Functions and linear mappings in $\mathbb{R}^n$

Lecture 7

Mathematics - 1st year, English

Faculty of Computer Science
Alexandru Ioan Cuza University of Iasi

e-mail: corina.forascu@gmail.com

facebook: Corina Forăscu

November 13, 2018

Outline of the lecture

1. Functions in Euclidean spaces
   - Functions
   - Functions of several variables

2. Linear maps
   - Matrices associated with linear operators
   - Adjoint operators
   - Eigenvalues and eigenvectors

Functions

Definition

Let A and B be sets. We say that a relation $f \subseteq A \times B$ is a *function from A to B* and we denote $f : A \to B$ if:

1. $\mathrm{Dom}\, f = A$;
2. $(x, y) \in f,\ (x, z) \in f \Rightarrow y = z,\ \forall x \in A,\ \forall y, z \in B$.

For $x \in A$ we denote by $f(x)$ the unique element $y \in B$ such that $(x, y) \in f$.

Proposition

i) *If $f : A \to B$ and $g : B \to C$ are functions, then $g \circ f$ is a function, $g \circ f : A \to C$ and*

$$(g \circ f)(x) = g(f(x)),\ \forall x \in A.$$

ii) *If $f : A \to B$, $g : B \to C$ and $h : C \to D$ are functions, then*

$$(h \circ g) \circ f = h \circ (g \circ f).$$

Definition

If $f : A \to B$ is a function and $E \subseteq A$, $F \subseteq B$, we denote:

- $f|_E := \{(x, f(x)) \mid x \in E\}$, the *restriction* of f to the subset E;
- $f[E] := \{f(x) \mid x \in E\}$, the *image* of f *through* the subset E;
- $\mathrm{Im}\, f := f[A]$, the *image* of f;
- $f^{-1}[F] := \{x \in A \mid f(x) \in F\}$, the the *preimage* or the *inverse image* of f *through* the subset F.

Of course:

- $\mathrm{Dom}\, f|_E = E$ and $f|_E(x) = f(x), \ \forall x \in E$;
- $f^{-1}[B] = \mathrm{Dom}\, f = A$ and $f[\emptyset] = f^{-1}[\emptyset] = \emptyset$.

Definition

A function $f : A \to B$ is called:

- *injective* or *one-to-one* if for any $x, y \in A$,

$$f(x) = f(y) \Rightarrow x = y;$$

- *surjective* or *onto* if $\mathrm{Im}\, f = B$, *i.e.*

$$\forall y \in B, \exists x \in A : f(x) = y;$$

- *bijective* if it is both injective and surjective;
- *invertible* if there exists $g : B \to A$ such that $f \circ g = 1_B$ and $g \circ f = 1_A$.

Proposition

A function $f : A \to B$ is bijective if and only if it is invertible. In this case, f^{-1} is a bijective function from B to A and

$$f \circ f^{-1} = 1_B, \quad f^{-1} \circ f = 1_A.$$

Functions of several variables

- Functions of the type

$$f : \mathrm{D}(f) \subseteq \mathbb{R}^n \to \mathbb{R}^m, \text{ where } m, n \in \mathbb{N}^*,$$

  are called *vector valued* (or $\mathbb{R}^m$*-valued*) *functions of n* (*real*) *variables*.
- In the case $m = 1$, we will simply call the function f a *real* (or *real-valued*) *function of n* (*real*) *variables*.
- In the case $m > 1$, for every $\mathbf{x} = (x_1, \ldots, x_n) \in \mathrm{D}(f)$, $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ has m components, that we will usually denote

$$f_1(\mathbf{x}) = f_1(x_1, \ldots, x_n),\ f_2(\mathbf{x}) = f_2(x_1, \ldots, x_n), \ldots,\ f_m(\mathbf{x}) = f_m(x_1, \ldots, x_n).$$

  Hence, we have m real functions of n variables, $f_j : \mathrm{D}(f) \to \mathbb{R}$, $1 \leq j \leq m$, such that

$$f(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), f_2(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)).$$

- Conversely, if $f_j : \mathrm{D}(f_j) \to \mathbb{R}$, $1 \leq j \leq m$ are m real functions of n variables, then we can define an $\mathbb{R}^m$-valued function of n variables by the above formula, where this time $\mathrm{D}(f) := \mathrm{D}(f_1) \cap \cdots \cap \mathrm{D}(f_m)$.

Examples

1. *Basic elementary functions:*

- the *constant* function:
  the function $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = c$, $\forall x \in \mathbb{R}$, where $c \in \mathbb{R}$;
- the *identity* function $1_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$ (recall that $1_{\mathbb{R}}(x) = x$, $\forall x \in \mathbb{R}$);
- the *exponential* function with *basis* $a > 0$: the function $f : \mathbb{R} \to \mathbb{R}$, $f(x) := a^x$, $\forall x \in \mathbb{R}$;
- the *logarithmic* function with *basis* $a > 0$, $a \neq 1$: $\log_a : (0, +\infty) \to \mathbb{R}$ is the inverse of the *exponential* function with *basis* $a > 0$;
- the *power function* with exponent $a \in \mathbb{R}$:
  $f : \mathrm{D}(f) \subseteq \mathbb{R} \to \mathbb{R}$, with $f(x) := x^a$, $\forall x \in \mathbb{R}$;
- the (*direct*) *trigonometric* functions: cos, sin, tg, ctg;
- the *inverse trigonometric* functions: arccos, arcsin, arctg, arcctg.

