Mathematics for Machine Learning Part I (Linear Algebra)

K.Bouanane

7 mars 2023

Reminder

Systems of Linear Equations

Systems of Linear Equations

Definition

a linear system of equations is made up of a finite number **m** of linear equations relating a finite number **n** of unknowns $x_1, x_2, \ldots, x_n \in \mathbb{R}$: it is written in the form

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

each linear system has either a unique solution, or an infinity of solutions or no solution

Examples I

Examples :

- The system of equations :

$$\begin{cases} x_1 + x_2 + x_3 = 3 \\ x_1 - x_2 + 2x_3 = 2 \\ 2x_1 + 3x_3 = 1 \end{cases}$$

  has no solutions

Examples II

- The system of equations :

$$\begin{cases} x_1 + x_2 + x_3 = 3 \\ x_1 - x_2 + 2x_3 = 2 \\ x_2 + x_3 = 2 \end{cases}$$

has a unique solution which is **(1,1,1)**, $x_1 = x_2 = x_3 = 1$

- The system of equations :

$$\begin{cases} x_1 + x_2 + x_3 = 3 & (1) \\ x_1 - x_2 + 2x_3 = 2 & (2) \\ 2x_1 + 3x_3 = 5 & (3) \end{cases}$$

Examples III

Note that the equation (3) is redundant since (3) = (1) + (2),so we can write $2_x1 = 5 - 3x_3$ and $2x_2 = 1 + x_3$ So if we put $x_3 = a \in \mathbb{R}$,we have $\left(\frac{5}{2} - \frac{3}{2}a; \frac{1}{2} + \frac{1}{2}a; a\right)$ is solution of the system. So we have an infinite number of solutions.

Systems of Linear Equations

A compact way to write a linear system is as follows :

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} \cdot x_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} \cdot x_2 + \cdots + \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \cdot x_n = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$\iff \underbrace{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{X} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}}_{b}$$

Matrices

Definition

Matrices I

Let $\mathbf{m}, \mathbf{n} \in \mathbb{N}$.
An $(\mathbf{m} \times \mathbf{n})$ matrix $\mathbf{A}$ is a collection of $(\mathbf{m} \times \mathbf{n})$ elements $a_{ij}, i = 1 \ldots m, j = 1 \ldots n$ consisting of $\mathbf{m}$ rows and $\mathbf{n}$ columns.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \quad \textit{With } a_{ij} \in \mathbb{R}$$

We can then write $\mathbf{A} \in \mathbb{R}^{m \times n}$
By convention, a $(1 \times n)$ matrix is called a row vector and an $(m \times 1)$ matrix is called a column vector

Addition and Multiplication of Matrices

Addition and Multiplication of Matrices I

1. Let $\mathbf{A}, \mathbf{B}$,two $(\mathbf{m} \times \mathbf{n})$ matrices,Then $\mathbf{A} + \mathbf{B}$ denotes the sum matrix of $\mathbf{A}$ and $\mathbf{B}$ such that

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{bmatrix}$$

2. For the matrices $\mathbf{A} \in \mathbb{R}^{m \times k}$ and $\mathbf{B} \in \mathbb{R}^{k \times n}$, the product of $\mathbf{A}$ and $\mathbf{B}$,Let $\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{m \times n}$ is defined with $c_{ij} = \sum_{l=1}^{k} a_{il} \cdot b_{lj}$ For $i = 1 \dots n, j = 1 \dots n$ We notice that $\mathbf{AB} \neq \mathbf{BA}$

Addition and Multiplication of Matrices II

3. the identity matrix, denoted $\mathbf{I} \in \mathbb{R}^{n \times n}$ is such that

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

the elements on the diagonal = 1 and 0 elsewhere

matrix properties

matrix properties I

- Associativity :
  $\forall \mathbf{A} \in \mathbb{R}^{m \times p}, \mathbf{B} \in \mathbb{R}^{p \times q}, \mathbf{C} \in \mathbb{R}^{q \times n}$, We have
  $(\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C})$
- Distributivity :
  $\forall \mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times p}, \mathbf{C}, \mathbf{D} \in \mathbb{R}^{p \times n}$ We have :

$$(\mathbf{A} + \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot \mathbf{C} + \mathbf{B} \cdot \mathbf{C}$$
$$\mathbf{A} \cdot (\mathbf{C} + \mathbf{D}) = \mathbf{A} \cdot \mathbf{C} + \mathbf{A} \cdot \mathbf{D}$$

span

- Multiplication by the identity matrix :
  $\forall \mathbf{A} \in \mathbb{R}^{m \times n}$, we have $\mathbf{I}_n \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{I}_n = \mathbf{A}$

Inverse and transpose of a matrix

Inverse and transpose of a matrix I

Definition

We consider a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$

Let the matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$ be such that $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n = \mathbf{B} \cdot \mathbf{A}$ $\mathbf{B}$ is called the inverse matrix of $\mathbf{A}$ is denoted $\mathbf{A}^{-1}$

Inverse and transpose of a matrix II

Remarks

1. The inverse matrix $\mathbf{A}^{-1}$ may not exist by a given square matrix $\mathbf{A}$
2. In the case where $\mathbf{A}^{-1}$ exists , $\mathbf{A}$ is said to be

   non-singular / invertible / regular

   Otherwise A is called non-invertible / singular
3. Consider the matrix $\mathbf{A} \in \mathbb{R}^{2\times 2}$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Inverse and transpose of a matrix III

Remarks

if we multiply **A** with

$$\grave{\mathbf{A}} = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

We obtain

$$\mathbf{A}\grave{\mathbf{A}} = \begin{bmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix} = (a_{11}a_{22} - a_{12}a_{21}) \cdot \mathbf{I}$$

Then

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

If and only if $a_{11}a_{22} - a_{12}a_{21} \neq 0$

Inverse and transpose of a matrix IV

Definition

For $\mathbf{A} \in \mathbb{R}^{m \times n}$, the matrix $\mathbf{B} \in \mathbb{R}^{n \times m}$ with $b_{ij} = a_{ji}$, is called the transpose of $\mathbf{A}$, we write $\mathbf{B} = \mathbf{A}^T$

Inverse and transpose of a matrix V

Definition

A matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric if $\mathbf{A} = \mathbf{A}^T$

Remark

1. The sum of two symmetric matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ is symmetric.
2. The product of two symmetric matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ is not always symmetric

Multiplication by a scalar

Multiplication by a scalar I

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\lambda \in \mathbb{R}$. So $\lambda \mathbf{A} = \mathbf{K}$ with $\mathbf{K}_{ij} = \lambda a_{ij}$ We have the following properties for $\lambda, \psi \in \mathbb{R}$, we have :

1. Associativity :
   1. $(\lambda\psi) \cdot \mathbf{C} = \lambda(\psi\mathbf{C}), with \mathbf{C} \in \mathbb{R}^{m \times n}$
   2. $\lambda(\mathbf{BC}) = (\lambda\mathbf{B}) \cdot \mathbf{C} = \mathbf{B}(\lambda\mathbf{C}) = (\mathbf{BC}) \cdot \lambda$ with $\mathbf{B} \in \mathbb{R}^{m \times k}; \mathbf{C} \in \mathbb{R}^{k \times n}$
   3. $(\lambda\mathbf{C} =^{T} = \mathbf{C}^{T}\lambda^{T} = \mathbf{C}^{T}\lambda = \lambda\mathbf{C}^{T}$ (since $\lambda = \lambda^{T}$)
2. Distributivity :
   1. $(\lambda + \psi)\mathbf{C} = \lambda\mathbf{C} + \psi\mathbf{C}, \mathbf{C} \in \mathbb{R}^{m \times n}$
   2. $\lambda(\mathbf{B} + \mathbf{C}) = \lambda\mathbf{B} + \lambda\mathbf{C}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{m \times n}$

