Linear Algebra: Basic Tools

Computational Linear Algebra for Large Scale Problems

Stefano Berrone, Francesco Della Santa and Ulderico Fugacci

stefano.berrone@polito.it

October 18, 2023

Basics on Linear Algebra

Outline:

Matrices

Definition

Given $m, n \in \mathbb{N}$, a **matrix** $A = (a_{i,j})$ of size $m \times n$ is a rectangular array of mn elements arranged in m rows and n columns.

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix}$$

We denote as

- $(a_{i,j}) \in \mathbb{R}^{m \times n}$ is a shortening of $(a_{i,j})_{1 \leq i \leq m,\ 1 \leq j \leq m,}$;
- $a_{i,j}$, with $1 \leq i \leq m$, $1 \leq j \leq n$, an **entry** of A;
- $M_{m,n}(\mathbb{R})$, or $\mathbb{R}^{m \times n}$, the set of matrices of size $m \times n$ with entries in $\mathbb{R}$ (If $m = n$, we denote it as $M_n(\mathbb{R})$ or $\mathbb{R}^{n \times n}$).

Matrices

Some Basic Notations:

Given a matrix $A = (a_{i,j}) \in \mathbb{R}^{m\times n}$, we define:

- the **transpose** of A, the matrix $A^\top := (b_{i,j}) \in \mathbb{R}^{n\times m}$ where $b_{i,j} = a_{j,i}$ (i.e., the rows of $A^\top$ are the columns of A):
$$\text{e.g., } A = \begin{bmatrix} 2 & 1 & 0 \\ 3 & 7 & 4 \end{bmatrix} \Rightarrow A^\top = \begin{bmatrix} 2 & 3 \\ 1 & 7 \\ 0 & 4 \end{bmatrix};$$
- **row** i, the row matrix/vector $A_{i,\cdot} := [a_{i,1}, \dots, a_{i,n}] =: \boldsymbol{a}_{i,\cdot} \in \mathbb{R}^{1\times n}$;
- **column** j, the column matrix/vector $A_{\cdot,j} := \begin{bmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{bmatrix} =: \boldsymbol{a}_{\cdot,j} \in \mathbb{R}^{m\times 1}$;
- the **main diagonal**, the collection of the entries of the kind $a_{i,i}$, for $i = 1, \dots, \min(m, n)$.

Sometimes we also use the notation (Matlab-style) $A_{i,:}$ and $A_{:,j}$ instead of $A_{i,\cdot}$ and $A_{\cdot,j}$, respectively.

Matrices

Some Basic Notations:

Let $A = (a_{i,j})$ be a **square** matrix, i.e., $A \in M_n(\mathbb{R})$. A is called:

- **upper/lower triangular**, if $a_{i,j} = 0$ for each $i > j / i < j$,

  e.g., $\begin{bmatrix} 4 & 0 & 1 \\ 0 & 5 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ is UT and $\begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \\ 1 & 2 & 1 \end{bmatrix}$ is LT;
- **diagonal**, if it is upper and lower triangular (i.e., $a_{i,j} = 0$ for $i \neq j$),

  e.g., $\begin{bmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 5 \end{bmatrix}$;
- **symmetric**, if $A = A^\top$;
- **identity/unit matrix** I_n **of size** n if it is the diagonal square with $a_{i,i} = 1$ for each $i = 1, \dots, n$;
- **null/zero matrix** 0_n **of size** n if it is the matrix in $\mathbb{R}^{n \times n}$ with all entries null (we can extend the definition to $0_{m,n} \in \mathbb{R}^{m \times n}$).

Matrices - Operations

Binary Operations

Let us endow the set $M_{m,n}(\mathbb{R})$ with some algebraic operations:

- $A + B$, **sum of matrices**;
- λA, **product of a matrix with a scalar**;
- AB, **product of matrices**.

Matrices - Operations

Sum of Matrices

Given $A = (a_{i,j})$ and $B = (b_{i,j})$ in $\mathbb{R}^{m\times n}$, $A + B$ is the matrix in $\mathbb{R}^{m\times n}$ defined as:

$$A + B := \begin{bmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & \dots & a_{1,n} + b_{1,n} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} & \dots & a_{2,n} + b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} + b_{m,1} & a_{m,2} + b_{m,2} & \dots & a_{m,n} + b_{m,n} \end{bmatrix}.$$

Main Properties:

- **Associativity.** If $A, B, C \in \mathbb{R}^{m\times n}$, then $(A + B) + C = A + (B + C)$.

- **Commutativity.** If $A, B \in \mathbb{R}^{m\times n}$, then $A + B = B + A$.

- **Identity Existence.** For each $A \in \mathbb{R}^{m\times n}$, $A + 0_{m,n} = A$.

- **Inverse Existence.** For each $A \in \mathbb{R}^{m\times n}$, denoting as $-A$ the matrix $(-a_{i,j}) \in \mathbb{R}^{m\times n}$, then $A + (-A) = 0_{m,n}$.

In a nutshell, $(\mathbb{R}^{m\times n}, +)$ is an **Abelian group** (see slide 52).

Matrices - Operations

Product of a Matrix with a Scalar

Given $A = (a_{i,j}) \in \mathbb{R}^{m\times n}$ and $\lambda \in \mathbb{R}$, λA is the matrix in $\mathbb{R}^{m\times n}$ defined as:

$$\lambda A := \begin{bmatrix} \lambda a_{1,1} & \lambda a_{1,2} & \dots & \lambda a_{1,n} \\ \lambda a_{2,1} & \lambda a_{2,2} & \dots & \lambda a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{m,1} & \lambda a_{m,2} & \dots & \lambda a_{m,n} \end{bmatrix}.$$

Main Properties:

- **Associativity.** If $A \in \mathbb{R}^{m\times n}$ and $\lambda, \mu \in \mathbb{R}$, then $(\lambda\mu)A = \lambda(\mu A)$.

- **Distributivity.** If $A, B \in \mathbb{R}^{m\times n}$ and $\lambda, \mu \in \mathbb{R}$, then $\lambda(A + B) = \lambda A + \lambda B$ and $(\lambda + \mu)A = \lambda A + \mu A$.

- **One-neutrality property.** If $A \in \mathbb{R}^{m\times n}$ and $\lambda \in \mathbb{R}$ are such that $\lambda A = A$, then $\lambda = 1$.

- **Zero-product property.** If $A \in \mathbb{R}^{m\times n}$ and $\lambda \in \mathbb{R}$ are such that $\lambda A = 0_{m,n}$, then $\lambda = 0$ or $A = 0_{m,n}$.

In a nutshell, $\mathbb{R}^{m\times n}$ is a **vector space over** $\mathbb{R}$ (see slide 53).

Matrices - Operations

Product of Matrices

Given $A = (a_{i,j}) \in \mathbb{R}^{m\times n}$ and $B = (b_{i,j}) \in \mathbb{R}^{n\times p}$, $C = AB$ is the matrix in $\mathbb{R}^{m\times p}$ defined as:

$$AB := (c_{i,j}),$$

where

$$c_{i,j} = a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \cdots + a_{i,n}b_{n,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j},$$

i.e., $c_{i,j}$ is obtained by multiplying term-by-term the entries of row $A_{i,\cdot}$ and column $B_{\cdot,j}$ and summing these n products (i.e. the scalar product of $\boldsymbol{a}_{i,\cdot} \cdot \boldsymbol{b}_{\cdot,j}$, see slide 73).

Matrices - Operations

Example:

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 0 & 4 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & -2 & 1 \\ 0 & -1 & 3 & 0 \\ -3 & 1 & 2 & 0 \end{bmatrix}$$

$$AB = \left(\sum_{k=1}^{n} a_{i,k} b_{k,j} \right) = \begin{bmatrix} -7 & 4 & -1 & 2 \\ -6 & -2 & 16 & 0 \end{bmatrix}$$

Matrices - Operations

Main Properties:

- **Associativity.** If $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$, $C \in \mathbb{R}^{p\times q}$, then
$$(AB)C = A(BC) \in \mathbb{R}^{m\times q}\,;$$
- **Distributivity (left).** If $A \in \mathbb{R}^{m\times n}$, $B, C \in \mathbb{R}^{n\times p}$, then
$$A(B + C) = AB + AC \in \mathbb{R}^{m\times p}\,.$$
- **Distributivity (right).** If $A, B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{n\times p}$, then
$$(A + B)C = AC + BC \in \mathbb{R}^{m\times p}\,.$$

Matrices - Operations

Some Properties Do NOT Hold in General.

- **Commutativity**: $A, B \in \mathbb{R}^{n\times n} \not\Rightarrow AB = BA$. Obvious for non-square matrices;
- **Zero-product property:** $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$ s.t. $AB = 0_{m,p} \not\Rightarrow A = 0_{m,n}$ or $B = 0_{n,p}$.

Examples

1. $\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} -2 & -4 \\ -1 & 2 \end{bmatrix}$ and $\begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$;
2. $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \Rightarrow A^2 = A \cdot A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

   A matrix $A \in \mathbb{R}^{n\times n}$ for which there exists $k \in \mathbb{N}$ such that $A^k = 0_n$ is called **nilpotent**.

   Nilpotent matrices are a subset of singular matrices: all nilpotent matrices are singular, but, not all singular matrices are nilpotent. A matrix is nilpotent if and only if zero is its only eigenvalue, i.e. its caracteristic polynomial is A^n, and $A^n = 0$ by the Cayley-Hamilton Theorem (see slides...).

Matrices - Operations

Binary Operations and Transposition

Given two matrices A, B of appropriate sizes and $\lambda \in \mathbb{R}$, we have that:

- $(A+B)^\top = A^\top + B^\top$;
- $(\lambda A)^\top = \lambda A^\top$;
- $(AB)^\top = B^\top A^\top$;
- $(A^\top)^\top = A$.
- $(A^{-1})^\top = (A^\top)^{-1}$ (see following slides).

Matrices - Operations

Focusing on square matrices in $\mathbb{R}^{n \times n}$, we have the following property: **Identity Existence:** for each $A \in \mathbb{R}^{n \times n}$,

$$I_n A = A I_n = A.$$

Invertibility of a Square Matrix:

A matrix $A \in \mathbb{R}^{n \times n}$ is defined **invertible** if and only if there exists a matrix $B \in \mathbb{R}^{n \times n}$ such that

$$AB = BA = I_n.$$

B is called **inverse** of A and is denoted by A^{-1}.

Matrices - Operations

Inverse of a Matrix and Properties:

- **Not** all the square matrices are invertible;
- If it exists, the inverse of a matrix $A \in M_n(\mathbb{R})$ is **unique**;
- Let $A, B \in \mathbb{R}^{n \times n}$ be two invertible matrices. Then, AB is invertible and

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

So, the subset of the invertible matrices of $\mathbb{R}^{n \times n}$, endowed with matrix multiplication, is a group called **general linear group** and it is denoted as $GL_n(\mathbb{R})$.

Examples

$$\begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} \in GL_2(\mathbb{R}), \qquad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \notin GL_2(\mathbb{R}).$$

Matrices - Determinant

Definition

Given $J_n = \{j_1, j_2, \dots, j_n\}$, let us define the set $S_n(J_n)$ of the **permutations** of J_n as the set of the bijective functions

$$\begin{aligned} \sigma : J_n &\to J_n \ . \\ j &\mapsto \sigma(j) \end{aligned}$$

From now on we will denote $S_n(J_n)$ as S_n for simplicity.

A permutation $\sigma \in S_n$ can be represented through the following table:

$$\sigma = \begin{pmatrix} j_1 & j_2 & \dots & j_n \\ \sigma(j_1) & \sigma(j_2) & \dots & \sigma(j_n) \end{pmatrix} .$$

The number of elements of S_n (i.e. the number of possible permutations) is $n! := n(n-1)\cdots 2$.

Matrices - Determinant

S_n endowed with the operation of composition of functions "$\circ$" is a (non-Abelian) **group**.
For simplicity, we denote the composition of two permutations $\rho, \sigma \in S_n$ as $\rho\sigma$ instead of $\rho \circ \sigma$. **Remember:** $\rho\sigma(j) = \rho\left(\sigma(j)\right)$.
Main Properties:

- **Associativity:** for each $\rho, \sigma, \tau \in S_n$, it holds $(\rho\sigma)\tau = \rho(\sigma\tau)$;
- **Identity Existence:** The identity function $\iota : J_n \to J_n$ such that $\iota(j) = j$ for each $j \in J_n$, is the identity element of the group $(S_n, \circ)$. Then

$$\iota\sigma = \sigma = \sigma\iota\,, \ \forall\sigma \in S_n\,.$$

- **Inverse Existence:** for each $\sigma \in S_n$, since σ is bijective it is invertible and $\sigma^{-1} \in S_n$.
- **Inverse of a Composition:** for each $\sigma, \rho \in S_n$, $(\sigma\rho)^{-1} = \rho^{-1}\sigma^{-1}$.

