Mathematics for Data Analysis
– Linear Algebra –

Matteo Gorgone

University of Messina, Department MIFT

email: mgorgone@unime.it

Why Linear Algebra?

Linear algebra is a central field of mathematics that is universally agreed to be a prerequisite to a deeper understanding of data analysis.
Linear algebra is the mathematics of data (vectors and matrices are the language of data).
Linear algebra is the study of lines and planes, vector spaces and mappings that are required for linear transformations.

Applications of Linear Algebra

- Matrices in Engineering, such as a line of springs.
- Graphs and Networks, such as analyzing networks.
- Markov Matrices, Population, and Economics, such as population growth.
- Linear Programming, the simplex optimization method.
- Fourier Series: Linear Algebra for functions, used widely in signal processing.
- Computer Graphics, such as the various translation, rescaling and rotation of images.
- Linear Algebra for statistics and probability, such as least squares for regression.

Linear Algebra and Statistics

Some clear fingerprints of linear algebra on statistics and statistical methods include:

- use of vector and matrix notation, especially with multivariate statistics;
- solutions to least squares and weighted least squares, such as for linear regression;
- estimates of mean and variance of data matrices;
- the covariance matrix that plays a key role in multinomial Gaussian distributions;
- Principal Component Analysis for data reduction that draws many of these elements together.

Linear Algebra and Data Analysis

Linear algebra is used in data preprocessing, data transformation, and model evaluation.
Then, we need to be familiar with:

- Vector spaces and subspaces;
- Euclidean spaces;
- Linear operators;
- Vector and matrix operations;
- Linear systems;
- Distances and Metrics;
- Eigenvalues and Eigenvectors;
- Quadratic forms.

Basic definitions

Group (additive notation)

A group G is a non–empty set of elements endowed with a binary law of composition

$$
\begin{aligned}
+ \colon G \times G &\to G \\
(x, y) &\mapsto x + y
\end{aligned}
$$

satisfying the following axioms:

1. the associative property holds:

$$\forall x, y, z \in G \qquad (x + y) + z = x + (y + z);$$

2. there exists the identity element:

$$\exists 0 \in G \quad \text{such that} \quad \forall x \in G \quad 0 + x = x + 0 = x;$$

3. every element has an opposite:

$$\forall x \in G \quad \exists -x \in G \quad \text{such that} \quad x + (-x) = -x + x = 0.$$

In this case, $(G, +)$ is called additive group.

Basic definitions

Group (multiplicative notation)

A group G is a non–empty set of elements endowed with a binary law of composition

$$\begin{aligned} \cdot : G \times G &\to G \\ (x, y) &\mapsto x \cdot y \end{aligned}$$

satisfying the following axioms:

1. the associative property holds:
$$\forall x, y, z \in G \qquad (x \cdot y) \cdot z = x \cdot (y \cdot z);$$
2. there exists the identity element:
$$\exists 1 \in G \quad \text{such that} \quad \forall x \in G \quad 1 \cdot x = x \cdot 1 = x;$$
3. every element has an inverse:
$$\forall x \in G \quad \exists x^{-1} \in G \quad \text{such that} \quad x \cdot x^{-1} = x^{-1} \cdot x = 1.$$

In this case, $(G, \cdot)$ is called multiplicative group.

Basic definitions

Remark

The identity element 0 (1) and the opposite element $-x$ (x^{-1}) of every element $x \in G$ are unique.

Commutative group (additive notation)

An additive group $(G, +)$ is Abelian if

$$x + y = y + x \qquad \forall x, y \in G.$$

Commutative group (multiplicative notation)

A multiplicative group $(G, \cdot)$ is Abelian if

$$x \cdot y = y \cdot x \qquad \forall x, y \in G.$$

Basic definitions

Exercises

Establish if the following sets with the corresponding operations are groups:

1. $(\mathbb{N}, +)$, $(\mathbb{N}, \cdot)$;
2. $(\mathbb{Z}, +)$, $(\mathbb{Z}, \cdot)$;
3. $(\mathbb{Q}, +)$, $(\mathbb{Q}, \cdot)$;
4. $(\mathbb{R}, +)$, $(\mathbb{R}, \cdot)$;
5. $(\mathbb{Z} - \{0\}, \cdot)$, $(\mathbb{Q} - \{0\}, \cdot)$, $(\mathbb{R} - \{0\}, \cdot)$.

Basic definitions

Subgroup (additive notation)

A subgroup H is a non–empty subset of a group $(G, +)$ which is itself a group with the operation inherited from that of G.

Equivalently, a subgroup H of G is a subset of G, denoted by $H \leq G$, such that

1. $0_G \in H$;
2. $-x \in H \quad \forall x \in H$;
3. H is closed with respect to the law of composition of G, *i.e.*,

$$\forall x, y \in H \quad x + y \in H.$$

Basic definitions

Subgroup (multiplicative notation)

A subgroup H is a non–empty subset of a group $(G, \cdot)$ which is itself a group with the operation inherited from that of G.
Equivalently, a subgroup H of G is a subset of G, denoted by $H \leq G$, such that

1. $1_G \in H$;
2. $x^{-1} \in H \quad \forall x \in H$;
3. H is closed with respect to the law of composition of G, *i.e.*,

$$\forall x, y \in H \quad x \cdot y \in H.$$

Subgroup: properties

Each group G contains at least two subgroups:

1. the group G itself;
2. the trivial subgroup $\{0_G\}$ ($\{1_G\}$) formed only by the identity element of G.

Proper subgroup

A subgroup H of a group G is called proper if H is a proper subset of G, *i.e.*, $H \subset G$.

Basic definitions

Example

The set of even numbers form a proper subgroup of $(\mathbb{Z}, +)$. In fact, the sum of two even numbers is even, 0 (the identity element) is even, and the opposite element of an even number is even too.

Example

The integers divisible by a fixed natural number $n \in \mathbb{N}$ (that is, the integers expressible as the product between n and a suitable integer) form a subgroup of $(\mathbb{Z}, +)$, denoted by $n\mathbb{Z} = \{nz \ : \ n \in \mathbb{N}, z \in \mathbb{Z}\}$. Therefore $n\mathbb{Z} \leq \mathbb{Z}$ for each $n \in \mathbb{N}$ (note that $0\mathbb{Z} = \{0\}$).

Example

Let $H = \{-1, 0, 1, 2, 3, 4, 5, 6\}$.
$(H, +)$ is not a subgroup of $\mathbb{Z}$! Also, it is neither a subgroup of $\mathbb{Q}$ nor a subgroup of $\mathbb{R}$.
The same holds for $(H, \cdot)$!

Basic definitions

Group operations

Let H and K be subgroups of a group G. Then:

- $H \cap K$ is a subgroup of G;
- $H \cup K$ is not, in general, a subgroup of G. $H \cup K$ is a subgroup of G if and only if $H \subseteq K$ or $K \subseteq H$.

Example

Consider the additive group $(\mathbb{Z}, +)$ and the two subgroups

$$A = 2\mathbb{Z}, \qquad B = 3\mathbb{Z}.$$

We have that $A \cup B$ is not a subgroup of $\mathbb{Z}$. In fact:

$$2 \in 2\mathbb{Z} \subseteq 2\mathbb{Z} \cup 3\mathbb{Z} \qquad \text{and} \qquad 3 \in 3\mathbb{Z} \subseteq 2\mathbb{Z} \cup 3\mathbb{Z};$$

but

$$2 + 3 = 5 \notin 2\mathbb{Z} \cup 3\mathbb{Z},$$

because

$$5 \notin 2\mathbb{Z} \qquad \text{and} \qquad 5 \notin 3\mathbb{Z}.$$

Basic definitions

Field

A field $\mathbb{K}$ is a non–empty set on which two binary laws of composition

$$+\colon \mathbb{K}\times\mathbb{K}\to\mathbb{K} \qquad \text{and} \qquad \cdot\colon \mathbb{K}\times\mathbb{K}\to\mathbb{K},$$

called respectively addition and multiplication, are defined such that

1. $\mathbb{K}$ is an abelian group with respect to the addition, *i.e.*,
   - $\forall x, y, z \in \mathbb{K} \qquad (x+y)+z = x+(y+z)$;
   - $\exists 0 \in \mathbb{K}$ such that $\forall x \in \mathbb{K} \quad 0+x = x+0 = x$;
   - $\forall x \in \mathbb{K} \quad \exists -x \in \mathbb{K}$ such that $x+(-x) = -x+x = 0$;
   - $\forall x, y \in \mathbb{K} \qquad x+y = y+x$;
2. $\mathbb{K}-\{0\}$ is an abelian group with respect to the multiplication, *i.e.*,
   - $\forall x, y, z \in \mathbb{K} \qquad (x\cdot y)\cdot z = x\cdot(y\cdot z)$;
   - $\exists 1 \in \mathbb{K}$ such that $\forall x \in \mathbb{K} \quad 1\cdot x = x\cdot 1 = x$;
   - $\forall x \in \mathbb{K} \quad \exists x^{-1} \in \mathbb{K}$ such that $x\cdot x^{-1} = x^{-1}\cdot x = 1$;
   - $x\cdot y = y\cdot x \qquad \forall x, y \in \mathbb{K}$;
3. addition and multiplication are connected by the distributive law

$$(x+y)\cdot z = (x\cdot z)+(y\cdot z) \qquad \forall x, y, z \in \mathbb{K}.$$

Basic definitions

Subfield

A subset H of a field $\mathbb{K}$ which is closed under addition and multiplication, and containing the opposite and the inverse of all its elements is a field too, and it is called a subfield of $\mathbb{K}$.

Examples

- $(\mathbb{Z}, +, \cdot)$ is not a field;
- $(\mathbb{Q}, +, \cdot)$ is a field;
- $(\mathbb{R}, +, \cdot)$ is a field;
- $(\mathbb{Q}, +, \cdot)$ is a subfield of $(\mathbb{R}, +, \cdot)$.

Fields and Vector Spaces

Fields are fundamental in the definition of vector spaces; most of the properties of the latter (existence of a basis, dimension, subspaces) do not depend on the particular field employed. Moreover, the possibility of defining a scalar product (and therefore a structure of Euclidean space) depends on the chosen field. Diagonalization of linear operators is related to the field of the vector space, as it is linked to the presence of roots of the characteristic polynomial.

Vector spaces

Vector space

A vector space V over the field $\mathbb{K}$ is a non–empty set of elements $\mathbf{v}_1, \mathbf{v}_2, \ldots$ called vectors, with the following algebraic structure:

1. there is a mapping
$$+ \colon V \times V \to V$$
$$(\mathbf{u}, \mathbf{v}) \mapsto \mathbf{u} + \mathbf{v}$$
such that $(V, +)$ is an additive abelian group;
2. there is a mapping
$$\cdot \colon \mathbb{K} \times V \to V$$
$$(\lambda, \mathbf{v}) \mapsto \lambda \mathbf{v}$$
which satisfies the axioms:
   - $\forall \lambda, \mu \in \mathbb{K}, \quad \forall \mathbf{v} \in V \qquad (\lambda\mu)\mathbf{v} = \lambda(\mu\mathbf{v});$
   - $\forall \lambda, \mu \in \mathbb{K}, \quad \forall \mathbf{v} \in V \qquad (\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v};$
   - $\forall \lambda \in \mathbb{K}, \quad \forall \mathbf{u}, \mathbf{v} \in V \qquad \lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v};$
   - $\exists 1 \in \mathbb{K} \ : \ \forall \mathbf{v} \in V \qquad 1\mathbf{v} = \mathbf{v}1 = \mathbf{v}.$

Real vector spaces

A non–empty set V is a vector space over the field $\mathbb{R}$ if two binary laws of composition

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

called respectively addition and field multiplication, are defined such that

1. $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V \qquad (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w});$
2. $\exists \mathbf{0} \in V$ such that $\forall \mathbf{u} \in V \quad \mathbf{0} + \mathbf{u} = \mathbf{u} + \mathbf{0} = \mathbf{u};$
3. $\forall \mathbf{u} \in V \quad \exists -\mathbf{u} \in \mathbf{u}$ such that $\mathbf{u} + (-\mathbf{u}) = -\mathbf{u} + \mathbf{u} = \mathbf{0};$
4. $\forall \mathbf{u}, \mathbf{v} \in V \qquad \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u};$
5. $\forall \lambda, \mu \in \mathbb{R}, \quad \forall \mathbf{u} \in V \qquad (\lambda\mu)\mathbf{u} = \lambda(\mu\mathbf{u});$
6. $\forall \lambda, \mu \in \mathbb{R}, \quad \forall \mathbf{u} \in V \quad (\lambda + \mu)\mathbf{u} = \lambda\mathbf{u} + \mu\mathbf{u};$
7. $\forall \lambda \in \mathbb{R}, \quad \forall \mathbf{u}, \mathbf{v} \in V \quad \lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v};$
8. $\exists 1 \in \mathbb{R} \ : \ \forall \mathbf{u} \in V \quad 1\mathbf{u} = \mathbf{u}1 = \mathbf{u}.$

Example

The set $\mathbb{R}^n$ of the n–tuples

$$\mathbf{u} = (u_1, \dots, u_n), \qquad u_i \in \mathbb{R}$$

is a real vector space with the vector addition $+\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ and field multiplication $\cdot\colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ defined as:

$$\mathbf{u} + \mathbf{v} = (u_1, \dots, u_n) + (v_1, \dots, v_n) = (u_1 + v_1, \dots, u_n + v_n),$$
$$\lambda \cdot \mathbf{u} = \lambda \cdot (u_1, \dots, u_n) = (\lambda u_1, \dots, \lambda u_n).$$

In this case, the identity element with respect to the addition is the *n*-tuple

$$(0, \dots, 0)$$

and the opposite is the *n*-tuple

$$(-u_1, \dots, -u_n).$$

Vector spaces

Example

Let C be the set of all continuous real-valued functions f in the interval $I \subseteq \mathbb{R}$,

$$f : I \to \mathbb{R}.$$

If f and g are two continuous functions, also the function $f + g$ defined by

$$(f + g)(t) = f(t) + g(t)$$

is continous. Moreover, for any real number λ, the function λf defined by

$$(\lambda f)(t) = \lambda f(t)$$

is also continuous. Then, C with the above operations is a real vector space.
In this case, the identity element with respect to the addition is the function 0 defined by

$$0(t) = 0,$$

and the opposite $-f$ is the function given by

$$(-f)(t) = -f(t).$$

Example

Let S be an arbitrary set and V a vector space. Consider all mappings $f: S \to V$ and the define the sum of two mappings f and g as

$$(f+g)(s) = f(s) + g(s) \qquad s \in S$$

and the field multiplication as

$$(\lambda f)(s) = \lambda f(s) \qquad s \in S.$$

Then, the set of all mappings $f: S \to V$ is a vector space. The identity element is the function f defined by

$$f(s) = 0, \qquad s \in S.$$

Example

The set of polynomials of degree at most n with real coefficients given by

$$\mathbb{R}_n[x] = \{a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + a_n x^n \ : \ a_i \in \mathbb{R} \ \forall i \in \{0, 1, \ldots, n\}\}$$

is a real vector space with the operations

$$\begin{aligned} p(x) + q(x) &= (a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + a_n x^n) + (b_0 + b_1 x + \ldots + b_{n-1}x^{n-1} + b_n x^n) = \\ &= (a_0 + b_0) + (a_1 + b_1)x + \ldots + (a_{n-1} + b_{n-1})x^{n-1} + (a_n + b_n)x^n, \\ \lambda \cdot p(x) &= \lambda \cdot (a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + a_n x^n) = \\ &= \lambda a_0 + \lambda a_1 x + \ldots + \lambda a_{n-1}x^{n-1} + \lambda a_n x^n. \end{aligned}$$

Exercise

Prove that $\mathbb{R}^3$ is a real vector space with the operations $+ \colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ and $\cdot \colon \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$ defined as:

$$\mathbf{u} + \mathbf{v} = (u_1, u_2, u_3) + (v_1, v_2, v_3) = (u_1 + v_1, u_2 + v_2, u_3 + v_3),$$
$$\lambda \cdot \mathbf{u} = \lambda \cdot (u_1, u_2, u_3) = (\lambda u_1, \lambda u_2, \lambda u_3).$$

Exercise

Establish if $\mathbb{R}^3$ with the operations $+ \colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ and $\cdot \colon \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$ defined as:

$$\mathbf{u} + \mathbf{v} = (u_1, u_2, u_3) + (v_1, v_2, v_3) = (u_1 + v_1, u_2 + v_2, u_3 + v_3),$$
$$\lambda \cdot \mathbf{u} = \lambda \cdot (u_1, u_2, u_3) = (\lambda u_1, \lambda u_2, 0)$$

is a vector space.

Exercise

Establish if $\mathbb{R}^2$ with the operations $+ \colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ and $\cdot \colon \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ defined as:

$$\mathbf{u} + \mathbf{v} = (u_1, u_2) + (v_1, v_2) = (u_1 v_2, u_2 v_1),$$
$$\lambda \cdot \mathbf{u} = \lambda \cdot (u_1, u_2) = (u_1^{\lambda}, u_2^{\lambda})$$

is a vector space.

Linear combination

Let V be a vector space over the field $\mathbb{K}$, and $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ a family of vectors in V. Then, a vector $\mathbf{v} \in V$ is called a linear combination of the vectors $\mathbf{v}_i \in V$ $(i = 1, \ldots, n)$ if there exist some scalars $\lambda_i \in \mathbb{K}$ such that

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

If $\lambda_i = 0$ $\forall i = 1, \ldots, n$, then $\mathbf{v}$ is a trivial linear combination.

Exercises

1) Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{w}_1 = (0, 0, 1), \qquad \mathbf{w}_2 = (12, 11, -1), \qquad \mathbf{w}_3 = (1, -2, 1)$$

and the scalars $\lambda_1 = 2$, $\lambda_2 = 1$, $\lambda_3 = 3$. Write some possibile linear combinations.

2) In $\mathbb{R}_2[x]$ consider the polynomials

$$p(x) = 3x^2 + 2, \qquad q(x) = x - 2$$

with scalars $\lambda_1 = 2$, $\lambda_2 = -1$. Write some possible linear combinations.

3) Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = (1, 2, 0), \qquad \mathbf{v}_2 = (3, -4, 2)$$

with the scalars $\lambda_1 = 2$, $\lambda_2 = -1$. Write all possible linear combinations.

Exercises

4) Write the vector $\mathbf{w} = (1, 0, 2)$ as linear combination of

$$\mathbf{v}_1 = (0, 2, 0), \qquad \mathbf{v}_2 = (1, 1, 0), \qquad \mathbf{v}_3 = (0, 1, 1)$$

(Solution: $\lambda_1 = -3/2$, $\lambda_2 = 1$, $\lambda_3 = 2$)

5) Compute the scalars λ_i $(i = 1, 2, 3)$ such that the polynomial

$$q(x) = 3x^2 + 4x + 2,$$

is linear combination of the polynomials

$$p_1(x) = 3x^2 + 2x + 1, \quad p_2(x) = -2x^2 + 3, \quad p_3(x) = 4x^2 + 3x.$$

(Solution: $\lambda_1 = -1$, $\lambda_2 = 1$, $\lambda_3 = 2$)

6) Establish if the vector $\mathbf{w} = (0, 0, 1)$ is linear combination of the vectors $\mathbf{v}_1 = (1, 1, 0)$ and $\mathbf{v}_2 = (1, 2, 0)$.

Vector spaces

Linear independence

A family $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of vectors is called linearly independent if

$$\sum_{i=1}^{n} \lambda_i \mathbf{v}_i = \mathbf{0}, \quad \lambda_i \in \mathbb{K},\ \mathbf{v}_i \in V \quad \Longrightarrow \quad \lambda_i = 0 \quad \forall i = 1, \dots, n.$$

The vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ are called linearly dependent in the opposite case, *i.e.*, if there exist some scalars λ_i such that

$$\sum_{i=1}^{n} \lambda_i \mathbf{v}_i = \mathbf{0}, \qquad \lambda_i \in \mathbb{K}, \quad \mathbf{v}_i \in V,$$

with at least one scalar $\lambda_i \neq 0$.
Then, at least one of the vectors $\mathbf{v}_i$ can be expressed as linear combination of the remaining ones.

Examples

- A vector $\mathbf{v} \in V$, with $\mathbf{v} \neq \mathbf{0}$, is linearly independent. In fact, the equation $\lambda \mathbf{v} = \mathbf{0}$, with $\lambda \in \mathbb{K}$, implies that $\lambda = 0$.
- Two vectors $\mathbf{x}, \mathbf{y} \in V$ are linearly dependent if and only if $\mathbf{y} = \lambda \mathbf{x}$ (or $\mathbf{x} = \lambda \mathbf{y}$) for some $\lambda \in \mathbb{K}$.

Examples

- In $\mathbb{R}^3$, the vectors
$$\mathbf{v}_1 = (1, 0, 0), \qquad \mathbf{v}_2 = (0, 1, 0), \qquad \mathbf{v}_3 = (0, 0, 1)$$
are linearly independent.
- In $\mathbb{R}^3$, the vectors
$$\mathbf{v}_1 = (1, 1, 0), \qquad \mathbf{v}_2 = (0, 0, 2), \qquad \mathbf{v}_3 = (0, 0, -3)$$
are linearly dependent.
- In $\mathbb{R}^3$, the vectors
$$\mathbf{v}_1 = (1, -1, 1), \qquad \mathbf{v}_2 = (3, 1, 2), \qquad \mathbf{v}_3 = (1, 3, 0)$$
are linearly dependent.
- Consider the vector space $\mathbb{R}_2[x]$ of polynomials of degree at most 2 with real coefficients. The polynomials
$$p(x) = x^2 + 1, \qquad q(x) = x + 3$$
are linearly independent.

Remark

A family consisting of only one vector $\mathbf{v} \in V$ is linearly dependent if and only if $\mathbf{v} = \mathbf{0}$. Thus, every family containing the zero vector is linearly dependent.

Proposition

Every non zero vector in a vector space V is linearly independent.

Proposition

If someone of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of a vector space V is a zero vector, then the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly dependent.

Proposition

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be vectors in a vector space V over a field $\mathbb{K}$. If k vectors of them, $\mathbf{v}_1, \ldots, \mathbf{v}_k$, with $k < n$, are linearly dependent, then all n vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly dependent.

Proposition

Every subfamily of a linearly independent family of vectors is linearly independent.

Proposition

A family $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of vectors is linearly independent if and only if every vector $\mathbf{w} \in V$ can be written at most in one way as a linear combination of the vectors $\mathbf{v}_i$, *i.e.*, if and only if for each linear combination

$$\mathbf{w} = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i, \qquad \lambda_i \in \mathbb{K}, \quad \mathbf{v}_i \in V$$

the scalars λ_i $(i = 1, \dots, n)$ are uniquely determined by $\mathbf{w}$.

Vector spaces

System of generators

Let V be a vector space over the field $\mathbb{K}$. A subset $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_s\} \subseteq V$ is called a system of generators for V if every vector $\mathbf{w} \in V$ is a linear combination of the vectors $\mathbf{v}_i$ of S, *i.e.*,

$$\mathbf{w} = \sum_{i=1}^{s} \lambda_i \mathbf{v}_i, \qquad \lambda_i \in \mathbb{K}.$$

Remark

- The whole space V is clearly a system of generators.
- If $S \subseteq V$ is a system of generators for V and $T \subseteq S$ is a system of generators for S, it follows that T is also a system of generators for V.
- For every vector space $V \neq \{\mathbf{0}\}$ there exist an infinite number of system of generators.

How to check if a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a system of generators?

Given an arbitrary $\mathbf{w} \in V$, we have to check if there exist n scalars $\lambda_i \in \mathbb{K}$ such that

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_n \mathbf{v}_n = \mathbf{w},$$

i.e., we have to check if this linear system in the unknowns $\lambda_1, \ldots, \lambda_n$ admits solution.

Example

Consider the set $A = \{(0,2),(1,0),(1,1)\} \subseteq \mathbb{R}^2$. It is a system of generators of $\mathbb{R}^2$.
For every vector $\mathbf{w} = (w_1, w_2) \in \mathbb{R}^2$ we can determine $\lambda_i \in \mathbb{R}$ $(i = 1,2,3)$ such that

$$(w_1, w_2) = \lambda_1(0,2) + \lambda_2(1,0) + \lambda_3(1,1).$$

For example, with $\mathbf{w} = (w_1, w_2) = (27,4)$ we can choice

$$\lambda_1 = 2, \qquad \lambda_2 = 27, \qquad \lambda_3 = 0.$$

Note that this choice is not unique!
In fact, by considering

$$\lambda_1 = 1, \qquad \lambda_2 = 25, \qquad \lambda_3 = 2,$$

we can have a linear combination that generates the vector $\mathbf{w} = (27,4)$.
If we want to check that A is a system of generators, we have to check if the linear system

$$\begin{aligned} \lambda_2 + \lambda_3 &= w_1, \\ 2\lambda_1 + \lambda_3 &= w_2 \end{aligned}$$

admits solution for all $\mathbf{w} \in V$.

