Mathematics for Machine Learning Part I (Linear Algebra)

October 5, 2023

Systems of Linear Equations

Systems of Linear Equations

Definition

A linear system of equations consists of a finite number **m** of linear equations involving 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, an infinite number of solutions or no solution at all

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 $x_1 = x_2 = x_3 = 1$

- for 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 $2x_1 = 5 - 3x_3$ and $2x_2 = 1 + x_3$

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 matrices $\mathbf{A} \in \mathbb{R}^{m \times k}$ and $\mathbf{B} \in \mathbb{R}^{k \times n}$, the product of $\mathbf{A}$ and $\mathbf{B}$, denoted $\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{m \times n}$, is defined with $c_{ij} = \sum_{l=1}^{k} a_{il} \cdot b_{lj}$, for $i = 1 \dots m, j = 1 \dots n$.
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 :

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

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

- 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 $\mathbf{A}$ and is denoted $\mathbf{A}^{-1}$

Inverse and transpose of a matrix II

Remarks

1. The inverse matrix $\mathbf{A}^{-1}$ may not exist for 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 $\mathbf{A}$ 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 $\mathbf{A}$ by

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

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 $\mathbf{B} = \mathbf{A}^\top$

Inverse and transpose of a matrix V

We then have the following important properties :

1. $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I} = \mathbf{A}^{-1}\mathbf{A}$
2. $(\mathbf{AB})^{-\mathbf{1}} = \mathbf{B}^{-\mathbf{1}}\mathbf{A}^{-\mathbf{1}}$
3. $(\mathbf{A}+\mathbf{B})^{-1} \neq \mathbf{A}^{-1} + \mathbf{B}^{-1}$
4. $(\mathbf{A}^\top)^\top = \mathbf{A}$
5. $(\mathbf{A}+\mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top$
6. $(\mathbf{AB})^\top = \mathbf{B}^\top \cdot \mathbf{A}^\top$

Inverse and transpose of a matrix VI

Definition

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

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

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\lambda \in \mathbb{R}$.

$\lambda \mathbf{A} = \mathbf{K}$ with $\mathbf{K}_{ij} = \lambda a_{ij}$.

For $\lambda, \psi \in \mathbb{R}$, we have the following properties :

- 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. $(\lambda\mathbf{C})^\top = \mathbf{C}^\top\lambda^\top = \mathbf{C}^\top\lambda = \lambda\mathbf{C}^\top$ (since $\lambda = \lambda^\top$)

Multiplication by a scalar

- Distributivity :
  1. $(\lambda + \psi)\mathbf{C} = \lambda\mathbf{C} + \psi\mathbf{C}, \mathbf{C} \in \mathbb{R}^{m\times n}$
  2. $\lambda(\mathbf{B} + \mathbf{C}) = \lambda\mathbf{B} + \lambda\mathbf{C}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{m\times n}$

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

$$(\lambda+\psi)\mathbf{C} = (\lambda+\psi)\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} (\lambda+\psi) & (\lambda+\psi)\times 2 \\ (\lambda+\psi)\times 3 & (\lambda+\psi)\times 4 \end{pmatrix}$$

$$= \begin{bmatrix} \lambda+\psi & 2\lambda+2\psi \\ 3\lambda+3\psi & 4\lambda+4\psi \end{bmatrix} = \begin{bmatrix} \lambda & 2\lambda \\ 3\lambda & 4\lambda \end{bmatrix} + \begin{bmatrix} \psi & 2\psi \\ 3\psi & 4\psi \end{bmatrix}$$

$$= \lambda \cdot \mathbf{C} + \psi \cdot \mathbf{C}$$

Solving Systems of Linear Equations

Notation

Consider a linear system of equations

$$\mathbf{A}\mathbf{X} = \mathbf{b}$$

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 $i-th$ column of the matrix $\mathbf{A}$
Notice that the system has two equations and four unknowns.
This means that there are infinite number of solutions.

Particular solution and general solution II

One solution can be trivially deduced by taking 42 for Column 1 (x_1) and 8 for Column 2 (x_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 $[42, 8, 0, 0]^\top$.

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

This solution is not unique.