2. *Elementary functions:* Any function which can be obtained by applying all or some of the four basic operations on basic elementary functions: *addition*, *multiplication*, *subtraction* and *division*.

3. *Special functions:*

- *floor* function: $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) := \lfloor x \rfloor = \sup \{n \in \mathbb{Z} \mid n \leq x\}$;
- *ceiling* function: $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) := \lceil x \rceil = \inf \{n \in \mathbb{Z} \mid n \geq x\}$;
- *sawtooth* function: $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) := \{x\} = x - \lfloor x \rfloor$;
- *sign* function: $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) := \operatorname{sgn} x = \begin{cases} -1, & x < 0; \\ 0, & x = 0; \\ 1, & x > 0; \end{cases}$
- *absolute value* function: $f : \mathbb{R} \to \mathbb{R}$ defined by
  $f(x) := |x| = \begin{cases} x, & x \geq 0; \\ -x, & x < 0; \end{cases}$
- *positive part* function: $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) := x^+ = \begin{cases} x, & x \geq 0; \\ 0, & x < 0; \end{cases}$
- *negative part* function: $f : \mathbb{R} \to \mathbb{R}$ defined by
  $f(x) := x^- = \begin{cases} 0, & x \geq 0; \\ -x, & x < 0; \end{cases}$

- *Heaviside* function: $f:\mathbb{R}\to\mathbb{R}$ defined by $f(x) := \begin{cases} 1, & x \geq 0; \\ 0, & x < 0; \end{cases}$
- *Dirichlet* function: $f:\mathbb{R}\to\mathbb{R}$ defined by $f(x) := \begin{cases} 1, & x \in \mathbb{Q}; \\ 0, & x \in \mathbb{R}\setminus\mathbb{Q}; \end{cases}$
- *Riemann* function: $f:[0,1]\to\mathbb{R}$ defined by
$$f(x) := \begin{cases} 0, & x = 0 \text{ or } x \in (0,1)\setminus\mathbb{Q}; \\ \dfrac{1}{q}, & x = \dfrac{p}{q} \text{ with } p \in \mathbb{N},\ q \in \mathbb{N}^*,\ (p,q) = 1. \end{cases}$$

Examples of real functions of several variables

1. $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$, defined by

$$f(x_1, x_2) := -\sqrt{\sin(x_1^2 + x_2^2)},\ (x_1, x_2) \in A,$$

where

$$\begin{aligned} A := & \quad \left\{ (x_1, x_2) \in \mathbb{R}^2 \mid \sin(x_1^2 + x_2^2) \geq 0 \right\} \\ = & \quad \left\{ (x_1, x_2) \in \mathbb{R}^2 \mid \exists k \in \mathbb{N} : 2k\pi \leq x_1^2 + x_2^2 \leq (2k+1)\pi \right\}. \end{aligned}$$

2. $f : A \subseteq \mathbb{R}^3 \to \mathbb{R}$, defined by

$$f(x_1, x_2, x_3) := \ln(1 - x_1 - x_2 - x_3) - (x_1 + x_3)^{x_2},\ (x_1, x_2, x_3) \in A,$$

where

$$A := \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 < 1,\ x_1 + x_3 > 0 \right\}.$$

3. The *polynomial function* $P : \mathbb{R}^n \to \mathbb{R}$ defined by

$$P(x_1, x_2, \ldots, x_n) := \sum_{i_1, i_2, \ldots, i_n = 0}^{k_1, k_2, \ldots, k_n} a_{i_1, i_2, \ldots, i_n} x_1^{i_1} \cdot x_2^{i_2} \cdot \cdots \cdot x_n^{i_n}, \ (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n.$$

- The real numbers $a_{i_1, i_2, \ldots, i_n}$ are called the *coefficients* of the polynomial P.
- Every term $a_{i_1, i_2, \ldots, i_n} x_1^{i_1} \cdot x_2^{i_2} \cdot \cdots \cdot x_n^{i_n}$ where $a_{i_1, i_2, \ldots, i_n} \neq 0$ is called a *monomial* (of P).
- The *degree* of this monomial is $i_1 + i_2 + \cdots + i_n$.
- We call the *degree* of the polynomial P the largest degree among all its monomials.
- We say that the polynomial P is *homogeneous* if all its monomials have the same degree.
  An example is the following polynomial of degree 1:

$$P(x_1, x_2, \ldots, x_n) := a_1 x_1 + a_2 x_2 + \cdots + a_n x_n, \ (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n.$$

- A polynomial P is called *symmetric polynomial* if for every *permutation* (i.e., bijective function) $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$, we have

$$\sum_{i_1,i_2,\ldots,i_n=0}^{k_1,k_2,\ldots,k_n} a_{i_1,i_2,\ldots,i_n} x_1^{i_1} \cdot x_2^{i_2} \cdot \cdots \cdot x_n^{i_n} = \sum_{i_1,i_2,\ldots,i_n=0}^{k_1,k_2,\ldots,k_n} a_{i_1,i_2,\ldots,i_n} x_{\sigma(1)}^{i_1} \cdot x_{\sigma(2)}^{i_2} \cdot \cdots \cdot x_{\sigma(n)}^{i_n}.$$

  For instance, $P(x_1, x_2) := ax_1^2 + bx_1x_2 + cx_2^2$, $(x_1, x_2) \in \mathbb{R}^2$ is a symmetric polynomial if and only if $a = c$.