Multiplication by a scalar II

Example (Distributivity) :

If we define $\mathbf{C} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ then for all $\lambda, \psi \in \mathbb{R}$, we get

$$
\begin{aligned}
(\lambda + \psi)\mathbf{C} &= (\lambda + \psi) \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \\
&= \begin{bmatrix} (\lambda + \psi) & (\lambda + \psi) \times 2 \\ (\lambda + \psi) \times 3 & (\lambda + \psi) \times 4 \end{bmatrix} = \begin{bmatrix} \lambda + \psi & 2\lambda + 2\psi \\ 3\lambda + 3\psi & 4\lambda + 4\psi \end{bmatrix} \text{span} \\
&= \begin{bmatrix} \lambda & 2\lambda \\ -3\lambda & 4\lambda \end{bmatrix} + \begin{bmatrix} \psi & 2\psi \\ 3\psi & 4\psi \end{bmatrix} \\
&= \lambda \cdot \mathbf{C} + \psi \cdot \mathbf{C}
\end{aligned}
$$

Solving systems of linear equations

Notation

$$\mathbf{AX} = \mathbf{b} with \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{X} \in \mathbb{R}^{n} and \mathbf{b} \in \mathbb{R}^{m}$$

Particular solution and general solution

Particular solution and general solution I

Consider the following system of equations :

$$\begin{bmatrix} 1 & 0 & 8 & -4 \\ 0 & 1 & 2 & 12 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 42 \\ 8 \end{bmatrix} \Leftrightarrow \sum x_i c_i = b$$

where C_i denotes the column of **A** A solution to this system can be directly deduced : taking 42 do this column 1 and 8 do column 2, so

$$b = \begin{bmatrix} 42 \\ 8 \end{bmatrix} = 42 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 8 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

So a solution is $[42, 8, 0, 0]^T$.

Particular solution and general solution II

This type of solution is called particular solution or special solution.

This solution is not unique. To formulate all the solutions of the system, we need to generate 0 in a non-trivial way using the columns of the matrix. To do this, we express column 3 using the first 2 columns

$$\begin{bmatrix} 8 \\ 2 \end{bmatrix} = 8 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Then $8C_1 + 2C_2 - 1 \times C_3 + 0 \times C_4 = 0$ with C_1, C_2, C_3, C_4 denote the 4 columns of the matrix and $(8, 2, -1, 0)^T$ is a

Particular solution and general solution III

solution Then $\lambda_1 \begin{bmatrix} 8 \\ 2 \\ -1 \\ 0 \end{bmatrix}$, $\forall \lambda_1 \in \mathbb{R}$, is also a system solution,because

$$\begin{bmatrix} 1 & 0 & 8 & -4 \\ 0 & 1 & 2 & 12 \end{bmatrix} \left(\lambda_1 \begin{bmatrix} 8 \\ 2 \\ -1 \\ 0 \end{bmatrix} \right) = \lambda_1 (8C_1 + 2C_2 - C_3) = 0$$

By analogy, we treat column 4 of the matrix in the same way, using the first two columns, we thus generate another set of

Particular solution and general solution IV

solutions :

$$-4\begin{bmatrix}1\\0\end{bmatrix}+12\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}-4\\12\end{bmatrix}$$

$$\Rightarrow -4C_1+12C_2+0\cdot C_3-C_4=0$$

So then we have :

$$\begin{bmatrix}1 & 0 & 8 & -4\\0 & 1 & 2 & 12\end{bmatrix}\left(\lambda_2\begin{bmatrix}-4\\12\\0\\-1\end{bmatrix}\right)=0$$

Particular solution and general solution V

For $\lambda_2 \in \mathbb{R}$. We can then write the set :

$$\left\{ x \in \mathbb{R}^n : x = \begin{bmatrix} 42 \\ 8 \\ 0 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 8 \\ 2 \\ -1 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} -4 \\ 12 \\ 0 \\ -1 \end{bmatrix}, \lambda_1, \lambda_2 \in \mathbb{R} \right\}.$$

which is the set of all solutions of the linear system called general solution

The adopted approach consists of :

- look for a solution specific to $\mathbf{AX} = \mathbf{b}$
- look for all solutions of $\mathbf{Ax} = 0$

Particular solution and general solution VI

- Combine the solutions obtained in 1/ and 2/ to form the general solution

Remark

1. Neither the particular solution nor the general solution is unique.
2. In general obtaining the general solution is not as simple as in the example
   Previous on the form of the matrix has made it possible to obtain a special solution and the general solution easily
   We present in what follows a technique which will allow to transform a linear system into a simpler form, Gaussian elimination

Elementary transformations

Elementary transformations I

Solving a system of linear equations involves a team of elementary transformations

It is :

1. Swap two equations (rows of matrix **A**)
2. Multiplication of an equation (line) by a constant $\lambda \in \mathbb{R} \backslash \{0\}$
3. Addition of 2 equations (2 lines)

Example :

For $a \in \mathbb{R}$, we seek all the relations of the system of equations :

$$\begin{cases} -2x_1 + 4x_2 - 2x_3 - x_4 + 4x_5 = -3 \\ 4x_1 - 8x_2 + 3x_3 - 3x_4 + x_5 = 2 \\ x_1 - 2x_2 + x_3 - x_4 + x_5 = 0 \\ x_1 - 2x_2 - 3x_4 + 4x_5 = a \end{cases}$$

Elementary transformations II

We define the augmented matrix of a system $[\mathbf{A}|\mathbf{b}]$:

$$\begin{array}{c} L_1 \\ L_2 \\ L_3 \\ L_4 \end{array} \left[\begin{array}{ccccc|c} -2 & 4 & -2 & -1 & 4 & -3 \\ 4 & -8 & 3 & -3 & 1 & 2 \\ 1 & -2 & 1 & -1 & 1 & 0 \\ 1 & -2 & 0 & -3 & 4 & a \end{array}\right]$$

- Change L_1 and L_3

$$\left[\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 0 \\ 4 & -8 & 3 & -3 & 1 & 2 \\ -2 & 4 & -2 & -1 & 4 & -3 \\ 1 & -2 & 0 & -3 & 4 & a \end{array}\right] \rightsquigarrow \begin{array}{r} L_1 \\ -4L_1 \\ +2L_1 \\ -L_1 \end{array} \left[\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & -1 & -2 & 3 & a \end{array}\right]$$

Elementary transformations III

$$\left[\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & 0 & 0 & 0 & a+1 \end{array}\right] \rightsquigarrow$$

$$\begin{array}{r} \\ \cdot(-1) \\ \cdot(-1/3) \\ -L_2 - L_3 \end{array} \left[\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & a+1 \end{array}\right]$$

The augmented matrix $[\mathbf{A}|\mathbf{b}]$ is said to be in echelon form. Only by $a = -1$, the system can be solved a particular solution is :

Elementary transformations IV

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}$$ The general solution is then given by

$$\left\{ x \in \mathbb{R}^5 : x = \begin{pmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 0 \\ -1 \\ 2 \\ 1 \end{pmatrix}, \lambda_1, \lambda_2 \in \mathbb{R} \right\}$$

Elementary transformations V

Remark

The first nonzero coefficient of a row is called pivot is called pivot and is always denoted by the pivot of the line which divides it.