Matrices - Determinant

2-Cycle Decomposition

A permutation $\sigma \in S_n$ is called a **2-cycle** if for some $h, k \in \{1, \dots, n\}$ we have that:

$$\sigma(j) = \begin{cases} j_h & \text{if } j = j_k, \\ j_k & \text{if } j = j_h, \\ j & \text{otherwise.} \end{cases} \quad \text{i.e. } \sigma = \begin{pmatrix} j_1 & \cdots & j_h & \cdots & j_k & \cdots & j_n \\ j_1 & \cdots & j_k & \cdots & j_h & \cdots & j_n \end{pmatrix}.$$

In such a case σ is denoted by (h, k).

Proposition

Each permutation $\sigma \in S_n$ admits a decomposition into a product of disjoint 2-cycles.

N.B. Such a decomposition is, in general, not unique.

E.g., $\sigma := \begin{pmatrix} j_1 & j_2 & j_3 \\ j_3 & j_1 & j_2 \end{pmatrix} = (1,2)(1,3) = (2,3)(1,2).$

Matrices - Determinant

Sign of a permutation:
Although the decomposition of a permutation $\sigma \in S_n$ is not unique, the **parity** of the number of 2-cycles in all decompositions of σ is the same: i.e. all the decompositions have either an odd or an even number of 2-cycles.

Definition

Given a permutation $\sigma \in S_n$, the **sign of** σ is defined as

$$sgn(\sigma) := (-1)^m$$

where m is the number of 2-cycles in a decomposition of σ.

- The identity permutation is an even permutation;
- the composition of two even permutations is even;
- the composition of two odd permutations is even;
- the composition of an odd and an even permutation is odd;
- the inverse of every even permutation is even;
- the inverse of every odd permutation is odd.

Matrices - Determinant

Definition (Leibniz formula)

Given a square matrix $A = (a_{i,j}) \in \mathbb{R}^{n\times n}$, the **determinant of** A is defined as

$$det(A) := \sum_{\sigma\in S_n} sgn(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}.$$

Sometimes, in the following, $det(A)$ will be also denoted by $|A|$. The Leibniz formula involves n! summands, each of which is a product of n entries of the matrix, The computational cost is $n! \times n$ products and $n!$ sums.

Examples

$$det\left(\begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}\right) := \left|\begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}\right| = 1\,, \qquad det\left(\begin{bmatrix} 2 & 1 & 3 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{bmatrix}\right) = 0\,.$$

Matrices - Determinant

Some Easy Observations:

Let $A = (a_{i,j})$ be a matrix in $\mathbb{R}^{n\times n}$.

- $det(A^T) = det(A)$.
- If A is upper/lower triangular, then $det(A)$ is the product of the diagonal elements:
$$det(A) = \prod_{i=1}^{n} a_{i,i} = a_{1,1} \cdot a_{2,2}, \cdots a_{n,n}$$
- If A has a row/column null, then $det(A) = 0$.
- If A has two proportional rows/columns, then $det(A) = 0$.
In general: $det(A) = 0$ if rows/columns are **not** lin. ind. (see slide 59);
- If A' is obtained from A swapping two rows/columns, then $det(A') = -det(A)$.
- If A' is obtained from A multiplying a row by $\lambda \in \mathbb{R}$, then $det(A') = \lambda det(A)$.
- If A' is obtained from A adding to a row another row multiplied by $\lambda \in \mathbb{R}$, $\lambda \neq 0$, then $det(A') = det(A)$ (same result with columns).

Matrices - Determinant

Proposition

Let A be a matrix in $\mathbb{R}^{n\times n}$. The following facts are equivalent.

- A is invertible (i.e., $A \in GL_n(\mathbb{R})$).
- $det(A) \neq 0$.
- the rows/columns of A are linearly independent (see slide 59), i.e.

$$\gamma_1 A_{1,\cdot} + \cdots + \gamma_n A_{n,\cdot} = [0,\dots,0] \;\Leftrightarrow\; \gamma_i = 0\,,\ \forall\ i = 1,\dots,n$$

  or, equivalently,

$$\gamma_1 A_{\cdot,1} + \cdots + \gamma_n A_{\cdot,n} = [0,\dots,0]^\top \;\Leftrightarrow\; \gamma_i = 0\,,\ \forall\ i = 1,\dots,n$$

Binet's Theorem

Let A, B be two matrices in $\mathbb{R}^{n\times n}$. Then,

$$det(AB) = det(A)det(B).$$

Matrices - Determinant

Definition

Let $A = (a_{i,j})$ be a matrix in $\mathbb{R}^{n\times n}$. We define the (i,j)-**cofactor** of A as

$$\alpha^{(i,j)} := (-1)^{i+j} det \left(A^{(i,j)} \right) ,$$

where $A^{(i,j)}$ is the matrix obtained from A removing row $A_{i,\cdot}$ and column $A_{\cdot,j}$ and is called (i,j) **minor**.

First Laplace's Theorem

For any fixed $i = 1, \ldots, n$, we have that:

$$det(A) = \sum_{j=1}^{n} a_{i,j}\, \alpha^{(i,j)} = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j}\, det \left(A^{(i,j)} \right) .$$

The above formula is called **Laplace expansion** of $det(A)$ with respect to row $A_{i,\cdot}$. Analogously, $det(A)$ can be expressed by the Laplace expansion w.r.t. column $A_{\cdot,j}$, i.e. $det(A) = \sum_{i=1}^{n} a_{i,j}\alpha^{(i,j)}$.

Matrices - Determinant

Example (determinant of $A \in \mathbb{R}^{2\times 2}$ and 1^{st} Lap. Th.):

$$det(A) = \left| \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \right| =$$
$$= (-1)^{(1+1)} a_{1,1} det([a_{2,2}]) + (-1)^{(1+2)} a_{1,2} det([a_{2,1}]) =$$
$$= a_{1,1}a_{2,2} - a_{1,2}a_{2,1} \,.$$

Example (determinant of $A \in \mathbb{R}^{3\times 3}$ and 1^{st} Lap. Th.):

$$det(A) = a_{1,1} \left| \begin{bmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{bmatrix} \right| - a_{1,2} \left| \begin{bmatrix} a_{2,1} & a_{1,3} \\ a_{3,1} & a_{3,3} \end{bmatrix} \right| + a_{1,3} \left| \begin{bmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,1} \end{bmatrix} \right|$$
$$\left(1. \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} ; 2. \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} ; 3. \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \right) .$$

Example

$$
\left|\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}\right| =
$$

$$
\begin{aligned}
&= a_{1,1}\left|\begin{bmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{bmatrix}\right| - a_{1,2}\left|\begin{bmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{bmatrix}\right| + a_{1,3}\left|\begin{bmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{bmatrix}\right| \\
&= a_{1,1}\left(a_{2,2}a_{3,3} - a_{3,2}a_{2,3}\right) - a_{1,2}\left(a_{2,1}a_{3,3} - a_{3,1}a_{2,3}\right) \\
&\quad + a_{1,3}\left(a_{2,1}a_{3,2} - a_{3,1}a_{2,2}\right) \\
&= a_{1,1}a_{2,2}a_{3,3} - a_{1,1}a_{3,2}a_{2,3} - a_{1,2}a_{2,1}a_{3,3} + a_{1,2}a_{3,1}a_{2,3} \\
&\quad + a_{1,3}a_{2,1}a_{3,2} - a_{1,3}a_{3,1}a_{2,2}
\end{aligned}
$$

In red the elements that in an upper triangular matrix are zero.
In blue the elements that in a lower triangular matrix are zero.

Matrices - Determinant

Definition

Let $A = (a_{i,j})$ be a matrix in $\mathbb{R}^{n\times n}$. We define the **cofactor matrix** of A as

$$cof(A) := \left(\alpha^{(i,j)}\right) \in \mathbb{R}^{n\times n}.$$

Second Laplace's Theorem

For any $1 \leq r, s \leq n$, with $r \neq s$, we have that:

$$0 = \sum_{j=1}^{n} a_{r,j}\alpha^{(s,j)}.$$

I.e. if we multipy the element of a row by the cofactors of a different row and summing the products we get zero.
As a straight consequence of the above two Laplace's theorems, we obtain that:

$$A \cdot cof(A)^{\top} = det(A)\, I_n$$

the first Laplace's Theorem defnes the non-vanishing elements, whereas the second Laplace's Theorem defines the zero elements.
So, if $A \in GL_n(\mathbb{R})$, it holds

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

Matrices - Determinant

Exercise:
Compute the determinant and the inverse of the following matrix.

$$A := \begin{bmatrix} 1 & 3 & 2 \\ 0 & 2 & 3 \\ 2 & 0 & 1 \end{bmatrix}$$

$$det(A) = 12$$

$$cof(A) = \begin{bmatrix} 2 & 6 & -4 \\ -3 & -3 & 6 \\ 5 & -3 & 2 \end{bmatrix}$$

$$A \cdot cof(A)^{\top} = 12 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Matrices - Rank

Definition

Let A be a matrix in $\mathbb{R}^{m\times n}$. A **minor of order** t of A is the determinant of a submatrix[a] $A' \in \mathbb{R}^{t\times t}$of A.
Then, the **rank** $rk(A)$ of A is defined as the maximal order of a non-null minor of A.

[a] a submatrix is obtained removing an arbitray number of rows/columns from A.

Main Properties:

- If $A \in \mathbb{R}^{m\times n}$, then $rk(A) \leq \min\{m, n\}$.
- If $A \in \mathbb{R}^{n\times n}$, then
$$rk(A) = n \Leftrightarrow det(A) \neq 0.$$
- If $A \in \mathbb{R}^{m\times n}$, $B \in GL_m(\mathbb{R})$, $C \in GL_n(\mathbb{R})$, then
$$rk(A) = rk(BAC).$$

Matrices - Rank

Kronecker's Theorem

Let A be a matrix in $\mathbb{R}^{m\times n}$. If there exists a non-null minor μ_t of order t such that any other minor μ_{t+1} of order $t+1$ given by a submatrix "containing" the matrix of μ_t is null, then $rk(A) = t$.

Exercise:

Compute the rank of the following matrix.

$$A := \begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 1 & 1 & -1 \\ 1 & 1 & -1 & 0 \\ 1 & 0 & -2 & 1 \end{bmatrix}$$

Proposition: let t be the maximum number of linear independent columns or rows in A (see slide 59), then $rk(A) = t$.

Linear Systems

Definition

A **linear system** of m equations and n unknowns with coefficients in $\mathbb{R}$ is a collection of linear (i.e. degree 1) equations of this form:

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2 \\ \quad \vdots \qquad\quad \vdots \qquad\qquad\qquad \vdots \qquad\quad \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n = b_m \end{cases},$$

where:

- x_j, for $i = 1, \ldots, n$, are the **unknowns**;
- $a_{i,j} \in \mathbb{R}$, for $i = 1, \ldots, m$ and $j = 1, \ldots, n$, are the **coefficients** of the system;
- $b_i \in \mathbb{R}$, for $i = 1, \ldots, m$, are the **right hand side** terms.

Linear Systems

Matrix Representation

A linear system can be expressed in the following matricial form:

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

where:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

with $A \in \mathbb{R}^{m\times n}$, $\boldsymbol{x} \in \mathbb{R}^{n\times 1}$ (i.e. $\mathbb{R}^n$), and $\boldsymbol{b} \in \mathbb{R}^{m\times 1}$ (i.e. $\mathbb{R}^m$).

Solution of a Linear System

Solving the above system means retrieving the set of all the $\boldsymbol{x}$ in $\mathbb{R}^n$ satisfying equation $A\boldsymbol{x} = \boldsymbol{b}$.

Linear Systems - Existence of Solutions

Rouché-Capelli's Theorem

The linear system $A\boldsymbol{x} = \boldsymbol{b}$ admits solutions if and only if

$$rk(A) = rk(A|\boldsymbol{b}),$$

where $A|\boldsymbol{b}$ is the **augmented matrix** obtained by appending the column vector $\boldsymbol{b}$ to the matrix A.
If $rk(A) = rk(A|\boldsymbol{b})$, the system admits

$$\infty^{n-rk(A)}$$

solutions. More properly, the set of solutions forms an affine subspace of $\mathbb{R}^n$ of dimension $n - rk(A)$.

