast to the intersection, the union of two linear subspaces of V is *not* a linear subspace of V, in general.

2. Linear combinations. Bases and dimension

Definition. Let $(V, +, \cdot)$ be a linear space. A *linear combination* of the vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$ is a vector $\mathbf{y} \in V$ which can be written as

$$\mathbf{y} = \alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n,$$

where $n \in \mathbb{N}^*$ and $\alpha_1, \dots, \alpha_n \in \mathbb{R}$.

Remark. If W is a linear subspace of V, any linear combination of the vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in W$ is again an element of W.

Definition. Let $(V, +, \cdot)$ be a linear space and U be a non-empty subset of V. The set of *all* linear combinations of elements of U,

$$\{\alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n \mid n \in \mathbb{N}^*,\ \alpha_1, \dots, \alpha_n \in \mathbb{R},\ \mathbf{x}_1, \dots, \mathbf{x}_n \in V\}$$

is called the *linear subspace generated* by U, denoted $\mathrm{Lin}(U)$.

It is easy to prove that $U \subseteq \mathrm{Lin}(U)$ and $\mathrm{Lin}(U)$ is a linear subspace of V (hence the name). Moreover, it can be shown that $\mathrm{Lin}(U)$ is the smallest linear subspace of V which contains U.

Example. If $V := \mathbb{R}^3$, the linear subspace generated by $U := \{(1, 3, 2)\}$ is the line $\{(\alpha, 3\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\}$.

Definition. Let $(V, +, \cdot)$ be a linear space.

a) For $n \in \mathbb{N}^*$, the vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$ are called *linearly dependent* if there exist $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, not all 0, such that

$$\alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n = 0.$$

Otherwise, $\mathbf{x}_1, \dots, \mathbf{x}_n$ are called *linearly independent.*

b) A subset U of V is called *linearly independent* if for any *distinct* vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in U$, $\mathbf{x}_1, \dots, \mathbf{x}_n$ are *linearly independent.*

Remarks.

- By the above definition, $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$ are linearly independent if and only if the equation

$$\alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n = \mathbf{0}$$

has as unique solution $\alpha_1 = \dots = \alpha_n = 0$.

- If $\mathbf{0} \in \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, then clearly $\mathbf{x}_1, \dots, \mathbf{x}_n$ are linearly dependent (we take all α_k, $1 \leq k \leq n$, to be 0, except the α_k corresponding to the $\mathbf{x}_k$ which is $\mathbf{0}$). Hence, if $U \subseteq V$ is linearly independent, $0 \notin U$.
- The necessary and sufficient condition for the vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$ to be linearly dependent is that we can write a vector among $\mathbf{x}_1, \dots, \mathbf{x}_n$ as a linear combination of the others. Indeed, if

$$\mathbf{x}_k = \alpha_1 \mathbf{x}_1 + \dots + \alpha_{k-1} \mathbf{x}_{k-1} + \alpha_{k+1} \mathbf{x}_{k+1} + \dots + \alpha_n \mathbf{x}_n,$$

then

$$\alpha_1 \mathbf{x}_1 + \dots + \alpha_{k-1} \mathbf{x}_{k-1} + \alpha_k \mathbf{x}_k + \alpha_{k+1} \mathbf{x}_{k+1} + \dots + \alpha_n \mathbf{x}_n = \mathbf{0},$$

where $\alpha_k = -1 \neq 0$. Conversely, if

$$\alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n = \mathbf{0}$$

for some $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, not all 0, let $k \in \{1, \dots, n\}$ such that $\alpha_k \neq 0$. Then

$$\mathbf{x}_k = \left(-\frac{\alpha_1}{\alpha_k}\right) \mathbf{x}_1 + \dots + \left(-\frac{\alpha_{k-1}}{\alpha_k}\right) \mathbf{x}_{k-1} + \left(-\frac{\alpha_{k+1}}{\alpha_k}\right) \mathbf{x}_{k+1} + \dots + \left(-\frac{\alpha_n}{\alpha_k}\right) \mathbf{x}_n.$$

Definition. Let $(V,+,\cdot)$ be a linear space. A subset $B \subseteq V$ is called an *algebraic basis* or *Hamel basis* (or simply, a *basis*) of V if B is linearly independent and $\mathrm{Lin}(B) = V$.