Definition

A matrix is said to be in echelon form if :

1. All rows with zero coefficients are in the lower part of the matrix
2. A row that contains at least one nonzero element precedes rows with zero coefficients
3. The pivot of a line with at least one nonzero element is to the right of the pivots of the line that precedes it.

Elementary transformations VI

Remark

1. The variables that correspond to the pivots of a ladder matrix are called base variables. the other variables are called non-base variables or free variables
2. The echelonnee form makes it easy to obtain a particular solution : We express The part denoted $\tilde{b}$ in the form $\sum_{i=1}^{p} \alpha_i P_i$ or P_i are the columns corresponding to the base variables .

We determine the coefficients α_i begin with the rightmost pivot column. for the previous example :
We search $\alpha_1, \alpha_2, \alpha_3$ where

$$\alpha_1 \begin{matrix} 1 \\ 0 \\ 0 \\ 0 \end{matrix} + \alpha_2 \begin{matrix} 1 \\ 1 \\ 0 \\ 0 \end{matrix} + \alpha_3 \begin{matrix} -1 \\ -1 \\ 1 \\ 0 \end{matrix} = \begin{matrix} 0 \\ -2 \\ 1 \\ 0 \end{matrix}$$

Elementary transformations VII

Which gives $\alpha_3 = 1, \alpha_2 = -1, \alpha_1 = 2$ Hence the particular solution $x = [1, 0, -1, 1, 0]^T$

Definition

A system of equations is said to be in reduced scale form if :

- it is written in staggered form
- each pivot element is equal to 1
- The pivot is the only nonzero element in its column

Remark

The reduced step form provides the general solution for a system of linear equations. To obtain the general solution, we solve the system $\mathbf{A}x = 0$. We will need to express the non-pivot columns as a linear combination of the pivot columns.

Elementary transformations VIII

Example :

Let The Matrix $\begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \end{bmatrix}$ Pivot columns are $P_1, P_3 and P_4$

$$P_2 = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 3 \cdot P_1$$

Elementary transformations IX

Then $3 \cdot P_1 - P_2 \Rightarrow \begin{pmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

$$P_5 = \begin{pmatrix} 3 \\ 9 \\ -4 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 9 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - 4 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\Rightarrow \; P_5 \; = \; 3 \cdot P_1 + 9P_3 - 4P_4 \; \Rightarrow \; 3P_1 + 9P_3 - 4P_4 - P_5 \; = \; 0$$

Elementary transformations X

Then we have $\begin{pmatrix} 3 \\ 0 \\ 9 \\ -4 \\ -1 \end{pmatrix}$
Hence the general solution is written

$$\left\{ x \in \mathbb{R}^5 / x = \lambda_1 \begin{bmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 3 \\ 0 \\ 9 \\ -4 \\ -1 \end{bmatrix}, \lambda_1, \lambda_2 \in \mathbb{R} \right\}$$

Elementary transformations XI

Remark

We consider a matrix **A** in reduced echelon form such that all its rows contain at least one nonzero element with **k** rows and **n** columns ($\mathbf{A} \in \mathbb{R}^{k\times n}$).

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

Let $J_1, J_2, \dots, J_k$ be the pivot columns We want to solve the system $\mathbf{A}x = 0$ with $x \in \mathbb{R}^n$ Note that $J_1, J_2, \dots, J_k$ stands for unit vectors. We extend **A** to an $n \times n$ matrix $\tilde{\mathbf{A}}$ by joining $n-k$ rows of the form $[0, \dots, 0, -1, 0, \dots, 0]$ such that the diagonal of $\tilde{\mathbf{A}}$ is only elements 1 or -1.

Elementary transformations XII

Then the columns of $\tilde{\mathbf{A}}$ which contains -1 as pivot element correspond to the solutions of the homogeneous system $\mathbf{A}x = 0$.
Example : Let's go back to the previous example

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ \color{red}{0} & \color{red}{-1} & \color{red}{0} & \color{red}{0} & \color{red}{0} \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \\ \color{red}{0} & \color{red}{0} & \color{red}{0} & \color{red}{0} & \color{red}{-1} \end{bmatrix}$$

We directly obtain the solutions of the system $\mathbf{A}x = 0$

$$\left\{ x \in \mathbb{R}^5 : x = \lambda_1 \begin{matrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \end{matrix} + \lambda_2 \begin{bmatrix} 3 \\ 0 \\ 9 \\ -4 \\ -1 \end{bmatrix}, \lambda_1, \lambda_2 \in \mathbb{R} \right\}$$

Elementary transformations XIII

Remark

The algorithm which makes it possible to obtain the reduced echelon form of a given matrix by a sequence of successive transformations is called **Gaussian elimination** or **Gauss's method**.

Calculation of the inverse

Calculation of the inverse I

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$, To calculate $\mathbf{A}^{-1}$, we seek a matrix $\mathbf{X}$ such that $\mathbf{AX} = \mathbf{I}_n$ We can therefore write the set of linear systems $\mathbf{AX} = \mathbf{I}_n$ where $\mathbf{X} = [x_1|.x_2|\ldots|x_n]$. Representing the system by an augmented matrix $[\mathbf{A}|\mathbf{I}_n]$, the idea is to proceed to a Gaussian elimination by transforming the augmented matrix $[\mathbf{A}|\mathbf{I}_n]$ to $[\mathbf{I}_n|\mathbf{A}]$ Example :

We want to calculate the inverse of the matrix

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

Calculation of the inverse II

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\rightsquigarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -1 & 2 & -2 & 2 \\ 0 & 1 & 0 & 0 & 1 & -1 & 2 & -2 \\ 0 & 0 & 1 & 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 1 & -1 & 0 & -1 & 2 \end{array}\right]$$

So we deonte the inverse Matrix as :

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

Vector spaces

Groups

Groups I

We consider a set $\mathcal{G}$ and a map $+$ such that $+ : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$.
Then $\mathbf{G} = (\mathcal{G}, +)$ is called a group if we have :

- Closure $\mathcal{G}$ under $"+"$: $\forall x, y \in \mathcal{G} : x + y \in \mathcal{G}$
- Associativity : $\forall x, y, z \in \mathcal{G} : (x + y) + z = x + (y + z)$
- Neutral element : $\exists e \in \mathcal{G}, \forall x \in \mathcal{G} : x + e = x$ and $e + x = x$
- Inverse element : $\forall x \in \mathcal{G}, \exists y \in \mathcal{G} : x + y = e$ and $y + x = e$

We denote the inverse element of x by x^{-1} If, moreover, we have the commutative property $\forall x, y \in \mathcal{G} : x + y = y + x$, then $\mathbf{G} = (\mathcal{G}, +)$ is an Abelian group

Examples :

- $(\mathbb{Z}, +)$is an abelian group
- $(\mathbb{R}, \cdot)$ is not a group (because 0 has no iverse element)

Groups II

- $(\mathbb{R}^{n\times n}, +)$ is an Abelian group
- We consider $(\mathbb{R}^{n\times n}, \cdot)$ the set of $n \times n$ matrices modified by the multiplication operation

the closure and associativity conditions follow from the definition of the multiplication operation.