Example

Consider the vector space of polynomials of degree at most 2 with real coefficients

$$\mathbb{R}_2[x] = \{a + bx + cx^2, \text{ with } a, b, c \in \mathbb{R}\}.$$

A system of generator can be the set

$$\{1, x, x^2\} \subseteq \mathbb{R}_2[x],$$

and an arbitrary polynomial $p(x)$ can be written as

$$p(x) = a + bx + cx^2.$$

Another system of generators can be the set

$$\{1, x, x^2, x + 3x^2\}.$$

In fact, an arbitrary polynomial $p(x)$ can be written as

$$p(x) = a + bx + cx^2 + 0(x + 3x^2).$$

Exercises

1. Establish if the family of vectors
$$\{(1,2,0),(2,5,-2)\}$$
is a system of generators of $\mathbb{R}^3$.
2. Establish if the family of vectors
$$\{(1,-1,0,1),(1,0,0,0),(0,1,0,0),(1,-1,1,2)\}$$
is a system of generators of $\mathbb{R}^4$.

Basis

Let V be a vector space over the field $\mathbb{K}$. A family $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of vectors is called a basis of V if it is a system of generators and $\mathbf{v}_i$ are linearly independent, *i.e.*, if and only if every vector $\mathbf{w} \in V$ can be written at most in one way as

$$\mathbf{w} = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i, \qquad \lambda_i \in \mathbb{K}, \quad \mathbf{v}_i \in V.$$

The scalars λ_i are called components of $\mathbf{w}$ with respect to the basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$.

Remark

A basis is a system of generators. The converse is not in general true! A system of generators is a basis if and only if the vectors are linearly independent.

Exercise

Let $V = \{(a, b, 0) \ : \ a, b \in \mathbb{R}\}$ be a vector space over $\mathbb{R}$. Verify that the set of vectors

$$\{(1, 1, 0), (1, 2, 0)\} \subset V$$

is a basis for V.

Proposition

Let $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a finite system of generators for V, and assume that the vectors $\mathbf{v}_1, \dots, \mathbf{v}_r$ ($r \leq n$) are linearly independent. Then, there exists a basis of V which contains the vectors $\mathbf{v}_1, \dots, \mathbf{v}_r$ and is contained in S.

Vector spaces

Properties

- Every vector space $V \neq \{\mathbf{0}\}$ admits at least one basis.
- Every vector space $V \neq \{\mathbf{0}\}$ admits infinite bases.
- All bases of a vector space V have the same cardinality (*i.e.*, the number of vectors).

Dimension

The dimension of a vector space V over the field $\mathbb{K}$, denoted by $\dim_{\mathbb{K}}(V)$, is the cardinality of a basis of V. The dimension of a vector space is uniquely defined!
Then, a vector space V, over a field $\mathbb{K}$, is finite-dimensional if $\dim_{\mathbb{K}}(V) < \infty$, otherwise is infinite-dimensional.

A more efficient method to check system of generators

Based on the concept of dimension of a vector space, we can consider a new method to check if a set of vectors is a system of generators.
Let V be a vector space over a field $\mathbb{K}$, with $\dim_{\mathbb{K}}(V) = n$ and $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subseteq V$. Then, S is a system of generators for V if and only if:

- $\dim_{\mathbb{K}}(V) = n \leq |S| = k$;
- $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ contains n linearly independent vectors.

Vector spaces

Example

The vector space $V = \{\mathbf{0}\}$ has dimension equal to zero, *i.e.*,

$$\dim_{\mathbb{K}}(\{\mathbf{0}\}) = 0.$$

Example

$\mathbb{R}^n$ is a finite–dimensional vector space over $\mathbb{R}$, *i.e.*, $\dim_{\mathbb{R}}(\mathbb{R}^n) = n$.
The standard canonical basis is given by the vectors

$$\begin{aligned}
\mathbf{e}_1 &= (1, 0, 0, \ldots, 0), \\
\mathbf{e}_2 &= (0, 1, 0, \ldots, 0), \\
&\vdots \\
\mathbf{e}_n &= (0, 0, \ldots, 0, 1).
\end{aligned}$$

In fact, every vector $\mathbf{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ can be written as

$$\begin{aligned}
(v_1, v_2, \ldots, v_n) &= v_1(1, 0, 0, \ldots, 0) + v_2(0, 1, 0, \ldots, 0) + \ldots + v_n(0, 0, 0, \ldots, 1) = \\
&= v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \ldots + v_n\mathbf{e}_n.
\end{aligned}$$

Vector spaces

Example

The set of continuous functions $f : \mathbb{R} \longrightarrow [a, b]$ is a infinite–dimensional vector space.
In fact, if we consider the monomials $x, x^2, \ldots, x^n$, the condition

$$\lambda_1 x + \lambda_2 x^2 + \cdots \lambda_n x^n = 0$$

is satisfied only when $\lambda_i = 0$ $(i = 1, \ldots, n)$ $\forall n$.

Example

The vector space of polynomials of degree at most n with real coefficients

$$\mathbb{R}_n[x] = \{a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + a_n x^n \ : \ a_i \in \mathbb{R} \ \forall i \in \{0, 1, \ldots, n\}\}$$

is finite–dimensional over $\mathbb{R}$, *i.e.*, $\dim_{\mathbb{R}}(\mathbb{R}_n[x]) = n + 1$.
The standard canonical basis is given by the set of $n + 1$ vectors (polynomials):

$$\{1, x, x^2, \ldots, x^n\}.$$

Vector spaces

Exercises

1. Find the coordinates of the vector $\mathbf{w} = (1, -1, 3, 5) \in \mathbb{R}^4$ with respect to the standard canonical basis $\mathcal{B}$ of $\mathbb{R}^4$.
2. Find the coordinates of the vector $\mathbf{w} = (3, 4) \in \mathbb{R}^2$ with respect to the basis $\mathcal{B}$ of $\mathbb{R}^2$, where
$$\mathcal{B} = \{(2, 1),\ (1, 2)\}.$$
3. Verify if the set of vectors
$$\{(1, 1),\ \ (2, 3)\} \subset \mathbb{R}^2$$
is a basis of $\mathbb{R}^2$.
4. Verify if the set of vectors
$$\{(1, 1),\ \ (2, 3),\ \ (0, 1)\} \subset \mathbb{R}^2$$
is a basis of $\mathbb{R}^2$.
5. Verify if the set of vectors
$$\{(1, 0, 1),\ \ (1, 2, 0),\ \ (2, 2, 1)\} \subset \mathbb{R}^3$$
is a basis of $\mathbb{R}^3$.

Vector spaces

Subspace

Let V be a vector space over the field $\mathbb{K}$. A non–empty subset $U \subseteq V$ is called a subspace of V if U inherits the structure of a vector space from V (*i.e.*, it is a vector space over the field $\mathbb{K}$ with the operations inherited from V).

Theorem [Characterization of subspaces]

Let V be a vector space over the field $\mathbb{K}$. A non–empty subset $U \subseteq V$ is a subspace of V if and only if U is closed under addition and field multiplication defined in V, *i.e.*,

$$\mathbf{u} + \mathbf{v} \in U, \qquad \forall \mathbf{u}, \mathbf{v} \in U, \qquad (*)$$

$$\lambda \cdot \mathbf{u} \in U, \qquad \forall \lambda \in \mathbb{K}, \quad \forall \mathbf{u} \in U. \qquad (**)$$

Conditions $(*)$ and $(**)$ are equivalent to

$$\lambda \mathbf{u} + \mu \mathbf{v} \in U, \qquad \forall \lambda, \mu \in \mathbb{K}, \quad \forall \mathbf{u}, \mathbf{v} \in U.$$

Necessary condition for subspaces

From $(**)$ it follows that $\mathbf{0}_V \in U$. In fact, given $\lambda = 0$ and $\mathbf{u} \in U$, we have

$$0 \cdot \mathbf{u} = \mathbf{0}_U \in U,$$

where $\mathbf{0}_U$ is the identity element in U with respect to the addition defined in V. But, in a vector space, the identity element is unique, then

$$\mathbf{0}_U = \mathbf{0}_V \Longrightarrow \mathbf{0}_V \in U.$$

Example

The sets $\{\mathbf{0}\}$ and V are subspaces of V. They are called trivial subspaces.

Property

If U is a subspace of V, then $\dim_{\mathbb{K}}(U) \leq \dim_{\mathbb{K}}(V)$.

Exercises

1. Let $S_1 = \{(x, y, z) \in \mathbb{R}^3 \ : \ x + 2y + 3z = 0;\ 2x + z = 0\} \subseteq \mathbb{R}^3$. Check if S is a subspace of $\mathbb{R}^3$.
2. Let $S_2 = \{(x, y, z) \in \mathbb{R}^3 \ : \ x + 3y + 5z = 1;\ 2x + 4z = 0\} \subseteq \mathbb{R}^3$. Check if S is a subspace of $\mathbb{R}^3$.

Example

Then:

- the solutions set of a homogeneous linear system with coefficients in a field $\mathbb{K}$ is a subspace of $\mathbb{K}^n$;
- the solutions set of a non–homogeneous linear system with coefficients in a field $\mathbb{K}$ is not a subspace of $\mathbb{K}^n$;
- if the subset is defined by a non–linear system, we cannot deduce a priori if it is a subspace.

Exercises

1. The set $S_3 = \{(x, y) \in \mathbb{R}^2 \ : \ y^2 = 0\}$ is a subspace of $\mathbb{R}^2$?
2. The set $S_4 = \{(x, y) \in \mathbb{R}^2 \ : \ x^2 + y^2 = 0\}$ is a subspace of $\mathbb{R}^2$?
3. The set $S_5 = \{(x, y) \in \mathbb{R}^2 \ : \ x \geq 0,\ y \geq 0\}$ is a subspace of $\mathbb{R}^2$?
4. The set $S_6 = \{(x, y) \in \mathbb{R}^2 \ : \ e^x = 1\}$ is a subspace of $\mathbb{R}^2$?
5. The set $S_7 = \{(x_1 + x_2, 2x_1, x_2, 1) \in \mathbb{R}^4 \ : \ x_1, x_2 \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^4$?

Vector spaces

Span

Let V be a vector space over the field $\mathbb{K}$, and $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ a system of generators for V.
Then, the set of all linear combinations of the vectors $\mathbf{v}_i$, *i.e.*,

$$\text{span}(S) = \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_n) = \left\{ \sum_{i=1}^{n} \lambda_i \mathbf{v}_i, \quad \lambda_i \in \mathbb{K}, \ \ i = 1, \dots, n \right\}$$

is called the span of $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ (or span of S).
$\text{span}(S)$ is a subspace of V, called the subspace generated by S, or the linear closure of S.

Exercises

1. Let $V = \mathbb{R}$ and $v \in \mathbb{R}$ a non zero vector. Compute $\text{span}(v)$.
2. Let $V = \mathbb{R}^2$ and $\mathbf{v}_1 = (1, 1)$, $\mathbf{v}_2 = (2, 2)$. Compute $\text{span}(\mathbf{v}_1, \mathbf{v}_2)$.
3. Let $V = \mathbb{R}^3$ and $\mathbf{v}_1 = (0, 0, 1)$, $\mathbf{v}_2 = (0, 1, 0)$. Compute $\text{span}(\mathbf{v}_1, \mathbf{v}_2)$.
4. Let $V = \mathbb{R}^3$ and $\mathbf{v}_1 = (1, 0, 2)$, $\mathbf{v}_2 = (0, -1, 0)$, $\mathbf{v}_3 = (2, -2, 4)$. Compute $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$.
5. Let $V = \mathbb{R}^4$ and $\mathbf{v}_1 = (7, -4, 1, 0)$, $\mathbf{v}_2 = (-5, 1, 0, 2)$. Verify if the vector $\mathbf{w} = (1, 0, 4, 8) \in \text{span}(\mathbf{v}_1, \mathbf{v}_2)$.

Vector spaces

Theorem

Let V be a vector space over the field $\mathbb{K}$, $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ a set of linearly independent vectors of V, and $S = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$.
If $\exists \mathbf{w} \in V$ such that $\mathbf{w} \notin S$, then $\{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}\}$ is a set linearly independent vectors.

Theorem

Consider $n+1$ vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}$ in a vector space V over a field $\mathbb{K}$.
Then:

- $\text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) \subseteq \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1})$;
- $\text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1})$ if and only if $\mathbf{v}_{n+1} \in \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$.

Remark

From last theorem, given $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1} \in V$, we have:

- if $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}$ are linearly independent $\Longrightarrow \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) \neq \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1})$;
- if $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}$ are linearly dependent $\Longrightarrow \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1})$.

Vector spaces

Example

Let us consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = (1, 0, 1), \qquad \mathbf{v}_2 = (-1, 3, 2), \qquad \mathbf{v}_3 = (-3, 3, 0).$$

It is immediate to verify that

$$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$$

is a set of linearly dependent vectors.
But, we have also that

$$\{\mathbf{v}_1, \mathbf{v}_2\}, \qquad \{\mathbf{v}_1, \mathbf{v}_3\}, \qquad \{\mathbf{v}_2, \mathbf{v}_3\}$$

are sets of linearly independent vectors.
From the previous theorem, it follows that

$$\mathrm{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \mathrm{span}(\mathbf{v}_1, \mathbf{v}_2) = \mathrm{span}(\mathbf{v}_1, \mathbf{v}_3) = \mathrm{span}(\mathbf{v}_2, \mathbf{v}_3).$$

Vector spaces

Problem: determine a basis from a system of generators

Let V be a vector space over a field $\mathbb{K}$, with $\dim_{\mathbb{K}}(V) = n$ and $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subseteq V$ a system of generators for V, with $k > n$.
In the case $k \leq n$, we can distinguish:

- if $\dim_{\mathbb{K}}(V) = n > k$, then the set $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is not a system of generators;
- if $\dim_{\mathbb{K}}(V) = n = k$, then the set $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is already a basis for V.

By excluding the above cases, our aim is to extract a basis from a system of generators, *i.e.*, we want to determine a maximal subset of linerly independent vectors from the generators.

Determine a basis from a system of generators: how to proceed

Let V be a vector space over a field $\mathbb{K}$, with $\dim_{\mathbb{K}}(V) = n$, and $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subseteq V$ a system of generators for V, with $k > n$.
Then:

1. Consider the first vector $\mathbf{v}_1 \in S$: if it is $\mathbf{v}_1 = \mathbf{0}$ we exclude it; otherwise, we keep it.
2. Consider the second vector $\mathbf{v}_2 \in S$. We keep it:
   - if $\mathbf{v}_2 \neq \mathbf{0}$ and $\mathbf{v}_1$ has been excluded;
   - if $\mathbf{v}_1$ have been kept, and $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent.

   If noone of these cases is satisfied, we have to exclude $\mathbf{v}_2$.
3. Consider the third vector $\mathbf{v}_3 \in S$: we keep it if $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent.
4. Continue until you run out all vectors in S.

Finally, the set of not excluded vectors forms an extracted basis from the system of generators.

Example

Given the vectors in $\mathbb{R}^3$

$$\mathbf{v}_1 = (2, -1, 6), \qquad \mathbf{v}_2 = (-6, 3, -18), \qquad \mathbf{v}_3 = (1, 0, 1), \qquad \mathbf{v}_4 = (1, 1, -3),$$

determine a basis for $S = \text{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$ and its dimension.

Solution

- Consider $\mathbf{v}_1$. We have that $\mathbf{v}_1 = (2, -1, 6) \neq (0, 0, 0)$, then we keep it.
- Consider now $\mathbf{v}_2 = (-6, 3, -18)$, that is not a zero vector. Since $\mathbf{v}_2 = -3\mathbf{v}_1$, it follows that $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly dependent. This implies that we have to exclude $\mathbf{v}_2$.
- Consider now $\mathbf{v}_3 = (1, 0, 1)$. Since $\mathbf{v}_1$ and $\mathbf{v}_3$ are linearly independent (check it!), we keep $\mathbf{v}_3$.
- Finally, considering $\mathbf{v}_4 = (1, 1, -3)$, we note that $\mathbf{v}_4 = 3\mathbf{v}_3 - \mathbf{v}_1$, *i.e.*, $\mathbf{v}_1, \mathbf{v}_3, \mathbf{v}_4$ are not linearly independent. Then, we have to exclude $\mathbf{v}_4$.

We can conclude that a basis for $S = \text{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$ is

$$\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_3\}, \qquad \text{with} \quad \dim_{\mathbb{R}}(S) = 2.$$

Vector spaces

Exercise

Given the vectors in $\mathbb{R}^3$

$$\mathbf{v}_1 = (1, 0, -3), \qquad \mathbf{v}_2 = (2, 1, -5), \qquad \mathbf{v}_3 = (0, 4, 4),$$

determine a basis for $S = \text{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ and its dimension.

Problem: determine a basis for a subspace defined by a homogeneous linear system

Let V be a vector space over a field $\mathbb{K}$, with $\dim_{\mathbb{K}}(V) = n$, and $U \subseteq V$ a subspace of V defined as

$$U = \{(x_1, \dots, x_n) \in \mathbb{K}^n \ : \ f_1(x_1, \dots, x_n) = 0, \dots, f_m(x_1, \dots, x_n) = 0\},$$

where $f_i = 0$ $(i = 1, \dots, m)$ are homogeneous linear equations.
We need to:

1. determine solutions of the homogeneous linear system $f_i = 0$ $(i = 1, \dots, m)$;
2. express the solutions as a linear combination of the involved parameters; the vectors appearing in this linear combination form a basis for the subspace U.

Example

Given the subspace V of $\mathbb{R}^3$ defined as

$$V = \{(x, y, z) \in \mathbb{R}^3 \ : \ x - y + 2z = 0\},$$

determine dimension and a basis for V.

Solution

From the equation defining the subspace V we can write

$$y = x + 2z.$$

Then, the solution set is made by

$$(x, x + 2z, z) = x(1, 1, 0) + z(0, 2, 1), \qquad \text{with } x, z \in \mathbb{R}.$$

A basis for V is

$$\mathcal{B}_V = \{(1, 1, 0), (0, 2, 1)\}$$

and

$$\dim_{\mathbb{R}}(V) = 2.$$

Vector spaces

Exercises

1. Given the vectors in $\mathbb{R}^5$

$$\mathbf{v}_1 = (1,1,1,1,1), \qquad \mathbf{v}_2 = (2,2,2,2,2), \qquad \mathbf{v}_3 = (0,1,-2,3,1), \qquad \mathbf{v}_4 = (1,2,-1,4,2),$$

   determine a basis for $S = \text{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$ and its dimension.

2. Determine a basis and the dimension of the subspace V of $\mathbb{R}^3$ defined as

$$V = \{(a,b,0), \quad a,b \in \mathbb{R}\}.$$

3. Determine a basis and the dimension of the subspace V of $\mathbb{R}^5$ defined as

$$V = \{(a-b, b-c, 0, a-c, a-2b+c), \quad a,b,c \in \mathbb{R}\}.$$

4. Determine a basis and the dimension of the subspace V of $\mathbb{R}^5$ defined as

$$V = \{(x_1,x_2,x_3,x_4,x_5) \in \mathbb{R}^5 \ : \ x_1 - x_3 + 2x_4 + x_5 = 0,\ x_2 + 3x_4 + x_5 = 0\}.$$

Vector spaces

Operations between subspaces: Sum

Let U and V be subspaces of a finite–dimensional vector space W. The sum of the subspaces U and V, denoted by $U + V$, is defined as

$$U + V = \{\mathbf{w} \in W \quad : \quad \mathbf{w} = \mathbf{u} + \mathbf{v}, \quad \text{with} \quad \mathbf{u} \in U, \ \mathbf{v} \in V\}.$$

The sum $U + V$ is a subspace of W, and contains U and V, as subspaces.

Operations between subspaces: Intersection

Let U and V be subspaces of a finite–dimensional vector space W. The intersection of the subspaces U and V, denoted by $U \cap V$, is defined as

$$U \cap V = \{\mathbf{w} \in W \quad : \quad \mathbf{w} \in U, \ \ \mathbf{w} \in V\}.$$

The intersection $U \cap V$ is a subspace of W.
Furthermore, $U \cap V$ is a subspace of U and a subspace of V.

Vector spaces

How to determine dimension and basis of the sum of two subspaces

Let W be a vector space over the field $\mathbb{K}$, U and V subspaces of W, with $\dim_{\mathbb{K}}(U) = s$ and $\dim_{\mathbb{K}}(V) = t$.
At first, we have to determine

$$\mathcal{B}_U = \{\mathbf{u_1}, \dots, \mathbf{u}_s\}, \qquad \mathcal{B}_V = \{\mathbf{v_1}, \dots, \mathbf{v}_t\},$$

i.e., the two basis for U and V, respectively.
Then, the set

$$\mathcal{B}_U \cup \mathcal{B}_V = \{\mathbf{u_1}, \dots, \mathbf{u}_s, \mathbf{v_1}, \dots, \mathbf{v}_t\}$$

is a system of generators for the subspace $U + V$, and we can extract a basis from it.
Furthermore, the number of elements of the extracted basis is the dimension of $U + V$.

Example

Determine a basis and the dimension of the sum of two subspaces U and V of $\mathbb{R}^4$ defined as the following systems of generators

$$U = \text{span}(\mathbf{u}_1, \mathbf{u}_2), \qquad V = \text{span}(\mathbf{v}_1, \mathbf{v}_2),$$

where

$$\mathbf{u}_1 = (3, 0, 0, 0), \qquad \mathbf{u}_2 = (0, 2, 0, 0), \qquad \mathbf{v}_1 = (0, 0, -1, 0), \qquad \mathbf{v}_2 = (0, 7, 1, 0).$$

Solution

.

We observe that $\mathbf{u}_1$ and $\mathbf{u}_2$ are linearly independent, and form a basis for U:

$$\mathcal{B}_U = \{(3,0,0,0),(0,2,0,0)\}.$$

Also the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent, and form a basis for V:

$$\mathcal{B}_V = \{(0,0,-1,0),(0,7,1,0)\}.$$

Then, consider the union of the two basis

$$\mathcal{B}_U \cup \mathcal{B}_V = \{(3,0,0,0),(0,2,0,0),(0,0,-1,0),(0,7,1,0)\},$$

that is a system of generators of $U+V$, and extract a basis from $\mathcal{B}_U \cup \mathcal{B}_V$.
It is

$$\mathcal{B}_{U+V} = \{(3,0,0,0),(0,2,0,0),(0,0,-1,0)\}.$$

Since $|\mathcal{B}_{U+V}| = 3$, we have $\dim_{\mathbb{R}}(U+V) = 3$.

Example

Determine a basis and the dimension of the sum of two subspaces U and V of $\mathbb{R}^3$ defined as

$$U = \{(x, y, z) \in \mathbb{R}^3 \ : \ -2x + y = 0, \ x + z = 0\},$$
$$V = \{(x, y, z) \in \mathbb{R}^3 \ : \ 2x + y - z = 0\}.$$

Solution

Compute a basis $\mathcal{B}_U$ for U and $\mathcal{B}_V$ for V.
Consider the system of equations defining U:

$$\begin{cases} -2x + y = 0, \\ x + z = 0, \end{cases}$$

whose a possible solution is

$$y = 2x, \qquad z = -x.$$

Then, a generic element of U can be written as

$$(x, 2x, -x) = x(1, 2, -1), \qquad x \in \mathbb{R}.$$

... Solution

It follows that a basis for U is

$$\mathcal{B}_U = \{(1, 2, -1)\}.$$

Also, a possible solution for the equation defining V is

$$z = 2x + y,$$

and a generic element of V can be written as

$$(x, y, 2x + y) = x(1, 0, 2) + y(0, 1, 1), \qquad x, y \in \mathbb{R}.$$

It follows that a basis for V is

$$\mathcal{B}_V = \{(1, 0, 2), (0, 1, 1)\}.$$

Then, we consider

$$\mathcal{B}_U \cup \mathcal{B}_V = \{(1, 2, -1), (1, 0, 2), (0, 1, 1)\}$$

that is a system of generators for the subspace $U + V$, and extract a basis for $U + V$. Since the above vectors are linearly independent, we have

$$\mathcal{B}_{U+V} = \{(1, 2, -1), (1, 0, 2), (0, 1, 1)\}.$$

It is $|\mathcal{B}_{U+V}| = 3$ and $\dim_{\mathbb{R}}(U + V) = 3$.

Vector spaces

Example

Determine a basis and the dimension of the sum of two subspaces U and V of $\mathbb{R}^3$ defined as

$$U = \text{span}((1,1,-1),(1,2,-2)),$$
$$V = \{(x,y,z) \in \mathbb{R}^3 \ : \ y - z = 0\}.$$

Solution

Compute a basis $\mathcal{B}_U$ for U and $\mathcal{B}_V$ for V.
The vectors $(1,1,-1),(1,2,-2)$ are linearly independent, then they are a basis for U, *i.e.*,

$$\mathcal{B}_U = \{(1,1,-1),(1,2,-2)\}.$$

Consider the equation defining V:

$$y - z = 0,$$

whose a possible solution is

$$y = z.$$

Then, a generic element of V can be written as

$$(x,z,z) = x(1,0,0) + z(0,1,1), \qquad x,z \in \mathbb{R}.$$

. . . Solution

It follows that a basis for V is

$$\mathcal{B}_V = \{(1,0,0),(0,1,1)\}.$$

Then, we consider

$$\mathcal{B}_U \cup \mathcal{B}_V = \{(1,0,0),(0,1,1),(1,1,-1),(1,2,-2)\}$$

that is a system of generators for the subspace $U+V$, and extract a basis for $U+V$.
We have that

$$(1,0,0),(0,0,1),(1,1,-1)$$

are linearly independent and

$$(1,2,-2) = 2(1,1,-1) - (1,0,0).$$

Then, a basis for $U+V$ is

$$\mathcal{B}_{U+V} = \{(1,0,0),(0,1,1),(1,1,-1)\}.$$

It is $|\mathcal{B}_{U+V}| = 3$ and $\dim_{\mathbb{R}}(U+V) = 3$.