Linear Systems - Existence of Solutions

Homogeneous Case:
The linear system $A\boldsymbol{x} = \boldsymbol{b}$ admits always solutions if $\boldsymbol{b}$ is null, i.e. if $\boldsymbol{b} = \boldsymbol{0}_m := 0_{m,1} \in \mathbb{R}^m$; indeed, $rk(A) = rk(A|\boldsymbol{0}_m)$.
The **trivial** solution $\boldsymbol{x} = \boldsymbol{0}_n$ always satisfies the system $A\boldsymbol{x} = \boldsymbol{0}_m$.
A system such that $\boldsymbol{b} = \boldsymbol{0}_m$ is defined **homogeneous**.
N.B.:
Given a linear system $A\boldsymbol{x} = \boldsymbol{b}$ admitting solutions, its solutions are obtained summing a particular solution $\bar{\boldsymbol{x}}$ with the ones of the homogeneous linear system $A\boldsymbol{x} = \boldsymbol{0}_m$; i.e., the set of solutions for $A\boldsymbol{x} = \boldsymbol{b}$ is

$$\mathcal{S}_{\boldsymbol{b}} = \{\bar{\boldsymbol{x}} + \boldsymbol{x}^* \mid \boldsymbol{x}^* \in \mathcal{S}_0\} \ ,$$

where $\mathcal{S}_0 = \ker(A) \subseteq \mathbb{R}^n$ is the set of solutions of the homogeneous linear system $A\boldsymbol{x} = \boldsymbol{0}_m$.
Let $\bar{\boldsymbol{x}}$ be an arbitrary solution of $A\boldsymbol{x} = \boldsymbol{b}$, then the set of solutions is given by

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \ker(A), \ \text{with} \ \dim(\ker(A)) = n - rk(A).$$

Linear Systems - Elementary Methods

Elementary Strategies for Solving a Linear System:

- By substitution;
- Gaussian elimination;
- Cramer's rule.

Linear Systems

Exercise (recap of substitution method):
Using the substitution method, find the set of real solutions $\boldsymbol{x} \in \mathbb{R}^n$ of the linear system $A\boldsymbol{x} = \boldsymbol{b}$ with:

1. $A := \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$, $\boldsymbol{b} := \begin{bmatrix} 1 \\ 2 \end{bmatrix}$;
2. $A := \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix}$, $\boldsymbol{b} := \begin{bmatrix} 2 \\ 1 \end{bmatrix}$;
3. $A := \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix}$, $\boldsymbol{b} := \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Linear Systems - Gaussian Elimination

Definition

A matrix $A \in \mathbb{R}^{m \times n}$ is in **row echelon form** (sometimes a.k.a. "staircase form") if the first non-zero entry from the left (a.k.a. the **pivot**) of row $A_{(i+1),\cdot}$ is strictly to the right of the pivot of row $A_{i,\cdot}$, for each $i = 1, \dots, (m-1)$.
I.e. the column index $j_{(i+1)}$ of the pivot of $A_{(i+1),\cdot}$ is greater than the column index j_i of the pivot of $A_{i,\cdot}$.
E.g., A is in the following form:

$$A = \begin{bmatrix} 0 & \cdots & 0 & a_{1,j_1} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & \cdots & 0 & a_{2,j_2} & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \cdots & 0 & a_{3,j_3} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}.$$

(observe: $j_1 < j_2 < j_3 < \cdots$)

Linear Systems - Gaussian Elimination

Notation (elementary matrix of "type" $E^{(i,j)}$):
For each $i, j = 1, \dots, n$ with $i \neq j$, we denote with $E^{(i,j)}$ the matrix in $GL_n(\mathbb{R})$ that is obtained from I_n switching row i and row j.
Example:
For $n = 4$,

$$E^{(2,3)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Some Properties:

- $E^{(i,j)\,-1} = E^{(i,j)}$.
- Given a matrix A, the product $(E^{(i,j)}A)$ is the matrix from A switching row i and row j.

Linear Systems - Gaussian Elimination

Notation (elementary matrix of "type" $E^{(i)}(\lambda)$):
For each $i = 1, \dots, n$ and for each $\lambda \in \mathbb{R} \setminus \{0\}$, we denote with $E^{(i)}(\lambda)$ the matrix in $GL_n(\mathbb{R})$ obtained from I_n multiplying the row i with λ.
Example:
For $n = 4$,

$$E^{(3)}(-2) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} .$$

Some Properties:

- $(E^{(i)}(\lambda))^{-1} = E^{(i)}(1/\lambda)$.
- Given a matrix A, the product $(E^{(i)}(\lambda)A)$ is the matrix from A multiplying row i with λ.

Linear Systems - Gaussian Elimination

Notation (elementary matrix of "type" $E^{(i,j)}(\lambda)$): For each $i, j = 1, \dots, n$ with $i \neq j$ and for each $\lambda \in \mathbb{R} \setminus \{0\}$, we denote with $E^{(i,j)}(\lambda)$ the matrix in $GL_n(\mathbb{R})$ obtained from I_n adding row j, multiplied by λ, to row i.

Example:

For $n = 4$,

$$E^{(2,3)}(-1) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} .$$

Some Properties:

- $(E_{(i,j)}(\lambda))^{-1} = E^{(i,j)}(-\lambda)$.
- Given a matrix A, the product $(E^{(i,j)}(\lambda)A)$ is the matrix from A where we added row j, multiplied by λ, to row i.

Linear Systems - Gaussian Elimination

Proposition (Gaussian Elimination - Row Echelon Form):

Given a matrix $A \in \mathbb{R}^{m \times n}$, there exists a matrix $E \in GL_m(\mathbb{R})$ product of elementary matrices such that EA is in row echelon form.

Proposition (Gaussian Elimination - Lin. Systems Solutions):

Let $A\boldsymbol{x} = \boldsymbol{b}$ be a linear system with $A \in \mathbb{R}^{m \times n}$ and let $E \in GL_m(\mathbb{R})$ be a product of elementary matrices. Then, the solutions of the linear system $A\boldsymbol{x} = \boldsymbol{b}$ are the same of the linear system $A'\boldsymbol{x} = \boldsymbol{b}'$, where $A' := EA$, $\boldsymbol{b}' = E\boldsymbol{b}$.

N.B.

- The row echelon form of a matrix obtained through left multiplication with a product of elementary matrices is **not** unique.
- If $(A'|\boldsymbol{b}')$ is in row echelon form, the solutions of the linear system $A'\boldsymbol{x} = \boldsymbol{b}'$ can be quickly retrieved by substitution.

Linear Systems - Gaussian Elimination I

Gaussian Elimination Algorithm: Let $A \in \mathbb{R}^{m\times n}$ be a matrix **not** in row echelon form. Than the following algorithm returns EA of A.

1. $p \leftarrow \min(m, n)$;
2. $\delta_i \leftarrow 0$;
3. $nzc \leftarrow 0$; *(non-null columns counter)*
4. $j \leftarrow 0$;
5. **while** $nzc \leq p$ and $j \leq n$ **do:**
    1. $k \leftarrow j - \delta_i$ ("working" row);
    2. $\widehat{a} \leftarrow$ element of $A_{(k,\dots,m),j}$ with maximum *absolute value*;
    3. $\widehat{i} \leftarrow$ row index of $\widehat{a}$ in A, i.e. such that $\widehat{a} = a_{\widehat{i},j}$;
    4. **if** $\widehat{a} \neq 0$ **do:**
        1. **swap** row k and row $\widehat{i}$ in A;
        2. **for** $i = (k+1), \dots, m$ **do:** add row $-(a_{i,j}/\widehat{a})A_{j,\cdot}$ to row $A_{i,\cdot}$ (i.e. transforming in zeros all the entries of A under $a_{k,j}$);
        3. $nzc = nzc + 1$;
    5. **else:** $\delta_i \leftarrow \delta_i + 1$;
    6. $j = j + 1$;
6. **return** A (transformed in row echelon form).

Where $A_{(i_1,\dots,i_k),j}$ is the subcolumn of $A_{\cdot,j}$ obtained from the row indexes $i_1, \dots, i_k$, for any $i_1, \dots, i_k \in \{1, \dots, m\}$.

Linear Systems - Gaussian Elimination

Examples

Apply Gaussian elimination to find the set of the real solutions of the linear systems having the following augmented matrices:

1. $(A|\boldsymbol{b}) := \left[\begin{array}{ccc|c} 1 & 3 & 2 & 0 \\ 0 & 2 & 3 & 1 \\ 2 & 0 & 1 & 0 \end{array}\right]$;
2. $(A|\boldsymbol{b}) := \left[\begin{array}{ccc|c} 1 & 1 & -3 & 4 \\ 2 & 1 & -1 & 2 \\ 3 & 2 & -4 & 6 \end{array}\right]$;
3. $(A|\boldsymbol{b}) := \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 5 \\ 1 & 1 & 5 & 2 & 7 \\ 1 & 2 & -2 & 5 & 11 \end{array}\right]$.

Linear Systems - Gaussian Elimination

Definition

A matrix $A \in \mathbb{R}^{m\times n}$ is in **reduced row echelon form** if:

- A is in row echelon form;
- the pivot of each row is 1;
- each column containing a pivot has zeros everywhere else.

E.g., A is in the following form:

$$A = \begin{bmatrix} 0 & \cdots & 0 & 1 & \cdots & \cdots & 0 & \cdots & \cdots & 0 & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & \cdots & 0 & 1 & \cdots & \cdots & 0 & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \cdots & 0 & 1 & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}.$$

Linear Systems - Gaussian Elimination

Proposition (Gaussian Elimination - Reduced Row Echelon Form):
Given a matrix $A \in \mathbb{R}^{m\times n}$, there exists a matrix $E \in GL_m(\mathbb{R})$ product of elementary matrices such that EA is in **reduced** row echelon form.
Moreover, the reduced row echelon form of a matrix is **unique**, i.e. it does not depend on the sequence of elementary matrices used to obtain it.

Linear Systems - Gaussian Elimination

Proposition: Let A be a matrix in $\mathbb{R}^{n\times n}$. The following facts are equivalent.

1. The reduced row echelon form EA of A is the identity I_n, i.e. $EA = I_n$;
2. A is product of elementary matrices;
3. A belongs to $GL_n(\mathbb{R})$;
4. The linear system $A\mathbf{x} = \mathbf{0}_n$ admits only the null solution.

N.B.:

For the items (2) and (3) of the proposition, particular we have:

3. $A^{-1} = E = E_1 \cdots E_k$;
2. $A = E^{-1} = E_k^{-1} \cdots E_1^{-1}$.

Linear Systems - Gaussian Elimination

Exercise: Apply Gaussian elimination to compute the inverse of the following matrix.

$$A := \begin{bmatrix} 1 & 3 & 2 \\ 0 & 2 & 3 \\ 2 & 0 & 1 \end{bmatrix} .$$

Linear Systems - Cramer's Rule

Cramer's Rule:
If $A \in GL_n(\mathbb{R})$, then the linear system $A\boldsymbol{x} = \boldsymbol{b}$ admits a unique solution

$$\boldsymbol{x}^* = A^{-1}\boldsymbol{b}\,.$$

By second Laplace's Theorem (see slide 26), we have that

$$\boldsymbol{x}^* = \frac{1}{det(A)} cof(A)^\top \boldsymbol{b}.$$

Let us note that $\boldsymbol{c} := (cof(A)^\top \boldsymbol{b}) = ((\alpha^{(i,j)})^\top \boldsymbol{b}) \in \mathbb{R}^{n\times 1}$, where the i^{th} element is

$$c_i = c_{i,1} = \sum_{s=1}^{n} \alpha^{(s,i)} b_s = \sum_{s=1}^{n} (-1)^{s+i} det(A^{(s,i)})\, b_s$$

and it coincides with the determinant of the matrix obtained from A replacing column $A_{\cdot,i}$ with $\boldsymbol{b}$.

Linear Systems - Cramer's Rule

Cramer's Theorem

Let $A\boldsymbol{x} = \boldsymbol{b}$ be a linear system with $A \in GL_n(\mathbb{R})$ (and, therefore, $\boldsymbol{b} \in \mathbb{R}^n$). Then, the system admits a unique solution $\boldsymbol{x}^* = [x_1^*, \dots, x_n^*]^\top$ such that:

$$x_i^* = \frac{det \begin{bmatrix} a_{1,1} & \dots & a_{1,i-1} & b_1 & a_{1,i+1} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,i-1} & b_n & a_{n,i+1} & \dots & a_{n,n} \end{bmatrix}}{det(A)},$$

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

Linear Systems - Cramer's Rule

Exercise:
Apply Cramer's rule to find the set of the real solutions of the linear system having the following augmented matrix.

$$(A|\boldsymbol{b}) := \left[\begin{array}{ccc|c} 1 & 3 & 2 & 0 \\ 0 & 2 & 3 & 1 \\ 2 & 0 & 1 & 0 \end{array}\right]$$

Linear Systems - Cramer's Rule

Cramer's Rule and non-invertible matrices:

Let $A \in \mathbb{R}^{m\times n}$ be such that $A \notin GL_n(\mathbb{R})$ but $t := rk(A) = rk(A|\boldsymbol{b})$; then the linear system $A\boldsymbol{x} = \boldsymbol{b}$ admits ∞^{n-t} solutions.

In order to find them with the Cramer's Rule:

1. Consider an invertible submatrix $A' \in GL_t(\mathbb{R})$ of A.
2. Define a new column vector $\boldsymbol{b}'$ consisting of the elements of $\boldsymbol{b}$ corresponding to the rows of A' and "moving" in it the elements of the $n-t$ columns of A not involved in A'. For example:

$$\underbrace{\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}}_{b} \mapsto \underbrace{\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}}_{A'} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} b_1 - a_{1,3}x_3 \\ b_2 - a_{2,3}x_3 \end{bmatrix}}_{\boldsymbol{b}'}$$

3. Find the solution of the linear system having A' as the matrix of the coeffients and $\boldsymbol{b}'$ as the matrix of the constant terms (Note that $A' \in GL_t(\mathbb{R})$).