1. The matrix $\mathbf{I}$ is the neutral element
2. If $\mathbf{A}$ is regular then $\mathbf{A}^{-1}$ is the inverse element of $\mathbf{A}$ and in this case only, $(\mathbb{R}^{n\times n}, \cdot)$ is a group called e general linear group note $GL(n, \mathbb{R})$ this group is not Abelian.

Vector spaces

Vector spaces

Vector spaces I

Definition

A vector space $\mathbf{V} = (\mathcal{V}, +, \cdot)$ is a set $\mathcal{V}$ with the operations :

$$+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$$

$$\cdot : \mathbb{R} \times \mathcal{V} \to \mathcal{V}$$

where

- $(\mathcal{V}, +)$ is an Abelian group
- Distributivity :
  - $\forall \lambda \in \mathbb{R}, x, y \in \mathcal{V} : \lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$
  - $\forall \lambda \psi \in \mathbb{R}, x \in \mathcal{V} : (\lambda + \psi) \cdot x = \lambda \cdot x + \psi \cdot x$
- Associativity of the operation $(\cdot)$:
  $\forall \lambda, \psi \in \mathbb{R}, x \in \mathcal{V}, \lambda \cdot (\psi \cdot x) = (\lambda \cdot \psi) \cdot x$

Vector spaces II

- The neutral element with respect to $(\cdot)$:
  $\forall x \in \mathcal{V} : 1 \cdot x = x$

1. The elements $x \in \mathcal{V}$ were called vectors
2. The neutral element of $(\mathcal{V}, +)$ is the vector
   $0 = [0, 0, \ldots, 0]^T$

Remark

In what follows, the vector space $(\mathbf{V}, +, \cdot)$ is denoted **V.** or the operations $+$ and $\cdot$ denote the operations of addition and multiplication by a scalar

Examples :

$\mathcal{V} = \mathbb{R}^n, n \in \mathbb{N}$ is a vector space with the operations

Vector spaces III

1. Addition : $x + y = (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$
2. Multiplication by scalar : $\lambda x = \lambda(x_1, \ldots, x_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$ for all $\lambda \in \mathbb{R}, x \in \mathbb{R}^n$
3. $\mathcal{V} = \mathbb{R}^{m \times n}, m, n \in \mathbb{N}$ is an V.S with :
    1. Addition : $\mathbf{A} + \mathbf{B}$
    2. Multiplication by scalar : $\lambda \mathbf{A}$

    Note that $\mathbb{R}^{m \times n}$ is equivalent to $\mathbb{R}^{mn}$
4. $\mathcal{V} = \mathbb{C}$

Vector spaces IV

Remark

The vector spaces $\mathbb{R}^n$, $\mathbb{R}^{n\times 1}$, $\mathbb{R}^{1\times n}$ are only different in the way we write vectors. In the following, we will not make a distinction between $\mathbb{R}^n$ and $\mathbb{R}^{n\times 1}$,which allows us to write n-tuples as column vectors $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

We also distinguish other $\mathbb{R}^{n\times 1}$ and $\mathbb{R}^{1\times n}$ (line vectors)

Vector Subspaces

Vector Subspaces I

Definition

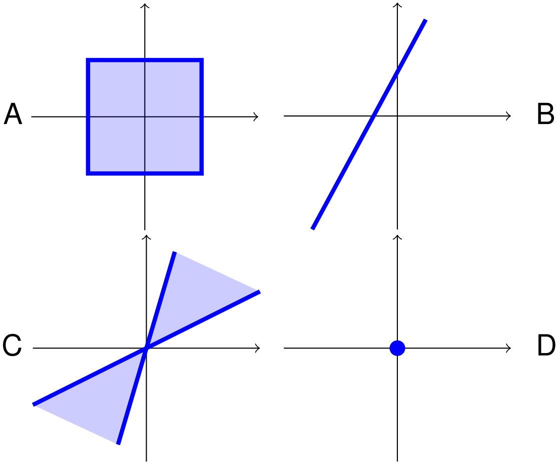
Let $\mathbf{V} = (\mathcal{V}, +, \cdot)$ be a V.S and $\mathcal{U} \subseteq \mathcal{V}$, $\mathcal{U} \neq \emptyset$. Then $\mathbf{U} = (\mathcal{U}, +, \cdot)$ is a vector subspace if $\mathbf{U}$ is a V.S with the operations $+$ and $\cdot$ restrictions to $\mathcal{U} \times \mathcal{U}$ and $\mathbb{R} \times \mathcal{U}$ In other words :

- $\mathcal{U} \neq \emptyset$: in particular $0 \in \mathbf{U}$
- Closure of $\mathbf{U}$:
  1. $\forall \lambda \in \mathbb{R}, \forall x \in \mathcal{U} : \lambda x \in \mathcal{U}$
  2. $\forall x, y \in \mathcal{U} : x + y \in \mathcal{U}$

Examples :

1. For any V.S $\mathbf{V}$; $\mathbf{V}$ and 0 are subspaces
2. The set $\mathbf{D}$ and a subspace of $\mathbb{R}^2$

Vector Subspaces II



Vector Subspaces III

1. The solution set of a homogeneous linear system $\mathbf{A}x = 0$ with n unknowns $\left\{x = [x_1, x_2, \ldots, x_n]^T : \mathbf{A}x = 0\right\}$ is a subspace of $\mathbb{R}^n$
2. The solution set of an inhomogeneous linear system $\mathbf{A}x = b$, $b \neq 0$ is not a subspace of $\mathbb{R}^n$
3. The intersection of several subspace is a subspace

Remark

Each Subspace $\mathbf{U} \subseteq \mathbf{V} = (\mathbb{R}^n, +, \cdot)$ is a solution of a system of linear equations $\mathbf{A}x = 0$ for $x \in \mathbb{R}n$

Linear Independence

Linear Independence I

Definition

We consider a V.S $\mathbf{V}$ and a finite number of vectors $x_1, \ldots, x_k \in \mathbf{V}$. Then each element $v \in \mathbf{V}$ such that

$$v = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k \text{ with } \lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$$

is a linear combination of $x_1, x_2, \ldots, x_k$

Linear Independence II

Definition

Let $\mathbf{V}$ be V.S and $x_1, \ldots, x_k \in \mathbf{V}$ with $k \in \mathbb{N}$ If there exists a linear combination such that $\sum \lambda_i x_i = 0$ with at least $\lambda_i \neq 0$ then $x_1, x_2, \ldots, x_n$ are said to be linearly dependent if $\sum \lambda_i x_i = 0$ implies that $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$ then the vectors $x_1, x_2, \ldots, x_k$ are said to be linearly independent

Properties :

1. If at least one vector $x_j = 0$ for $j \in \{1 \ldots k\}$ then $x_1, x_2, \ldots, x_k$ are linearly dependent
2. The same for $x_i = x_j$ by $i, j \in \{1, 2, \ldots, k\}$

Linear Independence III

3. The vectors $\{x_1, \ldots, x_k; x_i \neq 0, i = 1 \ldots k\}, k \geqslant 2$ are linearly dependent if and only if at least one of them can be written as a linear combination of the others. In particular if $x_i = \lambda x_j$, $\lambda \in \mathbb{R}$ then $\{x_1, x_2, \ldots, x_k\}$ are linearly dependent
4. One way to check if a set of vectors $x_1, x_2, \ldots, x_k \in \mathbf{V}$ is linearly independent is to use Gaussian elimination. It suffices to write $x_1, x_2, \ldots, x_k$ As columns of a matrix $\mathbf{A}$ that we write in its step form
5. Pivot columns correspond to vectors which are linearly independent, while non-pivot columns correspond to vectors which can be expressed as a linear combination of pivot columns.