Particular solution and general solution III

To formulate all the solutions of the system, we need to generate $\mathbf{0}$ 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 $(8, 2, -1, 0)^\top$ is a solution.

Particular solution and general solution IV

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

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

Particular solution and general solution V

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

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. It is called the *general solution*.

The approach we adopted consists in:

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

Particular solution and general solution VII

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

Particular solution and general solution VIII

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
   The form of the matrix has made it possible to obtain a special solution and the general solution easily
   In most cases, we have to proceed to some transformations that allow to transform a linear system into a simpler form.
   this technique is called Gaussian elimination

Elementary transformations

Elementary transformations I

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

That are:

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

Elementary transformations II

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 III

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

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

- Swap L_1 and L_3

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

Elementary transformations IV

$$\rightsquigarrow \begin{array}{c} L_1 \\ -L_2 \\ L_3 \\ L_4 - L_2 \end{array} \rightsquigarrow \left[\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & -2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & 0 & 0 & 0 & a-2 \end{array}\right]$$

$$\rightsquigarrow \begin{array}{c} L_1 \\ L_2 \\ \cdot(-\frac{1}{3})L_3 \\ L_4 - L_3 \end{array} \rightsquigarrow \left[\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & a+1 \end{array}\right]$$

The augmented matrix $[\mathbf{A}|\mathbf{b}]$ is said to be in echelon form.

Only by setting $a = -1$ that the system can be solved.

Elementary transformations V

A particular solution is $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}$.

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 VI

Remark

The first nonzero coefficient of a row is called pivot or the leading coefficient and the column that corresponds to that element is called the pivot column.

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 is before 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 VII

Remark

1. The variables that correspond to the pivots in the row echelon form are called *basic variables*. The other variables are called *free variables*.
2. The echelon form makes it easier to obtain a particular solution: we express the right-hand side of the equation system using the pivot columns, such that $\tilde{b} = \sum_{i=1}^{p} \alpha_i P_i$ with P_i the pivot columns corresponding to the basic variables .

We determine the coefficients α_i starting from the rightmost pivot column.

For the previous example:

Elementary transformations VIII

We seek $\alpha_1, \alpha_2, \alpha_3$ such that

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

Which gives $\alpha_3 = 1, \alpha_2 = -1, \alpha_1 = 2$

Hence the particular solution $x = [2, 0, -1, 1, 0]^\top$

Elementary transformations IX

Definition

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

- it is written in row echelon 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 X

Example :

Consider the matrix $\begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \end{bmatrix}$.

Pivot columns are $P_1, P_3 and P_4$ with

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

Hence $3 \cdot P_1 - P_2 = 0$ and we obtain $\begin{pmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

Elementary transformations XI

Similarly,

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

Hence

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

Then we have $\begin{pmatrix} 3 \\ 0 \\ 9 \\ -4 \\ -1 \end{pmatrix}$.

Elementary transformations XII

The general solution can be 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 XIII

The minus-1 Trick

Assume that a matrix $\mathbf{A}$ is in a reduced echelon form such that all its rows containing at least one nonzero element with ($\mathbf{A} \in \mathbb{R}^{k \times n}$).

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

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 $J_1, J_2, \dots, J_k$ are the standard unit vectors.

Elementary transformations XIV

We extend $\mathbf{A}$ to an $n \times n$ matrix $\tilde{\mathbf{A}}$ by adding $n - k$ rows of the form $[0, \ldots, 0, -1, 0, \ldots, 0]$ such that the diagonal of $\tilde{\mathbf{A}}$ is only elements 1 or -1.

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

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

Elementary transformations XV

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

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

Remark

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

Matrix Inverse Calculation

Calculation of the inverse I

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$.

To compute $\mathbf{A}^{-1}$, we seek a matrix $\mathbf{X}$ such that $\mathbf{AX} = \mathbf{I}_n$. We can therefore write the linear system $\mathbf{AX} = \mathbf{I}_n$, 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

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

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

Calculation of the inverse III

The inverse matrix is thus :

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

Vector spaces

Groups

Groups I

We consider a set $\mathcal{G}$ and an operation denoted +, such that $+ : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$.