Linear Systems - Cramer's Rule

Examples:
Apply Cramer's rule to find the set of the real solutions of the linear systems having the following augmented matrices:

1. $(A|\boldsymbol{b}) := \left[\begin{array}{ccc|c} 1 & 1 & -3 & 4 \\ 2 & 1 & -1 & 2 \\ 3 & 2 & -4 & 6 \end{array}\right]$,
2. $(A|\boldsymbol{b}) := \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 5 \\ 1 & 1 & 5 & 2 & 7 \\ 1 & 2 & -2 & 5 & 11 \end{array}\right]$.

Groups

Recall: Group

$(G, \oplus)$, a set G endowed with a binary operation $\oplus$, is a group if these properties hold

- **Associativity:** $(g_1 \oplus g_2) \oplus g_3 = g_1 \oplus (g_2 \oplus g_3)$, for each $g_1, g_2, g_3 \in G$;
- **Identity Existence:** exist an element $0_G \in G$ such that $g \oplus 0_G = g = 0_G + g$, for each $g \in G$;
- **Inverse Existence:** for each $g \in G$ exist an *inverse element* $g' \in G$ such that $g \oplus g' = 0_G$. We denote the inverse element g' as $-g$.

$(G, \oplus)$ is defined **abelian** or **commutative** if the **commutative property** holds, i.e.:

- **Commutativity:** $g_1 \oplus g_2 = g_2 \oplus g_1$, for each $g_1, g_2 \in G$;

Examples: $GL_n(\mathbb{R})$ (see slide 15); **Examples (Abelian):** $(\mathbb{R}, +)$ or $(\mathbb{R}_{>0}, \cdot)$.

Vector Spaces

Definition

A **vector space** $\mathcal{V}$ **over** $\mathbb{R}$:

- is an Abelian group $(\mathcal{V}, +)$
- has a "*product* **with** *a scalar*": $\mathbb{R} \times \mathcal{V} \to \mathcal{V}$ operation with the following properties:
  - **Distributivity ("right"):** $a(\boldsymbol{v} + \boldsymbol{w}) = a\boldsymbol{v} + a\boldsymbol{w}$, for each $a \in \mathbb{R}$ and $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$;
  - **Distributivity ("left"):** $(a + b)\boldsymbol{v} = a\boldsymbol{v} + b\boldsymbol{v}$, for each $a, b \in \mathbb{R}$ and $\boldsymbol{v} \in \mathcal{V}$;
  - **Associativity:** $a(b\boldsymbol{v}) = b(a\boldsymbol{v}) = (ab)\boldsymbol{v}$, for each $a, b \in \mathbb{R}$ and $\boldsymbol{v} \in \mathcal{V}$;
  - **"1-Neutrality":** $1\boldsymbol{v} = \boldsymbol{v}$, for each $v \in \mathcal{V}$.

Examples: $\mathbb{R}^n$ and $\mathbb{R}^{m \times n}$.
ATTENTION: "$+$"$\neq$"$+$" (sum in $\mathbb{R}$, sum in V). In general an abuse of notation is adopted, using for both operations the same symbol "$+$".

Vector Spaces - Definition

Definition

A triple $(\mathcal{V}, +, \cdot)$ is a vector space over the field $\mathbb{R}$ if:

1. **Associativity** $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$, $\forall v_1, v_2, v_3 \in \mathcal{V}$;
2. **Identity Existence** there exists a zero vector $0_{\mathcal{V}} \in \mathcal{V}$, such that $v + 0 = 0 + v = v$, $\forall v \in \mathcal{V}$;
3. **Inverse Existence** there exists $-v \in \mathcal{V}$ such that $-v + v = v + (-v) = 0$, $\forall v \in \mathcal{V}$;
4. **Commutativity** $v1 + v2 = v2 + v1$, $\forall v_1, v_2 \in \mathcal{V}$;
5. **Distributivity "right"** $c \cdot (v1 + v2) = c \cdot v_1 + c \cdot v_2$, $\forall v_1, v_2 \in \mathcal{V}$ and $\forall c \in \mathbb{R}$;
6. **Distributivity "left"** $(a + b) \cdot v = a \cdot v + b \cdot v$, $\forall v \in \mathcal{V}$ and $\forall a, b \in \mathbb{R}$;
7. **scalar Associativity** $(ab) \cdot v = a \cdot (b \cdot v) = b \cdot (a \cdot v)$, $\forall v \in \mathcal{V}$ and $\forall a, b \in \mathbb{R}$;
8. **1-Neutrality** $1 \cdot v = v$, $\forall v \in \mathcal{V}$.

The symbol $\cdot$ denoting the scalar-multiplication is usually omitted.

Vector Spaces - Definitions & Notations

Let us consider a vector space (v.s.) $\mathcal{V}$ over $\mathbb{R}$.

- $\mathbb{R}$ is the **field of scalars**[1] of $\mathcal{V}$. Elements of the "field of scalars" are called **scalars** and are denoted by lower-case letters (e.g. a, α);
- elements of $\mathcal{V}$ are called **vectors**; in the practice, vectors are **usually the elements of the v.s.** $\mathbb{R}^n$, with $n \geq 2$, for all the other v.s. we use the proper names of the objects (e.g. "matrices" for $\mathbb{R}^{m \times n}$);
- we denote the **vectors** of $\mathcal{V}$ with lower-case bold letters (e.g. $\boldsymbol{a}, \boldsymbol{\alpha}$); the elements of $\mathbb{R}^n$, with $n \geq 2$, as scalars, are denoted by the same letter and a positional index (e.g., a_i, α_i);
- we denote **sets or spaces** with upper-case letters, possibly in calligraphic style (e.g. $V, \mathcal{V}$);
- The **identity element** of $(\mathcal{V}, +)$ is denoted by $\mathbf{0}_{\mathcal{V}}$;
- we denote **matrices** with upper-case letters (e.g. A, Σ) and their elements as scalars denoted by the same lower-case letter and two positional indexes (e.g., $a_{i,j}$, $\sigma_{i,j}$).

[1] In general, a v.s. can be defined also over other fields, e.g. on $\mathbb{C}$.

Vector Spaces - Subspaces

Vector Subspaces:
A subset $\mathcal{W}$ of the vector space $\mathcal{V}$ is called **vector subspace of** $\mathcal{V}$ if it is a vector space w.r.t. the restrictions of the operations defined on $\mathcal{V}$. Equivalently, if and only if for any $\lambda_1, \lambda_2 \in \mathbb{R}$ and $\boldsymbol{w}_1, \boldsymbol{w}_2 \in \mathcal{W}$,

$$\lambda_1 \boldsymbol{w}_1 + \lambda_2 \boldsymbol{w}_2 \in \mathcal{W}.$$

Examples and Well Known results

- Given a vector space $\mathcal{V}$, $\{\mathbf{0}_{\mathcal{V}}\}$ and $\mathcal{V}$ are vector subspaces of $\mathcal{V}$;
- Let $\mathcal{W}$ be a vector subspace of $\mathcal{V}$, then $\mathbf{0}_{\mathcal{V}} \in \mathcal{W}$;
- Given two vector subspaces $\mathcal{W}$ and $\mathcal{Z}$ of $\mathcal{V}$, the **intersection** $\mathcal{W} \cap \mathcal{Z}$ is a vector subspace of $\mathcal{V}$;
- The vector subspaces of $\mathbb{R}^2$ are: $\{[0, 0]^\top\}$, $\mathbb{R}^2$ and all the lines passing through the origin;
- Given $A \in \mathbb{R}^{m \times n}$, the set of solutions of the homogeneous linear system $A\boldsymbol{x} = \mathbf{0}_m$ is a vector subspace of $\mathbb{R}^n$.
- Conversely, any vector subspace of $\mathbb{R}^n$ is set of solutions of a homogeneous linear system.

Vector Spaces - Subspaces

Sum of Vector Subspaces:
Given two vector subspaces $\mathcal{W}$ and $\mathcal{Z}$ of $\mathcal{V}$, we define the **sum** $\mathcal{W} + \mathcal{Z}$ as the **smallest vector subspace** of $\mathcal{V}$ containing $\mathcal{W} \cup \mathcal{Z}$, i.e.:

$$\mathcal{W} + \mathcal{Z} := \{\boldsymbol{w} + \boldsymbol{z} \,|\, \boldsymbol{w} \in \mathcal{W},\ \boldsymbol{z} \in \mathcal{Z}\}\,.$$

If $\mathcal{W} \cap \mathcal{Z} = \{\mathbf{0}_{\mathcal{V}}\}$, then the sum is called **direct** and denoted by $\mathcal{W} \oplus \mathcal{Z}$.
Observation:
The union ($\cup$) of two vector subspaces of $\mathcal{V}$ in general is **not** a vector subspace of $\mathcal{V}$.
E.g., $\mathcal{W} := \{\lambda[1,2]^\top \,|\, \lambda \in \mathbb{R}\}$, $\mathcal{Z} := \{\lambda[2,2]^\top \,|\, \lambda \in \mathbb{R}\}$ are vector subspaces of $\mathbb{R}^2$, but $[1,2]^\top + [2,2]^\top \notin \mathcal{W} \cup \mathcal{Z}$; then $\mathcal{W} \cup \mathcal{Z}$ is not a subspace of $\mathcal{V}$.

Vector Spaces - Generators

Systems of Generators:
Given a finite collection of vectors $W = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ of a vector space $\mathcal{V}$, we define the **subspace** $\mathcal{W}$ **generated by** $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$ the vector subspace of $\mathcal{V}$, denoted as $\mathcal{W} := \langle \boldsymbol{w}_1, \dots, \boldsymbol{w}_n \rangle$ (read "**span** of $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$"), whose elements are all the $\boldsymbol{v} \in \mathcal{V}$ that can be written as a sum of the form

$$\lambda_1 \boldsymbol{w}_1 + \dots + \lambda_n \boldsymbol{w}_n \,, \tag{1}$$

where $\lambda_1, \dots, \lambda_n \in \mathbb{R}$; i.e. $\mathcal{W} = \left\{ \boldsymbol{w} = \sum_{i=1}^{n} \lambda_i \boldsymbol{w}_i \mid \lambda_i \in \mathbb{R} \right\}$.
An expression as (1) is defined **linear combination of** $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$ **with coefficients** $\lambda_1, \dots, \lambda_n$.
A collection of vectors $W \subset \mathcal{V}$ is called **system of generators** (s.o.g) for any v.s. $\mathcal{Z}$ (even $\mathcal{V}$ itself) if

$$\mathcal{Z} = \langle \boldsymbol{w}_1, \dots, \boldsymbol{w}_n \rangle = \mathcal{W} \,;$$

Example: Vectors $[1,0]^\top, [0,1]^\top, [1,1]^\top$ are a system of generators for $\mathbb{R}^2$. In fact, e.g., $\forall\ [x_1, x_2]^\top \in \mathbb{R}^2$, it holds:

$$[x_1, x_2]^\top = (x_1 - 1)[1,0]^\top + (x_2 - 1)[0,1]^\top + [1,1]^\top .$$

Vector Spaces - Linear Independence

Linear Independence:

Vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathcal{V}$ are **linear independent** if

$$\lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_n \boldsymbol{v}_n = \mathbf{0}_{\mathcal{V}} \text{ with } \lambda_i \in \mathbb{R} \quad \Leftrightarrow \quad \lambda_1 = \cdots = \lambda_n = 0.$$

Remark

If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathcal{V}$ are linear independent, then each $\boldsymbol{v} \in \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \rangle$ can be written in a **unique** way as linear combination of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$.
More specifically, let $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ be such that $\boldsymbol{v} = \sum_{i=1}^{n} \lambda_i \boldsymbol{v}_i$, then any other combination of coefficients $\lambda'_1, \ldots, \lambda'_n \in \mathbb{R}$ such that $\boldsymbol{v} = \sum_{i=1}^{n} \lambda'_i \boldsymbol{v}_i$ **does not exist**.

Vector Spaces - Bases

Bases:

A collection $B = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ of vectors of the v.s. $\mathcal{V}$ is called a **basis of** $\mathcal{V}$ if:

- the vectors of B are a system of generators for $\mathcal{V}$, i.e., $\langle \boldsymbol{v}_1, \dots, \boldsymbol{v}_n \rangle = \mathcal{V}$;
- $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ are linearly independent.