Linear Independence IV

Example :

Consider $\mathbb{R}^4$ with $x_1 = \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix}$ $x_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}$ $x_3 = \begin{bmatrix} -1 \\ -2 \\ 1 \\ 1 \end{bmatrix}$ The we write $\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 = 0$

$$\Leftrightarrow \begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & -2 \\ -3 & 0 & 1 \\ 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Linear Independence V

$$\begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & -2 \\ -3 & 0 & 1 \\ 4 & 2 & 1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 3 & -2 \\ 0 & -2 & 5 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 1 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 5 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 1 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

The only solution to this system is I (each column is a pivot column) so x_1, x_2, x_3 are linearly independent

Linear Independence VI

Remark

We consider a V.S with k vectors $b_1, b_2, \ldots, b_k$ linearly independent, and m linear combinations :

$$
\begin{array}{c}
x_1 = \sum_{i=1}^{k} \lambda_{i1} \cdot b_i \\
x_2 = \sum_{i=1}^{k} \lambda_{i2} \cdot b_i \\
\vdots \\
x_m = \sum_{i=1}^{k} \lambda_{im} \cdot b_i
\end{array}
$$

Linear Independence VII

Remark

By defining $\mathbf{B} = [b_1, b_2, \ldots, b_k]$ we can write $x_j = \mathbf{B} \cdot \lambda_j$ with $\lambda_j = \begin{bmatrix} \lambda_{1j} \\ \vdots \\ \lambda_{kj} \end{bmatrix}$ Then :

$$\sum_{j=1}^{m} \psi_j x_j = \sum_{j=1}^{m} \psi_j \cdot \mathbf{B} \cdot \lambda_j = \mathbf{B} \cdot \sum_{j=1}^{m} \psi_j \cdot \lambda_j$$

Hence $\{x_1, x_2, \ldots, x_m\}$ are linearly independent if $\{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ are linearly independent

Example :

Linear Independence VIII

We consider $b_1, b_2, b_3, b_4 \in \Bbbk$ and

$$x_1 = b_1 - 2b_2 + b_3 - b_4$$
$$x_2 = -4b_1 - 2b_2 + 4b_4$$
$$x_3 = 2b_1 + 3b_2 - b_3 - 3b_4$$
$$x_4 = 17b_1 - 10b_2 + 11b_3 + b_4$$

$$\begin{bmatrix} 1 & -4 & 2 & 17 \\ -2 & -2 & 3 & -10 \\ 1 & 0 & -1 & 11 \\ -1 & 4 & -3 & 1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & -15 \\ 0 & 0 & 1 & -18 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Linear Independence IX

We can therefore write $x_4 = -7x_1 - 15x_2 - 18x_3$ Hence x_1, x_2, x_3 and x_4 are linearly dependent

Basis and Rank

Generating Set and Basis

Linear Independence I

Definition

(Generating Set) We consider a V.S $\mathbf{V} = (\mathcal{V}, +, \cdot)$ and a set of vectors $\mathcal{A} = \{x_1, \ldots, x_k\} \subseteq \mathcal{V}$. if any vector $v \in \mathcal{V}$ can be expressed as a linear combination of vectors of $\mathcal{A}$, then $\mathcal{A}$ is called the generating set of **V**.

1. The set of all linear combinations of the vectors of $\mathcal{A}$ is called the subspace spanned by $\mathcal{A}$ and is denoted by $span[\mathcal{A}]$
2. if the vector space **V** is spanned by $\mathcal{A}$ then $\mathbf{V} = span[\mathcal{A}]$

Linear Independence II

Definition

Basis Let $\mathbf{V} = (\mathcal{V}, +, \cdot)$ be a vector space and $\mathcal{A} \subseteq \mathcal{V}$. a generating set $\mathcal{A}$ of $\mathbf{V}$ is called minimal if *span*$[\mathcal{A}] = \mathbf{V}$ and $\forall \bar{A} \subseteq \mathcal{A}$, $\mathbf{V} \neq$ *span*$[\mathcal{A}]$ If $\mathcal{A}$ consists of linearly independent vectors then $\mathcal{A}$ is called the Basis of $\mathbf{V}$.

The following properties are then equivalent :

1. $\mathcal{B} \subseteq \mathcal{V}$ is a basis of $\mathbf{V}$
2. $\mathcal{B}$ is a minimal generator set of $\mathbf{V}$
3. $\mathcal{B}$ is a set of linearly independent vectors which is maximal
4. Each vector $x \in \mathbf{V}$ is a linear combination of the vectors of $\mathcal{B}$ and each linear combination is unique

Examples :

Linear Independence III

1. In $\mathbb{R}^3$, the set of vectors $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ forms a basis called the canonical/standard basis
2. The set $\mathbb{R}^3$, the set of vectors $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ forms another basis
3. The set $\mathcal{A} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ -4 \end{bmatrix} \right\}$ is linearly independent, but not a generating set (and no basis) of

Linear Independence IV

$\mathbb{R}^4$:For instance, the vector $[1, 0, 0, 0]^T$ cannot be obtained by a linear combination of elements in $\mathcal{A}$.

Remarks

1. Each vector space has at least one basis, can there be more than one ? However, all the bases are the m number of elements, called the demension of **V** and we write $dim(\mathbf{V})$
2. If $\mathbf{U} \subseteq \mathbf{V}$ a vector subspace of $\mathbf{V}$ then $dim(\mathbf{U}) \leqslant dim(\mathbf{V})$ and $dim(\mathbf{U}) = dim(\mathbf{V})$ if and only if $\mathbf{U} = \mathbf{V}$

Linear Independence V

Remarks

1. The dimension of a vector space is not necessarily the number of elements in a vector. The vector space $span[\begin{bmatrix} 0 \\ 1 \end{bmatrix}]$ is of dimension 1
2. We can obtain a basis pone a vector subspace $\mathbf{U} = span[x_1, x_2, \ldots, x_n] \subseteq \mathbb{R}^n$ by the following steps :
   - Write the matrix $\mathbf{A}$ to constitute vectors $x_1, x_2, \ldots, x_n$ as columns of $\mathbf{A}$.
   - Determine the echelon form of $\mathbf{A}$.
   - The pivot columns correspond to the basis vectors

Linear Independence VI

Example : Let the vectorial subspace $\mathbf{U} \subseteq \mathbb{R}^5$ generated by the vectors

$$x_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ -1 \\ -1 \end{bmatrix} \quad x_2 = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 2 \\ -2 \end{bmatrix} \quad x_3 = \begin{bmatrix} 3 \\ -4 \\ 3 \\ 5 \\ -3 \end{bmatrix} \quad x_4 = \begin{bmatrix} -1 \\ 8 \\ -5 \\ -6 \\ 1 \end{bmatrix}$$

Let The Matrix

$$\begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & -1 & -4 & 8 \\ -1 & 1 & 3 & -5 \\ -1 & 2 & 5 & -6 \\ -1 & -2 & -3 & 1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} \downarrow & \downarrow & & \downarrow \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Linear Independence VII

so x_1, x_2 and x_4 are linearly independent since
$\lambda_1 x_1 + \lambda_2 x_2 + \lambda_4 x_4 = 0 \;\Rightarrow\; \lambda_1 = \lambda_2 = \lambda_4 =$

Rank

Rank I

The number of linearly independent columns in a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is equal to the number of linearly independent rows and is called rank of $\mathbf{A}$, denoted $rk(\mathbf{A})$.