Theorem 2.1. *Let $n \in \mathbb{N}^*$. Then the set $\{\mathbf{e}_1, \dots, \mathbf{e}_n\} \subseteq \mathbb{R}^n$, where*

$$\mathbf{e}_k := (\underbrace{0,\dots,0}_{k-1 \text{ times}}, 1, 0, \dots, 0),\ 1 \le k \le n,$$

is a basis of $\mathbb{R}^n$, called the canonical basis *of $\mathbb{R}^n$.*

Proof. Let $\mathbf{x} = (x_1, \dots, x_n)$ an arbitrary vector of $\mathbb{R}^n$. Then

$$\begin{aligned} x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n &= x_1(1,0,\dots,0) + x_2(0,1,0,\dots,0) + \cdots + x_n(0,\dots 0,1) \\ &= (x_1,0,\dots,0) + (0,x_2,0,\dots,0) + \cdots + (0,\dots 0,x_n) = (x_1,x_2,\dots,x_n) = \mathbf{x}. \end{aligned}$$

This proves that $\mathrm{Lin}(\{\mathbf{e}_1,\dots,\mathbf{e}_n\}) = \mathbb{R}^n$.

On the other hand, if $\alpha_1,\dots,\alpha_n \in \mathbb{R}$ are such that $\alpha_1\mathbf{e}_1 + \cdots + \alpha_n\mathbf{e}_n = \mathbf{0}$, then we have $(\alpha_1,\dots,\alpha_n) = (0,\dots,0)$, *i.e.* $\alpha_1 = \cdots = \alpha_n = 0$. This means that the vectors $\mathbf{e}_1,\dots,\mathbf{e}_n$ are linearly independent. □

Definition. Let $(V,+,\cdot)$ be a linear space. We say that V is *finite-dimensional* if there exists a finite basis of V. Otherwise, V is called *infinite-dimensional*.

Theorem 2.2. *Let $(V,+,\cdot)$ be a finite-dimensional linear space, $n \in \mathbb{N}^*$ and $\{\mathbf{b}_1,\dots,\mathbf{b}_n\}$ a basis of V. Let $X = (\alpha_{i,j})_{\substack{1\le i\le n\\ 1\le j\le m}}$ be a matrix in $\mathscr{M}_{n,m}$. Then the m vectors*

$$\mathbf{x}_k := \alpha_{1,k}\mathbf{b}_1 + \cdots + \alpha_{n,k}\mathbf{b}_n,\ 1 \le k \le m$$

are linearly independent if and only if the rank of the matrix X is m.

Proof. The vectors $\mathbf{x}_1,\dots,\mathbf{x}_m$ are linearly independent if and only if the equation

$$\xi_1\mathbf{x}_1 + \cdots + \xi_m\mathbf{x}_m = \mathbf{0}$$

has only the trivial solution $\xi_1 = \cdots = \xi_m = 0$. This is equivalent with the fact that

$$(\alpha_{1,1}\xi_1 + \cdots + \alpha_{1,m}\xi_m)\mathbf{b}_1 + \cdots + (\alpha_{n,1}\xi_1 + \cdots + \alpha_{n,m}\xi_m)\mathbf{b}_n = 0$$

has just the trivial solution, *i.e.* the homogeneous system with n equations and m unknowns

$$\begin{cases} \alpha_{1,1}\xi_1 + \cdots + \alpha_{1,m}\xi_m & = & 0 \\ \dots\dots\dots\dots\dots\dots & & \vdots \\ \alpha_{n,1}\xi_1 + \cdots + \alpha_{n,m}\xi_m & = & 0 \end{cases}$$

has only the trivial solution. By the theory of linear systems in $\mathbb{R}$, this is equivalent to the fact that the matrix X has the rank m. □

Corollary. *Let $(V,+,\cdot)$ be a finite-dimensional linear space. If $n \in \mathbb{N}$, $m \in \mathbb{N}^*$, B is a basis of V with n elements and $\mathbf{x}_1,\dots,\mathbf{x}_m \in V$ are linearly independent vectors, then $m \le n$.*

Theorem 2.3. *Let $(V,+,\cdot)$ be a finite-dimensional linear space. Then there exists a unique $n \in \mathbb{N}$, called the* dimension *of V and denoted $\dim V$, such that every basis of V has precisely n elements.*

Proof. Let B be a finite basis of V having n elements, $n \in \mathbb{N}$ and B' another basis of V. Let us show that B' has n elements, too.

If B' had more than n elements, there would exist linearly independent vectors $\mathbf{x}_1,\dots,\mathbf{x}_n,\mathbf{x}_{n+1} \in B$, which contradicts the corollary of Theorem ??. Hence B' has m elements, with $m \le n$. Then, swaping the roles of B and B', we must have $m = n$. □

Remark. By Theorem **??**, the linear space $\mathbb{R}^n$ is finite dimensional and has dimension n.