Example:

Vectors $E = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\} \subset \mathbb{R}^n$ such that:

$$\boldsymbol{e}_1 := [1, 0, \dots, 0]^\top, \quad \boldsymbol{e}_2 := [0, 1, 0, \dots, 0]^\top, \quad \dots \quad \boldsymbol{e}_n := [0, \dots, 0, 1]^\top$$

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

Vector Spaces - Bases

Vector Representation with Different Bases:

- Let E be the canonical basis of $\mathbb{R}^n$ and let $\boldsymbol{v} \in \mathbb{R}^n$ be a vector such that its unique representation w.r.t. E is $\boldsymbol{v} = \sum_{i=1}^n \lambda_i \boldsymbol{e}_i$. Then, $\boldsymbol{v}$ is represented by the array
$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} \in \mathbb{R}^{n \times 1};$$
- Let $V = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ be a basis of $\mathcal{V}$ and let $\boldsymbol{v} \in \mathcal{V}$ be a vector such that its unique representation w.r.t. V is $\boldsymbol{v} = \sum_{i=1}^n \gamma_i \boldsymbol{v}_i$. Then, $\boldsymbol{v}$ is represented by the array
$$\boldsymbol{v}_V := \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix}_V \in \mathbb{R}^{n \times 1};$$

Observation: often the subscript "$[\cdot]_V$", denoting the basis used for writing $\boldsymbol{v}$, is omitted (as in the case of the canonical basis).

Vector Spaces - Bases

Example:
Let $\boldsymbol{v} = [2,3]^\top = 2\boldsymbol{e}_1 + 3\boldsymbol{e}_2 \in \mathbb{R}^2$ be a vector and let $V = \{\boldsymbol{v}_1 = [-1,0]^\top, \boldsymbol{v}_2 = [0,-1]^\top\}$, $W = \{\boldsymbol{w}_1 = [1,2]^\top, \boldsymbol{w}_2 = [-2,-1]^\top\}$ be two bases of $\mathbb{R}^2$.
Then $\boldsymbol{v} = -2\boldsymbol{v}_1 - 3\boldsymbol{v}_2 = \frac{4}{3}\boldsymbol{w}_1 - \frac{1}{3}\boldsymbol{w}_2$ and

$$\boldsymbol{v} = \begin{cases} \begin{bmatrix} 2 \\ 3 \end{bmatrix} & \text{(can. basis representation)} \\ \boldsymbol{v}_V = \begin{bmatrix} -2 \\ -3 \end{bmatrix}_V & (V \text{ basis representation}) \\ \boldsymbol{v}_W = \begin{bmatrix} (4/3) \\ -(1/3) \end{bmatrix}_W & (W \text{ basis representation}) \end{cases}$$

Vector Spaces - Bases

Exchange Lemma:
If $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ are a basis for $\mathcal{V}$ and $\boldsymbol{w} \in \mathcal{V}$ is such that $\boldsymbol{w} \notin \langle \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \rangle$, then $\boldsymbol{w}, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ are a basis for $\mathcal{V}$.

Basis Equipotence Theorem

Each basis of a vector space $\mathcal{V}$ consists of the same number of vectors.

Dimension of a Vector Space:
Given a vector space $\mathcal{V}$, we define the **dimension of** $\mathcal{V}$ the number $dim(\mathcal{V})$ of vectors of one of its bases.
Observation: any set of $k_1 > dim(\mathcal{V})$ vectors of $\mathcal{V}$ is **not** linearly independent; moreover, if a set of k_2 vectors is a s.o.g. for $\mathcal{V}$, then $k_2 \geq dim(\mathcal{V})$.

Vector Spaces

Extraction/Completion Lemma If $\mathcal{V} = \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \rangle$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ with $k \leq n$ are linearly independent, then it is possible to extract from $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ a basis for $\mathcal{V}$ containing $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$.

Existence of a Basis

Every non-null and finitely generated vector space admits a basis.

Existence of a Complement

If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ are linearly independent vectors of a v.s. $\mathcal{V}$ of dimension $n > k$, then there exist $\boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_n \in \mathcal{V}$ such that $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ is a basis of $\mathcal{V}$.
Equivalently: given a finitely generated v.s. $\mathcal{V}$, for any vector subspace $\mathcal{W}$ of $\mathcal{V}$, there exists a vector subspace $\mathcal{Z}$ of $\mathcal{V}$ such that $\mathcal{V} = \mathcal{W} \oplus \mathcal{Z}$.

Vector Spaces - Bases

Grassmann's Formula

Let $\mathcal{W}, \mathcal{Z}$ be two subspaces of a v.s. $\mathcal{V}$. We have that:

$$dim(\mathcal{W} + \mathcal{Z}) = dim(\mathcal{W}) + dim(\mathcal{Z}) - dim(\mathcal{W} \cap \mathcal{Z}).$$

In particular,

$$dim(\mathcal{W} \oplus \mathcal{Z}) = dim(\mathcal{W}) + dim(\mathcal{Z}).$$

Vector Spaces - Bases

Generators and Independence via Matrices:
Given a vector space $\mathcal{V} = \langle \boldsymbol{v}_1, \dots, \boldsymbol{v}_n \rangle$ and m vectors $\boldsymbol{w}_1, \dots, \boldsymbol{w}_m \in \mathcal{V}$, we define $A \in \mathbb{R}^{n\times m}$ as the matrix of size $n \times m$ such that, for each $i = 1, \dots, m$, $\boldsymbol{w}_i = \sum_{j=1}^{n} a_{j,i} \boldsymbol{v}_j$; i.e. the matrix such that

$$[\boldsymbol{w}_1 \mid \cdots \mid \boldsymbol{w}_m] = [\boldsymbol{v}_1 \mid \cdots \mid \boldsymbol{v}_n]\, A\,.$$

Properties:
Let $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ be a basis of $\mathcal{V}$, then:.

- $\langle \boldsymbol{w}_1, \dots, \boldsymbol{w}_m \rangle = \mathcal{V}$ if and only if A is right-invertible (i.e., there exists $A^{(r)}\mathbb{R}^{m\times n}$ such that $AA^{(r)} = I_n$). In particular, this implies that $m \geq n = rk(A)$;
- $\boldsymbol{w}_1, \dots, \boldsymbol{w}_m$ are linearly independent if and only if A is left-invertible (i.e., there exists $A^{(\ell)}\mathbb{R}^{m\times n}$ such that $A^{(\ell)}A = I_m$). This implies $m \leq n$.
  In particular, A is left-invertible if and only if $rk(A) = m$;
- **Direct consequence of previous properties:** $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$ is a basis of $\mathcal{V}$ if and only if $A \in GL_n(\mathbb{R})$ (obviously $m = n$).

Vector Spaces

Example:
If $\boldsymbol{w}_1 := [2,1]^\top$, $\boldsymbol{w}_2 := [3,0]^\top$, $\boldsymbol{w}_3 := [-1,2]^\top \in \mathcal{V} := \mathbb{R}^2$, we have that

$$\underbrace{\begin{bmatrix} 2 & 3 & -1 \\ 1 & 0 & 2 \end{bmatrix}}_{[\boldsymbol{w}_1 \;|\boldsymbol{w}_2 \;|\; \boldsymbol{w}_3]} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{[\boldsymbol{e}_1 \;|\; \boldsymbol{e}_2]} \underbrace{\begin{bmatrix} 2 & 3 & -1 \\ 1 & 0 & 2 \end{bmatrix}}_{A};$$

N.B.: since we are comparing $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3$ with the canonical basis, then $A = [\boldsymbol{w}_1 \;|\boldsymbol{w}_2 \;|\; \boldsymbol{w}_3]$ and $I_n = [\boldsymbol{e}_1 \;|\; \boldsymbol{e}_2]$.
Vectors $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3$ generate $\mathbb{R}^2$ because A is right-invertible. In fact,

$$\begin{bmatrix} 2 & 3 & -1 \\ 1 & 0 & 2 \end{bmatrix} \underbrace{\begin{bmatrix} 0 & 1/5 \\ 1/3 & 0 \\ 0 & 2/5 \end{bmatrix}}_{A^{(r)}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

However, $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3$ are not linear independent, indeed A is not left-invertible.

Vector Spaces - Norms

Definition

Given a vector space $\mathcal{V}$, a **norm** is a function $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ such that, for any $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$ and $\lambda \in \mathbb{R}$:

1. $\|\boldsymbol{v}\| \geq 0$ (non-negativity);
2. $\|\boldsymbol{v}\| = 0 \Leftrightarrow \boldsymbol{v} = \mathbf{0}_{\mathcal{V}}$;
3. $\|\lambda \boldsymbol{v}\| = |\lambda| \|\boldsymbol{v}\|$;
4. $\|\boldsymbol{v} + \boldsymbol{w}\| \leq \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$ (triangle inequality).

When endowed with a norm, $\mathcal{V}$ is called a **normed vector space**.

N.B.:
A norm induces a **distance** (or **metric**) $d : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ among the vectors of $\mathcal{V}$ making such that, for each $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$, it holds

$$d(\boldsymbol{v}, \boldsymbol{w}) := \|\boldsymbol{v} - \boldsymbol{w}\|.$$

A v.s. $\mathcal{V}$ endowed with a distance is a **metric space**.

Vector Spaces - Norms

Examples:

- For any $p \in \mathbb{R}$ with $p \geq 1$, $\mathbb{R}^n$ can be endowed with a norm $\|\cdot\|_p$ (called p-norm) such that

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

  for each $\boldsymbol{v} = [v_1, \dots, v_n]^\top \in \mathbb{R}^n$.
- the 2-norm is usually called **Euclidean Norm** (and the induced distance is called **Euclidean distance**);
- $\mathbb{R}^n$ can be endowed with the **infinite norm** $\|\cdot\|_\infty$ defined as

$$\|\boldsymbol{v}\|_\infty := \max_i |v_i|.$$

Vector Spaces - Norms

Definition

A norm $\|\cdot\|$ on the vector space $\mathbb{R}^{n\times n}$ is called **sub-multiplicative** if, for any $A, B \in \mathbb{R}^{n\times n}$, it holds

$$\|AB\| \leq \|A\|\|B\|.$$

Induced Matrix Norms

For each $p \in \mathbb{R}_{\geq 1} \cup \{+\infty\}$, the p-norms on $\mathbb{R}^m$ and $\mathbb{R}^n$ induce a "matrix p-norm" $\|\cdot\|_p$ on $\mathbb{R}^{m\times n}$ that is defined as

$$\|A\|_p := \sup_{\substack{\|\boldsymbol{v}\|_p=1 \\ \boldsymbol{v}\in\mathbb{R}^n}} \|A\boldsymbol{v}\|_p = \sup_{\substack{v\neq\boldsymbol{0} \\ \boldsymbol{v}\in\mathbb{R}^n}} \frac{\|Av\|_p}{\|v\|_p},$$

for any $A \in \mathbb{R}^{m\times n}$.

Vector Spaces - Norms

Some Properties of Induced Matrix Norms: for each $A \in \mathbb{R}^{m\times n}$, for each $p \in \mathbb{R}_{\geq 1} \cup \{+\infty\}$, we have that:

- $\|I_n\|_p = 1$;
- $\|A\boldsymbol{v}\|_p \leq \|A\|_p \|v\|_p$, for each $\boldsymbol{v} \in \mathbb{R}^n$ (**Attention:** $\mathbb{R}^m$, $\mathbb{R}^n$, $\mathbb{R}^{m\times n}$);
- if the two matrices are compatible, $\|\cdot\|_p$ on $\mathbb{R}^{m\times n}$ is sub-multiplicative;
- if $m = n$, $\|A\|_p \geq \rho(A)$, where the **spectral radius** $\rho(A)$ of a matrix A is the maximum of the absolute values of its eigenvalues (see slide 97);
- $\|A\|_2 = \sqrt{\rho(A^T A)}$;
- if $m = n$ and A is symmetric, then $\|A\|_2 = \rho(A)$;
- $\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{i,j}|$;
- $\|A\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{i,j}|$;
- if A is symmetric, then $\|A\|_1 = \|A\|_\infty$;
- $\|A\|_2 \leq \sqrt{\|A\|_1 \|A\|_\infty}$.

Vector Spaces - Norms

Entry-Wise Matrix Norms:
We define as **"entry-wise"** p**-norm** on $\mathbb{R}^{m\times n}$ the norm

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

for each $p \in \mathbb{R}_{\geq 1}$, and the norm $(\max_{i,j}|a_{i,j}|)$ for $p=\infty$.
Watch Out: often the symbol $\|\cdot\|_p$ is also used for denoting the "entry-wise" p-norm on $\mathbb{R}^{m\times n}$. However, here we will denote it with the symbol $\|\cdot\|_{(p)}$.
The norm $\|\cdot\|_{(2)}$ is also called **Frobenius norm**, denoted by $\|\cdot\|_F$.

Some Properties:

- For each $p \in \mathbb{R}_{\geq 1} \cup \{+\infty\}$, the $\|\cdot\|_{(p)}$ is sub-multiplicative.
- For each $A \in \mathbb{R}^{m\times n}$, $\|A\|_2 \leq \|A\|_F$.