Linear maps

Definition

Let $(V, +, \cdot)$ and $(W, +, \cdot)$ two linear spaces. A function $T : V \to W$ is called *linear* if:

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}),\ \forall \mathbf{u}, \mathbf{v} \in V$ (*additivity*);
2. $T(\alpha \cdot \mathbf{u}) = \alpha \cdot T(\mathbf{u}),\ \forall \alpha \in \mathbb{R},\ \forall \mathbf{u} \in V$ (*homogeneity*).

We use also the name *linear operator* or *linear map/mapping* for linear functions.

Example. All polynomials of degree 1 are linear mappings. Recall that a polynomial of degree 1 is a function $P : \mathbb{R}^n \to \mathbb{R}$ of the form

$$P(x_1, x_2, \ldots., x_n) := a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n,\ (x_1, x_2, \ldots., x_n) \in \mathbb{R}^n.$$

Proposition

Let $(V, +, \cdot)$ and $(W, +, \cdot)$ two linear spaces. The function $T : V \to W$ is a linear operator if and only if

$$T(\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v}) = \alpha \cdot T(\mathbf{u}) + \beta \cdot T(\mathbf{v}),\ \forall \alpha, \beta \in \mathbb{R},\ \forall \mathbf{u}, \mathbf{v} \in V.$$

Remarks

1. When the linear map $T : V \to W$ is bijective, T is called a *linear isomorphism* between V and W.
It is easy to prove that $T^{-1} : W \to V$ is also a linear isomorphism.
We say that two linear spaces V and W are *isomorphic* if there is at least a linear isomorphism between the two spaces.

2. If $V = W$, a linear map $T : V \to V$ is also called *linear endomorphism*.
The identity function 1_V is clearly a linear endomorphism on V.

3. Let $T : V \to W$ be a linear map. If $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and $\mathbf{u}_1, \ldots, \mathbf{u}_n \in V$, then

$$T(\alpha_1 \cdot \mathbf{u}_1 + \cdots + \alpha_n \cdot \mathbf{u}_n) = \alpha_1 \cdot T(\mathbf{u}_1) + \cdots + \alpha_n \cdot T(\mathbf{u}_n).$$

Of course, $T(\mathbf{0}) = \mathbf{0}$ ($T(\mathbf{0}_V) = \mathbf{0}_W$).

Remarks (contd)

4. If V and W are linear spaces, we denote $\mathrm{L}(V; W)$ the set of all linear maps between V and W.
It is clear that (see Lecture 5) $\mathrm{L}(V; W)$ is still a linear space (when endowed with the natural the addition and multiplication with scalars of functions).
If $V = W$ we simply denote $\mathrm{L}(V)$ instead $\mathrm{L}(V; V)$.

5. Let U, V and W be linear spaces. If $T : V \to W$ and $S : U \to V$ are linear operators, then $T \circ S$ is still a linear operator between V and W.
Therefore, the composition $\circ$ introduces a new internal operation on $\mathrm{L}(V)$, which is associative and has 1_V as neutral element.

Kernel and range of a linear operator

Definition

Let $(V, +, \cdot)$ and $(W, +, \cdot)$ two linear spaces and $T : V \to W$ a linear operator.

- The set
$$\ker T := \{\mathbf{u} \in V \mid T(\mathbf{u}) = \mathbf{0}_W\} = T^{-1}[\{\mathbf{0}_W\}].$$
is called the *kernel* or the *null space* of the operator T.
- The set $\mathrm{Im}\, T$ is sometimes called the *range* of T.

Proposition

Let $(V, +, \cdot)$ and $(W, +, \cdot)$ two linear spaces and $T : V \to W$ a linear operator.

i) $\ker T$ *is a linear subspace of V and $\mathrm{Im}\, T$ is a linear subspace of W.*

ii) *T is injective if and only if $\ker T = \{\mathbf{0}_V\}$.*

Dimension theorem

The next result is one of the fundamental results of linear algebra.

Theorem

Let $(V, +, \cdot)$ be a finite-dimensional linear space, $(W, +, \cdot)$ a linear space and $T : V \to W$ a linear operator. Then $\mathrm{Im}\, T$ is a finite-dimensional subspace of W and

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

- The above relation is called the *dimension formula*.

Let $T : V \to W$ be a linear operator between linear spaces.