We have seen, by the corollary of Theorem **??** that if $\dim V = n \in \mathbb{N}$, then there are no linearly independent vectors $\mathbf{x}_1,\dots,\mathbf{x}_m \in V$ with $m > n$. If $m < n$, the next result shows that we can add vectors to $\mathbf{x}_1,\dots,\mathbf{x}_m$ in order to obtain a basis.

Proposition 2.4. *Let $(V,+,\cdot)$ be a finite-dimensional linear space and $n := \dim V$. If $m \in \mathbb{N}^*$ with $m \le n$ and $\mathbf{x}_1,\dots,\mathbf{x}_m \in V$ are linearly independent vectors, then there exists a finite set $\tilde{B} \subseteq V$ with $n - m$ elements such that $\{\mathbf{x}_1,\dots,\mathbf{x}_m\} \cup \tilde{B}$ forms a basis of V.*

Proof. We will prove this result by mathematical induction on $p := n - m$.

I. If $p = 0$ *(i.e.* $n = m$), let us show that we can take $\tilde{B} = \varnothing$ (which has 0 elements). Indeed, let $\mathbf{x} \in V$ be an arbitrary vector; since $\dim V = n$, it is not possible that $\mathbf{x},\mathbf{x}_1,\dots,\mathbf{x}_n$ are linearly independent, by the corollary of Theorem **??**. Therefore, there exist $\alpha,\alpha_1,\dots,\alpha_n \in \mathbb{R}$ such that

$$\alpha\mathbf{x} + \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n = \mathbf{0}.$$

Since $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly independent, $\alpha \neq 0$. It follows that

$$\mathbf{x} = \left(-\frac{\alpha_1}{\alpha}\right)\mathbf{x}_1 + \cdots + \left(-\frac{\alpha_n}{\alpha}\right)\mathbf{x}_n;$$

hence $\mathbf{x} \in \operatorname{Lin}(\{\mathbf{x}_1, \ldots, \mathbf{x}_n\})$.

This proves that $\operatorname{Lin}(\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}) = V$, so $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is a basis for V.

II. Suppose now that the result is true for some $p \geq 0$ and let us show it for $p+1$. We assume then that $n - m = p + 1$. Let $W := \operatorname{Lin}(\{\mathbf{x}_1, \ldots, \mathbf{x}_m\})$; of course $W \subsetneq V$, otherwise we would have $\dim V = m$ (*i.e.* $p+1 = 0$!). Let $\mathbf{x} \in V \setminus W$. Since $\mathbf{x} \notin \operatorname{Lin}(\{\mathbf{x}_1, \ldots, \mathbf{x}_m\})$, from the same argument as in **I** (with n replaced by m), it is clear that $\mathbf{x}, \mathbf{x}_1, \ldots, \mathbf{x}_m$ are linearly independent.

By the induction hypothesis, there exists a set $\bar{B} \subseteq V$ with $n - (m+1) = p$ elements such that $\{\mathbf{x}, \mathbf{x}_1, \ldots, \mathbf{x}_m\} \cup \bar{B}$ forms a basis of V. Of course, $\bar{B} \cap \{\mathbf{x}, \mathbf{x}_1, \ldots, \mathbf{x}_m\} = \varnothing$, hence $\tilde{B} := \bar{B} \cup \{\mathbf{x}\}$ has $p + 1 = n - m$ elements; moreover, $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\} \cup \tilde{B}$ is a basis for V ($\{\mathbf{x}_1, \ldots, \mathbf{x}_m\} \cup \tilde{B} = \{\mathbf{x}, \mathbf{x}_1, \ldots, \mathbf{x}_m\} \cup \bar{B}$). □

Theorem 2.5. *Let W be a linear subspace of a finite-dimensional linear space $(V, +, \cdot)$. Then W is finite-dimensional and $\dim W \leq \dim V$.*

Proof. We will prove this theorem by mathematical induction on $n = \dim V$.

I. If $n = 0$, then $V = \{\mathbf{0}\}$; we necessarily have $W = \{\mathbf{0}\}$, hence $\dim W = \dim V = 0$.

II. Suppose now that the result is true for some $n \geq 0$ and let us show it for $n + 1$. We can suppose, without loss of the generality, that $W \neq V$. Let $\mathbf{x}_0 \in V \setminus W$; by Proposition **??**, we can find $\tilde{B} \subseteq V$ with n elements such that $\{\mathbf{x}_0\} \cup \tilde{B}$ is a basis of V ($\mathbf{x}_0 \neq \mathbf{0}$, so $\mathbf{x}$ is linearly independent).