Vector Spaces - Scalar Products

Definition

Given a vector space $\mathcal{V}$, a **scalar product** is a function "$\cdot$" : $\mathcal{V} \times \mathcal{V} \to \mathbb{R}$ such that, for any $\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{z} \in \mathcal{V}$ and any $\lambda \in \mathbb{R}$:

1. $\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{w} \cdot \boldsymbol{v}$ (symmetry);
2. $(\boldsymbol{v} + \boldsymbol{w}) \cdot \boldsymbol{z} = \boldsymbol{v} \cdot \boldsymbol{z} + \boldsymbol{w} \cdot \boldsymbol{z}$ (distributivity; valid also for $\boldsymbol{z} \cdot (\boldsymbol{v} + \boldsymbol{w})$);
3. $\lambda(\boldsymbol{v} \cdot \boldsymbol{w}) = (\lambda \boldsymbol{v}) \cdot \boldsymbol{w}$ (associativity w.r.t. a scalar);
4. $(\boldsymbol{v} \cdot \boldsymbol{v}) \geq 0$ and $(\boldsymbol{v} \cdot \boldsymbol{v}) = 0$ if and only if $\boldsymbol{v} = \boldsymbol{0}_{\mathcal{V}}$.

Some Properties:

- For each $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$ and any $\lambda \in \mathbb{R}$:
  1. $(\lambda \boldsymbol{v}) \cdot \boldsymbol{w} = \lambda(\boldsymbol{v} \cdot \boldsymbol{w}) = \boldsymbol{v} \cdot (\lambda \boldsymbol{w})$;
  2. $\boldsymbol{0}_{\mathcal{V}} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{0}_{\mathcal{V}} = 0$.
- A scalar product on a vector space $\mathcal{V}$ induces a **norm** $\|\cdot\|$ on $\mathcal{V}$ defined, for any $\boldsymbol{v} \in \mathcal{V}$, as $\|\boldsymbol{v}\| := \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$ and such that:
  1. for any $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$, $\boldsymbol{v} \cdot \boldsymbol{w} \leq \|\boldsymbol{v}\|\|\boldsymbol{w}\|$ (Cauchy-Schwarz inequality).

Vector Spaces - Scalar Products

Examples.

- We define the **standard scalar product** on $\mathbb{R}^n$ as:

$$\boldsymbol{v} \cdot \boldsymbol{w} := \underbrace{\boldsymbol{v}^\top \boldsymbol{w}}_{\substack{\text{matrix}\\ \text{product}}} = \sum_{i=1}^{n} v_i w_i \,.$$

- The standard scalar product on $\mathbb{R}^n$ induces the **Euclidian norm**.
- The Frobenius norm $\|\cdot\|_F$ on $\mathbb{R}^{m\times n}$ is induced by the scalar product

$$A \cdot B := tr(A^\top B)\,,$$

indeed, $\|A\|_F := (\sum_{i=1}^m \sum_{j=1}^m a_{i,j}^2)^{1/2} = \sqrt{tr(A^\top A)}$.

Notation: for each square matrix $A \in \mathbb{R}^{n\times n}$, the symbol $tr(A)$ denotes the **trace** of the matrix A, that is the sum of its diagonal elements; i.e.

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

Morphisms

Definition

Let $\mathcal{V}, \mathcal{W}$ be two vector spaces over $\mathbb{R}$. We call a function

$$F : \mathcal{V} \to \mathcal{W}$$

linear function or **homomorphism** of $\mathbb{R}$-vector spaces if the following properties hold:

1. $F(\boldsymbol{v}_1 + \boldsymbol{v}_2) = F(\boldsymbol{v}_1) + F(\boldsymbol{v}_2)$, for each $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathcal{V}$;
2. $F(\lambda \boldsymbol{v}) = \lambda F(\boldsymbol{v})$, for each $\lambda \in \mathbb{R}$, for each $\boldsymbol{v} \in \mathcal{V}$.

N.B.: $\mathcal{V}$ and $\mathcal{W}$ can have different dimensions.

Immediate Consequences:

- $F(\mathbf{0}_{\mathcal{V}}) = \mathbf{0}_{\mathcal{W}}$,
- $F(-\boldsymbol{v}) = -F(\boldsymbol{v})$, for each $\boldsymbol{v} \in \mathcal{V}$.

Morphisms

Definition

Given a linear function $F : \mathcal{V} \to \mathcal{W}$ we define:

- the **kernel** of F as the vector subspace of $\mathcal{V}$ such that:

$$\ker(F) := \{\boldsymbol{v} \in \mathcal{V} \,|\, F(\boldsymbol{v}) = \boldsymbol{0}_{\mathcal{W}}\};$$

- the **image** of F as the vector subspace of $\mathcal{W}$ such that:

$$\mathrm{Im}(F) := \{\boldsymbol{w} \in \mathcal{W} \,|\, \exists \boldsymbol{v} \in \mathcal{V},\ F(\boldsymbol{v}) = \boldsymbol{w}\}.$$

Properties:

- F is **injective** $\Leftrightarrow \ker(F) = \{\boldsymbol{0}_{\mathcal{V}}\}$; $(\boldsymbol{w} = F(\boldsymbol{v}) = F(\boldsymbol{v} + \boldsymbol{z}) = F(\boldsymbol{v}) + F(\boldsymbol{z})$, $\forall \boldsymbol{z} \in \ker(F))$
- F is **surjective** $\Leftrightarrow \mathrm{Im}(F) = \mathcal{W}$.

Morphisms

Proposition:

Let $F : \mathcal{V} \to \mathcal{W}$ be a linear function and let $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ be a system of generators of $\mathcal{V}$. Then, $\{\boldsymbol{w}_1 = F(\boldsymbol{v}_1), \dots, \boldsymbol{w}_n = F(\boldsymbol{v}_n)\}$ are a s.o.g. of $\mathrm{Im}(F)$.

In fact, for each $\boldsymbol{w} \in \mathrm{Im}(F)$ exists $\boldsymbol{v} \in \mathcal{V}$ such that $F(\boldsymbol{v}) = \boldsymbol{w}$ and, for the **linearity** of F, it holds

$$\boldsymbol{w} = F(\boldsymbol{v}) = F\left(\sum_{i=1}^{n} \lambda_i \boldsymbol{v}_i\right) = \sum_{i=1}^{n} \lambda_i F(\boldsymbol{v}_i) = \sum_{i=1}^{n} \lambda_i \boldsymbol{w}_i$$

ATTENTION: if $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ is a basis of $\mathcal{V}$, it is **not** guaranteed that $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ is a basis of $\mathcal{W}$ (we can guarantee only that it is a generator of $\mathrm{Im}(F)$). Examples are all the linear functions where $dim(\mathcal{V}) > dim(\mathcal{W})$.

Morphisms

Let $F : \mathcal{V} \to \mathcal{W}$ be a linear function such that we know only the values $\{\boldsymbol{w}_1 = F(\boldsymbol{v}_1), \dots, \boldsymbol{w}_n = F(\boldsymbol{v}_n)\}$ associated to a basis $B_{\mathcal{V}} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ of $\mathcal{V}$.
Can we compute the image $\boldsymbol{w} = F(\boldsymbol{v})$ for each $\boldsymbol{v} \in \mathcal{V}$, using just the information above?
ANSWER: Yes, because a basis admits unique representations $\boldsymbol{v} = \sum_{i=1}^{n} \lambda_i \boldsymbol{v}_i$ of vectors; therefore $F(\boldsymbol{v}) := \sum_{i=1}^{n} \lambda_i F(\boldsymbol{v}_i)$, with $\lambda_i \in \mathbb{R}$ uniquely defined, for each $\boldsymbol{v} \in \mathcal{V}$.
Then, F can be characterized ("fully described") knowing its action on the vectors a basis of $\mathcal{V}$.

Proposition

A linear function $F : \mathcal{V} \to \mathcal{W}$ is **uniquely determined** by the images of the elements of a basis of $\mathcal{V}$.

Morphisms - Isomorphisms

Definition

A linear function $F : \mathcal{V} \to \mathcal{W}$ is defined **isomorphism** of vector spaces if F is injective and surjective (i.e. is bijective).
Two vector spaces $\mathcal{V}$ and $\mathcal{W}$ are called **isomorphic** (also written as $\mathcal{V} \simeq \mathcal{W}$) if there exists an isomorphism between them.

Proposition:
A vector space $\mathcal{V}$ over $\mathbb{R}$ of dimension n is isomorphic to $\mathbb{R}^n$.
In particular:

Proposition

Each pair of vector spaces over $\mathbb{R}$ having the same dimension are isomorphic.

Morphisms

Rank-Nullity Theorem

If $F : \mathcal{V} \to \mathcal{W}$ is a linear function, then

$$dim(\mathcal{V}) = dim(\ker(F)) + dim(\mathrm{Im}(F)).$$

Corollary:
If $F : \mathcal{V} \to \mathcal{V}$ is a linear function, then

$$F \text{ is injective } (dim(\ker(F)) = 0) \ \Leftrightarrow$$
$$\Leftrightarrow \ F \text{ is surjective } (\mathrm{Im}(F) = \mathcal{V}) \ \Leftrightarrow \ F \text{ is an isomorphism}\,.$$

Observation: the name of the theorem is given by the result of the proposition at slide 85.

Morphisms - Linear Functions and Matrices

Definition

Let $F : \mathcal{V} \to \mathcal{W}$ be a linear function. Given a basis $V := \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ of $\mathcal{V}$ and a basis $W := \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$ of $\mathcal{W}$, we define **transformation matrix of F**, w.r.t. the bases V and W, the matrix $A_{VW}(F) = (a_{i,j}) \in \mathbb{R}^{m\times n}$ such that

$$F(\boldsymbol{v}_j) =: \widehat{\boldsymbol{w}}_j = \sum_{i=1}^{m} a_{i,j}\boldsymbol{w}_i\,,\ \forall\, j = 1, \dots, n\,.$$

In other words, the elements of the j-th column of $A_{VW}(F)$ are the coefficient of $\widehat{\boldsymbol{w}}_j = F(\boldsymbol{v}_j)$ w.r.t. the basis W of $\mathcal{W}$.

Observation: Recalling the notation of slide 61, we have that $F(\boldsymbol{v}_j)$ can be represented as the array $F(\boldsymbol{v}_j)_W = [a_{1,j}, \dots, a_{m,j}]_W^\top$, for each $j = 1, \dots, m$ and, therefore, it holds

$$A_{VW}(F) = [F(\boldsymbol{v}_1)_W \mid \cdots \mid F(\boldsymbol{v}_n)_W]$$

(i.e. the columns of $A_{VW}(F)$ are the images of the vectors in V, written w.r.t. the basis W).

Morphisms - Linear Functions and Matrices

Observation: let $\boldsymbol{v}_j$ be the j-th vector of the basis V. Then its representation w.r.t. V is $(\boldsymbol{v}_j)_V = [0, \dots, \underbrace{1}_{j\text{-th el.}}, \dots, 0]_V^\top$ and, therefore, it holds

$$F(\boldsymbol{v}_j)_W = A_{VW}(F)\,(\boldsymbol{v}_j)_V \tag{2}$$

Then, transformation matrices allow for describing linear functions in a **matricial form**. More specifically, given $\boldsymbol{v} \in \mathcal{V}$, let us compute $\boldsymbol{w} = F(\boldsymbol{v})$:

- $\boldsymbol{v}$ can be uniquely written w.r.t. V, i.e. as an array
$$\boldsymbol{v}_V = [\lambda_1, \dots, \lambda_n]_V^\top .$$
- $\boldsymbol{w} = F(\boldsymbol{v})$ can be uniquely written w.r.t. W, i.e. as an array
$$\boldsymbol{w}_W = [\gamma_1, \dots, \gamma_m]_W^\top .$$
- by definition of $A_{VW}(F)$ and (2), it holds $\boldsymbol{w}_W = A_{VW}(F)\boldsymbol{v}_V$ because

$$\begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{bmatrix}_W = F(\boldsymbol{v})_W = \sum_{j=1}^{n} \lambda_j F(\boldsymbol{v}_j)_W = A_{VW}(F) \left(\sum_{j=1}^{n} \lambda_j (\boldsymbol{v}_j)_V \right) = A_{VW}(F) \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}_V$$

Morphisms - Linear Functions and Matrices

Example:
Let $F : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear function defined as follows:

$$\begin{aligned} [1,0,0]^\top &\mapsto [0,2]^\top , \\ [0,1,0]^\top &\mapsto [1,-1]^\top , \\ [0,0,1]^\top &\mapsto [-1,3]^\top , \end{aligned}$$

assuming to consider as bases V, W the canonical bases of $\mathbb{R}^3$ and $\mathbb{R}^2$, respectively. Then the matrix $A := A_{VW}(F)$ is

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

Morphisms - Linear Functions and Matrices

Example (continued):
Given the vector $\boldsymbol{v} := [2, 3, -1]^\top \in \mathbb{R}^3$, let us compute $F(\boldsymbol{v})$.