- If $\ker T$ is finite-dimensional, the number $\dim(\ker T)$ is called the *nullity* of T and is denoted by $\mathrm{null}\, T$.
- If $\mathrm{Im}\, T$ is finite-dimensional, then $\dim(\mathrm{Im}\, T)$ is called the *rank* of T and is denoted by $\mathrm{rank}\, T$.
- The dimension formula becomes

$$\mathrm{null}\, T + \mathrm{rank}\, T = \dim V.$$

Proposition

Let $(V, +, \cdot)$ be a finite-dimensional linear space, $(W, +, \cdot)$ a linear space and $T : V \to W$ a linear operator. The following statements are equivalent:

1. *T is injective;*
2. $\text{rank}\, T = \dim V$;
3. $\text{null}\, T = 0$;
4. *for any linearly independent vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ in V, the vectors $T(\mathbf{u}_1), \ldots, T(\mathbf{u}_n)$ are linearly independent.*

Proposition

Let $(V, +, \cdot)$ be linear space, $(W, +, \cdot)$ a finite-dimensional linear space and $T : V \to W$ a linear operator. The following statements are equivalent:

1. *T is surjective;*
2. $\text{rank}\, T = \dim W$;
3. *for any vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ which generate V, the vectors $T(\mathbf{u}_1), \ldots, T(\mathbf{u}_n)$ generate W.*

Proposition

Let $(V, +, \cdot)$ and $(W, +, \cdot)$ be two finite-dimensional linear spaces and $T : V \to W$ a linear operator. The following statements are equivalent:

1. *T is bijective;*
2. $\operatorname{rank} T = \dim V = \dim W$;
3. *for any basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ of V, the set $T[B] = \{T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)\}$ is a basis of W.*

Matrices associated with linear operators

- Let $(V,+,\cdot)$, $(W,+,\cdot)$ be two finite-dimensional linear spaces with $\dim V = n$ and $\dim W = m$.
- Let $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$: a basis of V and $\bar{B} = \{\bar{\mathbf{b}}_1, \dots, \bar{\mathbf{b}}_m\}$: a basis of W.

1. Suppose that $T : V \to W$ is a linear operator.
a) For every $k \in \{1, \dots, n\}$ we can write

$$T(\mathbf{b}_k) = a_{1k}\bar{\mathbf{b}}_1 + \cdots + a_{mk}\bar{\mathbf{b}}_m,$$

i.e. $a_{1k}, \dots, a_{mk} \in \mathbb{R}$ are the coordinates of $T(\mathbf{b}_k)$ with respect to the basis $\bar{B}$. Then the matrix

$$A_{B,\bar{B}} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \in \mathscr{M}_{mn}$$

is called the *matrix associated* to the operator T with respect to the bases B, $\bar{B}$.

b) If $\mathbf{v} \in V$, let $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ be the coordinates of $\mathbf{v}$ with respect to B. Then

$$\begin{aligned} T(\mathbf{v}) &= T(\alpha_1 \mathbf{b}_1 + \cdots + \alpha_n \mathbf{b}_n) = \alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n) \\ &= \alpha_1(a_{11}\bar{\mathbf{b}}_1 + \cdots + a_{m1}\bar{\mathbf{b}}_m) + \cdots + \alpha_n(a_{1n}\bar{\mathbf{b}}_1 + \cdots + a_{mn}\bar{\mathbf{b}}_m) \\ &= (\alpha_1 a_{11} + \cdots + \alpha_n a_{1n})\bar{\mathbf{b}}_1 + \cdots + (\alpha_1 a_{m1} + \cdots + \alpha_n a_{mn})\bar{\mathbf{b}}_m. \end{aligned}$$

This means that if a vector $\mathbf{v} \in V$ has $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ as coordinates and $T(\mathbf{v}) \in W$ has $\beta_1, \ldots, \beta_m \in \mathbb{R}$ as coordinates, then

$$X_{\bar{B}} = A_{B,\bar{B}} \cdot X_B,$$

where $X_B := [\alpha_1, \ldots, \alpha_n]^T \in \mathscr{M}_{n1}$ and $X_{\bar{B}} := [\beta_1, \ldots, \beta_m]^T \in \mathscr{M}_{m1}$.

c) Let $r \in \{1, \dots, \min\{m, n\}\}$ be the the *rank* of the matrix $A_{B,\bar{B}}$.

- Since r is the maximal number of independent vectors among $T(\mathbf{b}_1), \dots, T(\mathbf{b}_n)$ (for that, see Theorem 2.2 in Lecture 5), let's say $T(\mathbf{b}_{k_1}), \dots, T(\mathbf{b}_{k_r})$, it follows that $\dim(\mathrm{Im}\, T) \geq r$.
- On the other hand, supposing that $\dim(\mathrm{Im}\, T) > r$, one can find $\mathbf{v} \in V$ such that $T(\mathbf{b}_{k_1}), \dots, T(\mathbf{b}_{k_r})$ and $T(\mathbf{v})$ are linear independent (Propostion 2.4 in Lecture 5).
- But $T(\mathbf{v})$ is a linear combination of $T(\mathbf{b}_1), \dots, T(\mathbf{b}_n)$. Since for every $k \notin \{\mathbf{b}_{k_1}, \dots, \mathbf{b}_{k_r}\}$, $T(\mathbf{b}_k)$ is a linear combination of $T(\mathbf{b}_{k_1}), \dots, T(\mathbf{b}_{k_r})$, it follows that $T(\mathbf{v})$ is a linear combination of $T(\mathbf{b}_{k_1}), \dots, T(\mathbf{b}_{k_r})$, which contradicts the linear independency of $T(\mathbf{b}_{k_1}), \dots, T(\mathbf{b}_{k_r})$ and $T(\mathbf{v})$.
- Therefore, $\dim(\mathrm{Im}\, T) = r$, *i.e.*