Now, if $\mathbf{x}$ is an arbitrary vector in W, there exist $\alpha_0, \alpha_1, \ldots, \alpha_n \in \mathbb{R}$ such that

$$\alpha_0\mathbf{x}_0 + \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n = \mathbf{x},$$

i.e.

$$\alpha_0\mathbf{x}_0 = \mathbf{x} - \alpha_1\mathbf{x}_1 - \cdots - \alpha_n\mathbf{x}_n \in W.$$

Since $\mathbf{x}_0 \notin W$, α_0 must be equal to 0 and therefore

$$\mathbf{x} = \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n.$$

This proves that $W \subseteq \operatorname{Lin}(\tilde{B})$. By the induction hypothesis, since $\dim \operatorname{Lin}(\tilde{B}) = n$, we have that W is finite-dimensional and

$$\dim W \leq \dim \operatorname{Lin}(\tilde{B}) = n < n + 1 = \dim V.$$

□

3. Change of coordinates

Proposition 3.1. *Let $(V, +, \cdot)$ be a finite-dimensional linear space with dimension $n \in \mathbb{N}^*$. If $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is a basis of V, then for every $\mathbf{x} \in V$ there exist and are unique $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ such that*

$$\mathbf{x} = \alpha_1\mathbf{b}_1 + \cdots + \alpha_n\mathbf{b}_n.$$

The scalars $\alpha_1, \ldots, \alpha_n$ are called the coordinates *of $\mathbf{x}$ with respect to $\mathbf{b}_1, \ldots, \mathbf{b}_n$.*

Remarks.

1. In $\mathbb{R}^n$, the coordinates of a vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ with respect to the elements of the canonical basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ are precisely $x_1, \ldots, x_n$ (*i.e.*, the coordinates of $\mathbf{x}$).

2. Let $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be a basis of $\mathbb{R}^n$ (not necessarily the canonical one), $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $\alpha_1, \ldots, \alpha_n$ the coordinates of $\mathbf{x}$ with respect to $\mathbf{b}_1, \ldots, \mathbf{b}_n$. Then the relation

$$\mathbf{x} = \alpha_1\mathbf{b}_1 + \cdots + \alpha_n\mathbf{b}_n$$

can be written in a matrix-way as

$$\mathbf{x}^T = \mathbf{B} \cdot X_B,$$

where

$$\mathbf{B} = \left[\mathbf{b}_1^T \ldots \mathbf{b}_n^T\right] \in \mathscr{M}_n$$

is the matrix having on the k-th column the coordinates of $\mathbf{b}_k$, while

$$\mathbf{X_B} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathscr{M}_{n,1}$$

is the column-matrix of the coordinates of of $\mathbf{x}$ with respect to $\mathbf{b}_1, \ldots, \mathbf{b}_n$.

Definition. Let $(V,+,\cdot)$ be a finite-dimensional linear space with dimension $n \in \mathbb{N}^*$, $B = \{\mathbf{b}_1,\dots,\mathbf{b}_n\}$ a basis of V and $B' = \{\mathbf{b}'_1,\dots,\mathbf{b}'_n\}$ a set of m vectors in V. We call the *transition matrix* from B to B' the matrix

$$\mathrm{S} = \begin{bmatrix} s_{1,1} & \dots & s_{1,m} \\ \vdots & & \vdots \\ s_{n,1} & \dots & s_{n,m} \end{bmatrix} \in \mathscr{M}_{n,m},$$

where, for $1 \le k \le m$, $s_{1,k},\dots,s_{n,k}$ are the coordinates of the vector $\mathbf{b}'_k$ with respect to $\mathbf{b}_1,\dots,\mathbf{b}_n$.

Formally, we can write

$$\mathbf{B}' = \mathbf{B}\cdot S,$$

where B and B' are the row-matrix formed with the elements of B and respectively B':

$$\mathbf{B} = [\mathbf{b}_1 \dots \mathbf{b}_n], \mathbf{B} = [\mathbf{b}'_1 \dots \mathbf{b}'_m]$$

Theorem 3.2. *Let $(V,+,\cdot)$ be a finite-dimensional linear space with dimension $n \in \mathbb{N}^*$. If $B = \{\mathbf{b}_1,\dots,\mathbf{b}_n\}$, $B' = \{\mathbf{b}'_1,\dots,\mathbf{b}'_n\}$ are two bases of V and S is the transition matrix from B to B', then the matrix S is non-singular and S^{-1} is the transition matrix from B' to B.*