- Express v as linear combination of the elements of basis V. Since it is the canonical basis, we have that

$$\boldsymbol{v}_v = \boldsymbol{v} = [2, 3, -1]^\top$$

- Compute $F(\boldsymbol{v})_W = F(\boldsymbol{v})$ (because W is the can. basis) as

$$A\boldsymbol{v} = \begin{bmatrix} 0 & 1 & -1 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \end{bmatrix} = F(\boldsymbol{v}) \, .$$

Morphisms - Linear Functions and Matrices

Proposition

Given a linear function $F : \mathcal{V} \to \mathcal{W}$, then, for any basis V of $\mathcal{V}$ and W of $\mathcal{W}$, it holds

$$dim(\mathrm{Im}(F)) = rk(A_{VW}(F)) \,.$$

$$dim(\mathrm{Im}(F)) = \min\{dim(\mathcal{W}),\ \#columns\ of\ A = A_{VW}(F)\ l.i.\} \,.$$

Morphisms - Linear Functions and Matrices

The Vector Space $Hom(V, W)$

Given two vector spaces $\mathcal{V}, \mathcal{W}$ over $\mathbb{R}$, we denote by $Hom(\mathcal{V}, \mathcal{W})$ the **set of linear functions from $\mathcal{V}$ to $\mathcal{W}$**(homomorphisms from $\mathcal{V}$ to $\mathcal{W}$). Further, we define the two following operations.

- the **sum**:$Hom(\mathcal{V}, \mathcal{W}) \times Hom(\mathcal{V}, \mathcal{W}) \to Hom(\mathcal{V}, \mathcal{W})$ such that, for each $F, G \in Hom(\mathcal{V}, \mathcal{W})$, the linear function $H := F + G$ is defined as
$$H(\boldsymbol{v}) := (F + G)(\boldsymbol{v}) = F(\boldsymbol{v}) + G(\boldsymbol{v}) \,, \quad \forall \ \boldsymbol{v} \in \mathcal{V} \,;$$
- the **product with a scalar**:$\mathbb{R} \times Hom(\mathcal{V}, \mathcal{W}) \to Hom(\mathcal{V}, \mathcal{W})$ such that for each $\lambda \in \mathbb{R}$, for each $F \in Hom(\mathcal{V}, \mathcal{W})$, the linear function $H := \lambda F$ is defined as
$$H(\boldsymbol{v}) := (\lambda F)(\boldsymbol{v}) = \lambda F(\boldsymbol{v}) \,, \quad \forall \ \boldsymbol{v} \in \mathcal{V} \,.$$

Proposition: The two above operations make $Hom(V, W)$ a **vector space over** $\mathbb{R}$.

Morphisms - Linear Functions and Matrices

Proposition - $Hom(\mathcal{V}, \mathcal{W})$ **and Transformation Matrices:**
Fixed two bases $V = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ and $W = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$ of $\mathcal{V}$ and $\mathcal{W}$, respectively, the **function** $A_{VW}(\cdot)$ that returns the transformation matrix for each $F \in Hom(\mathcal{V}, \mathcal{W})$, i.e. the function

$$\begin{aligned} A_{VW} : Hom(\mathcal{V}, \mathcal{W}) &\longrightarrow \mathbb{R}^{m \times n} \\ F &\mapsto A_{VW}(F)\,, \end{aligned}$$

is a **linear function** and, in particular, an **isomorphism** of vector spaces.

Morphisms - Linear Functions and Matrices

Since A_{VW} is an isomorphism, the following properties are true:

- **linearity:** $A_{VW}(F+G) = A_{VW}(F) + A_{VW}(G)$ and $A_{VW}(\lambda F) = \lambda A_{VW}(F)$, for each $F, G \in Hom(\mathcal{V}, \mathcal{W})$, for each $\lambda \in \mathbb{R}$;
- **uniquely identified:** Fixed two bases V, W of $\mathcal{V}$ and $\mathcal{W}$, respectively, for each $F \in Hom(\mathcal{V}, \mathcal{W})$ exist one and only one matrix $A \in \mathbb{R}^{m \times n}$ such that $A = A_{VW}(F)$, and viceversa;
- **composition of homomorphisms:** given three bases V, W, Z of three v.s. $\mathcal{V}, \mathcal{W}, \mathcal{Z}$, respectively, for each $F \in Hom(\mathcal{V}, \mathcal{W})$ and for each $G \in Hom(\mathcal{W}, \mathcal{Z})$ the transformation matrix of the homomorphism $H \in Hom(\mathcal{V}, \mathcal{Z})$ defined as

$$H(\boldsymbol{v}) := (G \circ F)(\boldsymbol{v}) = G\left(F(\boldsymbol{v})\right), \quad \forall\ \boldsymbol{v} \in \mathcal{V}$$

has a transformation matrix $A_{VZ}(H)$ such that

$$A_{VZ}(H) = \underbrace{A_{WZ}(G) A_{VW}(F)}_{\text{matrix product}} .$$

Morphisms - Linear Functions and Matrices

Proposition

Let $F : \mathcal{V} \to \mathcal{V}$ be a linear function, and let $A \in \mathbb{R}^{n \times n}$ be the transformation matrix of F w.r.t. two bases V and W of $\mathcal{V}$. Then,

$$F \text{ is an isomorphism } \Leftrightarrow A \text{ is invertible } \Leftrightarrow$$

$$\Leftrightarrow rk(A) = n = \dim(\mathcal{V}) = \dim(\mathrm{Im}(F)) \Leftrightarrow$$

$$\Leftrightarrow det(A) \neq 0 \Leftrightarrow \dim(\ker(F)) = 0 \Leftrightarrow \ker(F) = \{0_{\mathcal{V}}\}.$$

Morphisms - Linear Functions and Matrices

Change of Bases: Let V and W be two bases of a v.s. $\mathcal{V}$.
We define **matrix of change of basis from V to W** the transformation matrix associated with the identity map $I_{VW} : \mathcal{V} \to \mathcal{V}$ defined as

$$I_{VW}(\boldsymbol{v}_V) = \boldsymbol{v}_W \,, \quad \forall \boldsymbol{v} \in \mathcal{V} \,.$$

N.B.: $\boldsymbol{v}_V$ and $\boldsymbol{v}_W$ are the same vector, written in two different bases.
N.B.: $I_{VW} : \mathcal{V} \to \mathcal{V}$ is trivially an isomorphism as well as its inverse.
Then, we denote this matrix as M_{VW} (i.e. $M_{VW} := A_{VW}(I_{VW})$) and, remembering the observation of slide 81, we have that

$$M_{VW} = [I_{VW}(\boldsymbol{v}_1) \mid \cdots \mid I_{VW}(\boldsymbol{v}_n)] = [(\boldsymbol{v}_1)_W \mid \cdots \mid (\boldsymbol{v}_n)_W]$$

(i.e. the columns of M_{VW} are the vectors V written w.r.t. the basis W).

Inverse mapping and matrix

M_{VW} is **invertible** and its inverse is $M_{WV} = M_{VW}^{-1}$.

In order to prove that $M_{WV} = M_{VW}^{-1}$ we can simply consider the isomorphism $I_{VV} : \mathcal{V} \to \mathcal{V}$ whose matrix is the identity matrix $I := A_{VV}(I_{VV}) \in \mathbb{R}^{n \times n}$, we have that $I_{VV} = I_{VW} \circ I_{WV} = I_{WV} \circ I_{VW}$

$$I = M_{VW}\, M_{WV} = M_{WV}\, M_{VW} \Leftrightarrow M_{WV} = M_{VW}^{-1}$$

Morphisms - Linear Functions and Matrices

Equivalent Matrices

Two matrices $A, B \in \mathbb{R}^{m\times n}$ are called **equivalent** if there exist $P \in GL_m(\mathbb{R})$ and $Q \in GL_n(\mathbb{R})$ such that

$$B = PAQ.$$

Some Properties:

- Matrix equivalence defines an **equivalence relation**[a] on $\mathbb{R}^{m\times n}$.
- Given $A, B \in \mathbb{R}^{m\times n}$,

$$A, B \text{ are equivalent} \Leftrightarrow rk(A) = rk(B) \quad (\textit{see pag}.\ 28).$$

[a] "$\sim$" is an eq. rel. if: $x \sim x$; $x \sim y$ iff $y \sim x$; if $x \sim y$ and $y \sim z$, then $x \sim z$.

Morphisms - Linear Functions and Matrices

Linear Functions and Equivalent Matrices

Let $F : \mathcal{V} \to \mathcal{W}$ be a linear function, let V, V' be two bases of $\mathcal{V}$, and W, W' be two bases of $\mathcal{W}$.

Then, the two transformation matrices $A_{VW}(F)$, $A_{V'W'}(F)$ are **equivalent**. In fact,

$$A_{V'W'}(F) = M_{WW'} A_{VW}(F) M_{V'V} \,.$$

Conversely, let A, B be two equivalent matrices in $\mathbb{R}^{m \times n}$, and let $\mathcal{V}, \mathcal{W}$ be two vector spaces of dimension n and m, respectively.

Then, there exists a linear function $F : \mathcal{V} \to \mathcal{W}$ such that A and B are the **transformation matrices** of F under two different choices of bases of $\mathcal{V}$: V, V' and $\mathcal{W}$: W, W'.

Morphisms - Linear Functions and Matrices

Similar Matrices

Two matrices $A, B \in \mathbb{R}^{n \times n}$ are called **similar** if there exists $P \in GL_n(\mathbb{R})$ such that

$$B = P^{-1} A P.$$

Some Properties:

- Matrix similarity defines an **equivalence relation** on $\mathbb{R}^{n \times n}$.
- Given two similar matrices $A, B \in \mathbb{R}^{n \times n}$, then:
  1. A, B are equivalent;
  2. $det(A) = det(B)$ $(det(P^{-1}) = 1/det(P))$.

Morphisms - Linear Functions and Matrices

Linear Functions and Similar Matrices

Let $F : \mathcal{V} \to \mathcal{V}$ be a linear function, let V, W be two bases of $\mathcal{V}$.
Then, the two transformation matrices $A_{VV}(F)$, $A_{WW}(F)$ are **similar**. In fact,

$$A_{WW}(F) = M_{VW} A_{VV}(F) M_{WV} \,.$$

Conversely, let A, B be two similar matrices in $\mathbb{R}^{n \times n}$, and let $\mathcal{V}$ be a vector space of dimension n.
Then, there exists a linear function $F : \mathcal{V} \to \mathcal{V}$ such that A and B are the **transformation matrices** $A_{VV}(F)$, $A_{WW}(F)$ with respect to two different bases V and W of $\mathcal{V}$.

Morphisms - Diagonalization

A linear function $F : \mathcal{V} \to \mathcal{W}$, given two fixed bases V, W of $\mathcal{V}, \mathcal{W}$, respectively, can be represented as an equivalence class of matrices, i.e. the **set** of all the **matrices equivalent** to $A_{VW}(F)$ (**similar** matrices if $\mathcal{W} = \mathcal{V}$ and $V = W$).
Is it possible to choose among the elements of such a class the one providing the "simplest" representation of F?
ANSWER: Yes, in some cases.

Morphisms - Diagonalization

Definition

A linear function $F : \mathcal{V} \to \mathcal{V}$ is defined **diagonalizable** if there exists a basis V of $\mathcal{V}$ such that the transformation matrix $A_{VV}(F)$ is diagonal, i.e.

$$A_{VV}(F) = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} .$$

It means that exists a basis $V = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ of $\mathcal{V}$ and $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ such that, for each $i = 0, \dots, n$,

$$F(\boldsymbol{v}_i) = \lambda_i \boldsymbol{v}_i \, .$$

Morphisms - Diagonalization

Definition (continued)

A value $\lambda \in \mathbb{R}$ is an **eigenvalue** for F if there exists a non-zero vector $\boldsymbol{v} \in \mathcal{V}$ such that:

$$F(\boldsymbol{v}) = \lambda \boldsymbol{v}.$$

Given an eigenvalue for F, we define the **eigenspace** associated with λ as the vector subspace $\mathcal{V}_\lambda$ of $\mathcal{V}$:

$$\mathcal{V}_\lambda := \{\boldsymbol{v} \in \mathcal{V} \,|\, F(\boldsymbol{v}) = \lambda \boldsymbol{v}\}.$$

The vectors of $\mathcal{V}_\lambda$ are called **eigenvectors** of λ.

Some Properties: Let λ be an eigenvalue of $F : \mathcal{V} \to \mathcal{V}$,

- $\mathcal{V}_\lambda = \ker(F - \lambda Id)$, where Id is the identity (linear) function;
- $dim(\mathcal{V}_\lambda) > 0$;
- $\mathcal{V}_\lambda$ is F-invariant, i.e., $F(\mathcal{V}_\lambda) \subseteq \mathcal{V}_\lambda$.

Morphisms - Diagonalization

Proposition

Let $F : \mathcal{V} \rightarrow \mathcal{V}$ be linear function and let A be the transformation matrix $A_{VV}(F)$ of F w.r.t. a basis V of $\mathcal{V}$.
Given $\lambda \in \mathbb{R}$, the following facts are equivalent:

1. λ is an eigenvalue for F;
2. $\ker(F - \lambda Id) \neq \{\mathbf{0}_{\mathcal{V}}\}$;
3. $det(A - \lambda I_n) = 0$;
4. $det((A - \lambda I_n)^k) = 0, \ \forall k \in \mathbb{N}_+$.