Vector spaces

How to determine dimension and basis of the intersection of two subspaces defined by system of generators?

Let W be a vector space over the field $\mathbb{K}$, U and V subspaces of W, with $\dim_{\mathbb{K}}(U) = s$, and $\dim_{\mathbb{K}}(V) = t$.
At first, we have to extract

$$\mathcal{B}_U = \{\mathbf{u_1}, \dots, \mathbf{u}_s\}, \qquad \mathcal{B}_V = \{\mathbf{v_1}, \dots, \mathbf{v}_t\},$$

i.e., the two basis for U and V, respectively.
Then, we note that every vector $\mathbf{w} \in U \cap V$ if and only if $\mathbf{w} \in U$ and $\mathbf{w} \in V$. This means that $\mathbf{w}$ can be expressed as linear combination of both vectors of the bases $\mathcal{B}_U$ and $\mathcal{B}_V$, *i.e.*,

$$\mathbf{w} = \sum_{i=1}^{s} \alpha_i \mathbf{u}_i, \qquad \alpha_i \in \mathbb{K}, \qquad (*)$$

and

$$\mathbf{w} = \sum_{j=1}^{t} \beta_j \mathbf{v}_j, \qquad \beta_j \in \mathbb{K}.$$

...How to determine dimension and basis of the intersection of two subspaces defined by system of generators?

It follows that

$$\sum_{i=1}^{s} \alpha_i \mathbf{u}_i = \sum_{j=1}^{t} \beta_j \mathbf{v}_j$$

that can be written as

$$\sum_{i=1}^{s} \alpha_i \mathbf{u}_i - \sum_{j=1}^{t} \beta_j \mathbf{v}_j = \mathbf{0},$$

i.e., we have a linear system in the unknowns α_i ($i = 1, \dots, s$) and β_j ($j = 1, \dots, t$).
Now, determine the solutions ($s + t$ scalars)

$$\overline{\alpha}_1, \dots, \overline{\alpha}_s, \overline{\beta_1}, \dots, \overline{\beta}_t.$$

Then, we consider the s solutions $\overline{\alpha}_1, \dots, \overline{\alpha}_s$ and insert them into $(*)$, *i.e.*,

$$\mathbf{w} = \sum_{i=1}^{s} \overline{\alpha_i} \mathbf{u}_i.$$

A basis for $U \cap V$ is determined by expressing the vector $\mathbf{w}$ as a linear combination of the involved free parameters.

Vector spaces

Example

Determine a basis for the intersection of the following subspaces of $\mathbb{R}^4$:

$$U = \text{span}((1, 0, 3, 0), (0, 1, -1, 1)),$$
$$V = \text{span}((1, 1, 4, 1), (-1, 1, 2, 1), (0, 3, 5, 3)).$$

Solution

At first we note that the vectors in U and V are linearly independent (separately). Then, bases for U and V are

$$\mathcal{B}_U = \{(1, 0, 3, 0), (0, 1, -1, 1)\},$$
$$\mathcal{B}_V = \{(1, 1, 4, 1), (-1, 1, 2, 1), (0, 3, 5, 3)\}.$$

We have that every vector $\mathbf{w} \in \mathbb{R}^4$ belongs to $U \cap V$ if and only if $\mathbf{w} \in U$ and $\mathbf{w} \in V$. This implies that there exist the scalars $\alpha_1, \alpha_2, \beta_1, \beta_2, \beta_3 \in \mathbb{R}$ such that

$$\begin{aligned}
\mathbf{w} &= \alpha_1(1, 0, 3, 0) + \alpha_2(0, 1, -1, 1) = (\alpha_1, \alpha_2, 3\alpha_1 - \alpha_2, \alpha_2),\\
\mathbf{w} &= \beta_1(1, 1, 4, 1) + \beta_2(-1, 1, 2, 1) + \beta_3(0, 3, 5, 3) =\\
&= (\beta_1 - \beta_2, \beta_1 + \beta_2 + 3\beta_3, 4\beta_1 + 2\beta_2 + 5\beta_3, \beta_1 + \beta_2 + 3\beta_3).
\end{aligned}$$

Example

By equating both right sides, we obtain

$$\begin{aligned}
\alpha_1 &= \beta_1 - \beta_2 \\
\alpha_2 &= \beta_1 + \beta_2 + 3\beta_3 \\
3\alpha_1 - \alpha_2 &= 4\beta_1 + 2\beta_2 + 5\beta_3 \\
\alpha_2 &= \beta_1 + \beta_2 + 3\beta_3,
\end{aligned}$$

whose solution is

$$\begin{aligned}
\beta_1 &= -3\beta_2 - 4\beta_3, \\
\alpha_1 &= \beta_1 - \beta_2 = -4\beta_2 - 4\beta_3, \\
\alpha_2 &= \beta_1 + \beta_2 + 3\beta_3 = -2\beta_2 - \beta_3.
\end{aligned}$$

Now, substitute α_1 and α_2 into $\mathbf{w} = \alpha_1(1, 0, 3, 0) + \alpha_2(0, 1, -1, 1)$, that is

$$\begin{aligned}
\mathbf{w} &= (-4\beta_2 - 4\beta_3)(1, 0, 3, 0) + (-2\beta_2 - \beta_3)(0, 1, -1, 1) = \\
&= (-4\beta_2 - 4\beta_3, 0, -12\beta_2 - 12\beta_3, 0) + (0, -2\beta_2 - \beta_3, 2\beta_2 + \beta_3, -2\beta_2 - \beta_3) = \\
&= (-4\beta_2 - 4\beta_3, -2\beta_2 - \beta_3, -10\beta_2 - 11\beta_3, -2\beta_2 - \beta_3).
\end{aligned}$$

Example

Then, we have

$$\mathbf{w} = (-4\beta_2 - 4\beta_3, -2\beta_2 - \beta_3, -10\beta_2 - 11\beta_3, -2\beta_2 - \beta_3) = -\beta_2(4, 2, 10, 2) - \beta_3(4, 1, 11, 1).$$

It follows that a basis for $U \cap V$ is

$$\mathcal{B}_{U\cap V} = \{(4, 2, 10, 2), (4, 1, 11, 1)\},$$

with

$$\dim_{\mathbb{R}}(U \cap V) = 2.$$

How to determine dimension and basis of the intersection of two subspaces defined by homogeneous linear equations?

Let W be a vector space over the field $\mathbb{K}$, with $\dim_{\mathbb{K}}(W) = n$, U and V subspaces of W defined as

$$U = \{(x_1, \dots, x_n) \in \mathbb{K}^n \ : \ f_1(x_1, \dots, x_n) = 0, \dots, f_s(x_1, \dots, x_n) = 0\},$$
$$V = \{(x_1, \dots, x_n) \in \mathbb{K}^n \ : \ g_1(x_1, \dots, x_n) = 0, \dots, g_t(x_1, \dots, x_n) = 0\},$$

where $f_i = 0$ $(i = 1, \dots, s)$ and $g_j = 0$ $(j = 1, \dots, t)$ are homogeneous linear equations.
We need to:

1. construct the homogeneous linear system made by $s + t$ equations

$$f_1(x_1, \dots, x_n) = 0, \dots, \ f_s(x_1, \dots, x_n) = 0,$$
$$g_1(x_1, \dots, x_n) = 0, \dots, \ g_t(x_1, \dots, x_n) = 0,$$

   and determine a solution set;
2. extract a basis from the solution set of the homogeneous linear system (by expressing the solutions as a linear combination of the involved parameters; the vectors appearing in this linear combination form a basis).

Example

Determine a basis for the intersection of the following subspaces of $\mathbb{R}^3$:

$$U = \{(x, y, z) \in \mathbb{R}^3 \ : \ x - y + 2z = 0\},$$
$$V = \{(x, y, z) \in \mathbb{R}^3 \ : \ x - 3y = 0, \ x + 5y + 8z = 0\}.$$

Solution. Construct the linear system

$$x - y + 2z = 0,$$
$$x - 3y = 0,$$
$$x + 5y + 8z = 0.$$

We note that only the first and second equation are independent, then we have the system

$$x - y + 2z = 0, \qquad x - 3y = 0,$$

with two equations in three unknowns. A possible solution is

$$x = 3y, \qquad z = -y.$$

Then

$$(x, y, z) = (3y, y, -y) = y(3, 1, -1),$$

and a basis for $U \cap V$ is $\mathcal{B}_{U\cap V} = \{(3, 1, -1)\}$.

Vector spaces

Theorem [Grassmann]

Let W be a finite–dimensional vector space over a field $\mathbb{K}$, and U and V be subspaces of W. Then

$$\dim_{\mathbb{K}}(U+V) = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(V) - \dim_{\mathbb{K}}(U \cap V).$$

Direct sum

Let W be a finite–dimensional vector space over a field $\mathbb{K}$, and U and V be subspaces of W.
The vector space W is called direct sum of U and V, and is denoted by $W = U \bigoplus V$, if and only if

1. $U + V = W$;
2. $U \cap V = \{\mathbf{0}\}$.

Vector spaces

Theorem

Let W be a finite–dimensional vector space over a field $\mathbb{K}$, and U and V be subspaces of W, with W that is direct sum of U and V.
Then,

$$W = U \bigoplus V \Rightarrow \dim_{\mathbb{K}}(W) = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(V).$$

Proof

In fact, from Grassman theorem:

$$\begin{aligned}\dim_{\mathbb{K}}(W) = \dim_{\mathbb{K}}(U+V) &= \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(V) - \dim_{\mathbb{K}}(U \cap V) = \\ &= \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(V) - 0 = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(V).\end{aligned}$$

Remark

The converse is not in general true, *i.e.*,

$$\dim_{\mathbb{K}}(W) = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(V) \nRightarrow W = U \bigoplus V.$$

Example

Let $W = \mathbb{R}^3$ and the subspaces of $\mathbb{R}^3$

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

The bases for U and V are

$$\mathcal{B}_U = \{(1,2,0),(1,1,1)\}, \qquad \mathcal{B}_V = \{(0,0,3)\},$$

then $\dim_{\mathbb{R}}(U) = 2$ and $\dim_{\mathbb{R}}(V) = 1$. A system of generators for $U + V$ is

$$\mathcal{B}_U \cup \mathcal{B}_V = \{(1,2,0),(1,1,1),(0,0,3)\},$$

that are linearly independent. Then, a basis for $U + V$ is

$$\mathcal{B}_{U+V} = \{(1,2,0),(1,1,1),(0,0,3)\},$$

i.e., $\dim_{\mathbb{R}}(U + V) = 3$. Also, $U + V$ is a subspace of $\mathbb{R}^3$ with dimension 3, then $U + V = \mathbb{R}^3$. From Grassman formula, it is

$$\dim_{\mathbb{R}}(U \cap V) = \dim_{\mathbb{R}}(U) + \dim_{\mathbb{R}}(V) - \dim_{\mathbb{R}}(U + V) = 2 + 1 - 3 = 0,$$

i.e., $\mathbb{R}^3 = U \bigoplus V$ since $U \cap V = \{\mathbf{0}\}$ and $U + V = \mathbb{R}^3$.

Example

Let $W = \mathbb{R}^3$ and

$$U = \text{span}((1,0,0),(0,0,1)), \qquad V = \text{span}((0,0,3)).$$

The bases for U and V are

$$\mathcal{B}_U = \{(1,0,0),(0,0,1)\}, \qquad \mathcal{B}_V = \{(0,0,3)\},$$

then $\dim_{\mathbb{R}}(U) = 2$ and $\dim_{\mathbb{R}}(V) = 1$.
We have

$$3 = \dim_{\mathbb{R}}(\mathbb{R}^3) = \dim_{\mathbb{R}}(U) + \dim_{\mathbb{R}}(V) = 2 + 1,$$

but $\mathbb{R}^3$ is not direct sum of U and V, since $U \cap V \neq \{\mathbf{0}\}$. In fact, for $k \neq 0$ every vector $(0,0,k)$ belongs both to U and V. The same can be proved by using Grassman theorem!

. . . Example

In fact, considering

$$\mathcal{B}_U \cup \mathcal{B}_V = \{(1,0,0),(0,0,1),(0,0,3)\},$$

we can extract a basis for $U+V$, that is

$$\mathcal{B}_{U+V} = \{(1,0,0),(0,0,1)\},$$

i.e., $\dim_{\mathbb{R}}(U+V) = 2$.
Then, from Grassman formula

$$\dim_{\mathbb{R}}(U+V) = \dim_{\mathbb{R}}(U) + \dim_{\mathbb{R}}(V) - \dim_{\mathbb{K}}(U \cap V),$$

we have

$$2 = 2 + 1 - \dim_{\mathbb{R}}(U \cap V) \Rightarrow \dim_{\mathbb{R}}(U \cap V) = 1 \neq 0.$$

Vector spaces

Theorem

Let W be a finite–dimensional vector space over a field $\mathbb{K}$, and U and V be subspaces of W, with W that is direct sum of U and V. Then,

$$W = U \bigoplus V \iff \text{ every vector } \mathbf{w} \in W \text{ can be written in at most one way as}$$

$$\mathbf{w} = \mathbf{u} + \mathbf{v}, \quad \text{with } \mathbf{u} \in U, \ \mathbf{v} \in V.$$

Theorem

Let W be a finite–dimensional vector space over a field $\mathbb{K}$, and U and V be subspaces of W.
Let $\mathcal{B}_U = \{\mathbf{u_1}, \ldots, \mathbf{u}_s\}$ be a basis for U and $\mathcal{B}_V = \{\mathbf{v_1}, \ldots, \mathbf{v}_t\}$ a basis for V.
Then,

$$W = U \bigoplus V \iff \mathcal{B}_U \cup \mathcal{B}_V = \{\mathbf{u_1}, \ldots, \mathbf{u}_s, \mathbf{v_1}, \ldots, \mathbf{v}_t\} \text{ is a basis for } W.$$

Example

Verify if $W = U \bigoplus V$, with $W = \mathbb{R}^3$ and the subspaces of $\mathbb{R}^3$

$$U = \text{span}((1,0,0),(0,1,0)), \qquad V = \text{span}((2,0,0),(0,0,1)).$$

Solution

The bases for U and V are

$$\mathcal{B}_U = \{(1,0,0),(0,1,0)\}, \qquad \mathcal{B}_V = \{(2,0,0),(0,0,1)\},$$

then $\dim_{\mathbb{R}}(U) = \dim_{\mathbb{R}}(V) = 2$. A system of generators for $U + V$ is

$$\mathcal{B}_U \cup \mathcal{B}_V = \{(1,0,0),(0,1,0),(2,0,0),(0,0,1)\},$$

and extracting a basis from it, we have

$$\mathcal{B}_{U+V} = \{(1,0,0),(0,1,0),(0,0,1)\},$$

i.e., $\dim_{\mathbb{R}}(U+V) = 3$. Also, $U + V$ is a subspace of $\mathbb{R}^3$ with dimension 3, then $U + V = \mathbb{R}^3$.
But, $\mathbb{R}^3$ is not direct sum of U and V, since $U \cap V \neq \{\mathbf{0}\}$.
In fact, from Grassman formula

$$\dim_{\mathbb{R}}(U \cap V) = \dim_{\mathbb{R}}(U) + \dim_{\mathbb{R}}(V) - \dim_{\mathbb{R}}(U+V) = 2 + 2 - 3 = 1.$$

Example

Verify if $W = U \bigoplus V$, with $W = \mathbb{R}^4$ and the subspaces of $\mathbb{R}^4$

$$U = \text{span}((1,0,0,0)), \qquad V = \text{span}((0,-1,2,3)).$$

Solution

The bases for U and V are $\mathcal{B}_U = \{(1,0,0,0)\}$, $\mathcal{B}_V = \{(0,-1,2,3)\}$, then $\dim_{\mathbb{R}}(U) = \dim_{\mathbb{R}}(V) = 1$. From the previous theorem

$$4 = \dim_{\mathbb{R}}(\mathbb{R}^4) \neq \dim_{\mathbb{R}}(U) + \dim_{\mathbb{R}}(V) = 1 + 1 = 2,$$

and we can conclude that $\mathbb{R}^4$ is not direct sum of U and V. For completeness, compute a basis for $U + V$ and $U \cap V$. A system of generators for $U + V$ is made by the vectors

$$\mathcal{B}_U \cup \mathcal{B}_V = \{(1,0,0,0),(0,-1,2,3)\},$$

that are linearly independent. Then, $\mathcal{B}_{U+V} = \{(1,0,0,0),(0,-1,2,3)\}$, *i.e.*, $\dim_{\mathbb{R}}(U+V) = 2$. Then, $U + V$ is a subspace of $\mathbb{R}^4$ with dimension 2, and we have

$$4 = \dim_{\mathbb{R}}(\mathbb{R}^4) \neq \dim_{\mathbb{R}}(U+V) = 2 \Rightarrow U + V \neq \mathbb{R}^4.$$

We also note that $U \cap V = \{\mathbf{0}\}$. In fact, from Grassman formula

$$\dim_{\mathbb{R}}(U \cap V) = \dim_{\mathbb{R}}(U) + \dim_{\mathbb{R}}(V) - \dim_{\mathbb{R}}(U+V) = 1 + 1 - 2 = 0.$$

Example

Verify if $W = U \bigoplus V$, with $W = \mathbb{R}^2$ and the subspaces of $\mathbb{R}^2$

$$U = \{(x, y) \in \mathbb{R}^2 \ : \ x + y = 0\}, \qquad V = \{(x, y) \in \mathbb{R}^2 \ : \ x = 0\}.$$

Solution

Let us look for a basis of $U \cap V$. Consider the system

$$x + y = 0, \qquad x = 0.$$

It has only the solution $x = y = 0$, then $U \cap V = \{\mathbf{0}\}$ and $\dim_{\mathbb{R}}(U \cap V) = 0$. Then, look for bases of U and V. A solution to $x + y = 0$ is $y = -x$. An element of U can be written as

$$(x, -x) = x(1, -1).$$

Then

$$\mathcal{B}_U = \{(1, -1)\}, \qquad \dim_{\mathbb{R}}(U) = 1.$$

An element of V can be written as

$$(0, y) = y(0, 1).$$

Then

$$\mathcal{B}_V = \{(0, 1)\}, \qquad \dim_{\mathbb{R}}(V) = 1.$$

... Example

We have

$$2 = \dim_{\mathbb{R}}(\mathbb{R}^2) = \dim_{\mathbb{R}}(U) + \dim_{\mathbb{R}}(V) = 1 + 1 = 2,$$

and we can conclude that $\mathbb{R}^2 = U \bigoplus V$.
Also, using Grassman formula,

$$\dim_{\mathbb{R}}(U + V) = \dim_{\mathbb{R}}(U) + \dim_{\mathbb{R}}(V) - \dim_{\mathbb{R}}(U \cap V) = 1 + 1 - 0 = 2,$$

we confirmt that $U + V = \mathbb{R}^2$, and a basis for $U + V$ is determined extracting it from the system of generators

$$\mathcal{B}_U \cup \mathcal{B}_V = \{(1, -1), (0, 1)\}.$$

Since the two vectors are linearly independent, we have

$$\mathcal{B}_{U+V} = \{(1, -1), (0, 1)\}.$$

Exercises

1. Given the subspaces of $\mathbb{R}^4$

$$U = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \ : \ x_1 - x_4 = 0, \ \ x_2 + x_4 = 0\},$$
$$V = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \ : \ x_1 + 2x_2 + x_4 = 0, \ x_3 - x_4 = 0\},$$

   determine dimension and a basis for U, V, $U + V$ and $U \cap V$.

2. Given the standard canonical basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ of $\mathbb{R}^4$ and the subspaces

$$U = \text{span}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3),$$
$$V = \text{span}(\mathbf{e}_2 - \mathbf{e}_3, \mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_4),$$

   determine dimension and a basis for U, V, $U + V$ and $U \cap V$.

Vector spaces

Exercises

3. Given the subspaces of $\mathbb{R}^3$
$$U = \{(2a, b, a) \in \mathbb{R}^3 \; : \; a, b \in \mathbb{R}\},$$
$$V = \{(c, 0, c) \in \mathbb{R}^3 \; : \; c \in \mathbb{R}\},$$
determine dimension and a basis for U, V, $U + V$ and $U \cap V$.
4. Given the subspaces of $\mathbb{R}^3$
$$U = \mathrm{span}((1, 1, -1), (1, -1, 0)),$$
$$V = \{(x, y, z) \in \mathbb{R}^3 \; : \; x + z = 0\},$$
determine dimension and a basis for U, V, $U + V$ and $U \cap V$.
5. Let V and U be subspaces of $\mathbb{R}^4$ such that
$$\dim_{\mathbb{R}}(U) = 3, \qquad \dim_{\mathbb{R}}(V) = 2.$$
Is it possible that $U \cap V = \{\mathbf{0}\}$?

Supplementary subspaces

Let W be a finite–dimensional vector space over a field $\mathbb{K}$, and U and V be subspaces of W. The subspaces U and V are said to be supplementary (or complementary) in W if and only if

$$W = U \bigoplus V.$$

Operations between subspaces: Union

Let W be a finite–dimensional vector space over a field $\mathbb{K}$, and U and V be subspaces of W. Then, the union of the subspaces U and V, denoted by $U \cup V$, is defined as

$$U \cup V = \{\mathbf{w} \in W \quad : \quad \mathbf{w} \in U \ \text{ or } \ \mathbf{w} \in V\}.$$

In general, the union $U \cup V$ is not a subspace of W!

Remark

$U \cup V$ is subspace of $W \iff U \subseteq V$ or $V \subseteq U$.

Vector spaces

Example: union of subspaces

Let $W = \mathbb{R}^2$ and the subspaces of $\mathbb{R}^2$

$$U = \{(x, y) \in \mathbb{R}^2 \ : \ x = 0\}, \qquad V = \{(x, y) \in \mathbb{R}^2 \ : \ y = 0\}.$$

By definition,

$$U \cup V = \{\mathbf{w} \in W \ \ : \ \ \mathbf{w} \in U \ \text{ or } \ \mathbf{w} \in V\},$$

i.e.,

$$U \cup V = \{(x, y) \in \mathbb{R}^2 \ : \ x = 0 \ \text{ or } \ y = 0\}.$$

Verify if $U \cup V$ is subspace of $\mathbb{R}^2$.
Check the first property of subspaces. For all vectors $\mathbf{u}, \mathbf{v} \in U \cup V$, we want to prove that $\mathbf{u} + \mathbf{v} \in U \cup V$.
Let $\mathbf{u} = (0, 1) \in U$ and $\mathbf{v} = (1, 0) \in V$.
It is

$$(1, 0) + (0, 1) = (1, 1).$$

But $(1, 1) \notin U$ and $(1, 1) \notin V$; then, it follows that $(1, 1) \notin U \cup V$.
The set $U \cup V$ is not subspace of $\mathbb{R}^2$ since it is not closed with respect to the addition!

Bilinear form

Let V and W be vector spaces over a field $\mathbb{K}$. A bilinear form is a function

$$\begin{aligned} f\colon V \times W &\to \mathbb{K} \\ (\mathbf{v},\mathbf{w}) &\mapsto f(\mathbf{v},\mathbf{w}) \end{aligned}$$

that is linear in V and in W, *i.e.*,

- $f(\lambda\mathbf{u}+\mu\mathbf{v},\mathbf{w}) = \lambda f(\mathbf{u},\mathbf{w}) + \mu f(\mathbf{v},\mathbf{w}), \qquad \forall \mathbf{u},\mathbf{v}, \in V,\ \forall \mathbf{w} \in W,\ \forall \lambda,\mu \in \mathbb{K};$
- $f(\mathbf{v},\lambda\mathbf{u}+\mu\mathbf{w}) = \lambda f(\mathbf{v},\mathbf{u}) + \mu f(\mathbf{v},\mathbf{w}), \qquad \forall \mathbf{v}, \in V,\ \forall \mathbf{u},\mathbf{w} \in W,\ \forall \lambda,\mu \in \mathbb{K}.$

In particular, if $V = W$, the function $f\colon V \times V \to \mathbb{K}$ is called bilinear form over V.

Properties

If $f\colon V \times W \to \mathbb{K}$ is a bilinear form, then:

- $f(\mathbf{0}_V,\mathbf{w}) = 0 = f(\mathbf{v},\mathbf{0}_W), \qquad \forall \mathbf{v} \in V,\ \forall \mathbf{w} \in W.$

If $f\colon V \times V \to \mathbb{K}$ is a bilinear form over V, then:

- f is nondegenerate if $f(\mathbf{v},\mathbf{w}) = 0\ \ \forall \mathbf{w} \in V \Rightarrow \mathbf{v} = \mathbf{0}$;
- f is symmetric if $f(\mathbf{v},\mathbf{w}) = f(\mathbf{w},\mathbf{v}) \qquad \forall \mathbf{v},\mathbf{w} \in V$;
- f is skew–symmetric or antisymmetric if $f(\mathbf{v},\mathbf{w}) = -f(\mathbf{w},\mathbf{v}) \qquad \forall \mathbf{v},\mathbf{w} \in V$.