$$\mathrm{rank}\, A_{B,\bar{B}} = \mathrm{rank}\, T.$$

Change of bases

d) Let $B' = \{\mathbf{b}'_1, \ldots, \mathbf{b}'_n\}$ be another basis of V and $\bar{B}' = \{\bar{\mathbf{b}}'_1, \ldots, \bar{\mathbf{b}}'_m\}$ be another basis of W.
Let us denote $S = (s_{ij})_{1 \le i,j \le n} \in \mathscr{M}_n$ the transition matrix from B to B' and $\bar{S} = (\bar{s}_{ij})_{1 \le i,j \le m} \in \mathscr{M}_m$ the transition matrix from $\bar{B}$ to $\bar{B}'$.
This means that

$$\mathbf{b}'_k = s_{1k}\mathbf{b}_1 + \cdots + s_{nk}\mathbf{b}_n, \ \forall k \in \{1, \ldots, n\};$$
$$\bar{\mathbf{b}}'_j = \bar{s}_{1j}\bar{\mathbf{b}}_1 + \cdots + \bar{s}_{mj}\bar{\mathbf{b}}_m, \ \forall j \in \{1, \ldots, m\}.$$

Let $A_{B',\bar{B}'} := (a'_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in \mathscr{M}_{mn}$ be the matrix associated to the operator T with respect to the bases B', $\bar{B}'$. Then, for $1 \le k \le n$,

$$\begin{aligned} T(\mathbf{b}'_k) &= a'_{1k}\bar{\mathbf{b}}'_1 + \cdots + a'_{mk}\bar{\mathbf{b}}'_m \\ &= a'_{1k}(\bar{s}_{11}\bar{\mathbf{b}}_1 + \cdots + \bar{s}_{m1}\bar{\mathbf{b}}_m) + \cdots + a'_{mk}(\bar{s}_{1m}\bar{\mathbf{b}}_1 + \cdots + \bar{s}_{mm}\bar{\mathbf{b}}_m) \\ &= (a'_{1k}\bar{s}_{11} + \cdots + a'_{mk}\bar{s}_{1m})\bar{\mathbf{b}}_1 + \cdots + (a'_{1k}\bar{s}_{m1} + \cdots + a'_{mk}\bar{s}_{mm})\bar{\mathbf{b}}_m. \end{aligned}$$

On the other hand,

$$T(\mathbf{b}'_k) = (s_{1k}a_{11} + \cdots + s_{nk}a_{1n})\bar{\mathbf{b}}_1 + \cdots + (s_{1k}a_{m1} + \cdots + s_{nk}a_{mn})\bar{\mathbf{b}}_m.$$

Identifying the coordinates with respect to $\bar{B}$ we get

$$a'_{1k}\bar{s}_{j1} + \cdots + a'_{mk}\bar{s}_{jm} = s_{1j}a_{j1} + \cdots + s_{nk}a_{jn},\ \forall k \in \{1,\ldots,n\},\ \forall j \in \{1,\ldots,m\},$$

i.e.

$$\bar{S} \cdot A_{B',\bar{B}'} = A_{B,\bar{B}} \cdot S,$$

so we finally get

$$A_{B',\bar{B}'} = \bar{S}^{-1} \cdot A_{B,\bar{B}} \cdot S.$$

e) Suppose now that $(W', +, \cdot)$ is another finite-dimensional linear space with $\dim W' = m$ and $T' : W \to W'$ is a linear operator.

- If $\bar{B}' = \{\bar{\mathbf{b}}'_1, \ldots, \bar{\mathbf{b}}'_m\}$ is a basis of W' and $A_{\bar{B},\bar{B}'} \in \mathscr{M}_{mm'}$ is the matrix associated to T' with respect to $\bar{B}$ and $\bar{B}'$, then one can show in a similar way that $T' \circ T : V \to W'$ has $A_{\bar{B},\bar{B}'} \cdot A_{B,\bar{B}}$ as associated matrix with respect to B and $\bar{B}'$.
- A simple consequence is that $T \in \mathrm{L}(V)$ is bijective if and only if its associated matrix (with respect to any basis of V) is invertible.

2. Conversely, if $A=(a_{ij})_{\substack{1\le i\le m\\1\le j\le n}}$ is a matrix in $\mathscr{M}_{mn}$, then one can define a function $T:V\to W$ by the following formula

$$T(\mathbf{v}):=(\alpha_1 a_{11}+\cdots+\alpha_n a_{1n})\bar{\mathbf{b}}_1+\cdots+(\alpha_1 a_{m1}+\cdots+\alpha_n a_{mn})\bar{\mathbf{b}}_m,$$

where $\alpha_1,\ldots,\alpha_n\in\mathbb{R}$ are the coordinates of $\mathbf{v}$ with respect to the basis B.
It is easy to prove that T is a linear mapping, called the *linear operator associated to* A with respect to the bases B, $\bar{B}$.
The matrix associated to T with respect to the bases B, $\bar{B}$ is precisely A.