Then $\mathbf{G} = (\mathcal{G}, +)$ is called a *group* if:

- **Closure of $\mathcal{G}$ 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:

Groups II

- $(\mathbb{Z}, +)$ is an Abelian group
- $(\mathbb{R}, \cdot)$ is not a group (because 0 has no inverse element)
- $(\mathbb{R}^{n\times n}, +)$ is an Abelian group
- We consider $(\mathbb{R}^{n\times n}, \cdot)$ the set of $n \times n$ matrices with matrix multiplication as defined previously. 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 general linear group and is denoted $GL(n, \mathbb{R})$. This group is not Abelian.

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:
  - $\forall \lambda \in \mathbb{R}, x, y \in \mathcal{V} : \lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$
  - $\forall \lambda \psi \in \mathbb{R}, x \in \mathcal{V} : (\lambda + \psi) \cdot x = \lambda \cdot x + \psi \cdot x$

Vector Spaces II

- Associativity of the operation $(\cdot)$:
  $\forall \lambda, \psi \in \mathbb{R}, x \in \mathcal{V}, \lambda \cdot (\psi \cdot x) = (\lambda \cdot \psi) \cdot x$
- The neutral element with respect to $(\cdot)$:
  $\forall x \in \mathcal{V} : 1 \cdot x = x$

1. Elements $x \in \mathcal{V}$ are called vectors.
2. The neutral element of $(\mathcal{V}, +)$ is the vector $0 = [0, 0, \ldots, 0]^\top$.
3. The inner operation $+$ is called vector addition.
4. Elements $\lambda \in \mathbb{R}$ are called scalars and the outer operation $\cdot$ is called multiplication by scalars.

Vector Spaces III

Remark

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

Examples :

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

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 $\lambda \in \mathbb{R}, x \in \mathbb{R}^n$

Vector Spaces IV

$\mathcal{V} = \mathbb{R}^{m\times n}, m, n \in \mathbb{N}$ 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 $\mathbb{R}^{mn}$

Vector Spaces V

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 $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

However, we will distinguish between $\mathbb{R}^{n\times 1}$ and $\mathbb{R}^{1\times n}$ (row 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 $\cdot$ are restrictd to $\mathcal{U} \times \mathcal{U}$ and $\mathbb{R} \times \mathcal{U}$. In other words:

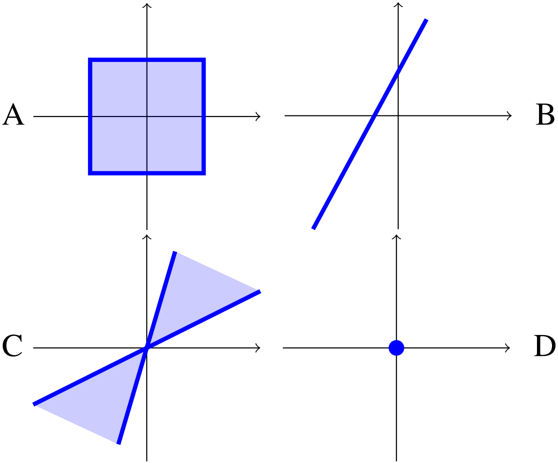
- $\mathcal{U} \neq \emptyset$: in particular $0 \in \mathbf{U}$
- Closure of $\mathbf{U}$:
  1. $\forall \lambda \in \mathbb{R}, \forall x \in \mathcal{U} : \lambda x \in \mathcal{U}$
  2. $\forall x, y \in \mathcal{U} : x + y \in \mathcal{U}$

Examples :

1. For any V.S $\mathbf{V}$, $\mathbf{V}$ and 0 are subspaces.

Vector Subspaces II

2. In the following figures, only the set **D** is a subspace of $\mathbb{R}^2$.



Vector Subspaces III

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

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

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

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 vectors $x_1, x_2, \ldots, x_n$ 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. 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 row echelon 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 $x_1 = \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix}$ $x_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}$ $x_3 = \begin{bmatrix} -1 \\ -2 \\ 1 \\ 1 \end{bmatrix}$

Hence we can 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

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

Notice that eac column is a pivot column and the only solution to this system is $\lambda_1 = \lambda_2 = \lambda_3 = 0$, thus 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 :

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

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 $b_1, b_2, b_3, b_4 \in \mathbb{R}^n$ 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

Generating Set I

Definition (Generating Set)

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 $\mathbf{V}$.