Scalar product

Let V be a vector space over the field $\mathbb{R}$, and consider the binary operation

$$\cdot : V \times V \to \mathbb{R}$$
$$(\mathbf{u}, \mathbf{v}) \mapsto \mathbf{u} \cdot \mathbf{v}$$

with the following properties:

(1) $(\lambda \mathbf{u} + \mu \mathbf{v}) \cdot \mathbf{w} = \lambda \mathbf{u} \cdot \mathbf{w} + \mu \mathbf{v} \cdot \mathbf{w}, \qquad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V,\ \forall \lambda, \mu \in \mathbb{R}$ (linearity);

(2) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}, \qquad \forall \mathbf{u}, \mathbf{v} \in V$ (symmetry);

(3) $\mathbf{u} \cdot \mathbf{u} \geq 0 \quad \forall \mathbf{u} \in V$, and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$ (positive–definiteness).

This operation is called scalar product (it can be alternatively denoted by $\langle \mathbf{u}, \mathbf{v} \rangle$), and the vector space V is said to be Euclidean. The above properties imply also that

- $\forall \mathbf{u} \in V \qquad \mathbf{u} \cdot \mathbf{0} = \mathbf{0} \cdot \mathbf{u} = 0$;
- $\mathbf{u} \cdot (\lambda \mathbf{v} + \mu \mathbf{w}) = \lambda \mathbf{u} \cdot \mathbf{v} + \mu \mathbf{u} \cdot \mathbf{w}, \qquad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V,\ \forall \lambda, \mu \in \mathbb{R}$;
- $(\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = \mathbf{u} \cdot \mathbf{u} + 2\mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{v}, \qquad \forall \mathbf{u}, \mathbf{v} \in V$.

Hence, a scalar product on a real vector space is a positive–definite symmetric bilinear form.

Euclidean spaces

Example: canonical scalar product

The binary operation $\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined as:

$$\mathbf{u} \cdot \mathbf{v} = (u_1, \ldots, u_n) \cdot (v_1, \ldots, v_n) = \sum_{i=1}^{n} u_i v_i = u_1 v_1 + \ldots + u_n v_n$$

is a scalar product in $\mathbb{R}^n$, and it is called the canonical (or euclidean) scalar product.

Exercises

Let $V = \mathbb{R}^3$. Verify if the following binary operations are scalar products:

1. $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$;
2. $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_1 v_2 - u_2 v_1 + u_2 v_2 + u_3 v_3$;
3. $\mathbf{u} \cdot \mathbf{v} = 2u_1 v_1 - u_1 v_2 - u_2 v_1 + 3u_2 v_2 + u_3 v_3$;
4. $\mathbf{u} \cdot \mathbf{v} = -u_1 v_1 + u_1 v_2 + u_2 v_1 - u_2 v_2 - u_3 v_3$;
5. $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_1 v_2 + u_2 v_1 + u_2 v_2 + u_3 v_3$;
6. $\mathbf{u} \cdot \mathbf{v} = 5u_1 v_1 + 4u_1 v_3 + 5u_2 v_2 + 2u_2 v_3 + 4u_3 v_1 + 2u_3 v_2 + 5u_3 v_3$.

Orthogonal vectors

Let $\cdot : V \times V \to \mathbb{R}$ be a scalar product in a real vector space V.
Two vectors $\mathbf{u}, \mathbf{v} \in V$ are said to be orthogonal with respect to the assigned scalar product in V if

$$\mathbf{u} \cdot \mathbf{v} = 0.$$

Set of orthogonal vectors

Let $\cdot : V \times V \to \mathbb{R}$ be a scalar product in a real vector space V.
A family of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subseteq V$ is said to be a set of orthogonal vectors with respect to the assigned scalar product $\cdot$ in V if

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \qquad \forall i \neq j \quad i, j = 1, \ldots, m.$$

Euclidean spaces

Orthogonal basis

Let $\cdot\colon V \times V \to \mathbb{R}$ be a scalar product in a real vector space V.
A set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ is an orthogonal basis for V if and only if:

1. $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis for V;
2. $\mathbf{v}_i \cdot \mathbf{v}_j = 0 \qquad \forall i \neq j \quad i, j = 1, \dots, n.$

Example

The standard canonical basis $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ of $\mathbb{R}^3$ is an orthogonal basis of $\mathbb{R}^3$ with respect to the canonical scalar product. In fact;

$$\mathbf{e}_1 \cdot \mathbf{e}_2 = (1,0,0) \cdot (0,1,0) = 1 \times 0 + 0 \times 1 + 0 \times 0 = 0,$$
$$\mathbf{e}_1 \cdot \mathbf{e}_3 = (1,0,0) \cdot (0,0,1) = 1 \times 0 + 0 \times 0 + 0 \times 1 = 0,$$
$$\mathbf{e}_2 \cdot \mathbf{e}_3 = (0,1,0) \cdot (0,0,1) = 0 \times 0 + 1 \times 0 + 0 \times 1 = 0.$$

Property

Let V be a real vector space with $\dim_{\mathbb{R}}(V) = n$, and $\cdot : V \times V \to \mathbb{R}$ be a scalar product in V. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subseteq V$ be a set of non-zero vectors, with $m \leq n$.

If $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is a set of orthogonal vectors $\Rightarrow \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is a set of linearly independent vectors.

Proof.

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ orthogonal non zero-vectors. We have to prove that

$$\lambda_1 \mathbf{v}_1 + \ldots + \lambda_m \mathbf{v}_m = \mathbf{0} \implies \lambda_i = 0 \;\; \forall i = 1, \ldots, m.$$

Multiplying both sides by $\mathbf{v}_i$:

$$\mathbf{v}_i \cdot (\lambda_1 \mathbf{v}_1 + \ldots + \lambda_m \mathbf{v}_m) = \mathbf{v}_i \cdot \mathbf{0},$$

i.e., due to linearity of scalar product,

$$\lambda_1 \mathbf{v}_i \cdot \mathbf{v}_1 + \ldots + \lambda_m \mathbf{v}_i \cdot \mathbf{v}_m = 0.$$

Since $\mathbf{v}_i \cdot \mathbf{v}_j = 0 \;\; \forall i \neq j$ and $\mathbf{v}_i \neq \mathbf{0} \;\; \forall i$, the previous relation reduces to

$$\lambda_i \mathbf{v}_i \cdot \mathbf{v}_i = \mathbf{0}.$$

Due to the positive definiteness of the scalar product, it follows that $\lambda_i = 0 \;\; \forall i = 1, \ldots, m$.

Remark

The converse is not in general true:

If $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ is a set of linearly independent vectors $\not\Rightarrow$ $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ is a set of orthogonal vectors.

In fact, consider $V = \mathbb{R}^3$ and the canonical scalar product in $\mathbb{R}^3$.
The vectors

$$\mathbf{v}_1 = (1, 3, 0), \qquad \mathbf{v}_2 = (0, 2, 1)$$

are linearly independent but not orthogonal.
It is

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = (1, 3, 0) \cdot (0, 2, 1) = 1 \times 0 + 3 \times 2 + 0 \times 1 = 6 \neq 0.$$

Remark

Given a scalar product in a real vector space and a set of linearly independent vectors, we will see how to construct a set of orthogonal vectors!

Property

Let V be a real vector space with $\dim_{\mathbb{R}}(V) = n$, and $\cdot\colon V \times V \to \mathbb{R}$ be a scalar product in V. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq V$ be a set of non-zero vectors.

$$\text{If } \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \text{ is a set of orthogonal vectors} \Rightarrow \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \text{ is an orthogonal basis of } V.$$

Example

Let $V = \mathbb{R}^3$ and the canonical scalar product $\cdot\colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ defined as

$$\mathbf{x} \cdot \mathbf{y} = (x_1, x_2, x_3) \cdot (y_1, y_2, y_3) = x_1 y_1 + x_2 y_2 + x_3 y_3.$$

Prove that the vectors

$$\mathbf{v}_1 = (1, -1, 0), \qquad \mathbf{v}_2 = (1, 1, 0), \qquad \mathbf{v}_3 = (0, 0, 1)$$

form an orthogonal basis of $\mathbb{R}^3$.

Solution. Since $\dim_{\mathbb{R}}(\mathbb{R}^3) = 3$ and we have 3 vectors, we only need to show that $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are orthogonal. It is:

$$\begin{aligned}
\mathbf{v}_1 \cdot \mathbf{v}_2 &= (1, -1, 0) \cdot (1, 1, 0) = 1 \times 1 + (-1) \times 1 + 0 \times 0 = 0, \\
\mathbf{v}_1 \cdot \mathbf{v}_3 &= (1, -1, 0) \cdot (0, 0, 1) = 1 \times 0 + (-1) \times 0 + 0 \times 1 = 0, \\
\mathbf{v}_2 \cdot \mathbf{v}_3 &= (1, 1, 0) \cdot (0, 0, 1) = 1 \times 0 + 1 \times 0 + 0 \times 1 = 0.
\end{aligned}$$

Euclidean spaces

Exercises

Let $V = \mathbb{R}^3$ and the following vectors

$$\mathbf{v}_1 = (1, 0, 0), \qquad \mathbf{v}_2 = (0, 1, 0), \qquad \mathbf{v}_3 = (0, 1/2, 1).$$

Then:

1. verify that $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis of $\mathbb{R}^3$;
2. verify that the binary operation $\cdot : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ defined as

   $$\mathbf{x} \cdot \mathbf{y} = (x_1, x_2, x_3) \cdot (y_1, y_2, y_3) = x_1 y_1 + 2x_2 y_2 - x_2 y_3 - x_3 y_2 + 2x_3 y_3$$

   is a scalar product in $\mathbb{R}^3$;
3. verify if $\mathcal{B}$ is an orthogonal basis with respect to the scalar product $\cdot$.

Euclidean spaces

Cauchy–Schwarz inequality

Let V be a real vector space and $\cdot\colon V \times V \to \mathbb{R}$ be a scalar product in V, *i.e.*, V is an Euclidean space. Then, for all $\mathbf{u}, \mathbf{v} \in V$ it is:

$$|\mathbf{u} \cdot \mathbf{v}|^2 \leq (\mathbf{u} \cdot \mathbf{u})(\mathbf{v} \cdot \mathbf{v}).$$

Proof.

Let $\mathbf{u} \neq 0$ and $\mathbf{v} \neq 0$ (if $\mathbf{u} = 0$ or $\mathbf{v} = 0$ the inequality is trivially satisfied). Then, for all $\lambda \in \mathbb{R}$ it is

$$(\mathbf{u} + \lambda\mathbf{v}) \cdot (\mathbf{u} + \lambda\mathbf{v}) = \lambda^2\mathbf{v} \cdot \mathbf{v} + 2\lambda\mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{u} \geq 0.$$

Since this is an algebraic second degree equation in λ that is non-negative, it follows that the discriminant have to be less than or equal to zero, *i.e.*,

$$|\mathbf{u} \cdot \mathbf{v}|^2 - (\mathbf{u} \cdot \mathbf{u})(\mathbf{v} \cdot \mathbf{v}) \leq 0.$$

Euclidean spaces

Vector norm

In a Euclidean space V, we can introduce the norm of a vector as a function

$$\begin{aligned} \|\cdot\|\colon V &\to \mathbb{R} \\ \mathbf{v} &\mapsto \|\mathbf{v}\| \end{aligned}$$

that satisfies the axioms:

- $\|\mathbf{v}\| \geq 0, \quad \forall \mathbf{v} \in V;$
- $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$;
- $\|\alpha \mathbf{v}\| = |\alpha| \, \|\mathbf{v}\|, \quad \forall \alpha \in \mathbb{R}, \forall \mathbf{v} \in V;$
- $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|, \quad \forall \mathbf{u}, \mathbf{v} \in V.$

Property

The triangle inequality is equivalent to

$$\|\mathbf{u}\| - \|\mathbf{v}\| \leq \|\mathbf{u} - \mathbf{v}\|, \quad \forall \mathbf{u}, \mathbf{v} \in V.$$

In fact:

$$\|\mathbf{u}\| = \|\mathbf{u} - \mathbf{v} + \mathbf{v}\| \leq \|\mathbf{u} - \mathbf{v}\| + \|\mathbf{v}\| \Rightarrow \|\mathbf{u}\| - \|\mathbf{v}\| \leq \|\mathbf{u} - \mathbf{v}\|.$$

Euclidean spaces

L_p norm

In a Euclidean space V, with $\dim(V)_{\mathbb{R}} = n$, a class called L_p norms (or Hölder norms) are well–defined for any parameter $p \in [1, \infty)$:

$$\|\mathbf{v}\|_p = \left(\sum_{i=1}^{n} |v_i|^p \right)^{\frac{1}{p}}.$$

Properties of L_p norms

- If $\dim(V)_{\mathbb{R}} = 1$ all L_p norms, with $p \geq 1$, are equal to the absolute value.
- For $p < 1$ the class L_p is not a norm, since the triangle inequality is violated.

Euclidean spaces

1-norm

For $p = 1$:

$$\|\mathbf{v}\|_1 = \sum_{i=1}^{n} |v_i|.$$

2-norm or Euclidean norm

For $p = 2$:

$$\|\mathbf{v}\|_2 = \sqrt{\sum_{i=1}^{n} v_i^2}.$$

∞-norm or max norm

For $p \to \infty$:

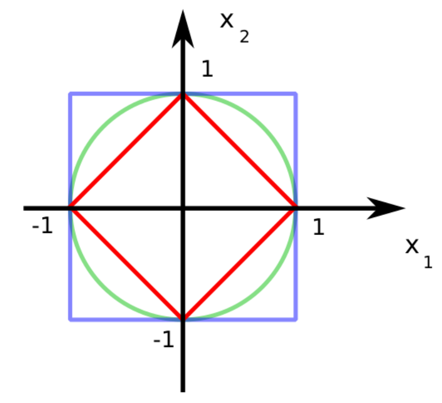
$$\|\mathbf{v}\|_\infty = \lim_{p\to\infty} \|\mathbf{v}\|_p = \max_{1\leq i\leq n} |v_i|.$$

Example

In the plane, a norm can be rapresented by a unitary ball. Let $V = \mathbb{R}^2$ be an Euclidean space, and

$$S = \{\mathbf{x} \in \mathbb{R}^2 \ : \ \|\mathbf{x}\| = 1\}$$

the unitary ball. Graphically, we have



with the unitary ball in 1-norm, the the unitary ball in 2-norm, and the unitary ball in max-norm.

Euclidean spaces

Induced–norm by a scalar product

Let V be a real vector space and $\cdot\colon V \times V \to \mathbb{R}$ be a scalar product in V, *i.e.*, V is an Euclidean space. Then, the induced-norm by a scalar product is the function $\|\cdot\|\colon V \to \mathbb{R}$ defined as

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}}.$$

Induced–norm: properties

- If $\|\cdot\|$ is an induced-norm, the Cauchy-Schwarz inequality can be written in the form

$$|\mathbf{u}\cdot\mathbf{v}| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|.$$

- The triangle inequality can be proved by means of the Cauchy-Schwarz inequality; in fact:

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) = \mathbf{u}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v}+2\mathbf{u}\cdot\mathbf{v} = \\ &= \|\mathbf{u}\|^2+\|\mathbf{v}\|^2+2\mathbf{u}\cdot\mathbf{v} \leq \|\mathbf{u}\|^2+\|\mathbf{v}\|^2+2\|\mathbf{u}\|\,\|\mathbf{v}\| = (\|\mathbf{u}\|+\|\mathbf{v}\|)^2.\end{aligned}$$

- The 2-norm is an induced-norm by the euclidean scalar product; in fact:

$$\|\mathbf{v}\|_2 = \sqrt{v_1^2+\ldots+v_n^2} = \sqrt{v_1v_1+\ldots+v_nv_n} = \sqrt{(v_1,\ldots,v_n)\cdot(v_1,\ldots,v_n)} = \sqrt{\mathbf{v}\cdot\mathbf{v}}.$$

Equivalent norms

Let V be an Euclidean space. Two norms $\|\cdot\|_1$, $\|\cdot\|_2$, defined in V, are said to be equivalent if there exist $\alpha, \beta \in \mathbb{R}$, with $\alpha > 0$ and $\beta > 0$, such that

$$\alpha\|\mathbf{v}\|_1 \leq \|\mathbf{v}\|_2 \leq \beta\|\mathbf{v}\|_1, \qquad \forall \mathbf{v} \in V.$$

Property

If $\dim_{\mathbb{R}}(V) = n < +\infty$ all norms that can be defined in V are equivalent; in particular, 1-norm, 2-norm and ∞-norm are equivalent.

Euclidean spaces

Example

Let $V = \mathbb{R}^3$ and consider the scalar product defined as

$$\mathbf{x} \cdot \mathbf{y} = 2x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2 + x_3y_3.$$

The induced-norm by this scalar product is given by

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{2x_1x_1 - x_1x_2 - x_2x_1 + 2x_2x_2 + x_3x_3} = \sqrt{2x_1^2 - 2x_1x_2 + 2x_2^2 + x_3^2}.$$

For example, if $\mathbf{x} = (1, 0, 1)$, it is

$$\|\mathbf{x}\| = \sqrt{2x_1^2 - 2x_1x_2 + 2x_2^2 + x_3^2} = \sqrt{2 \times 1^2 - 2 \times 1 \times 0 + 2 \times 0^2 + 1^2} = \sqrt{2+1} = \sqrt{3}.$$

Normalized vector

Consider an Euclidean space, and a norm $\|\cdot\|: V \to \mathbb{R}$.
If $\|\mathbf{v}\| = 1$, then $\mathbf{v}$ is called a normalized or unit vector.

How to normalize a vector?

The normalization of a vector $\mathbf{v} \in V$, with $\mathbf{v} \neq \mathbf{0}$, needs to determine a vector $\mathbf{u} \in V$ such that $\|\mathbf{u}\| = 1$ and $\mathbf{u}, \mathbf{v}$ are linearly dependent.
The normalized vector of $\mathbf{v}$ is given by

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}.$$

Example

Let $V = \mathbb{R}^2$ with the euclidean scalar product. Normalize the vector $\mathbf{v} = (3, -4)$.

Solution

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{(3, -4) \cdot (3, -4)} = \sqrt{3 \times 3 + (-4) \times (-4)} = \sqrt{9 + 16} = \sqrt{25} = 5.$$

Then,

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{1}{5}(3, -4) = \left(\frac{3}{5}, -\frac{4}{5}\right).$$

It is $\|\mathbf{u}\| = 1$. In fact:

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{\left(\frac{3}{5}, -\frac{4}{5}\right) \cdot \left(\frac{3}{5}, -\frac{4}{5}\right)} = \sqrt{\frac{9}{25} + \frac{16}{25}} = \sqrt{1} = 1.$$

Example

Let $V = \mathbb{R}^3$ and consider the scalar product defined as

$$\mathbf{x} \cdot \mathbf{y} = 2x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2 + x_3y_3.$$

Normalize the vector $\mathbf{v} = (1, 1, 0)$.

Solution

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{(1,1,0) \cdot (1,1,0)} = \sqrt{2 \times 1 \times 1 - 1 \times 1 - 1 \times 1 + 2 \times 1 \times 1 + 0} = \sqrt{2 - 2 + 2} = \sqrt{2}.$$

Then,

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{1}{\sqrt{2}}(1,1,0) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right).$$

It is $\|\mathbf{u}\| = 1$. In fact:

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right) \cdot \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)} = \sqrt{2\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} + 2\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} + 0} = 1.$$

Euclidean spaces

Set of orthonormal vectors

Let V be an Euclidean space and $\|\cdot\|\colon V \to \mathbb{R}$ be a vector norm.
A family of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subseteq V$ is said to be a set of orthonormal vectors if

1. $\mathbf{v}_i \cdot \mathbf{v}_j = 0 \qquad \forall i \neq j \quad i, j = 1, \ldots, m;$
2. $\|\mathbf{v}_i\| = 1 \quad \forall i = 1, \ldots, m.$

Orthonormal basis

Let V be an Euclidean space and $\|\cdot\|\colon V \to \mathbb{R}$ be a vector norm.
A set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq V$ is an orthonormal basis for V if and only if:

1. $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for V;
2. $\mathbf{v}_i \cdot \mathbf{v}_j = 0 \qquad \forall i \neq j \quad i, j = 1, \ldots, n;$
3. $\|\mathbf{v}_i\| = 1 \quad \forall i = 1, \ldots, m.$

Example

The standard canonical basis of $\mathbb{R}^n$

$$\mathcal{B} = \{(1, 0, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)\}$$

is an orthonormal basis with respect to the euclidean scalar product.

Example

The basis of $\mathbb{R}^3$

$$\mathcal{B} = \{(2,0,0),(0,2,0),(0,0,2)\}$$

is an orthogonal basis with respect to the euclidean scalar product, but is not orthonormal. We can obtain and orthonormal basis by dividing each vector by the corresponding norm. It is:

$$\|\mathbf{v}_1\| = \sqrt{2^2+0^2+0^2} = \sqrt{4} = 2;$$
$$\|\mathbf{v}_2\| = \sqrt{0^2+2^2+0^2} = \sqrt{4} = 2;$$
$$\|\mathbf{v}_3\| = \sqrt{0^2+0^2+2^2} = \sqrt{4} = 2.$$

Then, the corresponding orthonormal vectors are

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \left(\frac{2}{2},\frac{0}{2},\frac{0}{2}\right) = (1,0,0);$$
$$\mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(\frac{0}{2},\frac{2}{2},\frac{0}{2}\right) = (0,1,0);$$
$$\mathbf{u}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \left(\frac{0}{2},\frac{0}{2},\frac{2}{2}\right) = (0,0,1).$$

Exercises

1. Let $V = \mathbb{R}^3$, and the scalar product defined as
$$\mathbf{x} \cdot \mathbf{y} = 2x_1y_1 - x_1y_2 - x_2y_1 + 3x_2y_2 + x_3y_3.$$
Given the vectors $\mathbf{v}_1 = (1,1,1)$, $\mathbf{v}_2 = (2,0,-1)$, $\mathbf{v}_3 = (0,-3,3)$, and denoting with $\|\cdot\|$ the induced norm by the scalar product, compute $\|\mathbf{v}_1\|$, $\|\mathbf{v}_2\|$, $\|\mathbf{v}_3\|$.
2. Let $V = \mathbb{R}^2$ and consider the scalar product defined as
$$\mathbf{x} \cdot \mathbf{y} = 2x_1y_1 - x_1y_2 - x_2y_1 + 3x_2y_2.$$
Using the induced norm, normalize the vector $\mathbf{v} = (1,1)$.
3. Let $V = \mathbb{R}^4$ and consider the euclidean scalar product. Using the induced norm, normalize the vectors $\mathbf{v}_1 = (1,1,1,1)$, $\mathbf{v}_2 = (-1,0,2,-2)$, $\mathbf{v}_3 = (2,3\sqrt{2},0,\sqrt{3})$, and verify if $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are orthogonal.
4. Let $V = \mathbb{R}^2$, and the scalar product defined as
$$\mathbf{x} \cdot \mathbf{y} = 6x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2.$$
Prove that the vectors $\mathbf{v}_1 = \left(\frac{1}{3}, \frac{1}{3}\right)$ and $\mathbf{v}_2 = \left(-\frac{2}{3\sqrt{5}}, \frac{7}{3\sqrt{5}}\right)$ form an orthonormal basis in $\mathbb{R}^2$.

Exercises

5. Let $\mathcal{B}$ be a basis for $\mathbb{R}^3$ defined as

$$\mathcal{B} = \{(1, -2, 0), (-1, 0, -2), (0, -2, -1)\}.$$

Verify if $\mathcal{B}$ is an orthonormal basis with respect to the following scalar product:

$$\mathbf{x} \cdot \mathbf{y} = 5x_1y_1 + 4x_1y_3 + 5x_2y_2 + 2x_2y_3 + 4x_3y_1 + 2x_3y_2 + 5x_3y_3.$$

6. Let $\mathcal{B}$ be a basis for $\mathbb{R}^2$ defined as

$$\mathcal{B} = \{(1, 0), (2, 1)\},$$

and the scalar product

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 - 2x_1y_2 - 2x_2y_1 + 5x_2y_2.$$

Verify if $\mathcal{B}$ is an orthogonal basis with respect to the introduced scalar product and normalize it.

Problem: determining orthonormal basis

Let V be an Euclidean vector space, and $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ a basis for V. Given two vectors $\mathbf{u} = \sum_{i=1}^n u_i \mathbf{e}_i$ and $\mathbf{v} = \sum_{i=1}^n v_i \mathbf{e}_i$, their scalar product is

$$\mathbf{u} \cdot \mathbf{v} = \left(\sum_{i=1}^n u_i \mathbf{e}_i \right) \cdot \left(\sum_{j=1}^n v_j \mathbf{e}_j \right) = \sum_{i=1}^n \sum_{j=1}^n u_i v_j (\mathbf{e}_i \cdot \mathbf{e}_j).$$

It is always possible to construct a new basis $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ for V provided that

$$\mathbf{f}_i \cdot \mathbf{f}_j = \delta_{ij}, \qquad \forall i, j = 1, \ldots, n,$$

i.e., we are looking for an orthonormal basis $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$. In such a case, the scalar product becomes:

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^n \sum_{j=1}^n u_i v_j (\mathbf{f}_i \cdot \mathbf{f}_j) = \sum_{i=1}^n \sum_{j=1}^n u_i v_j \delta_{ij} = \sum_{i=1}^n u_i v_i,$$

i.e., it reduces to the euclidean scalar product, and the induced 2–norm (euclidean norm) is naturally recovered:

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{\sum_{i=1}^n \sum_{j=1}^n u_i u_j (\mathbf{f}_i \cdot \mathbf{f}_j)} = \sqrt{\sum_{i=1}^n \sum_{j=1}^n u_i u_j \delta_{ij}} = \sqrt{\sum_{i=1}^n u_i^2}.$$

Euclidean spaces

Gram-Schmidt orthonormalization

Let V be an Euclidean vector space, and $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ a basis for V. Then, we can construct an orthonormal basis $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$, *i.e.*, such that

$$\mathbf{f}_i \cdot \mathbf{f}_j = \delta_{ij}, \qquad \forall i, j = 1, \ldots, n.$$

Proof.