Moreover, if $\mathbf{x} \in V$ and $\alpha_1,\dots,\alpha_n$, $\alpha'_1,\dots,\alpha'_n$ are the coordinates of $\mathbf{x}$ with respect to $\mathbf{b}_1,\dots,\mathbf{b}_n$, respectively $\mathbf{b}'_1,\dots,\mathbf{b}'_n$, then

$$X_{B'} = \mathrm{S}^{-1}\cdot X_B,$$

where

$$\mathbf{X_B} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathscr{M}_{n,1}, \mathbf{X'_B} = \begin{bmatrix} \alpha'_1 \\ \vdots \\ \alpha'_n \end{bmatrix} \in \mathscr{M}_{n,1}$$

Definition. Let $(V,+,\cdot)$ be a finite-dimensional linear space with dimension $n \in \mathbb{N}^*$. Let $B = \{\mathbf{b}_1,\dots,\mathbf{b}_n\}$, $B' = \{\mathbf{b}'_1,\dots,\mathbf{b}'_n\}$ be two bases of V and S the transition matrix from B to B'. We say that B and B' have the *same orientation* if $\det \mathrm{S} > 0$.

4. Scalar products. Norms

Definition. Let $(V,+,\cdot)$ be a linear space. We say that a map $\langle\cdot,\cdot\rangle : V\times V \to \mathbb{R}$ is a *scalar product* on V if:

(i) $\langle\mathbf{x},\mathbf{x}\rangle \ge 0,\ \forall \mathbf{x}\in V$ (*positive definitness*);
(ii) $\langle\mathbf{x},\mathbf{x}\rangle = 0 \Rightarrow \mathbf{x} = 0,\ \forall\mathbf{x}\in V$;
(iii) $\langle\mathbf{x},\mathbf{y}\rangle = \langle\mathbf{y},\mathbf{x}\rangle,\ \forall\mathbf{x},\mathbf{y}\in V$ (*symmetry*);
(iv) $\langle\alpha\cdot\mathbf{x}+\beta\cdot\mathbf{y},\mathbf{z}\rangle = \alpha\langle\mathbf{x},\mathbf{z}\rangle + \beta\langle\mathbf{y},\mathbf{z}\rangle,\ \forall\alpha,\beta\in\mathbb{R},\ \forall\mathbf{x},\mathbf{y},\mathbf{z}\in V$ (*bilinearity*).

In this case, the quadruple $(V,+,\cdot,\langle\cdot,\cdot\rangle)$ is called a *prehilbertian space*.

For the sake of simplicity, we usually denote $(V,\langle\cdot,\cdot\rangle)$ instead of $(V,+,\cdot,\langle\cdot,\cdot\rangle)$, the operations $+$ and $\cdot$ being considered implicit.

Proposition 4.1. *Let $n\in\mathbb{N}^*$ and $\langle\cdot,\cdot\rangle : \mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ be defined as*

$$\langle(x_1,\dots,x_n),(y_1,\dots,y_n)\rangle := x_1y_1+\dots+x_ny_n,\ (x_1,\dots,x_n),(y_1,\dots,y_n)\in\mathbb{R}^n.$$

Then $\langle\cdot,\cdot\rangle$ is a scalar product on $\mathbb{R}^n$, called the Euclidean *(or* canonical*)* scalar product.

The proof of this result is immediate and left as exercice. Unless stated otherwise, we will always assume that $\mathbb{R}^n$ is endowed with the Euclidean scalar product.

Definition. Let $(V,\langle\cdot,\cdot\rangle)$ be a prehilbertian space.

a) We say that two vectors $\mathbf{x}\in V$ and $\mathbf{y}\in V$ are *orthogonal* and we denote $\mathbf{x}\perp\mathbf{y}$ if $\langle\mathbf{x},\mathbf{y}\rangle = 0$.
b) Let $\mathbf{x}\in V$ and U a non-empty subset of V. We say that x is orthogonal on U and we denote $\mathbf{x}\perp U$ if $\mathbf{x}\perp\mathbf{y}$ for every $\mathbf{y}\in U$.
c) If U is non-empty subset of V, we call U an *orthogonal system* if $\mathbf{x}\perp\mathbf{y}$ for any distinct $\mathbf{x},\mathbf{y}\in U$.
d) Let $U\subseteq V$. The *orthogonal complement* of U is the set

$$U^\perp := \{\mathbf{x}\in V \mid \mathbf{x}\perp U\}.$$

Remark. Let U be a non-empty subset of V. One can show that if $\mathbf{x} \in V$, then $\mathbf{x} \perp U$ if and only if $\mathbf{x} \perp \mathrm{Lin}(U)$. Therefore, $U^\perp = \mathrm{Lin}(U)^\perp$. It is also easy to prove that $U \cap U^\perp = \{\mathbf{0}\}$.