Property :

- $rk(\mathbf{a}) = rk(\mathbf{A}^T)$
- The columns of $\mathbf{A} \in \mathbb{R}^{m \times n}$ eugement a vector subspace $\mathbf{U} \subseteq \mathbb{R}^m$ with $dim(\mathbf{U}) = rk(\mathbf{A})$
- The lines of $\mathbf{A} \in \mathbb{R}^{m \times n}$ grow a vector subspace $\mathbf{W} \subseteq \mathbb{R}^n$ with $dim(\mathbf{W}) = rk(\mathbf{A})$. A basis can be obtained by applying a Gaussian elimination to $\mathbf{A}^T$
- For any $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{A}$ is regular if $rk(\mathbf{A}) = n$

Rank II

- For any matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^{m}$, The linear system $\mathbf{A}x = b$ can be solved if $rk(\mathbf{A}) = rk(\mathbf{A}|\mathbf{B})$ where $\mathbf{A}|\mathbf{B}$ is the augmented matrix

Rank III

Example :

1. Let $\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

   $\mathbf{A}$ has 2 linearly independent columns/rows so $rk(\mathbf{A}) = 2$

2. $\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ -2 & -3 & 1 \\ 3 & 5 & 0 \end{pmatrix}$

   By Gaussian elimination we have :

   $\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$. Then $rk(\mathbf{A}) = 2$

Linear Mappings

Linear Mappings I

Definition

Let $\mathbf{V}, \mathbf{W}$ be two Vector spaces. A map $\Phi : \mathbf{V} \rightarrow \mathbf{W}$ is called a linear map or linear transformation or homomorphism of vector spaces if $\forall x, y \in \mathbf{V}, \forall \lambda, \psi \in \mathbb{R}$:

$$\Phi(\lambda x + \psi y) = \lambda \Phi(x) + \psi \Phi(y)$$

Linear Mappings II

Definition

Let Φ be a map such as $\Phi : \mathbf{V} \rightarrow \mathbf{W}$. Then Φ is said :

1. Injective : if $\forall x, y \in \mathbf{V}, \Phi(x) = \Phi(y) \Rightarrow x = y$
2. Surjective : if $\Phi(\mathbf{V}) = \mathbf{W}$
3. Bijective : if Φ is injective and surjective

if Φ is bijective then there exists a map $\Phi^{-1} : \mathbf{W} \rightarrow \mathbf{V}$ such that $\Phi^{-1} \circ \Phi(x) = x$ is the inverse map of Φ

We then introduce :

1. Isomorphism : $\Phi : \mathbf{V} \rightarrow \mathbf{W}$ is linear bijective
2. Endomorphism : $\Phi : \mathbf{V} \rightarrow \mathbf{V}$ linear
3. Automorphism : $\Phi\mathbf{U} \rightarrow \mathbf{V}$ linear and bijective

Linear Mappings III

4. We define $Id_v : \mathbf{V} \to \mathbf{V}, x \mapsto x$ The identity map

Example :

The map $\Phi : \mathbb{R}^2 \to \mathbb{C}, \Phi(x) = x_1 + ix_2$ is a homomorphism, since :

$$\Phi\left(\lambda\begin{pmatrix}x_1\\x_2\end{pmatrix} + \psi\begin{pmatrix}y_1\\y_2\end{pmatrix}\right) = \Phi\begin{pmatrix}\lambda x_1 + \psi y_1\\ \lambda x_2 + \psi y_2\end{pmatrix}$$

$$= (\lambda x_1 + \psi y_1) + i(\lambda x_2 + \psi y_2)$$

$$= \lambda(x_1 + ix_2) + \psi(y_1 + iy_2) = \lambda\Phi(x) + \psi\Phi(y)$$

Linear Mappings IV

Theorem

Finite-dimensional vector spaces $\mathbf{V}$ *and* $\mathbf{W}$ *are isomorphic if and only if* $dim(\mathbf{V}) = dim(\mathbf{W})$

Linear Mappings V

Remarks

1. 15 hesitates considering $\mathbb{R}^{m \times n}$(The vector space of matrices $m \times n$) as $\mathbb{R}^{mn}$(The vector space of vectors of dimension mn) since their dimension is the same and there is a one-to-one linear map between $\mathbb{R}^{m \times n}$ and $\mathbb{R}^{mn}$
2. For two maps $\Phi : \mathbf{V} \rightarrow \mathbf{W}$ and $\Psi : \mathbf{W} \rightarrow \mathbf{X}$, the map $\Psi \circ \Phi : \mathbf{V} \rightarrow \mathbf{X}$ is also linear
3. If $\Phi : \mathbf{V} \rightarrow \mathbf{W}$ is an isomorphism, then $\Phi^{-1} : \mathbf{W} \rightarrow \mathbf{V}$ is also an isomorphism

Matrix representation of linear maps (transformation matrix)

Matrix representation of linear maps (transformation matrix) I

Definition

We consider a vector space **V** and a basis $\mathbf{B} = (b_1, b_2, \ldots, b_n)$ for each element $x \in \mathbf{V}$, we obtain a unique representation :

$$x = \alpha_1 b_1 + \alpha_2 b_2 + \cdots + \alpha_n b_n$$

of x with respect to **B.**

Then $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the coordinates of x with respect to the base **B**

Matrix representation of linear maps (transformation matrix) II

Definition

Let $\mathbf{V}, \mathbf{W}$ be two Vector spaces with respective bases $\mathbf{B} = (b_1, b_2, \ldots, b_n)$ and $\mathbf{C} = (c_1, c_2, \ldots, c_m)$ and let $\Phi : \mathbf{V} \to \mathbf{W}$. For $j \in \{1, 2, \ldots, n\}$

$$\Phi(b_j) = \alpha_{1j} c_1 + \alpha_{2j} c_2 + \cdots + \alpha_{mj} c_m = \sum_{i=1}^{m} \alpha_{ij} c_i$$

Matrix representation of linear maps (transformation matrix) III

Definition

is the unique representation of $\Phi(b_j)$ with respect to **C**. Then we call the matrix $m \times n$, $\mathbf{A}_\Phi$ whose elements are given by $\mathbf{A}_\Phi(c, j) = \alpha_{ij}$ The matrix transformation of Φ with respect to the bases **B** of **V** and **C** of **W**

The coordinates of $\Phi(b_j)$ with respect to the base **C** of $>$ are the J^{th} column of $\mathbf{A}_\Psi$ For an element $x \in \mathbf{V}$ and its image $y \in \Phi(x) \in \mathbf{W}$, if $\hat{x}$ is the vector of coordinates of x with respect to base **B** and if $\hat{y}$ is the vector of coordinates of y with respect to base **C**, then

$$\hat{y} = \mathbf{A}_\Phi \hat{x}$$

Matrix representation of linear maps (transformation matrix) IV

The matrix $\mathbf{A}_\Phi$ can be used to join the coordinates with respect to a basis of $\mathbf{V}$ to the coordinates with respect to a basis of $\mathbf{W}$
Example : Consider the homomorphism $\Phi : \mathbf{V} \to \mathbf{W}$ and $\mathbf{B} = (b_1, b_2, b_3)$ of $\mathbf{V}$, $\mathbf{C} = (c_1, c_2, c_3, c_4)$ of $\mathbf{W}$ with

$$\Phi(b_1) = 1c_1 - c_2 + 3c_3 - c_4$$
$$\Phi(b_2) = 2c_1 + c_2 + 7c_3 + 2c_4$$
$$\Phi(b_3) = 3c_2 + c_3 + 4c_4$$