Let

$$\begin{aligned}
\mathbf{f}_1 &= \frac{\mathbf{e}_1}{\|\mathbf{e}_1\|}, \\
\mathbf{f}_2 &= \frac{\mathbf{e}_2 - (\mathbf{e}_2 \cdot \mathbf{f}_1)\mathbf{f}_1}{\|\mathbf{e}_2 - (\mathbf{e}_2 \cdot \mathbf{f}_1)\mathbf{f}_1\|}, \\
\mathbf{f}_3 &= \frac{\mathbf{e}_3 - (\mathbf{e}_3 \cdot \mathbf{f}_1)\mathbf{f}_1 - (\mathbf{e}_3 \cdot \mathbf{f}_2)\mathbf{f}_2}{\|\mathbf{e}_3 - (\mathbf{e}_3 \cdot \mathbf{f}_1)\mathbf{f}_1 - (\mathbf{e}_3 \cdot \mathbf{f}_2)\mathbf{f}_2\|}, \\
&\ldots \\
\mathbf{f}_n &= \frac{\mathbf{e}_n - (\mathbf{e}_n \cdot \mathbf{f}_1)\mathbf{f}_1 - (\mathbf{e}_n \cdot \mathbf{f}_2)\mathbf{f}_2 - \ldots - (\mathbf{e}_n \cdot \mathbf{f}_{n-1})\mathbf{f}_{n-1}}{\|\mathbf{e}_n - (\mathbf{e}_n \cdot \mathbf{f}_1)\mathbf{f}_1 - (\mathbf{e}_n \cdot \mathbf{f}_2)\mathbf{f}_2 - \ldots - (\mathbf{e}_n \cdot \mathbf{f}_{n-1})\mathbf{f}_{n-1}\|}.
\end{aligned}$$

$\square$

Exercises

1. Given the vectors in $\mathbb{R}^3$

$$\mathbf{v}_1 = (1, 1, 0), \qquad \mathbf{v}_2 = (2, 0, 0),$$

determine an orthonormal basis for $S = \text{span}(\mathbf{v}_1, \mathbf{v}_2)$ with respect to the euclidean scalar product.

2. Given the vectors in $\mathbb{R}^4$

$$\mathbf{v}_1 = (1, 1, 0, 0), \qquad \mathbf{v}_2 = (1, 0, 0, 0), \qquad \mathbf{v}_3 = (0, 0, 0, 1),$$

determine an orthonormal basis for $S = \text{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ with respect to the euclidean scalar product.

3. Orthonormalize the basis $\mathcal{B}$ of $\mathbb{R}^3$

$$\mathcal{B} = \{(-1, 0, 1), (1, 0, 2), (0, 1, 0)\}$$

with respect to the euclidean scalar product.

4. Orthonormalize the basis $\mathcal{B}$ of $\mathbb{R}^3$

$$\mathcal{B} = \{(0, 0, 1), (2, 0, 0), (0, -1, 0)\}$$

with respect to the scalar product defined as

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_1 y_3 + x_2 y_2 + x_3 y_1 + 2 x_3 y_3.$$

Linear operators

Linear operator

Let U and V be vector spaces over the field $\mathbb{K}$, with $\dim_{\mathbb{K}}(U) = n$ and $\dim_{\mathbb{K}}(V) = m$.
Then, a mapping Φ, say

$$\begin{aligned} \Phi \colon U &\to V \\ \mathbf{u} &\mapsto \Phi(\mathbf{u}) \end{aligned}$$

is a linear operator if

$$\begin{aligned} \Phi(\mathbf{u} + \mathbf{v}) &= \Phi(\mathbf{u}) + \Phi(\mathbf{v}), \qquad \forall \mathbf{u}, \mathbf{v} \in U, \\ \Phi(\lambda \mathbf{u}) &= \lambda \Phi(\mathbf{u}), \qquad \forall \lambda \in \mathbb{K}, \quad \forall \mathbf{u} \in U. \end{aligned}$$

Theorem [Characterization of linear operators]

Let U and V be vector spaces over the field $\mathbb{K}$, with $\dim_{\mathbb{K}}(U) = n$ and $\dim_{\mathbb{K}}(V) = m$.
Then, a mapping $\Phi \colon U \to V$ is a linear operator if and only if

$$\Phi(\lambda \mathbf{u} + \mu \mathbf{v}) = \lambda \Phi(\mathbf{u}) + \mu \Phi(\mathbf{v}), \qquad \forall \lambda, \mu \in \mathbb{K}, \quad \forall \mathbf{u}, \mathbf{v} \in U.$$

This means that a linear operator preserves vector space operations.

Linear operators

Necessary condition for linear operators

From

$$\Phi(\lambda \mathbf{u}) = \lambda \Phi(\mathbf{u}), \qquad \forall \lambda \in \mathbb{K}, \quad \forall \mathbf{u} \in U,$$

it follows that

$$\Phi(\mathbf{0}_U) = \mathbf{0}_V.$$

In fact, given $\lambda = 0$ and $\mathbf{u} \in U$, we have

$$\Phi(\mathbf{0}_U) = \Phi(0\mathbf{u}) = 0\,\Phi(\mathbf{u}) = \mathbf{0}_V,$$

where $\mathbf{0}_U$ and $\mathbf{0}_V$ are the identity elements in U and V with respect to the addition.

Linear operators

Exercises

Verify if the following mappings are linear operators:

1. $\Phi\colon \mathbb{R}^3 \to \mathbb{R}^3$, with
$$\Phi(x, y, z) = (2x + y, 3y, z + 5);$$
2. $\Phi\colon \mathbb{R}^2 \to \mathbb{R}^2$, with
$$\Phi(x, y) = (x, 3y);$$
3. $\Phi\colon \mathbb{R}^2 \to \mathbb{R}$, with
$$\Phi(x, y) = x^2 + y;$$
4. $\Phi\colon \mathbb{R}^3 \to \mathbb{R}^3$, with
$$\Phi(x, y, z) = (x + 2y, x + 4z, y - 3z);$$
5. $\Phi\colon \mathbb{R}^3 \to \mathbb{R}$, with
$$\Phi(x, y, z) = x + 3z.$$

Linear operators

Kernel of a linear operator

Let $\Phi\colon U \to V$ be a linear operator, with $\dim_{\mathbb{K}}(U) = n$ and $\dim_{\mathbb{K}}(V) = m$.
The kernel of Φ, denoted by $\ker\Phi$, is a subspace of U defined as

$$\ker\Phi = \{\mathbf{u} \in U : \Phi(\mathbf{u}) = \mathbf{0}\}.$$

Image of a linear operator

Let $\Phi\colon U \to V$ be a linear operator, with $\dim_{\mathbb{K}}(U) = n$ and $\dim_{\mathbb{K}}(V) = m$.
The image of Φ, denoted by $\mathrm{Im}\Phi$, is a subspace of V defined as

$$\mathrm{Im}\Phi = \{\mathbf{v} \in V : \mathbf{v} = \Phi(\mathbf{u}), \text{ with } \mathbf{u} \in U\}.$$

Properties

Let $\Phi\colon U \to V$ be a linear operator, with $\dim_{\mathbb{K}}(U) < \infty$.

- The linear operator Φ is injective if and only if $\ker\Phi = \{\mathbf{0}\}$.
- The linear operator Φ is surjective if and only if $\mathrm{Im}\Phi = V$.
- $\dim_{\mathbb{K}}(U) = \dim_{\mathbb{K}}(\ker\Phi) + \dim_{\mathbb{K}}(\mathrm{Im}\Phi)$.

Linear operators

Linear isomorphism

Let $\Phi\colon U \to V$ be a linear operator, with $\dim_{\mathbb{K}}(U) = n$ and $\dim_{\mathbb{K}}(V) = m$.
If Φ is bijective:

- Φ is called linear isomorphism;
- $\Phi^{-1}\colon V \to U$ is a linear isomorphism, and Φ^{-1} is called inverse isomorphism;
- U and V are isomorphic vector spaces;
- the image under Φ of a basis in U is a basis in V.

Definitions

- If $\Phi\colon U \to U$ is a linear operator, Φ is called endomorphism of U.
- If $\Phi\colon U \to U$ is a linear isomorphism, Φ is called automorphism of U.

Linear operators

Representation of linear operators

Let $\Phi\colon U \to V$ be a linear operator, with $\dim_{\mathbb{K}}(U) = n$ and $\dim_{\mathbb{K}}(V) = m$. Let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ and $\{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ bases for U and V, respectively. Then, given $\mathbf{u} \in U$ and $\mathbf{v} \in V$, we have

$$\mathbf{u} = \sum_{j=1}^{n} u_j \mathbf{e}_j, \qquad u_j \in \mathbb{K},$$

$$\mathbf{v} = \sum_{i=1}^{m} v_i \mathbf{f}_i, \qquad v_i \in \mathbb{K},$$

and, from $\Phi(\mathbf{u}) = \mathbf{v}$, it follows that

$$\Phi(\mathbf{u}) = \Phi\left(\sum_{j=1}^{n} u_j \mathbf{e}_j\right) = \sum_{j=1}^{n} u_j \Phi(\mathbf{e}_j) = \sum_{i=1}^{m} v_i \mathbf{f}_i = \mathbf{v}.$$

But, every vector $\Phi(\mathbf{e}_j)$ can be written as a linear combination of the vectors $\mathbf{f}_i$ $(i = 1, \dots, m)$, *i.e.*,

$$\Phi(\mathbf{e}_j) = \sum_{i=1}^{m} T_{ij} \mathbf{f}_i, \qquad j = 1, \dots, n.$$

Linear operators

Representation of linear operators

We obtain

$$\sum_{j=1}^{n}\sum_{i=1}^{m} u_j T_{ij}\mathbf{f}_i = \sum_{i=1}^{m} v_i\mathbf{f}_i \Longrightarrow \sum_{i=1}^{m}\left(v_i - \sum_{j=1}^{n} T_{ij}u_j\right)\mathbf{f}_i = \mathbf{0}.$$

In other words, the components of a vector $\mathbf{u} \in U$ are transformed according to the law

$$v_i = \sum_{j=1}^{n} T_{ij}u_j, \qquad i = 1, \ldots, m.$$

The linear operator Φ determines a $m \times n$ matrix $T = (T_{ij})$, where i counts the rows and j counts the columns. Conversely, every $m \times n$ matrix T determines a linear operator.
Thus, after choosing bases of U and V, a matrix T defines the action of the linear operator Φ on any vector, with the rule

$$\Phi(\mathbf{e}_j) = \sum_{i=1}^{m} T_{ij}\mathbf{f}_i, \qquad j = 1, \ldots, n.$$

The matrix T is called the representation of the linear operator Φ.

Linear operators

How to determine the linear operator associated to a matrix

A linear operator $\Phi\colon U \to V$ maps a column vector $\mathbf{u} = (u_1, \dots, u_n)$ into the column vector

$$\mathbf{v} = T\mathbf{u} = \begin{pmatrix} T_{11}u_1 + T_{12}u_2 + \cdots + T_{1n}u_n \\ T_{21}u_1 + T_{22}u_2 + \cdots + T_{2n}u_n \\ \vdots \\ T_{m1}u_1 + T_{m2}u_2 + \cdots + T_{mn}u_n \end{pmatrix} \in \mathbb{K}^m,$$

where the matrix T is

$$T = \begin{pmatrix} T_{11} & T_{12} & \cdots & T_{1n} \\ T_{21} & T_{22} & \cdots & T_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ T_{m1} & T_{m2} & \cdots & T_{mn} \end{pmatrix} \in \mathbb{K}^{m \times n}.$$

Then, every matrix $T \in \mathbb{K}^{m\times n}$ defines a linear operator $\Phi\colon \mathbb{K}^n \to \mathbb{K}^m$ such that for every vector $\mathbf{u} \in \mathbb{K}^n$ it is $\Phi(\mathbf{u}) = T\mathbf{u} \in \mathbb{K}^m$.

Remark

The number of columns of T is egual to $\dim(\mathbb{K}^n) = n$, whereas the number of rows of T is egual to $\dim(\mathbb{K}^m) = m$.

Linear operators

Example

Determine the linear operator defined by the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{pmatrix} \in \mathbb{R}^{3\times 2}.$$

Solution

The number of columns of A is egual to $\dim(\mathbb{R}^2) = 2$, whereas the number of rows of A is egual to $\dim(\mathbb{R}^3) = 3$. This means we have to consider a generic vector $\mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2$ and compute $A\mathbf{x}$:

$$A\mathbf{x} = \begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 4x_1 + 5x_2 \\ 7x_1 + 8x_2 \end{pmatrix}.$$

Then, we obtain the linear operator $\Phi\colon \mathbb{R}^2 \to \mathbb{R}^3$ defined by the matrix A:

$$\Phi(x_1, x_2) = (x_1 + 2x_2, 4x_1 + 5x_2, 7x_1 + 8x_2).$$

Linear operators

Exercise

Determine the linear operator defined by the matrix

$$A = \begin{pmatrix} 2 & 2 & -1 \\ 1 & 0 & 3 \\ 0 & 4 & 6 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Linear operators

How to determine the matrix associated to a linear operator

Let $\Phi\colon U \to V$ a linear operator, with $\dim_{\mathbb{K}}(U) = n$ and $\dim_{\mathbb{K}}(V) = m$, and the basis $\mathcal{B}_U = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ and $\mathcal{B}_V = \{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ for U and V, respectively.
It is useful to remember the following steps for constructing the matrix corresponding to a linear operator in an assigned basis:

- compute $\Phi(\mathbf{e}_j)$, *i.e.*, the images of the vectors of the basis $\mathcal{B}_U$.
- the images $\Phi(\mathbf{e}_j)$ must be expressed as a linear combination of the vectors of the basis $\mathcal{B}_V$,*i.e.*,

$$\Phi(\mathbf{e}_j) = \sum_{i=1}^{m} T_{ij}\mathbf{f}_i;$$

- the coordinates of the obtained vectors are the columns of the associated matrix.

Remarks

- The matrices associated to a linear operator are as many as the bases, *i.e.*, they are infinite!
- Linear operators act on vectors, whereas the associated matrices act on the coordinates with respect to the bases.
- The associated matrix to $\Phi\colon U \to V$ has $\dim_{\mathbb{K}}(V) = m$ rows and $\dim_{\mathbb{K}}(U) = n$ columns.

Exercises

Let $\Phi\colon \mathbb{R}^3 \to \mathbb{R}^2$ be a linear operator defined by

$$\Phi(x, y, z) = (x + y, z).$$

Determine the matrix associated to Φ with respect to the basis:

1. $$\mathcal{B}_{\mathbb{R}^3} = \{(1, 0, 1), (1, 0, 0), (1, 1, 1)\},$$
   $$\mathcal{B}_{\mathbb{R}^2} = \{(0, 1), (1, 1)\};$$

2. $$\mathcal{B}_{\mathbb{R}^3} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\},$$
   $$\mathcal{B}_{\mathbb{R}^2} = \{(1, 0), (0, 1)\}.$$

Linear operators

Exercises

Let $\Phi\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a linear operator defined by

$$\Phi(x, y) = (x - y, 2x + y).$$

Determine the matrix associated to Φ with respect to the basis:

1. $$\mathcal{B}_{\mathbb{R}^2} = \{(1, 2), (1, 0)\},$$
   $$\mathcal{B}'_{\mathbb{R}^2} = \{(1, 1), (0, 3)\};$$
2. $$\mathcal{B}_{\mathbb{R}^2} = \{(1, 0), (0, 1)\},$$
   $$\mathcal{B}'_{\mathbb{R}^2} = \{(1, 2), (0, 1)\}.$$

Linear operators

Linear operators defined by image of vectors

Let U and V be vector spaces over the field $\mathbb{K}$, and $\Phi\colon U \to V$ a linear operator.
The linear operator Φ is defined by the images of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_n\} \in U$ if it is written as

$$\Phi(\mathbf{u}_i) = \mathbf{v}_i \qquad \forall i = 1, \dots, n.$$

Example

The linear operator $\Phi\colon \mathbb{R}^2 \to \mathbb{R}^3$ with

$$\Phi(1,0) = (1,2,3), \qquad \Phi(0,5) = (4,0,7),$$

is defined by images of vectors.

Theorem

Let U and V be vector spaces over the field $\mathbb{K}$, with $\dim_{\mathbb{K}}(U) = n$ and $\dim_{\mathbb{K}}(V) = m$, and $\mathcal{B}_U = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ a basis for U, and $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ vectors of V. Then, there exists a unique linear operator $\Phi\colon U \to V$ such that

$$\Phi(\mathbf{e}_i) = \mathbf{v}_i \qquad \forall i = 1, \dots, n.$$

Linear operators

How to determine the matrix associated to a linear operator defined by image of vectors

Let U and V vector spaces over the field $\mathbb{K}$, with the basis $\mathcal{B}_U = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $\mathcal{B}_V = \{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ for U and V, respectively. Consider a linear operator $\Phi \colon U \to V$ defined by the images of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$, say

$$\Phi(\mathbf{u}_i) = \mathbf{v}_i \qquad \forall i = 1, \ldots, n.$$

Suppose that $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is a basis for U, so that we are sure there exists a unique linear operator $\Phi \colon U \to V$. We need to compute $\Phi(\mathbf{e}_j)$. This task can be done in the following way:

- express vectors of the basis of U as a linear combination of the vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$, *i.e.*,

$$\mathbf{e}_j = \sum_{i=1}^{n} T_{ij}\mathbf{u}_i, \qquad j = 1, \ldots, n;$$

- apply Φ to the vectors $\mathbf{e}_j$:

$$\Phi(\mathbf{e}_j) = \Phi\left(\sum_{i=1}^{n} T_{ij}\mathbf{u}_i\right) = \sum_{i=1}^{n} T_{ij}\Phi(\mathbf{u}_i) = \sum_{i=1}^{n} T_{ij}\mathbf{v}_i, \qquad j = 1, \ldots, n;$$

- finally, by expressing $\Phi(\mathbf{e}_j)$ as linear combinations of the vectors $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$, we obtain the coordinates entering the j–th column of the associated matrix.

Linear operators

Exercises

Let $\Phi\colon \mathbb{R}^2 \to \mathbb{R}^3$ be a linear operator defined by the image of vectors

$$\Phi(1,1) = (1,2,0),$$
$$\Phi(2,1) = (-1,3,1).$$

Determine the matrix associated to Φ with respect to the basis:

1. $$\mathcal{B}_{\mathbb{R}^2} = \{(1,0),(0,1)\},$$
   $$\mathcal{B}_{\mathbb{R}^3} = \{(1,1,1),(1,0,0),(0,-1,1)\};$$

2. $$\mathcal{B}_{\mathbb{R}^2} = \{(1,0),(0,1)\},$$
   $$\mathcal{B}_{\mathbb{R}^3} = \{(1,0,0),(0,1,0),(0,0,1)\};$$

Linear operators

Exercises

Let $\Phi\colon \mathbb{R}^2 \to \mathbb{R}^3$ be a linear operator defined by the image of vectors

$$\Phi(3,-1) = (1,-1,2),$$
$$\Phi(1,2) = (-2,-1,2).$$

Determine the matrix associated to Φ with respect to the basis:

1. $$\mathcal{B}_{\mathbb{R}^2} = \{(4,1),(-3,8)\},$$
   $$\mathcal{B}_{\mathbb{R}^3} = \{(1,0,0),(0,1,0),(0,0,1)\};$$

2. $$\mathcal{B}_{\mathbb{R}^2} = \{(1,0),(0,1)\},$$
   $$\mathcal{B}_{\mathbb{R}^3} = \{(1,0,0),(0,1,0),(0,0,1)\};$$

Operations with linear operators

Sum of linear operators

If $f\colon U \to V$ and $g\colon U \to V$ are two linear operators, then also their sum $f+g\colon U \to V$ is a linear operator, which is defined by

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

to which corresponds the sum matrix of the matrices of f and g.

Product of linear operators with scalars

If $f\colon U \to V$ is linear operator and $\lambda \in \mathbb{K}$, then the map $\lambda f\colon U \to V$, defined by $(\lambda f)(x) = \lambda(f(x))$, is a linear operator, to which corresponds the product of the matrix of f with the scalar λ.

Composition of linear operators

If $f\colon U \to V$ and $g\colon V \to W$ are two linear operators, then also their composition $g \circ f\colon U \to W$ is a linear operator, which is defined by

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

The matrix of the operator $g \circ f$ is the matrix product of the corresponding matrices g and f.

Examples

- The null matrix corresponds to the null operator $0\colon \mathbf{x} \to \mathbf{0}$.
- The identity matrix corresponds to the identity operator $I\colon \mathbf{x} \to \mathbf{x}$.

Matrices - Operations

Definition

Given a matrix $A = (a_{ij}) \in \mathbb{K}^{m\times n}$ $(1 = 1, \ldots, m,\ j = 1, \ldots n)$, $m \times n$ is the size of A. If $m = n$, A is a square matrix and n is the order of A.

Addition

The sum of two matrices $A, B \in \mathbb{K}^{m\times n}$ is the matrix $C \in \mathbb{K}^{m\times n}$, $C = A + B$, given by the elements

$$c_{ij} = a_{ij} + b_{ij}, \qquad i = 1, \ldots, m, \quad j = 1, \ldots, n.$$

Properties:

- Addition is commutative: $A + B = B + A,\ \forall A, B \in \mathbb{K}^{m\times n}$;
- Addition is associative: $(A + B) + C = A + (B + C),\ \forall A, B, C \in \mathbb{K}^{m\times n}$.

Product with scalars

The product of a scalar λ and a matrix $A \in \mathbb{K}^{m\times n}$ is given a matrix with the elements

$$(\lambda a)_{ij} = \lambda a_{ij}, \qquad i = 1, \ldots, m, \quad j = 1, \ldots, n.$$

Vector spaces of matrices

The set of all matrices of a fixed size forms a vector space!
Let $\mathbb{K}^{m \times n}$ denote the set of $m \times n$ matrices with entries in the field $\mathbb{K}$. Then, $\mathbb{K}^{m \times n}$ is a vector space over $\mathbb{K}$, with the matrix addition and product with scalars.

Properties

- The identity element is the zero matrix.
- $\dim_{\mathbb{K}}(\mathbb{K}^{m \times n}) = m \times n$.
- The concepts of linear independence, system of generators and basis do not change; we have to take into account that every element of the vector space $\mathbb{K}^{m \times n}$ is a matrix.

Dimensions and canonical basis for vector space of matrices

Let $\mathbb{K} = \mathbb{R}$. Then, the canonical basis of $\mathbb{R}^{m \times n}$ is the set given by the $m \times n$ matrices

$$E_{ij} = (e_{ij}) \in \mathbb{R}^{m \times n}$$

such that all elements are zero except the element e_{ij} which is equal to 1.

Example

Consider the vector space $\mathbb{R}^{3\times 2}$ (its elements are the rectangular matrices with 3 rows and 2 columns).
We have $\dim_{\mathbb{R}}(\mathbb{R}^{3\times 2}) = 3 \times 2 = 6$.
The canonical basis is formed by the 6 matrices

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix},$$

$$E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad E_{31} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad E_{32} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Matrix operations

Matrix-matrix product

The product of two matrices $A \in \mathbb{K}^{m\times p}$, $B \in \mathbb{K}^{p\times n}$ is the matrix $C \in \mathbb{K}^{m\times n}$, $C = AB$, given by the elements

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}, \qquad i = 1, \dots, m, \quad j = 1, \dots, n.$$

Properties:

- Matrix product is associative: $(AB)C = A(BC)$;
- Matrix product is distributive: $(A + B)C = AC + BC$ and $C(A + B) = CA + CB$;
- Matrix product is (in general) non commutative: $AB \neq BA$.

Matrix-vector product

The product of a matrix $A \in \mathbb{R}^{m\times p}$ and a vector $\mathbf{u} \in \mathbb{R}^{p}$ is the vector $\mathbf{v} \in \mathbb{R}^{m}$, $\mathbf{v} = A\mathbf{u}$, given by the elements

$$v_i = \sum_{k=1}^{p} a_{ik} u_k, \qquad i = 1, \dots, m.$$

Matrix operations

Transposition of a matrix

The transpose of a matrix $A \in \mathbb{K}^{m\times n}$ is the matrix $A^T \in \mathbb{K}^{n\times m}$ such that $a_{ij}^T = a_{ji}$ $(i = 1, \dots, m,$ $j = 1, \dots, n)$.

Properties:

- $(A^T)^T = A, \qquad \forall A \in \mathbb{K}^{m\times n}$;
- Linearity: $(\lambda A + \mu B)^T = \lambda A^T + \mu B^T, \qquad \forall A, B \in \mathbb{K}^{m\times n}, \ \forall \lambda, \mu \in \mathbb{K}$;
- $(AB)^T = B^T A^T, \qquad \forall A \in \mathbb{K}^{m\times p}, \forall B \in \mathbb{K}^{p\times n}$.