Definition. Let $(V, \langle\cdot,\cdot\rangle)$ be a prehilbertian space. For $\mathbf{x}, \mathbf{y} \in V \setminus \{\mathbf{0}\}$, we call the *angle* between $\mathbf{x}$ and $\mathbf{y}$ the number

$$\widehat{(\mathbf{x}, \mathbf{y})} = \sphericalangle(\mathbf{x}, \mathbf{y}) := \arccos \frac{\langle \mathbf{x}, \mathbf{y}\rangle}{\sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}\sqrt{\langle \mathbf{y}, \mathbf{y}\rangle}}.$$

It is clear that $\widehat{(\mathbf{x}, \mathbf{y})} = \widehat{(\mathbf{y}, \mathbf{x})} \in [0, \pi]$, $\forall \mathbf{x}, \mathbf{y} \in V \setminus \{\mathbf{0}\}$. Moreover, if $\mathbf{x}, \mathbf{y} \in V \setminus \{\mathbf{0}\}$, $\widehat{(\mathbf{x}, \mathbf{y})} = \pi/2$ if and only if $\mathbf{x} \perp \mathbf{y}$.

Definition. Let $(V, +, \cdot)$ be a linear space. We say that a map $\|\cdot\| : V \to \mathbb{R}$ is a *norm* on V if:

(i) $\|\mathbf{x}\| \geq 0$, $\forall \mathbf{x} \in V$;
(ii) $\|\mathbf{x}\| = 0 \Rightarrow \mathbf{x} = 0$, $\forall \mathbf{x} \in V$;
(iii) $\|\lambda\mathbf{x}\| = |\lambda| \|\mathbf{x}\|$, $\forall \lambda \in \mathbb{R}$, $\forall \mathbf{x} \in V$ (*homogeneity*);
(iv) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$, $\forall \mathbf{x}, \mathbf{y} \in V$ (*triangle property*).

In this case, the quadruple $(V, +, \cdot, \|\cdot\|)$ is called a *normed space*.

Also for simplicity, we usually denote $(V, \|\cdot\|)$ instead of $(V, +, \cdot, \|\cdot\|)$.

Proposition 4.2. *Let $(V, \langle\cdot,\cdot\rangle)$ be a prehilbertian space. Then the mapping $\|\cdot\| : V \to \mathbb{R}$ defined by*

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

is a norm on V, called the norm induced *by the scalar product $\langle\cdot,\cdot\rangle$.*

Proof. The first three properties in the norm definition are obvious, so we only prove the fourth.
For $\mathbf{x}, \mathbf{y} \in V$, we have, by the bilinearity and symmetry of the scalar product,

$$\langle \mathbf{x} + \lambda\mathbf{y}, \mathbf{x} + \lambda\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{x}\rangle + 2\lambda \langle \mathbf{x}, \mathbf{y}\rangle + \lambda^2 \langle \mathbf{y}, \mathbf{y}\rangle = \|\mathbf{x}\|^2 + 2\lambda \langle \mathbf{x}, \mathbf{y}\rangle + \lambda^2 \|\mathbf{y}\|^2, \ \forall \lambda \in \mathbb{R}. \tag{1}$$

But $\langle \mathbf{x} + \lambda\mathbf{y}, \mathbf{x} + \lambda\mathbf{y}\rangle \geq 0$, $\forall \lambda \in \mathbb{R}$. This implies that

$$\langle \mathbf{x}, \mathbf{y}\rangle^2 - \|\mathbf{x}\| \cdot \|\mathbf{y}\| \leq 0$$

(otherwise, the equation $\|\mathbf{x}\|^2 + 2\lambda \langle \mathbf{x}, \mathbf{y}\rangle + \lambda^2 \|\mathbf{y}\|^2 = 0$ would have two different solutions and $\lambda \mapsto \|\mathbf{x}\|^2 + 2\lambda \langle \mathbf{x}, \mathbf{y}\rangle + \lambda^2 \|\mathbf{y}\|^2$ would take negative values), *i.e.*

$$\langle \mathbf{x}, \mathbf{y}\rangle^2 \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$$

(this is known as *Schwarz inequality*). Taking now $\lambda = 1$ in (??), we obtain

$$\|\mathbf{x} + \mathbf{y}\|^2 = \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}\rangle = \|\mathbf{x}\|^2 + 2 \langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2 \leq \|\mathbf{x}\|^2 + 2 \|\mathbf{x}\| \cdot \|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2,$$

which clearly implies

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|.$$

□

There are norms which are not induced by any scalar product, as we will see in the next lecture.

Definition. Let $n \in \mathbb{N}^*$. The norm induced by the Euclidean scalar product on $\mathbb{R}^n$ is called the *Euclidean norm* and is denoted $\|\cdot\|_2$.