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

Generating Set 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 $\forall \bar{\mathcal{A}} \subseteq \mathcal{A}, \mathbf{V} \neq span[\mathcal{A}]$.

If $\mathcal{A}$ consists of linearly independent vectors then $\mathcal{A}$ is called *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 maximal set of linearly independent vectors,

Generating Set III

4. Each vector $x \in \mathbf{V}$ is a linear combination of the vectors of $\mathcal{B}$ and each linear combination is unique

Examples :

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

Generating Set IV

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

Generating Set V

Remarks

1. Each vector space possesses at least one basis. However, all the bases have the same number of elements, called the *dimension* of $\mathbf{V}$, denoted $\dim(\mathbf{V})$
2. If $\mathbf{U} \subseteq \mathbf{V}$ is 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 $\mathbf{U} = \mathbf{V}$

Generating Set VI

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 for a vector subspace $\mathbf{U} = span[x_1, x_2, \ldots, x_n] \subseteq \mathbb{R}^n$ by applying the following process :
   - 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.

Generating Set VII

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

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

Generating Set VIII

Consider the matrix

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

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

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 $\mathrm{rk}(\mathbf{A})$.

Property :

- $\mathrm{rk}(\mathbf{A}) = \mathrm{rk}(\mathbf{A}^\top)$
- The columns of $\mathbf{A} \in \mathbb{R}^{m\times n}$ span a vector subspace $\mathbf{U} \subseteq \mathbb{R}^m$ with $\dim(\mathbf{U}) = rk(\mathbf{A})$
- The rows of $\mathbf{A} \in \mathbb{R}^{m\times n}$ span 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 $\mathbf{A}^\top$
- For any $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{A}$ is regular if $\mathrm{rk}(\mathbf{A}) = n$

Rank II

- For any matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^{m}$, the system of linear equations $\mathbf{A}x = b$ can be solved if $\mathrm{rk}(\mathbf{A}) = \mathrm{rk}(\mathbf{A}|\mathbf{B})$, 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 $\mathrm{rk}(\mathbf{A}) = 2$

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 Maps

Linear maps I

Definition

Let $\mathbf{V}, \mathbf{W}$ be two Vector spaces. A map $\Phi : \mathbf{V} \to \mathbf{W}$ 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 maps II

Definition

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

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 $\Phi^{-1} : \mathbf{W} \longrightarrow \mathbf{V}$ such that $\Phi^{-1} \circ \Phi(x) = x$ is the inverse map of Φ

We then introduce:

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

Linear maps III

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

Example :

The map $\Phi : \mathbb{R}^2 \longrightarrow \mathbb{C}, \Phi(x) = x_1 + ix_2$ 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 maps IV

Theorem

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

Linear maps V

Remarks

1. This theorem is very powerful since it allows us to consider $\mathbb{R}^{m\times n}$ (vector space of matrices $m \times n$) as $\mathbb{R}^{mn}$(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 $\Phi : \mathbf{V} \to \mathbf{W}$ and $\Psi : \mathbf{W} \to \mathbf{X}$, the map $\Psi \circ \Phi : \mathbf{V} \to \mathbf{X}$ is also linear
3. If $\Phi : \mathbf{V} \to \mathbf{W}$ is an isomorphism, then $\Phi^{-1} : \mathbf{W} \to \mathbf{V}$ is also an isomorphism

Matrix Representation of Linear maps (transformation matrix)

Matrix representation of linear maps (transformation matrix) I

Definition

Consider a vector space $\mathbf{V}$ and a basis $\mathbf{B} = (b_1, b_2, \ldots, b_n)$.
each element $x \in \mathbf{V}$ has a unique representation:

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

with respect to $\mathbf{B}$.
Then $\alpha_1, \alpha_2, \ldots, \alpha_n$ are called *the coordinates of x with respect to the base* $\mathbf{B}$

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 $j \in \{1, 2, \ldots, n\}$

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

is the unique representation of $\Phi(b_j)$ with respect to $\mathbf{C}$.