Identity matrix

$\mathbb{I} \in \mathbb{K}^{n\times n} : \mathbb{I} = (\delta_{ij}) \quad i, j = 1, \dots, n.$

Null matrix

$O \in \mathbb{K}^{n\times n}$ such that all entries are 0.

Matrices

Upper triangular matrix

$A \in \mathbb{K}^{n\times n}$ is an upper triangular matrix if $a_{ij} = 0 \quad \forall i > j, \quad i,j = 1,\dots,n.$

Lower triangular matrix

$A \in \mathbb{K}^{n\times n}$ is a lower triangular matrix if $a_{ij} = 0 \quad \forall i < j, \quad i,j = 1,\dots,n.$

Diagonal matrix

$A \in \mathbb{K}^{n\times n}$ is a diagonal matrix if $a_{ij} = 0 \quad \forall i \neq j, \quad i,j = 1,\dots,n.$

Submatrix

A submatrix of a matrix is obtained by deleting any collection of rows and/or columns.

Matrices

Symmetric matrix

$A \in \mathbb{K}^{n\times n}$ is a symmetric matrix if $a_{ij} = a_{ji}$ $\forall i, j = 1, \dots, n$, *i.e.*, $A = A^T$.

Skew-symmetric matrix

$A \in \mathbb{K}^{n\times n}$ is a skew-symmetric matrix if $a_{ij} = -a_{ji}$ $\forall i, j = 1, \dots, n$, *i.e.*, $A = -A^T$.

Invertible matrix

$A \in \mathbb{K}^{n\times n}$ is an invertible matrix if there exists $B \in \mathbb{K}^{n\times n}$ such that $AB = BA = \mathbb{I}$.
B is called the inverse matrix of A and is denoted by $B = A^{-1}$.

Trace

The trace of a matrix $A \in \mathbb{K}^{n\times n}$ is given by

$$\text{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Properties:

- the set of matrices $A \in \mathbb{K}^{n\times n}$ such that $\text{tr}(A) = 0$ is a vector space;
- $\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B), \qquad \forall A, B \in \mathbb{K}^{n\times n}$;
- $\text{tr}(\lambda A) = \lambda\text{tr}(A), \qquad \forall\lambda \in \mathbb{K},\ \forall A \in \mathbb{K}^{n\times n}$;
- $\text{tr}(A) = \text{tr}(A^T), \qquad \forall A \in \mathbb{K}^{n\times n}$;
- $\text{tr}(AB) = \text{tr}(BA) \qquad \forall A \in \mathbb{K}^{m\times n},\ \forall B \in \mathbb{K}^{n\times m}$.

Determinant

To each matrix $A \in \mathbb{K}^{n\times n}$ we can associate the determinant as a unique function $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ such that:

- $\det(AB) = \det(A)\det(B), \qquad \forall A, B \in \mathbb{K}^{n\times n}$;
- $\det(\mathbb{I}) = 1$;
- $\det(A) \neq 0$ if and only if A is invertible.

Laplace expansion

The determinant of a matrix $A \in \mathbb{K}^{n\times n}$ can be computed by the recursive formula

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij}, \qquad \forall j = 1, \dots, n,$$

or, equivalently

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}, \qquad \forall i = 1, \dots, n,$$

where M_{ij} is the (i,j)-minor, *i.e.*, the determinant of the submatrix of A obtained by removing the i-th row and the j-th column of A. The term $C_{ij} = (-1)^{i+j} M_{ij}$ is called the cofactor of a_{ij} in A.
The computational cost is $O(n!)$.

Determinant: properties

Given $A \in \mathbb{K}^{n \times n}$:

- $\det(\lambda A) = \lambda^n \det(A), \qquad \forall \lambda \in \mathbb{K}$;
- $\det(A^T) = \det(A)$;
- $\det(A^{-1}) = \dfrac{1}{\det(A)}$;
- if A is a (lower or upper) triangular matrix then $\det(A) = \Pi_{i=1}^{n} a_{ii}$;
- interchanging any pair of rows or columns of a matrix multiplies its determinant by -1;
- adding a scalar multiple of one row (column) to another row (column) does not change the value of the determinant;
- $\det(A) = 0$ if
  - some (row) column is such that all its entries are zero;
  - two rows (columns) are proportional;
  - some row (column) can be expressed as a linear combination of the other rows (columns).

Minor

Let $A \in \mathbb{K}^{m \times n}$ and k an integer with $0 < k \leq m$, and $k \leq n$. A minor or order k of A is the determinant of a $k \times k$ submatrix obtained from A by deleting $m - k$ rows and $n - k$ columns.

Cofactor matrix

Given $A \in \mathbb{K}^{n\times n}$, the cofactor matrix C of A is the matrix whose entries are the cofactors of A, *i.e.*,

$$C_{ij} = (-1)^{i+j} M_{ij}, \qquad i,j = 1,\dots,n.$$

Adjugate matrix

The adjugate (or classical adjoint) matrix of $A \in \mathbb{K}^{n\times n}$ is the transpose of the cofactor matrix C, *i.e.*, $\text{adj}(A) = C^T$, with components

$$\text{adj}(A)_{ij} = C_{ji} = (-1)^{i+j} M_{ji}, \qquad i,j = 1,\dots,n.$$

Inverse matrix

A matrix $A \in \mathbb{K}^{n\times n}$ is invertible if and only if $\det(A) \neq 0$.
The inverse of the matrix A is the adjugate matrix of A times the reciprocal of the determinant of A, *i.e.*,

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

Property: $(AB)^{-1} = B^{-1}A^{-1}$.

Orthogonal matrix

$A \in \mathbb{K}^{n\times n}$ is an orthogonal matrix if $A^T = A^{-1}$, that means $AA^T = A^TA = I$.
Properties: $\det(A) = 1$ or $\det(A) = -1$.

Rank

The rank of a matrix $A \in \mathbb{K}^{m\times n}$, denoted by $\operatorname{rank}(A)$, is the maximum order of the not null minors of A, *i.e.*, the maximum order of the square submatrices which can be extracted from A so that their determinant is not null.
Equivalently:

- $\operatorname{rank}(A)$ is the maximum number of linearly independent rows of A;
- $\operatorname{rank}(A)$ is the maximum number of linearly independent columns of A;
- $\operatorname{rank}(A) = \dim_{\mathbb{K}}(\operatorname{Im}\Phi)$, with $\Phi\colon \mathbb{K}^n \to \mathbb{K}^m$ such that $\Phi(\mathbf{u}) = A\mathbf{u}$.

It follows that

$$0 \le \operatorname{rank}(A) \le \min\{m, n\}.$$

Properties:

- $\operatorname{rank}(A) = 0$ if and only if A is the null matrix;
- $\operatorname{rank}(A) = \operatorname{rank}(A^T)$;
- if $\operatorname{rank}(A) = \min\{m, n\}$, A is said to be full rank;
- if $m = n$ and $\det(A) \neq 0$, then $\operatorname{rank}(A) = n$, *i.e.*, the rank coincides with the order of A.

Rank-nullity theorem

Let $\Phi\colon U \to V$ be a linear operator between two vector spaces U and V over a field $\mathbb{K}$, with $\dim_{\mathbb{K}}(U) < \infty$, *i.e.*, U is finite-dimensional.
It is

$$\dim_{\mathbb{K}}(U) = \dim_{\mathbb{K}}(\ker\Phi) + \dim_{\mathbb{K}}(\mathrm{Im}\Phi).$$

In the vector space of matrices $\mathbb{K}^{m\times n}$, we associate to each matrix $A \in \mathbb{K}^{m\times n}$ a linear operator $\Phi\colon \mathbb{K}^n \to \mathbb{K}^m$ such that $\Phi(\mathbf{u}) = A\mathbf{u}$.
Since $\mathrm{rank}(A) = \dim_{\mathbb{K}}(\mathrm{Im}\Phi)$ and $n = \dim_{\mathbb{K}}(\mathbb{K}^n)$ is the number of columns of A, the rank-nullity theorem reads

$$n = \dim_{\mathbb{K}}(\ker\Phi) + \mathrm{rank}(A).$$

Rank of a matrix by means of determinants

Given $A \in \mathbb{K}^{m \times n}$:

- let $i_1 = \min\{m, n\}$. Compute all minors of order i_1. If there exists at least a non-zero minor of order i_1, then $\text{rank}(A) = i_1$; otherwise, it is $\text{rank}(A) < i_1$ and go to next step;
- let $i_2 = i_1 - 1$, and compute all minors of order i_2. If there exists at least a non-zero minor of order i_2, then $\text{rank}(A) = i_2$. If this is not true, it is $\text{rank}(A) < i_2$, and we need to reiterate the process;
- the algorithm stops when we find a non-zero minor, and the rank is egual to the order of the non-zero minor.

Remark

The number of all minors of different orders can be great even for a matrix of not large dimension. For example, a matrix of size 4×5 has 5 minors of the fourth order, 40 minors of the third order, 60 minors of the second order and 20 minors of the first order (minors of the first order coincide with elements of a matrix), *i.e.*, the matrix has 125 minors. However, such a method of determining the rank of a matrix is useless due to a great number of computations!

Bordering minor

Let $A \in \mathbb{K}^{m\times n}$ and consider a submatrix M of A of order k, with $k \leq \min\{m, n\}$.
A bordering minor of A is the determinant of every submatrix of A of order $k+1$ obtained by adding a row and a column to the submatrix M.

Bordering theorem (Kronecker)

Let $A \in \mathbb{K}^{m\times n}$ and $k \leq \min\{m, n\}$. We have:
$\text{rank}(A) = k$ if and only if there exists a non-zero minor of A of order k and all its bordering minors of order $k+1$ are equal to zero.

Bordering algorithm

- Find a non-zero minor M of order k.
- Compute all its bordering minors: if all bordering minors are equal to zero, the algorithm stops, and the rank of the matrix is equal to the order of the minor M. Otherwise, go to next step.
- If we find a non-zero bordering minor M' of order $k+1$, then compute all bordering minors of order $k+2$, *i.e.*, repeat the described cycle of computations.

There will be finitely many such cycles and the number of such cycles doesn't exceed the number of rows and columns.

Leading coefficient

For each row in a matrix, if the row does not consist of only zeros, then the leftmost non-zero entry is called the leading coefficient (or pivot) of that row.

Row echelon form

A matrix is in row echelon form if:

- all rows consisting of only zeros are at the bottom;
- the leading coefficient of a non-zero row is always strictly to the right of the leading coefficient of the row above it.

Reduced row echelon form

A matrix is in reduced row echelon form if:

- it is in row echelon form;
- all the leading coefficients are equal to 1;
- in every column containing a leading coefficient, all of the other entries in that column are zero.

Examples

This matrix is not in row echelon form:

$$\begin{bmatrix} 1 & a_0 & a_1 & a_2 & a_3 \\ 0 & a_1 & 2 & a_4 & a_5 \\ 0 & 1 & 0 & 1 & a_6 \\ 0 & 0 & a_7 & 0 & 0 \end{bmatrix}$$

This matrix is in row echelon form:

$$\begin{bmatrix} 1 & a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & 2 & a_4 & a_5 \\ 0 & 0 & 0 & 1 & a_6 \\ 0 & 0 & 0 & 0 & a_7 \end{bmatrix}$$

This matrix is in reduced row echelon form:

$$\begin{bmatrix} 1 & 0 & a_1 & 0 & b_1 \\ 0 & 1 & a_2 & 0 & b_2 \\ 0 & 0 & 0 & 1 & b_3 \end{bmatrix}$$

Gaussian elimination

Gaussian elimination consists of a sequence of operations performed on the rows of a matrix in order to transform it into an upper triangular form, *i.e.*, in row echelon form.
This method can also be used to compute:

- the rank of a matrix;
- the determinant of a square matrix;
- the inverse of a matrix (if it is invertible).

There are three types of elementary row operations:

1. swap the positions of two rows;
2. multiply a row by a non-zero scalar;
3. replace a row with the sum of that row and a multiple of another row.

Gauss algorithm

Let $A \in \mathbb{K}^{m\times n}$.

- Choose the leftmost column C_k, with $1 \le k \le n$ with at least a non-zero entry. Set as a pivot the element
$$\max_{1\le i\le m} |a_{ik}|.$$
- If the pivot is not in the first row R_1, swap R_1 with the row containing the pivot.
- Substitute the R_i $(i > 1)$ row with a linear combination of the row R_i and that containing the pivot in such a way the entries below the pivot are all zero, *i.e.*,
$$R_i \longrightarrow R_i + m_{ik} R_k,$$
where $m_{ik} = -\dfrac{a_{ik}}{a_{kk}}$.
- Repeat the process considering the submatrix obtained by removing the row and column containing the pivot, until a row echelon form is reached.

Determinant with Gauss elimination

After reducing a matrix $A \in \mathbb{K}^{n \times n}$ into a upper triangular matrix B, we have to take into account:

- each swap of rows changes the sign of the determinant;
- if a row is multiplied by a scalar λ, then the determinant of the reduced matrix is equal to the determinant of A multiplied by the scalar λ;
- substituting a row with a linear combination of that row with another one does not change the value of the determinant.

Let d be the product of the scalars by which the determinant has been multiplied, using the above rules. Then, the determinant of A is the quotient by d of the product of the elements of the diagonal of B, *i.e.*,

$$\det(A) = \frac{\prod_{i=1}^{n} b_{ii}}{d}.$$

The computation of the determinant using this procedure is the least demanding system known from the computational point of view (polynomial growth $O(n^3)$ instead of factorial $O(n!)$).

Rank with Gauss elimination

Let $A \in \mathbb{K}^{m \times n}$, and reduce it to a row echelon form. The rank of the matrix A is equal to the number of pivots of the obtained row echelon matrix.

Inverse matrix with Gauss-Jordan elimination

Given $A \in \mathbb{K}^{n\times n}$, row reduction can be used to compute its inverse matrix:

- create an augmented matrix with the left side being the matrix A and the right side the $n \times n$ identity matrix $\mathbb{I}$, *i.e.*, an $n \times 2n$ block matrix $[A|\mathbb{I}]$;
- using row operations, determine the reduced row echelon form of the $n \times 2n$ block matrix $[A|\mathbb{I}]$;
- the matrix A is invertible if and only if the left block can be reduced to the identity matrix $\mathbb{I}$; in this case, the right block of the final matrix is A^{-1}, *i.e.*, the result is a matrix $[\mathbb{I}|A^{-1}]$.

If the algorithm is not able to reduce the left block to $\mathbb{I}$, then A is not invertible.

Matrix norm

Norm

In the vector space $\mathbb{K}^{m\times n}$, we can introduce a matrix norm as a function

$$\begin{aligned} \|\cdot\| \colon \mathbb{K}^{m\times n} &\to \mathbb{R} \\ A &\mapsto \|A\| \end{aligned}$$

that satisfies the axioms:

- $\|A\| \geq 0, \quad \forall A \in \mathbb{K}^{m\times n}$;
- $\|A\| = 0$ if and only if $A = O$;
- $\|\lambda A\| = |\lambda|\,\|A\|, \quad \forall \lambda \in \mathbb{K},\ \forall A \in \mathbb{K}^{m\times n}$;
- $\|A + B\| \leq \|A\| + \|B\|, \quad \forall A, B \in \mathbb{K}^{m\times n}$;
- $\|AB\| \leq \|A\|\,\|B\|, \quad \forall A \in \mathbb{K}^{m\times p},\ \forall B \in \mathbb{K}^{p\times n}$.

Matrix norm

Operator norm

Let $\|\cdot\|_\alpha$ be a vector norm on $\mathbb{K}^n$ and $\|\cdot\|_\beta$ be a vector norm on $\mathbb{K}^m$. Since any $m \times n$ matrix $A \in \mathbb{K}^{m\times n}$ induces a linear operator $\Phi\colon \mathbb{K}^n \to \mathbb{K}^m$, we can define the induced norm (or operator norm) in $\mathbb{K}^{m\times n}$ as

$$\|A\|_{\alpha,\beta} = \sup\left\{\frac{\|A\mathbf{x}\|_\beta}{\|\mathbf{x}\|_\alpha} \;:\; \mathbf{x} \in \mathbb{K}^n \text{ with } \mathbf{x} \neq \mathbf{0}\right\}.$$

This norm measures how much the mapping induced by A can stretch vectors.

Matrix norms induced by vector L_p-norms

The vector L_p-norms induce the operator norms $\|\cdot\|_p$ defined as

$$\|A\|_p = \sup\left\{\frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} \;:\; \mathbf{x} \in \mathbb{K}^n \text{ with } \mathbf{x} \neq \mathbf{0}\right\}.$$

Matrix norms induced by vector L_p-norms

Matrix norm induced by 1-norm

$$\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|.$$

Matrix norm induced by ∞-norm

$$\|A\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|.$$

Properties:

- $\|A\|_1 = \|A^T\|_\infty$;
- if $A = A^T$, then $\|A\|_1 = \|A\|_\infty$.

Matrix norm induced by 2-norm (Spectral norm)

$$\|A\|_2 = \sqrt{\rho(A^* A)},$$

where ρ is the largest eigenvalue of the matrix $A^* A$, and A^* denotes the conjugate transpose (the transpose if $A \in \mathbb{R}^{n \times n}$).

Matrix norms induced by vector L_p-norms

Example: matrix norms induced by vector L_p-norms

Given the matrix

$$A = \begin{bmatrix} -3 & 5 & 7 & 1 \\ 2 & 6 & 4 & 2 \\ 0 & 2 & 8 & -1 \end{bmatrix},$$

we have:

$$\begin{aligned} \|A\|_1 &= \max_{1\le j\le 4} \sum_{i=1}^{3} |a_{ij}| = \\ &= \max(|-3| + 2 + 0; 5 + 6 + 2; 7 + 4 + 8; 1 + 2 + |-1|) = \max(5; 13; 19; 4) = 19, \\ \|A\|_\infty &= \max_{1\le i\le 3} \sum_{j=1}^{4} |a_{ij}| = \\ &= \max(|-3| + 5 + 7 + 1; 2 + 6 + 4 + 2; 0 + 2 + 8 + |-1|) = \max(16; 14; 11) = 16. \end{aligned}$$

Matrix norm

Consistent matrix norms

A matrix norm $\|\cdot\|$ on $\mathbb{K}^{m\times n}$ is called consistent with a vector norm $\|\cdot\|_\alpha$ on $\mathbb{K}^n$ and a vector norm $\|\cdot\|_\beta$ on $\mathbb{K}^m$ if

$$\|A\mathbf{x}\|_\beta \leq \|A\| \, \|\mathbf{x}\|_\alpha, \qquad \forall A \in \mathbb{K}^{m\times n}, \ \forall \mathbf{x} \in \mathbb{K}^n.$$

In the special case with $m = n$ and $\alpha = \beta$, the norm $\|\cdot\|$ is also called compatible with $\|\cdot\|_\alpha$.

Remark

All induced norms are consistent by definition!

Entry-wise matrix norms

Entry-wise matrix norms p-norm

Entry-wise matrix norms treat an $m \times n$ matrix as a vector of size $m \cdot n$, and use one of the familiar vector norms.

For $p \in [1, \infty)$:

$$\|A\|_p = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^p \right)^{\frac{1}{p}}.$$

Remark

Entry-wise p-norms are different from the induce L_p-norms!

Entry-wise matrix norms

Entry-wise 1-norm

For $p = 1$, we have the entry-wise 1-norm defined by:

$$\|A\|_1 = \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|.$$

Frobenius norm

For $p = 2$, we have the Frobenius norm defined by:

$$\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2}.$$

Entry-wise ∞-norm (max norm)

For $p \to \infty$, we have the entry-wise ∞-norm or max norm:

$$\|A\|_{\max} = \max_{1 \le i \le m, 1 \le j \le n} |a_{ij}|.$$

Entry-wise matrix norms

Example: entry-wise matrix norms

Given the matrix

$$A = \begin{bmatrix} -3 & 5 & 7 & 1 \\ 2 & 6 & 4 & 2 \\ 0 & 2 & 8 & -1 \end{bmatrix},$$

we have:

$$\|A\|_1 = \sum_{i=1}^{3}\sum_{j=1}^{4} |a_{ij}| = |-3| + 5 + 7 + 1 + 2 + 6 + 4 + 2 + 0 + 2 + 8 + |-1| = 41,$$

$$\|A\|_F = \sqrt{\sum_{i=1}^{3}\sum_{j=1}^{4} a_{ij}^2} = \sqrt{3^2 + 5^2 + 7^2 + 1^2 + 2^2 + 6^2 + 4^2 + 2^2 + 0^2 + 2^2 + 8^2 + 1^2} = \sqrt{213} \approx 14.6,$$

$$\|A\|_{\max} = \max_{1\leq i\leq 3, 1\leq j\leq 4} |a_{ij}| = 8.$$

Linear systems

Definition

A linear system is a collection of linear equations, involving the same variables, that must be verified all together.

General form

A general system of m linear equations with unknowns $x_j \in \mathbb{K}$ ($j = 1, \ldots, n$) can be written as

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases}$$

where $a_{ij} \in \mathbb{K}$ ($i = 1, \ldots, m$, $j = 1, \ldots, n$) are the coefficients of the system, and $b_i \in \mathbb{K}$ ($i = 1, \ldots, m$) are the constant terms. In components, the linear system reads

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \qquad i = 1, \ldots, m.$$

Linear systems

Matrix form

The linear system can we written in the matrix form

$$A\mathbf{x} = \mathbf{b},$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{K}^{m\times n}, \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n, \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^m.$$

A is called coefficient matrix or incomplete matrix, whereas $\mathbf{x}$ and $\mathbf{b}$ are the vectors of the unknowns and constant terms, respectively. We can associate to the system $A\mathbf{x} = \mathbf{b}$ also the complete matrix (or associated matrix, or augmented matrix) $[A|\mathbf{b}]$:

$$[A|\mathbf{b}] = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right) \in \mathbb{K}^{m\times(n+1)}.$$

Linear systems

Solution of a linear system

A n-tuple $(x_1, \ldots, x_n) \in \mathbb{K}^n$ is a solution of the system if it satisfies all m equations. The set of all possible solutions is called the solution set.

For a linear system three cases may occur:

- the system has infinitely many solutions;
- the system has a single unique solution;
- the system has no solution.

Geometric interpretation

The geometric interpretation of a linear system depend on the number of the unknowns:

- $n = 2 \Rightarrow$ each linear equation determines a line in the two-dimensional space. Because a solution of a linear system must satisfy all the equations, the solution set is the intersection of these lines, and is hence either a line, a single point, or the empty set;
- $n = 3 \Rightarrow$ each linear equation determines a plane in three-dimensional space, and the solution set is the intersection of these planes. The solution set may be a plane, a line, a single point, or the empty set;
- $n > 3 \Rightarrow$ each linear equation determines a hyperplane in the n-dimensional space. The solution set is the intersection of these hyperplanes, and is a flat, which may have any dimension lower than n.

Linear systems

Indipendence

The equations of a linear system are independent if none of the equations can be derived from the others. When the equations are independent, each equation contains new information about the variables, and removing any of the equations increases the size of the solution set. For linear equations, logical independence is the same as linear independence.

Consistency

If a linear system admits at least one solution, then it is called consistent (or compatible); otherwise, it is inconsistent (or incompatible, or impossible).

Equivalent systems

Two systems are said to be equivalent if and only if they have the same solution set.

Homogeneous linear system

If $b_i = 0 \;\; \forall i = 1, \dots, m$, the linear system is called homogeneous.