If $(x_1, \ldots, x_n) \in \mathbb{R}^n$, then $\|(x_1, \ldots, x_n)\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$.

Definition. Let $(V, \|\cdot\|)$ be a normed space. A vector $\mathbf{x} \in V$ such that $\|\mathbf{x}\| = 1$ is called a *versor*.

Definition. Let $(V, \langle\cdot,\cdot\rangle)$ be a prehilbertian space.

a) A non-empty subset $U \subseteq V$ is called an *orthonormal system* if U is an orthogonal system and every element of U is a versor.
b) If B is a basis of V and B is an orthogonal system, then B is called an *orthogonal basis*.
c) If B is a basis of V and B is an orthonormal system, then B is called an *orthonormal basis*.

In other words, U is an orthonormal system if and only if for any $\mathbf{x}, \mathbf{y} \in U$ we have

$$\langle \mathbf{x}, \mathbf{y}\rangle = \begin{cases} 0, & \mathbf{x} \neq \mathbf{y}; \\ 1, & \mathbf{x} = \mathbf{y}. \end{cases}$$

Of course, the canonical basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ in $\mathbb{R}^n$ is an orthonormal basis.

Definition. Let $(V, \langle\cdot,\cdot\rangle)$ be a finite-dimensional prehilbertian space with dimension $n \in \mathbb{N}^*$ and $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ a basis of V. We call the *Gram determinant* associated with the basis B the number $\det \mathrm{G} \in \mathbb{R}$, where

$$\mathrm{G} := \begin{bmatrix} \langle \mathbf{b}_1, \mathbf{b}_1\rangle & \langle \mathbf{b}_1, \mathbf{b}_2\rangle & \ldots & \langle \mathbf{b}_1, \mathbf{b}_n\rangle \\ \langle \mathbf{b}_2, \mathbf{b}_1\rangle & \langle \mathbf{b}_2, \mathbf{b}_2\rangle & \ldots & \langle \mathbf{b}_2, \mathbf{b}_n\rangle \\ \vdots & \vdots & & \vdots \\ \langle \mathbf{b}_n, \mathbf{b}_1\rangle & \langle \mathbf{b}_n, \mathbf{b}_2\rangle & \ldots & \langle \mathbf{b}_n, \mathbf{b}_n\rangle \end{bmatrix} \in \mathscr{M}_n$$

Of course, G is a symmetric matrix. Moreover, G is non-singular. Indeed, if we denote $g_{ij} := \langle \mathbf{b}_i, \mathbf{b}_j\rangle$ for $1 \le i, j \le n$, then solving the homogeneous system with n equations and n unknowns

$$\left\{ \begin{array}{lcl} g_{11}x_1 + \cdots + g_{1n}x_n & = & 0 \\ \ldots\ldots\ldots\ldots\ldots & & \vdots \\ g_{n1}x_1 + \cdots + g_{nn}x_n & = & 0 \end{array} \right. \tag{2}$$

is equivalent to finding $\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n$ such that $\langle \mathbf{b}_1, \mathbf{x}\rangle = 0, \ldots, \langle \mathbf{b}_n, \mathbf{x}\rangle = 0$, *i.e.* $\mathbf{x} \perp B$. Since $B^\perp = \mathrm{Lin}(B)^\perp = V^\perp = \{\mathbf{0}\}$, we have that $\mathbf{x}$ has to be $\mathbf{0}$, which means, by the linear independency of $\mathbf{b}_1, \ldots, \mathbf{b}_n$, that the system (??) has only the trivial solution. By the theory of linear systems in $\mathbb{R}$, G is invertible.

If $\mathbf{x}, \mathbf{y} \in V$ and $\alpha_1, \ldots, \alpha_n$, $\beta_1, \ldots, \beta_n$ are the coordinates of $\mathbf{x}$, respectively $\mathbf{y}$, with respect to $\mathbf{b}_1, \ldots, \mathbf{b}_n$, then

$$\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\beta_j \langle \mathbf{b}_i, \mathbf{b}_j\rangle = X_B^{\mathrm{T}} \cdot \mathrm{G} \cdot Y_B,$$

where $X_B = [\alpha_1, \ldots, \alpha_n]^T \in \mathscr{M}_{n,1}$ and $Y_B = [\beta_1, \ldots, \beta_n]^T \in \mathscr{M}_{n,1}$.

One last remark is that the basis B is orthogonal or orthonormal if and only if G is a *diagonal matrix*, respectively $\mathrm{G} = \mathrm{I}_n$.