Matrix representation of linear maps (transformation matrix) V

Then the transformation matrix $\mathbf{A}_\Phi$ with respect to $\mathbf{B}$ and $\mathbf{C}$ is given by $\begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{pmatrix}$

Basic change

Basic change I

Let $\mathbf{V}, \mathbf{W}$ be two Vector space and $\Phi : \mathbf{V} \to \mathbf{W}$ is a linear map, We consider two bases of $\mathbf{V}$

$$\mathbf{B} = (b_1, b_2, \ldots, b_n), \hat{\mathbf{B}} = (\hat{b_1}, \hat{b_2}, \ldots, \hat{b_n})$$

and two bases of $\mathbf{W}$

$$\mathbf{C} = (c_1, c_2, \ldots, c_m), \hat{\mathbf{C}} = (\hat{c_1}, \hat{c_2}, \ldots, \hat{c_m})$$

and $\mathbf{A}_\Phi$ is the transformation matrix of Φ(with $\mathbf{A} \in \mathbb{R}^{m \times n}$) with respect to a $\mathbf{B}$ and $\mathbf{C}$ and $\hat{\mathbf{A}_\Phi}$ is the transformation matrix of phi with respect to a $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$.

Basic change II

Theorem

The transformation matrix $\hat{\mathbf{A}_\Phi}$ *with respect to* $\hat{\mathbf{B}}$ *and* $\hat{\mathbf{C}}$ *can be written*

$$\hat{\mathbf{A}_\Phi} = \mathbf{T}^{-1}\mathbf{A}_\Phi\mathbf{A}$$

Where $\mathbf{S} \in \mathbb{R}^{n\times n}$ *is the transformation matrix of* Id_V *which maps the coordinates with respect to* $\hat{\mathbf{B}}$ *to the coordinates with respect to* $\mathbf{B}$.
$\mathbf{T} \in \mathbb{R}^{m\times m}$ *is the transformation matrix of* Id_W *which associates the coordinates with respect to* $\hat{\mathbf{C}}$ *with the coordinates with respect to* $\mathbf{C}$.

Proof I

We can write $\hat{\mathbf{B}}$ as a linear combination of $\mathbf{B}$, such that :

$$\hat{b}_j = s_{1j}b_1 + s_{2j}b_2 + \cdots + s_{nj}b_n = \sum_{i=1}^{n} s_{ij}b_i$$

Similarly, we can write $\hat{\mathbf{C}}$ as a linear combination of $\mathbf{C}$:

$$\hat{c}_k = t_{1k}c_1 + t_{2k}c_2 + \cdots + s_{mk}c_m = \sum_{l=1}^{m} t_{lk}c_k$$

By defining $\mathbf{S} = (s_{ij}) \in \mathbb{R}^{n\times n}$ as the matrix of transformations which associates the coordinates with respect to $\hat{\mathbf{B}}$ with the coordinates with respect to $\mathbf{B}$. And $\mathbf{T} = (t_{lk}) \in \mathbb{R}^{m}$ the

Proof II

transformation matrix that associates the coordinates with respect to $\hat{\mathbf{C}}$ to the coordinates with respect to $\mathbf{C}$. We have :

$$\Phi(\hat{b}_j = \sum_{k=1}^{m} \hat{a}_{kj}\hat{c}_k = \sum_{k=1}^{m} \hat{a}_{kj}(\sum_{l=1}^{m} t_{lk}c_l) = \sum_{l=1}^{m}(\sum_{k=1}^{m} t_{lk}\hat{a}_{kj}) \cdot c_l \qquad (4)$$

We can also express $\hat{b}_j \in \mathbf{V}$ as a linear combination of $b_j \in \mathbf{V}$

$$\Phi(\hat{b}_j) = \Phi(\sum_{i=1}^{n} s_{ij}b_i) = \sum_{i=1}^{n} s_{ij}\Phi(b_i) = \sum_{i=1}^{n} s_{ij} \cdot \sum_{l=1}^{m} a_{li}c_l = \sum_{l=1}^{m}(\sum_{i=1}^{n} a_{li}s_{i}j) \cdot$$

Proof III

By comparing (4) and (4), we have :

$$\sum_{k=1}^{m} t_{lk} \cdot \hat{a}_{kj} = \sum_{i=1}^{n} a_{li} \cdot s_{ij}$$

Then

$$\mathbf{T} \cdot \hat{\mathbf{A}}_{\Phi} = \mathbf{A}_{\phi}\mathbf{S}$$

Then

$$\hat{\mathbf{A}}_{\Phi} = \mathbf{T}^{-1}\mathbf{A}_{\Phi}\mathbf{S}$$

Basic change I

Example : We consider a linear map $\Phi : \mathbb{R}^3 \to \mathbb{R}^4$ with

$$\mathbf{A}_\Phi = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{pmatrix}$$

compared to standard bases

$$\mathbf{B} = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right); \mathbf{C} = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

Basic change II

We want to calculate the transformation matrix $\hat{\mathbf{A}}_\Phi$ with respect to the bases

$$\hat{\mathbf{B}} = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right); \hat{\mathbf{C}} = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right)$$

Then

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \mathbf{T} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Basic change III

On the ith column of **S** is the coordinate vector of $\hat{b}_i$ with respect to the base **B**, and the *jth* column of T represents the coordinate vector of $\hat{c}_j$ with respect to the base **C.** We then obtain

$$\hat{\mathbf{A}}_\Phi = \mathbf{T}^{-1}\mathbf{A}_\Phi\mathbf{S} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 3 & 2 & 1 \\ 0 & 4 & 2 \\ 10 & 8 & 4 \\ 1 & 6 & 3 \end{bmatrix}$$

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

Image and kernels

Image and kernels I

Definition

Let $\Phi : \mathbf{V} \to \mathbf{W}$. We define the

1. Kernel of Φ :

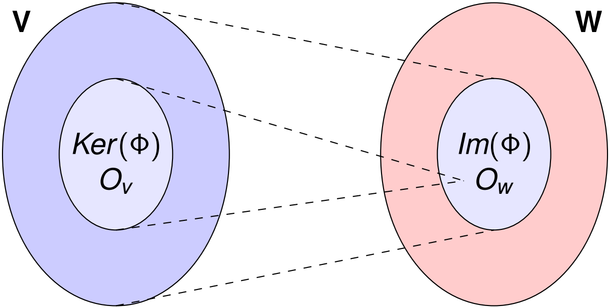
$$Ker(\Phi) = \Phi^{-1}(0_W) = \{v \in \mathbf{V} : \Phi(v) = 0_w\}$$

2. Image of Φ :

$$Im(\Phi) = \Phi(\mathbf{V}) = \{w \in \mathbf{W} | \exists v \in \mathbf{V} : \Phi(v) = w\}$$

We also call **V** and **W**, the domain and codomain of phi , respectively

Image and kernels II



$$\Phi : \mathbf{V} \rightarrow \mathbf{W}$$

Image and kernels III

Remark 1

Consider a linear map $\Phi : \mathbf{V} \to \mathbf{W}$, where $\mathbf{V}, \mathbf{W}$ are Vector spaces. We still have $\Phi(0_v) = 0_w$, so $0_v \in Ker(\Phi), Im(\Phi) \leq \mathbf{W}$ is a vector subspace of $\mathbf{W}$ and $Ker(\Phi)$ is a vector subspace of $\mathbf{V}$.