Linear systems

Solution set for homogeneous systems

- A homogeneous linear system is always compatible, in fact the n-tuple $(x_1, \ldots, x_n) = (0, \ldots, 0)$ is a solution, called trivial solution.
- The solution set of a homogeneous linear system with n unknowns and coefficients in $\mathbb{K}$ is a subspace of $\mathbb{K}^n$. In fact:
  1. if $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{K}^n$ and $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{K}^n$ are two vectors representing solutions to a homogeneous system, then the vector sum $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n) \in \mathbb{K}^n$ is also a solution to the system;
  2. if $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{K}^n$ is a vector representing a solution to a homogeneous system, and $\lambda \in \mathbb{K}$ is any scalar, then $\lambda\mathbf{x} = \lambda(x_1, \ldots, x_n) \in \mathbb{K}^n$ is also a solution to the system.
- Let $\Phi\colon \mathbb{K}^n \to \mathbb{K}^m$ defined as
$$\Phi(\mathbf{x}) = A\mathbf{x}.$$
We have that $\ker(\Phi) = \{\mathbf{x} \in \mathbb{K}^n : \ \Phi(\mathbf{x}) = \mathbf{0}\}$ is the solution set for the homogeneous system $A\mathbf{x} = \mathbf{0}$ and $\mathrm{Im}(\Phi) = \{\mathbf{y} \in \mathbb{K}^m : \ \mathbf{y} = \Phi(\mathbf{x}), \ \text{with} \ \mathbf{x} \in \mathbb{K}^n\}$ is generated by the columns of A. Due to rank-nullity theorem, the dimension of the solution set for homogeneous systems is
$$\dim_{\mathbb{K}}(\ker(\Phi)) = n - \mathrm{rank}(A).$$

Linear systems

Relation to nonhomogeneous systems

There is a close relationship between the solutions to a linear system $A\mathbf{x} = \mathbf{b}$ and the solutions to the corresponding homogeneous system $A\mathbf{x} = \mathbf{0}$. If $\mathbf{u}$ is any specific solution to the linear system $A\mathbf{x} = \mathbf{b}$, then the entire solution set can be described as

$$\{\mathbf{u} + \mathbf{v} : \ \mathbf{v} \text{ is any solution to } A\mathbf{x} = \mathbf{0}\}.$$

Geometrically, the solution set for $A\mathbf{x} = \mathbf{b}$ is a translation of the solution set for $A\mathbf{x} = \mathbf{0}$.
This reasoning only applies if the system $A\mathbf{x} = \mathbf{b}$ admits solution. This occurs if and only if the vector $\mathbf{b} \in \text{Im}(\Phi)$, where $\Phi\colon \mathbb{K}^n \to \mathbb{K}^m$ is the linear operator defined as

$$\Phi(\mathbf{x}) = A\mathbf{x}.$$

In fact, $\text{Im}(\Phi)$ is generated by the columns of A, and therefore $\mathbf{b} \in \text{Im}(\Phi)$ if and only if the span of the columns of A contains $\mathbf{b}$, *i.e.*, if and only if the space generated by the columns of A equals the space generated by the columns of $[A|\mathbf{b}]$. This is equivalent to require that the matrices A and $[A|\mathbf{b}]$ have the same rank (that is the Rouché-Capelli theorem). In such a way, since the the solution set for $A\mathbf{x} = \mathbf{b}$ is obtained from a translation of the solution set for $A\mathbf{x} = \mathbf{0}$, the dimension of the solution set of the system $A\mathbf{x} = \mathbf{b}$ is equal to the dimension of the solution set of the associated homogeneous system $A\mathbf{x} = \mathbf{0}$, *i.e.*, the dimension of both solution set is $\dim_{\mathbb{K}}(\ker(\Phi)) = n - \text{rank}(A)$.

Rouché-Capelli theorem

A linear system $A\mathbf{x} = \mathbf{b}$, with $A \in \mathbb{K}^{m \times n}$, $\mathbf{x} \in \mathbb{K}^n$ and $\mathbf{b} \in \mathbb{K}^m$, admits solution (is compatible) if and only if

$$\text{rank}(A) = \text{rank}(A|\mathbf{b}),$$

i.e., the rank of its coefficient matrix A is equal to the rank of its augmented matrix $[A|\mathbf{b}]$. The solution set is a subspace of $\mathbb{K}^n$ with dimension $n - \text{rank}(A)$.
In particular:

- if $\text{rank}(A) = \text{rank}(A|\mathbf{b}) = n$, the system admits a unique solution;
- if $\text{rank}(A) = \text{rank}(A|\mathbf{b}) < n$, the system admits $\infty^{n-\text{rank}(A)}$ solutions depending on $n - \text{rank}(A)$ parameters.

Methods for solving linear systems

Elimination of variables

Let $A\mathbf{x} = \mathbf{b}$ be a compatible linear system, with $A \in \mathbb{K}^{m\times n}$, $\mathbf{x} \in \mathbb{K}^n$ and $\mathbf{b} \in \mathbb{K}^m$. The simplest method for solving the system $A\mathbf{x} = \mathbf{b}$ is to repeatedly eliminate variables.

This method can be described as follows:

1. solve the first equation with respect to a variable in terms of the other variables;
2. substitute this expression into the remaining equations; this yields a system of equations with $m-1$ equations and $n-1$ unknowns;
3. repeat until the system is reduced to a single linear equation; two situations may occur:
   - if $\text{rank}(A) = \text{rank}(A|\mathbf{b}) = n$, the system $A\mathbf{x} = \mathbf{b}$ admits only one solution, and the obtained linear equation involves only one unknown; solve this equation with respect to the involved unknown, and then the remaining unknowns are determined by using backward substitutions;
   - if $\text{rank}(A) = \text{rank}(A|\mathbf{b}) < n$, the system $A\mathbf{x} = \mathbf{b}$ admits $\infty^{n-\text{rank}(A)}$ solutions depending on $n - \text{rank}(A)$ parameters; in the last equations there are $n - \text{rank}(A)$ unknowns that cannot be eliminated and the remaining ones can be expressed in terms of these $n - \text{rank}(A)$ unknowns. Choosing as free parameters a set of $n - \text{rank}(A)$ unknowns, the $\infty^{n-\text{rank}(A)}$ solutions can be determined by means of backward substitutions.

Methods for solving linear systems

Cramer's rule for square systems

Consider a system of n linear equations with n unknowns, represented in matrix form as follows:

$$A\mathbf{x} = \mathbf{b},$$

with

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in \mathbb{K}^{n\times n}, \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n, \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{K}^n.$$

If $\det(A) \neq 0$, the system $A\mathbf{x} = \mathbf{b}$ has a unique solution, whose values for the unknowns are given by:

$$x_i = \frac{\det(A_i)}{\det(A)}, \qquad i = 1, \ldots, n,$$

where A_i is the matrix formed by replacing the i-th column of A by the column vector $\mathbf{b}$.

Methods for solving linear systems

Remark

Cramer's method is suitable for computing the solution of $n \times n$ linear systems only if n is very small. In practice, the method requires the computation of $n+1$ determinants of $n \times n$ matrices. Applying Laplace expansion, each of these requires $n!$ multiplications, for a total of $(n+1)!$ multiplications. Alternatively, the $n+1$ determinants can be computed by means of the Gauss algorithm.

Cramer's rule for rectangular systems

Consider a compatible system of m linear equations with n unknowns, represented in matrix form as

$$A\mathbf{x} = \mathbf{b},$$

where $A \in \mathbb{K}^{m \times n}$, $\mathbf{x} \in \mathbb{K}^n$, and $\mathbf{b} \in \mathbb{K}^m$.
If $\text{rank}(A) = \text{rank}(A|\mathbf{b}) = r < n$, the system admits ∞^{n-r} solutions that can be computed by means of Cramer's rule with the following steps:

- let A' be the submatrix of A associated to a non-zero minor of order r;
- delete from the original system the equations corresponding to the rows of A which are non contained in A', and assign as free parameters the $n-r$ unknowns corresponding to the columns of A that are not contained in A';
- thus, we obtain a square system that can be solved applying the Cramer's rule.

Methods for solving linear systems

Gauss elimination

Let $A\mathbf{x} = \mathbf{b}$ be a linear system of m equations with n unknowns, where $A \in \mathbb{K}^{m\times n}$, $\mathbf{x} \in \mathbb{K}^n$ and $\mathbf{b} \in \mathbb{K}^m$. In Gaussian elimination, the linear system $A\mathbf{x} = \mathbf{b}$ is represented by the augmented matrix $[A|\mathbf{b}]$, that will be modified by using elementary row operations until it reaches a row echelon form. Because these operations are reversible, the obtained augmented matrix $[A|\mathbf{b}]$ always represents a linear system that is equivalent to the original one. Applying the Gauss algorithm, the augmented matrix $[A|\mathbf{b}]$ is reduced to a row echelon form, and two situations may occur:

1. if one or more rows of the reduced matrix are in the form $(0\dots0|k)$, with $k \neq 0$, the system is incompatible;
2. otherwise, the system is compatible and it admits ∞^{n-r} solutions, where n is the number of the unknowns, and r is the number of pivots of the reduced matrix.

If the system is compatible, we need to:

1. construct the linear system corresponding to the obtained row echelon matrix;
2. assign the role of free parameter to the $n-r$ unknowns that do not correspond to the pivots;
3. determine the solutions of the system proceeding with the backward substitutions.

Methods for solving linear systems

Backward substitution for triangular systems

Backward substitution is a procedure of solving a linear system $A\mathbf{x} = \mathbf{b}$ of n equations with n unknowns, with $A \in \mathbb{K}^{n\times n}$, $\mathbf{x} \in \mathbb{K}^n$, $\mathbf{b} \in \mathbb{K}^n$, and A is an upper triangular matrix (row echelon form) whose diagonal elements are not equal to zero. Since the matrix A is triangular, this procedure of solving a linear system is a modification of the general substitution method and can be described using simple formulas.
A similar procedure of solving a linear system with a lower triangular matrix is called forward substitution. The backward substitution can be considered as a part of the Gaussian elimination method for solving linear systems.

Backward substitution: algorithm

The backward substitution algorithm can be represented as

$$x_n = \frac{b_n}{a_{nn}},$$

$$x_i = \frac{b_i - \sum_{j=i+1}^{n} a_{ij}x_j}{a_{ii}}, \qquad \forall i = n-1, \ldots, 1.$$

Methods for solving linear systems

Forward substitution for triangular systems

Forward substitution is a procedure of solving a linear system $A\mathbf{x} = \mathbf{b}$ of n equations with n unknowns, with $A \in \mathbb{K}^{n\times n}$, $\mathbf{x} \in \mathbb{K}^n$, $\mathbf{b} \in \mathbb{K}^n$, and A is a lower triangular matrix whose diagonal elements are not equal to zero. Since the matrix A is triangular, this procedure of solving a linear system is a modification of the general substitution method and can be described using simple formulas.
Nevertheless, the structure of the forward substitution for a lower triangular matrix is similar to the structure of the backward substitution.

Forward substitution: algorithm

The forward substitution algorithm can be represented as

$$x_1 = \frac{b_1}{a_{11}},$$
$$x_i = \frac{b_i - \sum_{j=1}^{i-1} a_{ij}x_j}{a_{ii}}, \qquad \forall i = 2, \dots, n.$$

Methods for solving linear systems

How to proceed - Homogeneous linear systems

The homogeneous linear system $A\mathbf{x} = \mathbf{0}$, with $A \in \mathbb{K}^{m\times n}$ and $\mathbf{x} \in \mathbb{K}^n$, is always compatible (the complete matrix coincides with the incomplete matrix), and we have to compute $\text{rank}(A)$.
Then:

1. if $\text{rank}(A) = n$ (full rank), the system admits only one solution, *i.e.*, the trivial solution $\mathbf{x} = (0, \dots, 0)$;
2. if $\text{rank}(A) = r < n$, the system admits ∞^{n-r} solutions that can be determined with the suitable known methods.

How to proceed - Non-homogeneous linear systems

For a non-homogeneous linear system $A\mathbf{x} = \mathbf{b}$, with $A \in \mathbb{K}^{m\times n}$, $\mathbf{x} \in \mathbb{K}^n$ and $\mathbf{b} \in \mathbb{K}^m$, we have to check if the system is compatible by using the Rouché-Capelli theorem.
Then:

1. if $\text{rank}(A) \neq \text{rank}(A|\mathbf{b})$, the system does not admit solution;
2. if $\text{rank}(A) = \text{rank}(A|\mathbf{b}) = n$ (full rank), the system admits only one solution;
3. if $\text{rank}(A) = \text{rank}(A|\mathbf{b}) = r < n$, the system admits ∞^{n-r} solutions.

In the cases (2) and (3), the solutions are determined by using the suitable known methods.

Linear operators

Dimension for kernel and image

Let $\Phi\colon U \to V$ a linear operator between the vector spaces U and V over the field $\mathbb{K}$, with $\dim_{\mathbb{K}}(U) = n$ and $\dim_{\mathbb{K}}(V) = m$. The representation of Φ is the matrix $A \in \mathbb{K}^{m\times n}$ such that

$$\Phi(\mathbf{x}) = A\mathbf{x}, \qquad \mathbf{x} \in \mathbb{K}^n.$$

We have that $\ker(\Phi) = \{\mathbf{x} \in U \ : \ \Phi(\mathbf{x}) = \mathbf{0}\}$ is the solution set for the homogeneous system $A\mathbf{x} = \mathbf{0}$. Choose a basis for U and V, respectively,

$$\mathcal{B}_U = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}, \qquad \mathcal{B}_V = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}.$$

Due to rank-nullity theorem, the dimension of the kernel of Φ is equal to the dimension of the solution set for homogeneous systems, *i.e.*,

$$\dim_{\mathbb{K}}(\ker(\Phi)) = n - \operatorname{rank}(A).$$

and the dimension of the image of Φ is given by

$$\dim_{\mathbb{K}}(\operatorname{Im}(\Phi)) = \operatorname{rank}(A).$$

Basis for kernel

- If $\text{rank}(A) = n$, then $\dim_{\mathbb{K}}(\ker(\Phi)) = 0$, *i.e.*, $\ker(\Phi) = \{\mathbf{0}\}$, and we do not need to find a basis;
- If $\text{rank}(A) = r < n$, then $\dim_{\mathbb{K}}(\ker(\Phi)) = n - r$, and we determine ∞^{n-r} solutions of the system $A\mathbf{x} = \mathbf{0}$ which depend on $n - r$ free parameters; the recovered solutions must be expressed as linear combinations whose coefficients are the free parameters. The vectors of such linear combinations are a basis $\mathcal{B}_{sol} = \{\mathbf{s}_1, \ldots, \mathbf{s}_k\}$, with $k \leq n$, for the solution set of the homogeneous system.
  Two cases need to be distinguished:
  1. $U = \mathbb{R}^n$ and $\mathcal{B}_{\mathbb{R}^n}$ is the canonical basis of $\mathbb{R}^n \Rightarrow \mathcal{B}_{sol}$ is a basis for $\ker(\Phi)$;
  2. $U \neq \mathbb{R}^n$ or $\mathcal{B}_{\mathbb{R}^n}$ is not the canonical basis of $\mathbb{R}^n \Rightarrow$ the i-th vector basis of $\ker(\Phi)$ is determined by a linear combination of the vectors $\mathcal{B}_U = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ whose coefficients are the components of $\mathbf{s}_i$.

Basis for image

Compute $\text{rank}(A) = k$. Then, construct the set $\mathcal{B}_C = \{\mathbf{c}_1, \ldots, \mathbf{c}_k\}$, with $k \leq n$, where $\mathbf{c}_k$ are the linearly independent columns of A.
Two cases need to be distinguished:

1. $V = \mathbb{R}^m$ and $\mathcal{B}_{\mathbb{R}^m}$ is the canonical basis of $\mathbb{R}^m \Rightarrow \mathcal{B}_C$ is a basis for $\text{Im}(\Phi)$;
2. $V \neq \mathbb{R}^m$ or $\mathcal{B}_{\mathbb{R}^m}$ is not the canonical basis of $\mathbb{R}^m \Rightarrow$ the i-th vector basis of $\text{Im}(\Phi)$ is determined by a linear combination of the vectors $\mathcal{B}_V = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ whose coefficients are the components of $\mathbf{c}_i$.

Change of basis matrix

Let V be a vector space over the field $\mathbb{K}$, with $\dim_{\mathbb{K}}(V) = n$, and two bases $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, and $\mathcal{B}' = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\}$. Then, $\forall \mathbf{w} \in V$, we have

$$\mathbf{w} = \sum_{j=1}^{n} w_j \mathbf{v}_j = \sum_{i=1}^{n} w'_i \mathbf{v}'_i, \qquad w_j, w'_i \in \mathbb{K},$$

where w_j and w'_i are the components of $\mathbf{w}$ with respect to the bases $\mathcal{B}$ and $\mathcal{B}'$, respectively.
Since the vectors $\mathbf{v}_j \in \mathcal{B}$ are elements of V, we can express them as linear combinations of the vectors $\mathbf{v}'_i$ belonging to the basis $\mathcal{B}'$, *i.e.*,

$$\mathbf{v}_j = \sum_{i=1}^{n} A_{ij} \mathbf{v}'_i.$$

From

$$\mathbf{w} = \sum_{j=1}^{n} w_j \mathbf{v}_j = \sum_{j=1}^{n} w_j \left(\sum_{i=1}^{n} A_{ij} \mathbf{v}'_i \right) = \sum_{i=1}^{n} w'_i \mathbf{v}'_i,$$

it follows that

$$\sum_{i=1}^{n} \left(\sum_{j=1}^{n} A_{ij} w_j - w'_i \right) \mathbf{v}'_i = \mathbf{0}.$$

… Change of basis matrix

Due to the linear independence of vectors $\mathbf{v}'_i$, we have

$$w'_i = \sum_{j=1}^{n} A_{ij} w_j, \qquad i = 1, \dots, n,$$

or, in matrix form

$$\mathbf{w}' = A\mathbf{w},$$

where the i-th column of matrix A is made by the components of vectors $\mathbf{v}_i$ with respect to the basis $\mathcal{B}'$.
$A \in \mathbb{K}^{n\times n}$ is called change of basis matrix from the basis $\mathcal{B}$ to $\mathcal{B}'$.
Vice versa, if we want to determine the change of basis matrix $B \in \mathbb{K}^{n\times n}$ from the basis $\mathcal{B}'$ to $\mathcal{B}$, with the same procedure, we obtain

$$\mathbf{w} = B\mathbf{w}',$$

where the i-th column of matrix B is made by the components of vectors $\mathbf{v}'_i$ with respect to the basis $\mathcal{B}$.
From the relation

$$\mathbf{w}' = A\mathbf{w} = AB\mathbf{w}' \Longrightarrow (AB - \mathbb{I})\mathbf{w}' = \mathbf{0} \Longrightarrow AB = \mathbb{I},$$

i.e.,

$$B = A^{-1}.$$

Change of basis for linear operators

Let $\Phi\colon U \to V$ be a linear operator, with $\dim_{\mathbb{K}}(U) = n$ and $\dim_{\mathbb{K}}(V) = m$. Let $\mathcal{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, $\mathcal{E}' = \{\mathbf{e}'_1, \ldots, \mathbf{e}'_n\}$ bases for U, and $\mathcal{F} = \{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$, $\mathcal{F}' = \{\mathbf{f}'_1, \ldots, \mathbf{f}'_m\}$ bases for V. For all $\mathbf{u} \in U$ and $\mathbf{v} \in V$, it is

$$\mathbf{u} = \sum_{j=1}^{n} u_j \mathbf{e}_j = \sum_{i=1}^{n} u'_i \mathbf{e}'_i, \qquad \mathbf{v} = \sum_{j=1}^{m} v_j \mathbf{f}_j = \sum_{i=1}^{m} v'_i \mathbf{f}'_i, \qquad u_j, u'_i, v_j, v'_i \in \mathbb{K},$$

where u_j, u'_i (v_j, v'_i) are the components of $\mathbf{u}$ ($\mathbf{v}$) with respect to the bases $\mathcal{E}, \mathcal{E}'$ ($\mathcal{F}, \mathcal{F}'$), respectively. By means of the change of basis, we get

$$\mathbf{u}' = A\mathbf{u}, \qquad \mathbf{v}' = B\mathbf{v},$$

where $A \in \mathbb{K}^{n\times n}$ is the change of basis matrix from the basis $\mathcal{E}$ to $\mathcal{E}'$, and $B \in \mathbb{K}^{m\times m}$ is the change of basis matrix from the basis $\mathcal{F}$ to $\mathcal{F}'$. From the representation of a linear operator, we have

$$\mathbf{v} = T\mathbf{u}, \qquad \mathbf{v}' = T'\mathbf{u}',$$

where $T \in \mathbb{K}^{m\times n}$ is the associated matrix with respect to the bases $\mathcal{E}$ and $\mathcal{F}$, and $T' \in \mathbb{K}^{m\times n}$ is the associated matrix with respect to the bases $\mathcal{E}'$ and $\mathcal{F}'$. By using the above relations, we obtain:

$$T'\mathbf{u}' = \mathbf{v}' = B\mathbf{v} = BT\mathbf{u} = BTA^{-1}\mathbf{u}' \;\Rightarrow T' = BTA^{-1}.$$

Change of basis for endomorphism

Let $\Phi\colon V \to V$ be a linear operator, with and $\dim_{\mathbb{K}}(V) = n$. Let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $\mathcal{B}' = \{\mathbf{v}'_1, \dots, \mathbf{v}'_n\}$ bases for V, $T \in \mathbb{K}^{n\times n}$ the associated matrix with respect to the same basis $\mathcal{B}$, $T' \in \mathbb{K}^{n\times n}$ the associated matrix with respect to the same basis $\mathcal{B}'$, and $A \in \mathbb{K}^{n\times n}$ the change of basis matrix from the basis $\mathcal{B}$ to $\mathcal{B}'$. It is

$$T' = ATA^{-1},$$

or, equivalently,

$$T' = P^{-1}TP,$$

where $P \in \mathbb{K}^{n\times n}$ is the change of basis matrix from the basis $\mathcal{B}'$ to $\mathcal{B}$.

Eigenvalues and eigenvectors

Eigenvalues and eigenvectors

Let V be a vector space over the field $\mathbb{K}$, with $\dim_{\mathbb{K}}(V) = n$, and $\Phi\colon V \to V$ a linear operator with $A \in \mathbb{K}^{n\times n}$ its associated matrix.
A scalar λ is called eigenvalue (or characteristic value) of the matrix A if there exists a non-zero vector $\mathbf{v} \in V$ such that

$$A\mathbf{v} = \lambda\mathbf{v}.$$

The vector $\mathbf{v} \in V$ is called eigenvector (or characteristic vector) of the matrix A corresponding to the eigenvalue λ.

Interpretation

An eigenvector $\mathbf{v}$ of a linear operator Φ is a non-zero vector that, when the Φ is applied to it, does not change direction but it is scaled by a scalar factor.
The corresponding eigenvalue is the factor by which the eigenvector is scaled.

Remark

In a finite-dimensional vector space, it is equivalent to define eigenvalues and eigenvectors in terms of matrices or linear operators!

Eigenvalues and eigenvectors

Eigenvalues and eigenvector - Properties

- If $\mathbf{v} \in V$ is an eigenvector corresponding to the eigenvalue λ, then also $\alpha\mathbf{v}$, with $\alpha \in \mathbb{K}$ non-zero scalar, is an eigenvector.
- The set of the eigenvalues of a linear operator is called spectrum.
- The largest absolute value of the eigenvalues of linear operator (with associated matrix A) is called spectral radius and is denoted by $\rho(A)$.
- If $\mathbf{v}_1, \ldots, \mathbf{v}_k$, with $k = 1, \ldots, n$, are eigenvectors associated to distinct eigenvalues of a matrix A, then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent.

Eigenvalues and eigenvectors

Eigenspace

The set of all eigenvectors of a linear operator $\Phi\colon V \longrightarrow V$ (with associated matrix $A \in \mathbb{K}^{n\times n}$) corresponding to the same eigenvalue λ, together with the zero-vector, is a subspace of V that is called eigenspace (or characteristic space) of Φ associated to the eigenvalue λ, and is denoted by V_λ, say

$$V_\lambda = \{\mathbf{v} \in V \text{ such that } A\mathbf{v} = \lambda\mathbf{v}\}.$$

Eigensystem

The set of all eigenvectors of a linear operator $\Phi\colon V \longrightarrow V$, each paired with its corresponding eigenvalue, is called the eigensystem of Φ.

Eigenbasis

If a set of eigenvectors of a linear operator forms a basis in V, then this basis is called eigenbasis.

Characteristic polynomial

Let $\Phi\colon V \to V$ be a linear operator, with $\dim_{\mathbb{K}}(V) = n$, and $A \in \mathbb{K}^{n\times n}$ its associated matrix.
A scalar λ is an eigenvalue of the matrix A if there exists a non zero-vector $\mathbf{v} \in V$ such that

$$A\mathbf{v} = \lambda\mathbf{v},$$

that is equivalent to

$$(A - \lambda\mathbb{I})\mathbf{v} = \mathbf{0}, \qquad (*)$$

where $\mathbb{I}$ is $n \times n$ identity matrix and $\mathbf{0}$ is the zero-vector.
The result $(*)$ is a homogeneous linear system which will admit non-trivial solutions if and only if $\mathrm{rank}(A - \lambda\mathbb{I}) < n$, *i.e.*,

$$\det(A - \lambda\mathbb{I}) = 0,$$

that is called characteristic equation of A. The term

$$p(\lambda) = \det(A - \lambda\mathbb{I})$$

is called characteristic polynomial in the unknown λ associated to A, and is a polynomial of degree n.

Eigenvalues and eigenvectors

Characteristic polynomial - Properties

- Since the characteristic polynomial of a $n \times n$ matrix A is a polynomial of degree n, then it admits at most n distinct roots, *i.e.*, at most n distinct eigevalues, and it can be factored into the product of n linear terms,
$$\det(A - \lambda\mathbb{I}) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda),$$
where each λ_i may be real but in general is a complex number. The scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$, which may not all have distinct values, are the roots of the polynomial and are the eigenvalues of A.
- If $A \in \mathbb{R}^{n\times n}$, the coefficients of the characteristic polynomial will be real numbers, but the eigenvalues may still have non-zero imaginary parts. Therefore, the components of the corresponding eigenvectors may also have non-zero imaginary parts.
- If $A \in \mathbb{R}^{n\times n}$ and n is odd, the characteristic polynomial has odd degree, and at least one the roots is real; the remaining non-real roots are grouped into pairs of complex conjugates, namely with the two members of each pair having imaginary parts that differ only in sign and the same real part. Therefore, any real matrix with odd order has at least one real eigenvalue, whereas a real matrix with even order may not have any real eigenvalues. The eigenvectors associated with these complex eigenvalues are also complex and also appear in complex conjugate pairs.
- If the characteristic polynomial of the matrix A is full factorizable, then A is triangolable, *i.e.*, there exists a basis such that the associated matrix is in triangular form.