Theorem 4.3 (Gram-Schmidt orthonormalization procedure). *Let $(V, \langle\cdot,\cdot\rangle)$ be a finite-dimensional prehilbertian space with dimension $n \in \mathbb{N}^*$. If $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ a basis of V, there exists an orthonormal basis $B' = \{\mathbf{b}'_1, \ldots, \mathbf{b}'_n\}$ such that $\mathrm{Lin}(\{\mathbf{b}'_1, \ldots, \mathbf{b}'_k\}) = \mathrm{Lin}(\{\mathbf{b}_1, \ldots, \mathbf{b}_k\})$ for every $k \in \{1, \ldots, n\}$.*

One important aspect of this result is that every finite-dimensional prehilbertian space has an orthonormal basis.

Proof. It is enough to relax the result by demanding the existence of an orthogonal basis with the required properties instead of an orthonormal one (if $\{\mathbf{b}'_1, \ldots, \mathbf{b}'_n\}$ is an orthogonal basis, then $\left\{\frac{\mathbf{b}'_1}{\|\mathbf{b}'_1\|}, \ldots, \frac{\mathbf{b}'_n}{\|\mathbf{b}'_n\|}\right\}$ is an orthonormal basis).

Step 1. We take $\mathbf{b}'_1 = \mathbf{b}_1$.

Step 2. Suppose that, for $k < n$, we have already found $\mathbf{b}'_1, \ldots, \mathbf{b}'_k$ with $\{\mathbf{b}'_1, \ldots, \mathbf{b}'_k\}$ an orthogonal system such that $\mathrm{Lin}(\{\mathbf{b}'_1, \ldots, \mathbf{b}'_k\}) = \mathrm{Lin}(\{\mathbf{b}_1, \ldots, \mathbf{b}_k\})$. We determine $\mathbf{b}'_{k+1} = \lambda_1\mathbf{b}'_1 + \cdots + \lambda_k\mathbf{b}'_k + \mathbf{b}_{k+1}$ such that $\mathbf{b}'_{k+1} \perp \mathbf{b}'_j$, $\forall j \in \{1, \ldots, k\}$. This means that

$$\lambda_j \left\|\mathbf{b}'_j\right\|^2 + \left\langle \mathbf{b}_{k+1}, \mathbf{b}'_j\right\rangle = 0, \ \forall j \in \{1, \ldots, k\},$$

i.e. $\lambda_j = -\frac{\langle \mathbf{b}_{k+1}, \mathbf{b}'_j\rangle}{\left\|\mathbf{b}'_j\right\|^2}$ for $j \in \{1, \ldots, k\}$. In conclusion, we have found

$$\mathbf{b}'_{k+1} = \mathbf{b}_{k+1} - \frac{\langle \mathbf{b}_{k+1}, \mathbf{b}'_1\rangle}{\left\|\mathbf{b}'_1\right\|^2}\mathbf{b}'_1 - \cdots - \frac{\langle \mathbf{b}_{k+1}, \mathbf{b}'_k\rangle}{\left\|\mathbf{b}'_k\right\|^2}\mathbf{b}'_k.$$

Step 3. We repeat **Step 2** until we arrive to $k + 1 = n$. □

The algorithm described in the proof of this theorem is called the *Gram-Schmidt orthonormalization procedure.*

Selective bibliography

[1] V. Borcea, C. Davideanu, Corina Forăscu, *Probleme de Algebră Liniară*, Gh. Asachi Publishing House, Iaşi, 2000.
[2] S. Corbu, *Algebră liniară.Geometrie analitică (elemente de teorie şi aplicaţii)*, Editura "Fair Partners", Bucureşti, 2011.
[3] N. Cotfas, *Elemente de algebră liniară*, Editura Universităţii din Bucureşti, 2009.
[4] M. Gorunescu, F. Gorunescu, A. Prodan, *Matematici superioare*, Biostatistică şi Informatică, Editura Albastră, Cluj-Napoca, 2002.
[5] J. Hefferon, *Linear Algebra*, Saint Michael's College, 2014.
[6] Ion D. Ion, R. Nicolae, *Algebră*, Editura Didactică şi Pedagogică, Bucureşti, 1981.
[7] V. Postolică, *Eficienţă prin matematica aplicată*, Editura Matrix Rom, Bucureşti, 2006.
[8] A. Precupanu, *Bazele analizei matematice*, Editura Polirom, Iaşi, 1998.
[9] R. Singh, M. Bhatt, *Linear Algebra and Vector Calculus*, Mc Graw Hill Ed., 2016.
[10] S. Treil, *Linear Algebra Done Wrong*, Brown University, 2013.