Morphisms - Diagonalization

Characterization of the Eigenvalues:
Then, the eigenvalues of F are the roots in $\mathbb{R}$ of the polynomial

$$p_F(x) := det(A - xI_n)\,.$$

N.B.: Given two similar matrices $A, B \in \mathbb{R}^{n \times n}$, we have that $A - xI_n$, $B - xI_n \in \mathbb{R}^{n \times n}$ are similar as well and $det(A - xI_n) = det(B - xI_n)$. Therefore, $p_F(x)$ is called **characteristic polynomial of** F and it is **well-defined** (uniquely defined) and **invariant** to the choice of the matrix in the class of the similar matrices to A.

Morphisms - Diagonalization

If $\lambda \in \mathbb{R}$ is an eigenvalue of F, then

$$p_F(x) = (x - \lambda)^{m(\lambda)} q(x),$$

with $q(x)$ non-null polynomial.
Given an eigenvalue λ, the integer $m(\lambda)$ is called the **multiplicity** of λ as root of $p_F(x)$.

Theorems

1. A linear function $F : \mathcal{V} \to \mathcal{V}$ of vector spaces over $\mathbb{R}$ is diagonalizable (in $\mathbb{R}$) **if and only if** all the roots of $p_F(x)$ are in $\mathbb{R}$ and, for each root λ of $p_F(x)$, we have that $m(\lambda) = dim(\mathcal{V}_\lambda)$.
2. Let n be the dimension of $\mathcal{V}$. If $p_F(x)$ has n distinct solutions, then F is diagonalizable.

Morphisms - Diagonalization

Definition

A matrix $A \in \mathbb{R}^{n\times n}$ is called **diagonalizable** if A is similar to a diagonal matrix.

Proposition

Let A be a matrix in $\mathbb{R}^{n\times n}$, let $\mathcal{V}$ be a v.s. over $\mathbb{R}$ of dimension n. If V is a basis of $\mathcal{V}$ and F is a linear function $F : \mathcal{V} \to \mathcal{V}$ such that $A_{VV}(F) = A$, then:

$$F \text{ is diagonalizable} \Leftrightarrow A \text{ is diagonalizable.}$$

Morphisms - Diagonalization

Example: Consider the canonical basis E of $\mathbb{R}^3$ and the linear map $F : \mathbb{R}^3 \to \mathbb{R}^3$ having as transformation matrix $A_{EE}(F)$ the matrix:

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & 4 & -2 \end{bmatrix} .$$

The characteristic polynomial is

$$p_F(x) = det(A - xI_3) = det \begin{bmatrix} 2-x & 0 & 0 \\ 0 & 3-x & -1 \\ 0 & 4 & -2-x \end{bmatrix} = \cdots = -(x-2)^2(x+1) ;$$

then the eigenvalues are $\lambda_1 = 2, \lambda_2 = -1$ with multiplicities $m(\lambda_1) = 2$, $m(\lambda_2) = 1$.

Morphisms - Diagonalization

Example (continued):
Solving the linear systems $(A - \lambda_1 I_3)\boldsymbol{v} = \mathbf{0}$ and $(A - \lambda_2 I_3)\boldsymbol{v} = \mathbf{0}$ (see first item of the properties in slide 97), we find a basis for each one of the eigenspaces:

- $\{[1,0,0]^\top, [0,1,1]^\top\}$ for $\mathcal{V}_2$, then $dim(\mathcal{V}_2) = m(2)$;
- $\{[0,1,4]^\top\}$ for $\mathcal{V}_{-1}$, then $dim(\mathcal{V}_{-1}) = m(-1)$.

Then A is diagonalizable and we can write:

$$\underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}}_{A_{VV}(F)} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 4/3 & -1/3 \\ 0 & -1/3 & 1/3 \end{bmatrix}}_{M_{EV}} \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & 4 & -2 \end{bmatrix}}_{A=A_{EE}(F)} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 4 \end{bmatrix}}_{M_{VE}}$$

where $V := \{\underbrace{[1,0,0]^\top, [0,1,1]^\top}_{\text{basis of } \mathcal{V}_2}, \underbrace{[0,1,4]^\top}_{\text{bais of } \mathcal{V}_{-1}}\}$.

Morphisms - Diagonalization

Example:
Differently, the matrix

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

is not diagonalizable.
Indeed, the real root of the characteristic polynomial

$$det(A - xI_2) = det \begin{bmatrix} -x & 1 \\ 0 & -x \end{bmatrix} = x^2$$

is $\lambda_1 = 0$ with multiplicity 2 and $\{\boldsymbol{e}_1\}$ is a basis of $\mathcal{V}_0$ (i.e. $(A - xI_2)\boldsymbol{e}_1 = 0$).
Then,

$$m(0) = 2 \neq 1 = dim(\mathcal{V}_0).$$

Moreover, solving $(A - xI_2)\boldsymbol{v} = \boldsymbol{e}_1$ we find the generalized eigenvector $\boldsymbol{v} = \boldsymbol{e}_2$
Note that $(A - xI_2)^2\boldsymbol{e}_2 = (A - xI_2)\boldsymbol{e}_1 = 0$, i.e. a basis of $\mathbb{R}^2$ can be obtained finding a basis for $\ker((A - xI_2)^2)$ (generalized eigenspaces).
Note that $\langle \boldsymbol{e}_1, \boldsymbol{e}_2 \rangle = A\langle \boldsymbol{e}_1, \boldsymbol{e}_2 \rangle$, i.e. the generalized eigenspace for λ is A-invariant.

Morphisms - Diagonalization

Example:
The matrix

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 6 & 3 & 2 & 0 & 0 \\ 10 & 6 & 3 & 2 & 0 \\ 15 & 10 & 6 & 3 & 2 \end{bmatrix}$$

is not diagonalizable.
Compute, eigenvalues and generalized eigvectors. Let M the matrix with generalized eigenvectors as columns, compute $M^{-1}AM$.

- $\lambda(A) = 2,\ m(2) = 3,\ rk(A - 2I_5) = 4,\ dim(\mathcal{V}_2) = 1.$
- v_2-eigenvector of 2, $(A - 2I_5)w_2 = v_2,\ (A - 2I_5)z_2 = w_2.$
- $\mathcal{W}_2 =< v_2, w_2, z_2 >= A < v_2, w_2, z_2 >,\ < v_2, w_2, z_2 >= ker((A - 2I_5)^3),$
  $Av_2 = 2v_2,\ Aw_2 = 2w_2 + v_2,\ Az_2 = 2z_2 + w_2,$
  $A(\alpha v_2 + \beta w_2 + \gamma z_2) = \alpha 2v_2 + \beta(2w_2 + v_2) + \gamma(2z_2 + w_2).$
- $\lambda(A) = 1,\ m(1) = 2,\ rk(A - 1I_5) = 4,\ dim(\mathcal{V}_1) = 1.$
- v_1-eigenvector of 1, $(A - 1I_5)w_1 = v_1.$
- $\mathcal{W}_1 < v_1, w_1 >= A < v_1, w_1 >,\ < v_1, w_1 >= ker((A - 1I_5)^2).$
- $M = [v_2, w_2, z_2, v_1, w_1].$
- $$M^{-1}AM = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$
- The simple structure follows from a decomposition of $\mathbb{R}^n$ in subspaces A-invariant: $\mathbb{R}^n = \mathcal{W}_1 \oplus \mathcal{W}_2.$

Morphisms - Diagonalization

Diagonalization, Symmetric Matrices and Orthogonal Matrices:

Definition

A matrix $P \in \mathbb{R}^{n \times n}$ is called **orthogonal** if

$$P^\top P = PP^\top = I_n$$

(i.e. P orthogonal $\Leftrightarrow P \in GL_n(\mathbb{R})$ and $P^\top = P^{-1}$).

Spectral Theorem

Let A be a matrix in $\mathbb{R}^{n \times n}$. A is symmetric **if and only if** A is diagonalizable and there exists an orthogonal matrix $P \in \mathbb{R}^{n \times n}$ such that $P^\top AP$ is diagonal.

Symmetric matrices, orthogonal eigenvectors, different eigenvalues

Eigenvectors corresponding to distinct eigenvalues of a symmetric matrix must be orthogonal to each other.
Let $A = A^T$, $Ax = \lambda x$ and $Ay = \mu y$, with $\lambda \neq \mu$:

$$x^T \mu y = x^T A y = x^T A^T y = y^T A x = y^T \lambda x \to (\lambda - \mu) x^T y = 0.$$

Symmetric matrices, orthogonal eigenvectors, same eigenspace

Eigenvectors corresponding to the same eigenvalue need not be orthogonal to each other (they are somehow arbitrary in the same eigenspace). However, since every subspace has an orthonormal basis, you can find orthonormal bases for each eigenspace.
Let $A = A^T$, x and y unit vectors such that $Ax = \lambda x$, $Ay = \lambda y$, with $x \nparallel y$, and $v = \alpha x + \beta y \; \forall \alpha, \beta \in \mathbb{R}$:

$$Av = A(\alpha x + \beta y) = \alpha \lambda x + \beta \lambda y = \lambda v,$$

Let us take $z = y - (y^T x)x$, $z \perp x$,

$$Az = Ay - (y^T x)Ax = \lambda y - (y^T x)\lambda x = \lambda(y - (y^T x)x) = \lambda z.$$

Morphisms - Diagonalization

Diagonalization and Hermitian Matrices:

Definitions

- A matrix $A \in \mathbb{C}^{n \times n}$ is called **Hermitian**, or **self-adjoint**, if $A = A^*$, where A^* denotes the **conjugate**[a] **transpose** of A, i.e., the matrix $A^* := \bar{A}^\top$ such that $a^*_{i,j} = \bar{a}_{j,i}$.
- A matrix $U \in \mathbb{C}^{n \times n}$ is called **unitary** if $U^* U = UU^* = I_n$ (i.e. unitary matrices are the complex generalization of orthogonal matrices)

[a] $z = a + ib \in \mathbb{C}$, then the conjugate is $\bar{z} = a - ib$.

Spectral Theorem (Extension)

Let A be a matrix in $\mathbb{C}^{n \times n}$. A is Hermitian **if and only if** A is diagonalizable and there exists an unitary matrix $U \in \mathbb{C}^{n \times n}$ such that $U^* AU$ is diagonal and with **real** eigenvalues.

Morphisms - SVD

Theorem

Given $A \in \mathbb{C}^{m\times n}$, there exist three matrices $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$ and $\Sigma \in \mathbb{R}^{m\times n}$ such that

$$A = U\Sigma V^*,$$

and:

- U, V are **unitary**,
- Σ has non-null values only on the diagonal. In particular, the elements $\sigma_i := \sigma_{i,i}$ on the diagonal are such that $\sigma_i \in \mathbb{R}_{\geq 0}$.

N.B.: If A is matrix in $\mathbb{R}^{m\times n}$, then matrices U, V are **orthogonal** and

$$A = U\Sigma V^\top .$$

Morphisms - SVD

Definition A factorization $A = U\Sigma V^*$ is called the **singular value decomposition (SVD)** of A. Moreover:

- the non-null diagonal entries σ_i of Σ, for each $i = 1, \dots, \min(m, n)$, are called **singular values** of A;
- the columns $U_{\cdot,i} = \boldsymbol{u}_{\cdot,i}$ of U, for each $i = 1, \dots, m$, are called the **left-singular vectors** of A;
- the columns $V_{\cdot,j} = \boldsymbol{v}_{\cdot,j}$ of V (i.e. conjugate of rows of V^*), for each$j = 1, \dots, n$, are called **right-singular vectors** of A.

Morphisms - SVD

Example:
Let $A \in \mathbb{R}^{3\times 2}$ be the matrix

$$A := \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Then, a singular value decomposition of A is

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} \sqrt{2}/2 & 0 & \sqrt{2}/2 \\ \sqrt{2}/2 & 0 & -\sqrt{2}/2 \\ 0 & 1 & 0 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{bmatrix}}_{V^\top}.$$

Morphisms - SVD

Some SVD Properties: Let $A = U\Sigma V^*$ be an SVD of a matrix $A \in \mathbb{C}^{m\times n}$. We have that:

1. the singular values of A are the square roots of the eigenvalues of both AA^* and A^*A;
2. the left-singular vectors of A, i.e., the columns $\boldsymbol{u}_{\cdot,i}$ of U, are eigenvectors for AA^*;
3. the right-singular vectors of A, i.e., the columns $\boldsymbol{v}_{\cdot,i}$ of V, are eigenvectors for A^*A;
4. Let assume that $\sigma_1 \geq \cdots \geq \sigma_p \geq 0$ with $p = \min(m, n)$.
   If exists $k \in \mathbb{N}$ is such that $\sigma_k > 0$ and $\sigma_{k+1} = 0$, then:
   - $rk(A) = k$;
   - $\{\boldsymbol{u}_{\cdot,1}, \ldots, \boldsymbol{u}_{\cdot,k}\}$ is a basis of $\mathrm{Im}(A)$;
   - $\{\boldsymbol{v}_{\cdot,k+1}, \ldots, \boldsymbol{v}_{\cdot,p}\}$ is a basis of $\ker(A)$.