Φ is injective if $Ker(\Phi) = \{0\}$

Image and kernels IV

Remark 2

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ and Φ be a linear map

$$\Phi : \mathbb{R}^n \to \mathbb{R}^m$$
$$x \mapsto \mathbf{A}x$$

1. For $\mathbf{A} = [a_1, \ldots, a_n]$, with a_i column of $\mathbf{A}$, $i = \overline{1, n}$ We have

$$Im(\Phi) = \{\mathbf{A}x : x \in \mathbb{R}^n\} = \sum_{i=1}^{n} x_i a_i : x_1, \ldots, x_n \in \mathbb{R} = span[a_1, a_2, \ldots, a_n]$$

   In other words, the image of Φ is the subspace generated by the column vectors of $\mathbf{A}$. This subspace is called Column Space. So $rk(\mathbf{A}) = dim(Im(\Phi))$

Image and kernels V

Remark 2

1. The kernel $Ker(\Phi)$ is the general solution of the homogeneous system of linear equations $\mathbf{A}x = 0$.
2. The kernel is a vector subspace of $\mathbb{R}^n$, where n is the number of columns of $\mathbf{A}$

Example : We consider the application

$$\Phi : \mathbb{R}^4 \to \mathbb{R}^2, \begin{bmatrix} x1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} =$$

$$\begin{bmatrix} x_1 + 2x_2 - x_3 \\ x_1 + x_4 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Image and kernels VI

Φ is linear

1. To determine $Im(\Phi)$, it suffices to consider

$$Im(\Phi) = Span\left[\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}2\\0\end{bmatrix}, \begin{bmatrix}-1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right]$$

2. To find the kernel $Ker(\Phi)$, We answer $\mathbf{A}x = 0$, By Gaussian elimination $\mathbf{A}$ is put in reduced scaled form

$$\begin{bmatrix}1 & 2 & -1 & 0\\1 & 0 & 0 & 1\end{bmatrix} \rightsquigarrow \begin{bmatrix}1 & 0 & 0 & 1\\0 & 1 & -\frac{1}{2} & -\frac{1}{2}\end{bmatrix}$$

Image and kernels VII

Then

$$Ker(\Phi) = Span\left[\begin{bmatrix}0\\ \frac{1}{2}\\ 1\\ 0\end{bmatrix}, \begin{bmatrix}-1\\ \frac{1}{2}\\ 0\\ 1\end{bmatrix}\right]$$

Image and kernels VIII

Theorem

Fundamental of linear maps for two vector spaces $\mathbf{V}$ *and* $\mathbf{W}$ *and a linear map* $\Phi : \mathbf{V} \to \mathbf{W}$*, we have*

$$dim(Ker(\Phi)) + dim(Im(\Phi)) = dim(\mathbf{V})$$

We then have the consequences of the theorem :

1. *If* $dim(Im(\Phi)) < dim(\mathbf{V})$ *then* $Ker(\Phi)$ *is non-trivial (it contains at least one deferent element of* $\mathbf{O}_v$*)*
2. *If* $\mathbf{A}_\Phi$ *is the transformation matrix of* Φ *with respect to a base and* $dim(Im(\Phi)) < dim(\mathbf{V})$ *then the system of linear equations* $\mathbf{A}_\Phi x = 0$ *has m finitely many solutions*

Image and kernels IX

Theorem

If $dim(\mathbf{V}) = dim(\mathbf{W})$ *, then the following three-way equivalence holds :*

1. Φ *is injective*
2. Φ *is surjective*
3. Φ *is bijective*

since $Im(\Phi) \subseteq \mathbf{W}$

Affine Spaces

Affine Spaces

Affine Spaces

Affine Spaces I

Definition

Let $\mathbf{V}$ be a vector space $x \in \mathbf{V}$ and $\mathbf{U} \subseteq \mathbf{V}$ a subspace, then the set

$$\begin{aligned}\mathbf{L} = x_0 + \mathbf{U} &= \{x_0 + u : u \in \mathbf{U}\} \\ &= \{v \in \mathbf{V} / \exists u \in \mathbf{U} : v = x_0 + u\} \subseteq \mathbf{V}\end{aligned}$$

is called an affine space of $\mathbf{V}$.
$\mathbf{U}$ is called direction or space direction and x_0 is called support point

Examples of affine subspaces :
lines, planes of $\mathbb{R}^3$ which do not necessarily pass through the origin

Remarks I

Consider two affine subspaces

1. $\mathbf{L} = x_0 + \mathbf{U}$ and $\tilde{\mathbf{L}} = \tilde{x}_0 + \tilde{\mathbf{U}}$ of a vector space $\mathbf{V}$ Then $\mathbf{L} \subseteq \tilde{\mathbf{L}}$ Let $\mathbf{U} \subseteq \tilde{\mathbf{U}}$ and $x_0 - \tilde{x}_0 \in \tilde{\mathbf{U}}$
2. If $(b_1, b_2, \ldots, b_k)$ is a basis of $\mathbf{U}$, then each element $x \in \mathbf{L}$ can be written :

$$x = x_0 + \lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_k b_k$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ This representation is called parametric equation of $\mathbf{L}$ with the directional vectors $b_1, b_2, \ldots, b_k$ and parameters $\lambda_1, \lambda_2, \ldots, \lambda_k$

Remarks II

3. In $\mathbb{R}^n$, an $(n-1)$ affine subspace is called a hyperplane and a as a parametric equation $y = x_0 + \sum_{i=1}^{n-1} \lambda_i x_i$ where $x_1, x_2, \ldots, x_{n-1}$ forms a basis for a subspace $\mathbf{U} \subseteq \mathbb{R}^n$ of dimension $(n-1)$
4. For $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, the solution to the linear system of equations $\mathbf{A}x = b$ is either an empty set or an $n - rk(\mathbf{A})$ dimensional affine subspace of $\mathbb{R}^n$.
   In particular, the solution to the equation $\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n = b$ with $(\lambda_1, \lambda_2, \ldots, \lambda_n) \neq (0, 0, \ldots, 0)$ is a hyperplane of $\mathbb{R}^n$

Affine applications

Affine applications I

Definition

For two vector spaces $\mathbf{V}, \mathbf{W}$, a linear map $\Phi : \mathbf{V} \to \mathbf{W}$ and $a \in \mathbf{W}$, the map

$$\begin{aligned} \Psi : \mathbf{V} &\to \mathbf{W} \\ x &\mapsto a + \Phi(x) \end{aligned} \tag{6}$$

is an affine map from **V** to **W.**
The vector a is called the translation vector of Ψ

1. Each Affine Map $\Phi : \mathbf{V} \to \mathbf{W}$ is a composition of a linear map Φ and a translation $\mathcal{T} : \mathbf{W} \to \mathbf{W}$ in $\mathbf{W}$ such that $\Phi = \mathcal{T} \circ \Phi$
2. The composition $\grave{\Phi} \circ \Phi$ Of affine maps is also affine.