3. Suppose now that $V = \mathbb{R}^n$, $W = \mathbb{R}^m$ and B, $\bar{B}$ are the canonical bases in $\mathbb{R}^n$, respectively $\mathbb{R}^m$.

- Then
$$T(\mathbf{v}) = A_{B,\bar{B}} \cdot \mathbf{v},\ \forall \mathbf{v} \in \mathbb{R}^n.$$
where we have identified vectors in $\mathbb{R}^n$ with column-matrices in $\mathscr{M}_{n1}$ and vectors in $\mathbb{R}^m$ with column-matrices in $\mathscr{M}_{m1}$.
- If we identify vectors in Euclidean spaces with column-matrices, then we can rewrite the above formula as
$$T(\mathbf{v}) = A_{B,\bar{B}} \cdot \mathbf{v},\ \forall \mathbf{v} \in \mathbb{R}^n.$$

Adjoint operators

Definition

Let $(V, \langle\cdot,\cdot\rangle_V)$, $(W, \langle\cdot,\cdot\rangle_W)$ be two prehilbertian spaces and $T : V \to W$ a linear operator

- An operator $T^* : W \to V$ satisfying

$$\langle T^*(\mathbf{w}), \mathbf{v}\rangle_V = \langle T(\mathbf{v}), \mathbf{w}\rangle_W , \ \forall \mathbf{v} \in V, \ \forall \mathbf{w} \in W$$

  is called the *adjoint operator* of T.
- If $(W, \langle\cdot,\cdot\rangle) = (V, \langle\cdot,\cdot\rangle)$, the operator T is called *autoadjoint* or *symmetric* if $T = T^*$, *i.e.*

$$\langle T(\mathbf{w}), \mathbf{v}\rangle_V = \langle T(\mathbf{v}), \mathbf{w}\rangle_W , \ \forall \mathbf{v}, \mathbf{w} \in V.$$

- If $(W, \langle\cdot,\cdot\rangle) = (V, \langle\cdot,\cdot\rangle)$, the operator T is called *antisymmetric* if $T = -T^*$, *i.e.*

$$\langle T(\mathbf{w}), \mathbf{v}\rangle_V = -\langle T(\mathbf{v}), \mathbf{w}\rangle_W , \ \forall \mathbf{v}, \mathbf{w} \in V.$$

1. The adjoint of an operator is unique. Indeed, if T^* and $\tilde{T}^*$ are adjoints of T, then

$$\left\langle T^*(\mathbf{w}) - \tilde{T}^*(\mathbf{w}), \mathbf{v} \right\rangle_V = 0, \ \forall \mathbf{v} \in V, \ \forall \mathbf{w} \in W,$$

i.e. $T^*(\mathbf{w}) - \tilde{T}^*(\mathbf{w}) \in V^{\perp}$, for every $\mathbf{w} \in W$. Since $V^{\perp} = \{\mathbf{0}_V\}$, it follows that $\tilde{T}^* = T^*$.

2. If $(V, \langle \cdot, \cdot \rangle_V)$, $(W, \langle \cdot, \cdot \rangle_W)$ are finite-dimensional, then the adjoint of a linear operator $T : V \to W$ always exists.

- Indeed, by the Gram-Schmidt orthonormalization procedure, there exist orthonormal bases $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\bar{B} = \{\bar{\mathbf{b}}_1, \ldots, \bar{\mathbf{b}}_m\}$ in V, respectively W.
- Let $A_{B,\bar{B}}$ be the matrix associated to the operator T with respect to the bases B and $\bar{B}$.
- If $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and $\beta_1, \ldots, \beta_m \in \mathbb{R}$ are the coordinates of two vectors $\mathbf{v} \in V$ and $\mathbf{w} \in W$ with respect to B, respectively B', then we obtain

$$\begin{aligned}
&\langle T(\mathbf{v}), \mathbf{w} \rangle_W \\
&= \left\langle (\alpha_1 a_{11} + \cdots + \alpha_n a_{1n})\bar{\mathbf{b}}_1 + \cdots + (\alpha_1 a_{m1} + \cdots + \alpha_n a_{mn})\bar{\mathbf{b}}_m, \beta_1 \bar{\mathbf{b}}_1 + \cdots + \beta_n \bar{\mathbf{b}}_n \right\rangle_W \\
&= (\alpha_1 a_{11} + \cdots + \alpha_n a_{1n})\beta_1 + \cdots + (\alpha_1 a_{m1} + \cdots + \alpha_n a_{mn})\beta_m = \sum_{k=1}^{n} \sum_{j=1}^{m} \alpha_k \beta_j a_{jk}.
\end{aligned}$$

- If we define $T^* : W \to V$ as the linear operator associated with $A^{\mathrm{T}}_{B,\bar{B}} \in \mathscr{M}_{mn}$, then we see (by interchanging the roles of V and W) that

$$\langle T^*(\mathbf{w}), \mathbf{v} \rangle_V = \sum_{j=1}^{m} \sum_{k=1}^{n} \beta_j \alpha_k a_{jk},$$

  hence $\langle T^*(\mathbf{w}), \mathbf{v} \rangle_V = \langle T(\mathbf{v}), \mathbf{w} \rangle_W$. This proves that T^* is the adjoint of T.
- Clearly, T is autoadjoint or antisymmetric if and only if the matrix $A_{B,B}$ is *symmetric* ($A^{\mathrm{T}}_{B,B} = A_{B,B}$), respectively *antisymmetric* ($A^{\mathrm{T}}_{B,B} = -A_{B,B}$).