Eigenvalues and eigenvectors

How to compute eigenvalues and eigenvectors

1. Firstly, compute the eigenvalues of the matrix A, *i.e.*, the roots of the characteristic polynomial, say
$$\det(A - \lambda\mathbb{I}) = 0.$$
2. Collect $\lambda_1, \dots, \lambda_m$, with $1 \leq m \leq n$, the recovered distinct eigenvalues.
3. For each eigenvalue λ_i $(i = 1, \dots, m)$, compute the corresponding eigenspace V_{λ_i}, *i.e.*, the set of all eigenvectors associated to λ_i; this is done by determining a basis of the eigenspace V_{λ_i}, *i.e.*, by looking for the solution set of the linear homogeneous system
$$(A - \lambda_i\mathbb{I})\mathbf{v} = \mathbf{0}.$$
4. The eigenvectors corresponding to λ_i are the non-zero vectors of the subspace generated by the vectors of the basis $\mathcal{B}_{V_{\lambda_i}}$ of the eigenspace
$$V_\lambda = \{\mathbf{v} \in V \text{ such that } (A - \lambda\mathbb{I})\mathbf{v} = \mathbf{0}\}.$$

Eigenvalues and eigenvectors

Algebraic multiplicity

The algebraic multiplicity, denoted by $\mu_a(\lambda)$, of an eigenvalue λ of a matrix $A \in \mathbb{K}^{n\times n}$ is its multiplicity as a root of the characteristic polynomial.

Suppose the matrix A has $m \leq n$ distinct eigenvalues. The characteristic polynomial can be written as the product of m terms each corresponding to a distinct eigenvalue λ_i and raised to the power of the algebraic multiplicity, *i.e.*,

$$\det(A - \lambda\mathbb{I}) = (\lambda_1 - \lambda)^{\mu_a(\lambda_1)}(\lambda_2 - \lambda)^{\mu_a(\lambda_2)} \cdots (\lambda_m - \lambda)^{\mu_a(\lambda_m)}.$$

The eigenvalue's algebraic multiplicity is related to the dimension n as

$$1 \leq \mu_a(\lambda_i) \leq n, \qquad i = 1, \ldots, m$$
$$\sum_{i=1}^{m} \mu_a(\lambda_i) \leq n.$$

If $\mu_a(\lambda_i) = 1$, then λ_i is said to be a simple eigenvalue.

Eigenvalues and eigenvectors

Geometric multiplicity

The geometric multiplicity, denoted by $\mu_g(\lambda)$, of an eigenvalue λ of a matrix $A \in \mathbb{K}^{n\times n}$ is the dimension of the associated eigenspace V_λ, or, equivalently, the maximum number of linearly independent eigenvectors associated with λ. Since the eigenspace

$$V_\lambda = \{\mathbf{v} \in V \text{ such that } (A - \lambda\mathbb{I})\mathbf{v} = \mathbf{0}\}$$

is precisely the kernel of the matrix $(A - \lambda\mathbb{I})$, the rank-nullity theorem implies that

$$\mu_g(\lambda) = \dim_{\mathbb{K}}(V_\lambda) = n - \text{rank}(A - \lambda\mathbb{I}).$$

Relation between algebraic and geometric multiplicity

For the algebraic and geometric multiplicity of an eigenvalue λ of a matrix $A \in \mathbb{K}^{n\times n}$, the following relation holds:

$$1 \leq \mu_g(\lambda) \leq \mu_a(\lambda) \leq n.$$

Eigenvalues and eigenvectors

Additional properties of eigenvalues

Let $A \in \mathbb{K}^{n \times n}$ and λ_i $(i = 1, \ldots, n)$ eigenvalues of the matrix A.
The following properties hold:

- $\det(A) = \Pi_{i=1}^{n} \lambda_i$;
- $\mathrm{tr}(A) = \sum_{i=1}^{n} A_{ii} = \sum_{i=1}^{n} \lambda_i$;
- A is invertible if and only if all its eigenvalues are non-zero;
- if A is invertible, then the eigenvalues of A^{-1} are $\frac{1}{\lambda_i}$ and each eigenvalue's geometric multiplicity coincides. Moreover, since the characteristic polynomial of the inverse matrix is the reciprocal polynomial of the original, the eigenvalues share the same algebraic multiplicity; also, the eigenvectors of A are the same as the eigenvectors of A^{-1};
- if A is unitary, every eigenvalue has absolute value $|\lambda_i| = 1$.

Eigenvalues and eigenvectors

Similar matrices

Two matrices $A, B \in \mathbb{K}^{n\times n}$ are called similar if there exists an invertible matrix $P \in \mathbb{K}^{n\times n}$ such that

$$B = P^{-1}AP.$$

Similar matrices: properties

- Similar matrices represent the same linear transformation under two (possibly) different bases, with P being the change of basis matrix.
- A transformation $B \mapsto P^{-1}AP$ is called similarity transformation.
- Two similar matrices have the same rank, determinant, trace and characteristic polynomial.

Eigenvalues and eigenvectors

Diagonalizable matrix

Let V be a vector space over the field $\mathbb{K}$, with $\dim_{\mathbb{K}}(V) = n$, and $\Phi\colon V \to V$ a linear operator with $A \in \mathbb{K}^{n\times n}$ its associated matrix.
The matrix $A \in \mathbb{K}^{n\times n}$ is diagonalizable if it is similar to a diagonal matrix $\Lambda \in \mathbb{K}^{n\times n}$, *i.e.*, there exists an invertible matrix $P \in \mathbb{K}^{n\times n}$ such that

$$\Lambda = P^{-1}AP, \qquad \text{or, equivalently,} \qquad A = P\Lambda P^{-1},$$

where

- the i-th column of P is the i-th eigenvector of A;
- the diagonal elements of Λ are the corresponding eigenvalues, *i.e.*, $\Lambda_{ii} = \lambda_i$;
- the eigenvectors of A form a basis for V.

The number of linearly independent eigenvectors with non-zero eigenvalues is equal to the rank of the matrix A.

Remark

In an equivalent way, the linear operator $\Phi\colon V \to V$ is diagonalizable if there exists a basis of V made of the eigenvectors of the associated matrix A!

Eigenvalues and eigenvectors

Eigendecomposition

If $A \in \mathbb{K}^{n\times n}$ can be decomposed into a matrix $P \in \mathbb{K}^{n\times n}$ composed of its linearly independent eigenvectors, a diagonal matrix $\Lambda \in \mathbb{K}^{n\times n}$ with its eigenvalues along the diagonal, and the inverse of the matrix P of eigenvectors, say

$$A = P\Lambda P^{-1},$$

this procedure is called eigendecomposition and it is a similarity transformation.
The matrix P is the change of basis matrix of the similarity transformation. Essentially, the matrices A and Λ represent the same linear transformation expressed in two different bases.

Theorem

Let V be a vector space over the field $\mathbb{K}$, with $\dim_{\mathbb{K}}(V) = n$, and $\Phi\colon V \to V$ a linear operator with $A \in \mathbb{K}^{n\times n}$ its associated matrix.
The matrix $A \in \mathbb{K}^{n\times n}$ is diagonalizable if and only if:

1. $\lambda_i \in \mathbb{K} \qquad \forall i = 1, \dots, n$;
2. $\sum_{i=1}^{n} \mu_a(\lambda_i) = n$;
3. $\mu_a(\lambda_i) = \mu_g(\lambda_i) \qquad \forall i = 1, \dots, n$.

Eigenvalues and eigenvectors

Properties

- If $A \in \mathbb{K}^{n \times n}$ is symmetric $\Rightarrow$ A is diagonalizable.
- If $A \in \mathbb{K}^{n \times n}$ has n distinct eigenvalues $\lambda_i \in \mathbb{K}$ $\Rightarrow$ A is diagonalizable.
- If $A \in \mathbb{K}^{n \times n}$ can be eigendecomposed and if none of its eigenvalues are zero $\Rightarrow$ A is invertible and its inverse is given by
$$A^{-1} = P \Lambda^{-1} P^{-1}.$$
Furthermore, since Λ is a diagonal matrix, its inverse is easy to compute $\Lambda_{ii}^{-1} = \frac{1}{\lambda_i}$.
- If $A \in \mathbb{R}^{n \times n}$ is symmetric $\Rightarrow$ all eigenvalues of A are real.
- If $A \in \mathbb{R}^{n \times n}$ is symmetric $\Rightarrow$ eigenvectors corresponding to distinct eigenvalues are orthogonal.
- If $A \in \mathbb{R}^{n \times n}$ is symmetric $\Rightarrow$ the eigenvalues are real and the eigenvectors can be chosen real and orthonormal. Thus a real symmetric matrix A can be decomposed as
$$A = Q \Lambda Q^T \qquad (\text{equivalently}, \quad \Lambda = Q^T A Q),$$
where Q is an orthogonal matrix ($Q^{-1} = Q^T$) whose columns are the real orthonormal eigenvectors of A.
- If $A \in \mathbb{R}^{n \times n}$ is a symmetric matrix positive definite, positive semi-definite, negative definite, or negative semi-definite $\Rightarrow$ each eigenvalue is positive, non negative, negative, or non positive.

Quadratic forms

Quadratic form

Let V be a vector space over $\mathbb{K}$, with $\dim_{\mathbb{K}}(V) = n$, and $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ a basis for V.
An n–ary quadratic form over a field $\mathbb{K}$ is a homogeneous polynomial $q(x_1, \dots, x_n)$ of degree 2 in n variables with coefficients in $\mathbb{K}$:

$$q(\mathbf{x}) = q(x_1, \dots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j,$$

where $\mathbf{x} = (x_1, \dots, x_n) \in V$, x_i are the components of $\mathbf{x}$ with respect to the basis $\mathcal{B}$, and $a_{ij} \in \mathbb{K}$ are called coefficients of the quadratic form.
Equivalently, in matrix form

$$q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x},$$

where $A \in \mathbb{K}^{n \times n}$ is the symmetric matrix associated to $q(\mathbf{x})$.

Quadratic forms

Associated symmetric matrix

Any symmetric matrix $A \in \mathbb{K}^{n\times n}$ determines a unique quadratic form $q(x_1, \dots, x_n)$ in n variables by

$$q(x_1, \dots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j = \mathbf{x}^T A \mathbf{x}.$$

Vice versa, a quadratic form $q(x_1, \dots, x_n)$ in n variables, defined by $A \in \mathbb{K}^{n\times n}$, determines a unique matrix $B \in \mathbb{K}^{n\times n}$, say

$$b_{ij} = \frac{a_{ij} + a_{ji}}{2},$$

that is symmetric, defines the quadratic form $q(x_1, \dots, x_n)$ as the matrix A and is the unique symmetric matrix that defines $q(x_1, \dots, x_n)$.

Remark

Over the real numbers, there is a one-to-one correspondence between quadratic forms and symmetric matrices that determine them.

Quadratic forms

Example

Consider the matrix $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. The quadratic form $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ in the variables $\mathbf{x} = (x_1, x_2, x_3)$ associated to the matrix A is

$$q(\mathbf{x}) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} =$$
$$= a_{11}x_1^2 + (a_{12} + a_{21})x_1x_2 + (a_{13} + a_{31})x_1x_3 + a_{22}x_2^2 + (a_{23} + a_{32})x_2x_3 + a_{33}x_3^2.$$

So, two matrices define the same quadratic form if and only if they have the same elements on the diagonal and the same values for the sums $a_{12} + a_{21}$, $a_{13} + a_{31}$ and $a_{23} + a_{32}$. In particular, the quadratic form $q(\mathbf{x})$ is defined by a unique symmetric matrix

$$A = \begin{pmatrix} a_{11} & \frac{a_{12} + a_{21}}{2} & \frac{a_{13} + a_{31}}{2} \\ \frac{a_{12} + a_{21}}{2} & a_{22} & \frac{a_{23} + a_{32}}{2} \\ \frac{a_{13} + a_{31}}{2} & \frac{a_{23} + a_{32}}{2} & a_{33} \end{pmatrix}.$$

Quadratic forms

Equivalent quadratic forms

Let V be a vector space over $\mathbb{K}$, with $\dim_{\mathbb{K}}(V) = n$, $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ and $p(\mathbf{x}) = \mathbf{x}^T B\mathbf{x}$ be n-ary quadratic forms over the field $\mathbb{K}$ with associated symmetric matrices $A, B \in \mathbb{K}^{n\times n}$, respectively.
The quadratic forms $p(\mathbf{x})$ and $q(\mathbf{x})$ are called equivalent if there exists an invertible matrix $C \in \mathbb{K}^{n\times n}$ such that

$$p(\mathbf{x}) = q(C\mathbf{x}).$$

It follows that

$$\mathbf{x}^T B\mathbf{x} = p(\mathbf{x}) = q(C\mathbf{x}) = (C\mathbf{x})^T A(C\mathbf{x}) = \mathbf{x}^T C^T AC\mathbf{x},$$

i.e., the symmetric matrices A of $q(\mathbf{x})$ and B of $p(\mathbf{x})$ are related as

$$B = C^T AC.$$

Quadratic forms

Real quadratic forms

Let V be a real vector space, with $\dim_{\mathbb{R}}(V) = n$, and $q(\mathbf{x})$ be an n–ary quadratic form over the field $\mathbb{R}$. In this case, $q(\mathbf{x})$ is called real quadratic form.

Equivalence of real quadratic forms

Every n-ary real quadratic form $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, with associated symmetric matrix $A \in \mathbb{R}^{n\times n}$, is equivalent to a diagonal form

$$p(\mathbf{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2 = \mathbf{x}^T \Lambda \mathbf{x},$$

where Λ is a diagonal matrix whose entries are the real eigenvalues λ_i of A.
In fact, since A is a real symmetric matrix, A can be eigendecomposed and we have

$$\Lambda = Q^T A Q$$

where Q is an orthogonal matrix whose columns are the orthonormal eigenvectors of A. It is

$$p(\mathbf{x}) = \mathbf{x}^T \Lambda \mathbf{x} = \mathbf{x}^T Q^T A Q\mathbf{x} = (Q\mathbf{x})^T A(Q\mathbf{x}) = q(Q\mathbf{x}).$$

The quadratic form $p(\mathbf{x})$ is represented in diagonal form with respect to the orthonormal basis $\mathcal{B}'$ of eigenvectors of A.
Classification of all real quadratic forms up to equivalence can be reduced to the case of diagonal forms!

Definiteness of quadratic forms

Let V be a real vector space, with $\dim_{\mathbb{R}}(V) = n$, and $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ be an n–ary real quadratic form with associated symmetric matrix $A \in \mathbb{R}^{n\times n}$.

A definite quadratic form is a real quadratic form over the vector space V that has the same sign for every non-zero vector of V.

In general, a real quadratic form $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ (or, equivalently, the associated symmetrix matrix $A \in \mathbb{R}^{n\times n}$) is:

- positive definite if $\mathbf{x}^T A\mathbf{x} > 0 \quad \forall \mathbf{x} \in V, \ \mathbf{x} \neq \mathbf{0}$;
- positive semi-definite if $\mathbf{x}^T A\mathbf{x} \geq 0 \quad \forall \mathbf{x} \in V$;
- negative definite if $\mathbf{x}^T A\mathbf{x} < 0 \quad \forall \mathbf{x} \in V, \ \mathbf{x} \neq \mathbf{0}$;
- negative semi-definite if $\mathbf{x}^T A\mathbf{x} \leq 0 \quad \forall \mathbf{x} \in V$;
- indefinite if $\exists \mathbf{x}, \mathbf{y} \in V \ : \ \mathbf{x}^T A\mathbf{x} > 0, \ \mathbf{y}^T A\mathbf{y} < 0$.

Quadratic forms

Leading principal minor

The k-th leading principal minor of a matrix $A \in \mathbb{K}^{n \times n}$ is the determinant of the $k \times k$ sub-matrix obtained from A by deleting the last $n-k$ rows and the last $n-k$ columns.

Principal minor

The k-th principal minor of a matrix $A \in \mathbb{K}^{n \times n}$ is the determinant of the $k \times k$ sub-matrix obtained from A by deleting $n-k$ rows and columns with the same indices.

Remark

For an $n \times n$ square matrix there are n leading principal minors.

Remark

If M is a leading principal minor of a matrix $A \in \mathbb{K}^{n \times n} \Rightarrow M$ is a principal minor of a matrix $A \in \mathbb{K}^{n \times n}$. The converse is not in general true!

Example: leading principal minors

The matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

has 1-th, 2-th and 3-th leading principal minors

$$M_1 = \det\begin{pmatrix} a_{11} \end{pmatrix} = a_{11}, \qquad M_2 = \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad M_3 = \det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

respectively.

Quadratic forms

Example: principal minors

The matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

has the principal minors

$$\text{1-th} \longrightarrow M_1 = \det \begin{pmatrix} a_{11} \end{pmatrix} = a_{11}, \qquad M_2 = \det \begin{pmatrix} a_{22} \end{pmatrix} = a_{22}, \qquad M_3 = \det \begin{pmatrix} a_{33} \end{pmatrix} = a_{33},$$

$$\text{2-th} \longrightarrow M_4 = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad M_5 = \det \begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix}, \qquad M_6 = \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix},$$

$$\text{3-th} \longrightarrow M_7 = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Quadratic forms

Definiteness of quadratic forms - Sylvester's criterion

Let V be a vector space over $\mathbb{R}$, with $\dim_{\mathbb{R}}(V) = n$, and $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ be an n–ary real quadratic form with associated symmetric matrix $A \in \mathbb{R}^{n\times n}$. The definiteness of the quadratic form (equivalently, the sign of the matrix A) is characterized by the sign of leading principal minors and the principal minors of A. Then:

- A is positive definite if and only if its all leading principal minors are positive;
- A is positive semi–definite if and only if its all principal minors are non–negative;
- A is negative definite if and only if its all leading principal minors of odd order are negative, and all leading principal minors of even order are positive;
- A is negative semi–definite if and only if its all principal minors of odd order are non-positive, and all principal minors of even order are non-negative;
- A is indefinite in all the remaining cases.

Quadratic forms

Gauss elimination for definiteness of quadratic forms

Let V be a vector space over $\mathbb{R}$, with $\dim_{\mathbb{R}}(V) = n$, and $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ be an n–ary real quadratic form with associated symmetric matrix $A \in \mathbb{R}^{n\times n}$.
By means of Gauss elimination, the associated symmetric matrix A can be reduced to an upper triangular form, by preserving the sign of its determinant during the pivoting process. The product of the elements of the diagonal (the pivots) is the determinant and, since the k-th leading principal minor of a triangular matrix is the product of its diagonal elements up to row k, Sylvester's criterion for definetess of quadratic forms (or sign of the matrix) is equivalent to checking the sign of the diagonal elements. This condition can be checked each time a new row k of the triangular matrix is obtained during the Gauss elimination. Then:

- A is positive definite if and only if all pivots are positive;
- A is positive semi–definite if and only if all pivots are non–negative;
- A is negative definite if and only if all pivots are negative;
- A is negative semi–definite if and only if all pivots are non–positive;
- A is indefinite if and only if there are positive and negative pivots.

Quadratic forms

Definiteness of quadratic forms - Characterization in terms of eigenvalues

Let V be a vector space over $\mathbb{R}$, with $\dim_{\mathbb{R}}(V) = n$, and $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ be an n–ary real quadratic form with associated symmetric matrix $A \in \mathbb{R}^{n\times n}$.
Since A is a real symmetric matrix, all eigenvalues of A are real, and their sign characterize the definiteness of the quadratic form (equivalently, the sign of the matrix A).
Then:

- A is positive definite if and only if its all eigenvalues are positive;
- A is positive semi–definite if and only if its all eigenvalues are non–negative;
- A is negative definite if and only if its all eigenvalues are negative;
- A is negative semi–definite if and only if its all eigenvalues are non–positive;
- A is indefinite if and only if it admits both positive and negative eigenvalues.

Quadratic forms

Descartes' rule of signs

The sign of eigenvalues can be checked using Descartes' rule of alternating signs when the characteristic polynomial of a real symmetric matrix $M \in \mathbb{R}^{n\times n}$ is available. In such a case, all eigenvalues will be real. As well known, the characteristic polynomial of M is a polynomial of degree n that, ordered by descending variable exponent, can be written as

$$p(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0, \qquad a_i \in \mathbb{R}.$$

Descartes' rule:

- if $a_0 \neq 0$, there are non-zero roots, and the number of positive roots (a root of multiplicity k is counted as k roots) of the polynomial is equal to the number of sign changes between consecutive (non-zero) coefficients; for negative roots, we have

$$\text{number of negative roots} = n - (\text{number of positive roots});$$

- if $a_0 = 0$, collect the polynomial by a common factor; then, there are zero roots counted with their multiplicity, the number of positive roots can be computed for the obtained polynomial after collection as in the previous case; for negative roots, we have

$$\text{number of negative roots} = n - (\text{number of zero roots}) - (\text{number of positive roots}).$$

Quadratic forms

Example

Consider the quadratic form $q(x_1, x_2, x_3, x_4, x_5)$ associated to the real symmetric matrix

$$A = \begin{pmatrix} 4.85559 & 4.15077 & 2.86879 & 2.71129 & 2.42802 \\ 4.15077 & 4.96922, & 3.58577 & 3.05249 & 3.04842 \\ 2.86879 & 3.58577 & 4.48631 & 3.41923 & 3.8877 \\ 2.71129 & 3.05249 & 3.41923 & 4.72888 & 3.45937 \\ 2.42802 & 3.04842 & 3.8877 & 3.45937 & 3.93873 \end{pmatrix}$$

The characteristic polynomial of A is

$$p(\lambda) = -\lambda^5 + 22.9787\lambda^4 - 101.928\lambda^3 + 149.154\lambda^2 - 77.3386\lambda + 11.691.$$

Since in $p(\lambda)$ there are 5 sign changes, the matrix A admits 5 positive eigenvalues $\Longrightarrow$ A, and so the quadratic form $q(x_1, x_2, x_3, x_4, x_5)$, is positive definite.
In fact, the eigenvalues of A are

$$\lambda_1 = 0.258844, \qquad \lambda_2 = 0.612167, \qquad \lambda_3 = 1.35774, \qquad \lambda_4 = 3.07432, \qquad \lambda_5 = 17.6756.$$

Quadratic forms

Example

Consider the quadratic form $q(x_1, x_2, x_3, x_4, x_5)$ associated to the real symmetric matrix

$$A = \begin{pmatrix} 3.67822 & 4.32808 & 3.69609 & 3.1335 & 3.21014 \\ 4.32808 & 2.10251 & 3.48111 & 3.57061 & 3.5567 \\ 3.69609 & 3.48111 & 1.53572 & 3.49061 & 3.13288 \\ 3.1335 & 3.57061 & 3.49061 & 3.27227 & 2.8812 \\ 3.21014 & 3.5567 & 3.13288 & 2.8812 & 3.31134 \end{pmatrix}$$

The characteristic polynomial of A is

$$p(\lambda) = -\lambda^5 + 13.9001\lambda^4 + 44.72\lambda^3 + 10.3076\lambda^2 - 30.6237\lambda + 7.1184.$$

Since in $p(\lambda)$ there are 3 sign changes, the matrix A admits 3 positive eigenvalues and 2 negative eigenvalues $\Longrightarrow$ A, and so the quadratic form $q(x_1, x_2, x_3, x_4, x_5)$, is indefinite.
In fact, the eigenvalues of A are

$$\lambda_1 = 0.317928, \qquad \lambda_2 = 0.444959, \qquad \lambda_3 = -1.65589, \qquad \lambda_4 = -1.82825, \qquad \lambda_5 = 16.6213.$$

Quadratic forms

Properties

- Let $M \in \mathbb{R}^{n\times n}$ be symmetric. M is positive definite $\Leftrightarrow$ M can be decomposed as $M = A^T A$, with $A \in \mathbb{R}^{n\times n}$ invertible.
- Let $M \in \mathbb{R}^{n\times n}$ be symmetric. M is positive definite $\Leftrightarrow$ M can be decomposed in a unique way as $M = LL^T$ (Cholesky decomposition), where $L \in \mathbb{R}^{n\times n}$ is a lower triangular matrix with positive diagonal entries. If M is only positive semi–definite, then the Cholesky decomposition of the form $M = LL^T$ still holds where the diagonal entries of L are allowed to be zero, and this decomposition needs not be unique.
- $M \in \mathbb{R}^{n\times n}$ is negative (semi) definite $\Leftrightarrow$ $-M$ is positive (semi) definite.
- Every positive definite matrix $M \in \mathbb{R}^{n\times n}$ is invertible and its inverse M^{-1} is also positive definite.
- If $M \in \mathbb{R}^{n\times n}$ is positive definite $\Rightarrow$ $\text{rank}(M) = n$.
- If $M \in \mathbb{R}^{n\times n}$ is positive definite and $r > 0$ is a real number $\Rightarrow$ rM is positive definite.
- If $M, N \in \mathbb{R}^{n\times n}$ are positive (semi) definite $\Rightarrow$ $M + N$ is positive (semi) definite.
- If $M, N \in \mathbb{R}^{n\times n}$ are positive definite and $MN = NM$ $\Rightarrow$ MN is positive definite.
- If $M, N \in \mathbb{R}^{n\times n}$ are positive definite $\Rightarrow$ MNM and NMN are positive definite.
- If $M \in \mathbb{R}^{n\times n}$ is positive semi–definite $\Rightarrow$ $A^T MA$ is positive semi–definite for any (possibly rectangular) matrix A. If M is positive definite and A has full rank, then $A^T MA$ is positive definite.