Matrix representation of linear maps (transformation matrix) III

Definition

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 $\mathbf{B}$ of $\mathbf{V}$ and $\mathbf{C}$ of $\mathbf{W}$

The coordinates of $\Phi(b_j)$ with respect to the base $\mathbf{C}$ of > are the J^{th} column of $\mathbf{A}_\Psi$ For an element $x \in \mathbf{V}$ and its image $y = \Phi(x) \in \mathbf{W}$, if $\hat{x}$ is the vector of coordinates of x with respect to base $\mathbf{B}$ and if $\hat{y}$ is the vector of coordinates of y with respect to base $\mathbf{C}$, 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 correspond 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 $\Phi : \mathbf{V} \rightarrow \mathbf{W}$ 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 $\begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{pmatrix}$

Basis change

Basis change I

Let $\mathbf{V}, \mathbf{W}$ be two vector spaces and $\Phi : \mathbf{V} \to \mathbf{W}$ is a linear map. Let us consider two bases of $\mathbf{V}$

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

and two bases of $\mathbf{W}$

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

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

Basis 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 transformation

Proof II

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 $b_j \in \mathbf{V}$

$$\Phi(\hat{b}_j) = \Phi(\sum_{i=1}^{n} s_{ij}b_i) = \sum_{i=1}^{n} s_{ij}\Phi(b_i)$$

$$= \sum_{i=1}^{n} s_{ij} \cdot \sum_{l=1}^{m} a_{li}c_l = \sum_{l=1}^{m}(\sum_{i=1}^{n} a_{li}s_ij) \cdot c_l, j = 1, \ldots, n \qquad (5)$$

Proof III

By comparing (4) and (5), we have:

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

Thus

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

Therefore

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

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

1. Kernel of Φ:

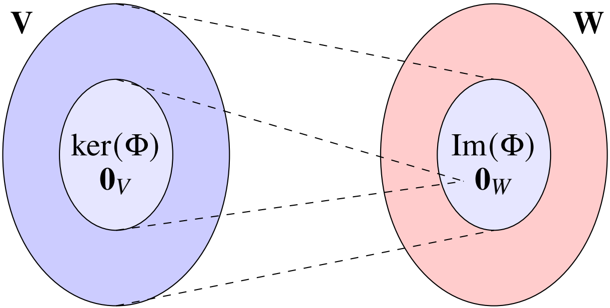
$$\ker(\Phi) = \Phi^{-1}(0_W) = \{v \in \mathbf{V} : \Phi(v) = 0_w\}$$

2. Image of Φ:

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

We also call $\mathbf{V}$ and $\mathbf{W}$, the domain and codomain of the map Φ , respectively

Image and kernels II



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

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 $0_v \in \ker(\Phi)$.

$\mathrm{Im}(\Phi) \leq \mathbf{W}$ is a vector subspace of $\mathbf{W}$ and $\ker(\Phi)$ is a vector subspace of $\mathbf{V}$.

Φ is injective if $\ker(\Phi) = \{\mathbf{0}\}$

Image and kernels IV

Remark 2

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

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

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

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

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 $\mathrm{rk}(\mathbf{A}) = \dim(\mathrm{Im}(\Phi))$

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

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

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

Φ is linear

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

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

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

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

Then

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

Theorem (Fundamental Theorem of Linear Algebra)

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

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

We then have the following:

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

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

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:

Affine Spaces II

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 $\Phi : \mathbf{V} \longrightarrow \mathbf{W}$ and $a \in \mathbf{W}$, the map

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

is an affine map from $\mathbf{V}$ to $\mathbf{W}$.

The vector a is called the translation vector of Ψ

1. Each Affine Map $\Phi : \mathbf{V} \longrightarrow \mathbf{W}$ is a composition of a linear map Φ and a translation $\mathcal{T} : \mathbf{W} \longrightarrow \mathbf{W}$ in $\mathbf{W}$ such that $\Phi = \mathcal{T} \circ \Phi$

Affine applications II

2. The composition $\grave{\Phi} \circ \Phi$ Of affine maps is also affine.