Orthogonal mappings

Definition

- Let (X, d), (Y, d') be metric spaces. We say that a mapping $f : X \to Y$ is an *isometry* (with respect to d and d') if

$$\mathrm{d}'(f(x), f(y)) = \mathrm{d}(x, y), \ \forall x, y \in X.$$

- Let $(V, \langle \cdot, \cdot \rangle)$ be a prehilbertian space and $T : V \to V$ a linear endomorphism. We say that T is *orthogonal* if

$$\|T(\mathbf{u})\| = \|\mathbf{u}\|, \ \forall \mathbf{u} \in V,$$

where $\|\cdot\|$ is the norm induced by the scalar product $\langle \cdot, \cdot \rangle$.

Remarks.

1. It is clear that a linear endomorphism $T \in \mathrm{L}(V)$ is an isometry if and only if T is orthogonal.

2. Suppose that V is finite-dimensional and $T \in \mathrm{L}(V)$ is orthogonal. Let us denote $\bar{T} := T^* \circ T$. Then

$$\langle \bar{T}(\mathbf{u}), \mathbf{v} \rangle = \langle (T^* \circ T)(\mathbf{u}), \mathbf{v} \rangle = \langle T(\mathbf{u}), T(\mathbf{v}) \rangle, \ \forall \mathbf{u}, \mathbf{v} \in V.$$

Hence

$$\langle \bar{T}(\mathbf{u}), \mathbf{u} \rangle = \| T(\mathbf{u}) \|^2 = \|\mathbf{u}\|^2, \ \forall \mathbf{u} \in V$$

and $\bar{T}$ is autoadjoint, *i.e.* $\langle \bar{T}(\mathbf{u}), \mathbf{v} \rangle = \langle \bar{T}(\mathbf{v}), \mathbf{u} \rangle$, $\forall \mathbf{u}, \mathbf{v} \in V$. Consequently,

$$\begin{aligned} 4 \langle \bar{T}(\mathbf{u}), \mathbf{v} \rangle &= \langle \bar{T}(\mathbf{u}+\mathbf{v}), \mathbf{u}+\mathbf{v} \rangle - \langle \bar{T}(\mathbf{u}-\mathbf{v}), \mathbf{u}-\mathbf{v} \rangle \\ &= \|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2 = 4 \langle \mathbf{u}, \mathbf{v} \rangle, \ \forall \mathbf{u}, \mathbf{v} \in V. \end{aligned}$$

Therefore, $\bar{T}(\mathbf{u}) - \mathbf{u} \in V^{\perp} = \{\mathbf{0}\}$, *i.e.* $\bar{T} = 1_V$. This shows that T^* is the inverse of the linear operator T, so T has to be bijective.
If $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is an orthonormal basis of V, we can show that the matrix $A := A_{B,B}$ associated to V with respect to B is *orthonormal*, *i.e.*

$$A^{\mathrm{T}} A = A A^{\mathrm{T}} = \mathrm{I}_n .$$

This implies that A is invertible, $A^{-1} = A^{\mathrm{T}}$ and $\det A \in \{-1, 1\}$.

Eigenvalues and eigenvectors

Definition

Let $(V, +, \cdot)$ be a linear space and $T \in \mathrm{L}(V)$.

- A vector $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ such that there exists $\lambda \in \mathbb{R}$ satisfying $T(\mathbf{v}) = \lambda \cdot \mathbf{v}$ is called an *eigenvector* of T, while the corresponding scalar λ is called an *eigenvalue* of T.
- If $\lambda \in \mathbb{R}$ is an eigenvalue of T, the linear subspace $\ker(T - \lambda \cdot 1_V)$ is called the *eigenspace* or *characteristic space* associated with λ.

Remarks.

1. The eigenspace associated with an eigenvalue $\lambda \in \mathbb{R}$ is the subspace of all eigenvectors corresponding to λ, so it is a subspace larger than $\{\mathbf{0}\}$. As a consequence, there are more than one (in fact, much more) eigenvectors corresponding to an eigenvalue (but only one eigenvalue corresponding to an eigenvector).

2. The eigenspace V_λ associated with an eigenvalue λ is invariant with respect to T, *i.e.* $T[V_\lambda] \subseteq V_\lambda$. Indeed, if $\mathbf{v} \in V_\lambda$, then

$$T(T(\mathbf{v})) = T(\lambda \cdot \mathbf{v}) = \lambda \cdot T(\mathbf{v}),$$

so $T(\mathbf{v}) \in V_\lambda$.

3. Of course, if λ_1 and λ_2 are two distinct eigenvalues, then $V_{\lambda_1} \cap V_{\lambda_2} = \{\mathbf{0}\}$. Actually, the following result states more.

Proposition

Let $(V, +, \cdot)$ be a linear space and $T \in \mathrm{L}(V)$. If $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ are distinct eigenvalues of T and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are corresponding eigenvectors, then $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent.

Characteristic polynomial

- Suppose now that V is a finite-dimensional linear space and $T \in \mathrm{L}(V)$.
- If $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is a basis of V and $A \in \mathscr{M}_n$ is the matrix associated to T with respect to B, then every eigenvalue $\lambda \in \mathbb{R}$ satisfies the equation
$$\det(A - \lambda\, \mathrm{I}_n) = 0.$$
We recall that the polynomial function $\lambda \mapsto \det(A - \lambda\, \mathrm{I}_n)$ is called the *characteristic polynomial* of A; we will also call it the *characteristic polynomial* of T, since this polynomial is invariant to changes of basis.
- Therefore, the eigenvalues of T are the real roots of the characteristic polynomial of T.
- If $\lambda \in \mathbb{R}$ is an eigenvalue of T, $\mathrm{null}(T - \lambda\, \mathrm{I}_n) = \dim \ker(T - \lambda \cdot 1_V)$ is called the *geometric multiplicity* of λ.
- If $\lambda \in \mathbb{R}$ is a root of a polynomial $P \in \mathbb{R}[X]$, we call *algebraic multiplicity* of λ the greatest $m \in \mathbb{N}^*$ such that $(X - \lambda)^m$ is a divisor of $P(X)$.
- The geometric multiplicity of an eigenvalue λ is smaller than the algebraic multiplicity of λ with respect to the characteristic polynomial of T.
- Therefore, if λ has algebraic multiplicity 1, then the geometric multiplicity of λ has to be 1 (*i.e.* $\ker(T - \lambda \cdot 1_V)$ has dimension 1).

Diagonalizable endomorphisms

Definition

Let $(V, +, \cdot)$ be a finite-dimensional linear space with $\dim V = n$ and $T \in \mathrm{L}(V)$. We say that T is *diagonalizable* if there exists a basis B of V such that the matrix associated to T with respect to B is a diagonal matrix, *i.e.* there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that $A_{B,B} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Remark. If an endomorphism T is autoadjoint, then it is diagonalizable.

Theorem

Let $(V, +, \cdot)$ be a finite-dimensional linear space and $T \in \mathrm{L}(V)$. Then T is diagonalizable if and only if the set of all eigenvectors generate V.

In the case $V = \mathbb{R}^n$, there is a practical method for determining if a an endomorphism $T : \mathbb{R}^n \to \mathbb{R}^n$ is diagonalizable.

1) We consider the canonical basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ of $\mathbb{R}^n$. With respect to this basis, we find the matrix A associated to T and the characteristic polynomial

$$P_A(\lambda) := \det(A - \lambda\, \mathrm{I}_n),\ \lambda \in \mathbb{R}.$$

2) We determine the eigenvalues of T by determinating the real roots of P_A. If all the n roots of P_A are real, we can continue. If not, T is not diagonalizable and we stop here.
3) For each eigenvalue λ we calculate $r_\lambda := \operatorname{rank}(A - \lambda\, \mathrm{I}_n)$. If $r_\lambda = n - m_\lambda$, for every eigenvalue λ, where m_λ is the algebraic multiplicity of λ in P_A, then we can conclude that T is diagonalizable. Otherwise, it is not and we stop here.
4) For each eigenvalue λ we solve the equation $A\mathbf{v} = \lambda\mathbf{v}$, where the vectors $\mathbf{v} \in \mathbb{R}^n$ are considered as colum matrices. Since $\operatorname{rank}(A - \lambda\, \mathrm{I}_n) = r_\lambda$ we can find linearly independent and orthonormal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_{r_\lambda}$ solving the equation (by Gram-Schmidt orthonormalization procedure).
5) The basis B of V for which the matrix associated to T is diagonal is then the set of all $\mathbf{v}_1, \ldots, \mathbf{v}_{r_\lambda}$, for all eigenvalues λ. The transition matrix S from $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ to B is the matrix which diagonalize A, *i.e.*

$$\operatorname{diag}(\lambda_1, \ldots, \lambda_n) = S^{-1}AS.$$

- S. Axler, *Linear Algebra Done Right,* Springer International Publishing AG, 2015.
- E. Cioară, M. Postolache, *Capitole de analiză matematică,* Ed. "Fair Partners", Buc., 2010.
- D. Drăghici, *Algebră*, Editura Didactică și Pedagogică, București, 1972.
- G. Galbură, F. Radó, *Geometrie*, Ed. Didactică și Pedag., București, 1979.
- J. Hefferon, *Linear Algebra*, Saint Michael's College, 2014.
- S. Heilman, *Linear Transformations and Matrices*, UCLA Department of Mathematics, Los Angeles, 2016.
- F. Iacob, E. Macovei, *Matematică (pentru anul I - ID, Informatică)*, Editura Universității "Al. I. Cuza", 2005-2006.
- I. D. Ion, R. Nicolae, *Algebră*, Editura Didactică și Pedagogică, București, 1981.
- K. Kuttler, *Linear Algebra, Theory And Applications,* The Saylor Foundation, 2013.
- I. Radomir, *Elemente de algebră vectorială, geometrie și calcul diferențial*, Editura Albastră, Cluj-Napoca, 2000.