Linear Algebra

Peter Minary and Jonathan Whiteley

Michaelmas Term 2018

Linear Algebra

concerns

- Vector spaces and subspaces, $\boldsymbol{S} \subset \boldsymbol{V}$

  e.g. plane, $\mathbb{R}^2$, 3D space, $\mathbb{R}^3$ and $\mathbb{R}^n$

  $$\mathbb{R}^2 \subset \mathbb{R}^3 \subset \mathbb{R}^{n>3},\ n \in \mathbb{N}$$

- Linear mappings, $\boldsymbol{f} : \boldsymbol{V} \to \boldsymbol{W}$

  e.g. $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{V}$, $\boldsymbol{\alpha} \in \mathbb{R}$

  -$\boldsymbol{f}(\boldsymbol{x} + \boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{y})$

  -$\boldsymbol{f}(\boldsymbol{\alpha x}) = \boldsymbol{\alpha f}(\boldsymbol{x})$

LA Applications

- Image restoration
- Discrete mathematical/computational models
- Networks
- Google search engine
- Markov Processes/financial modelling
- Computer graphics

Some major LA Applications

concern the solution of

- Systems of linear equations

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

- Eigenvalue problems

$$\boldsymbol{A}\mathbf{x} = \boldsymbol{\alpha}\mathbf{x},\ \boldsymbol{\alpha} \in \mathbb{R}$$

- Geometric Transformations

$$\mathbf{x}' = \boldsymbol{T}\mathbf{x}$$

Course Aims

- Lay down basic concepts of linear algebra used in subsequent study

  Vectors and geometry in two and three space dimensions. Algebraic properties. Dot products and the norm of a vector. Vector spaces, subspaces and vector space axioms. Linear independence of vectors. Basis and dimension of a vector space. Orthogonal vectors and subspaces. The Gram-Schmidt algorithm. Column and row space. Range and null space. Rank of a matrix. Matrix operations. Determinant and inverse. Elementary matrices. Gaussian elimination and pivoting. Row echelon form. Elementary matrix factorisations. One-to-one and onto transformations. Similarity and diagonalization. Systems of linear differential equations. The methods of Jacobi, Gauss-Seidel, successive over relaxation, and steepest descent. The QR factorisation, least squares problems.

Course Aims

- Lay down basic concepts of linear algebra used in subsequent study
- Demonstrate the wide applicability of this discipline within the scientific field
- Practice ways to conduct proofs that require algebraic manipulations, geometry and numerics
- Provide insight into linear algebra theorems and they impact on both the scientific field and everyday life

Role of LA in CS Studies

- Core areas

  internet search, graph analysis, graphics, machine learning, compilers, parallel computing

- Interdisciplinary areas

  scientific computing, bioinformatics, data mining, speech recognition, computer vision

- Related courses in our Department

  Machine Learning, Computer Graphics, Quantum Computer Science, Geometric Modelling

Linear Algebra Course Outline

- Lectures 1-3: Vectors and vector spaces
- Lectures 4-6: Independence and orthogonality
- Lectures 7-9: Matrices
- Lecture 10-12: Systems of linear equations
- Lectures 13-15: Elementary matrix factorisations, determinants
- Lectures 16-18: Linear transformations
- Lectures 19-21: Eigenvalues and eigenvectors
- Lectures 22-23: Iterative methods for solving linear equations
- Lectures 24: Overdetermined systems

Linear Algebra Course Material

https://www.cs.ox.ac.uk/teaching/materials18-19/linearalgebra/

Contains:

- Lecture notes & slides, published within about a week's delay
  David Kay's lecture notes form the basis of most of the lecture slides.
- Problem sheets, released every week
  These include additional problems marked as '(optional)'.

Related:

- Recommended reading
  Gilbert Strang, Introduction to Linear Algebra (3rd Edition)

Linear Algebra Course, Part I

Peter Minary, office: Room 367
email: peter.minary@cs.ox.ac.uk

- Lectures 1-3: Vectors and vector spaces

  Problem Sheet 1

- Lectures 4-6: Independence and orthogonality

  Problem Sheet 2

- Lectures 7-9: Matrices

  Problem Sheet 3

- Lectures 10-12: Systems of linear equations

  Problem Sheet 4

Linear Algebra-Part I

Peter Minary, office: Room 367
email: peter.minary@cs.ox.ac.uk

- Chapter 1: Vectors and vector spaces

  (Week 1, Lectures 1-3)

- Chapter 2: Independence and orthogonality

  (Week 2, Lectures 4-6)

- Chapter 3: Matrices

  (Week 3, Lectures 7-9)

- Chapter4: Systems of linear equations

  (Week 4, Lectures 10-12)

Chapter 1

Vectors and Vector Spaces

1.1. Vectors

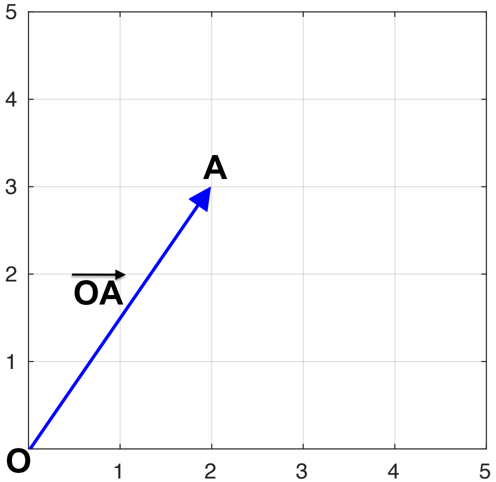
1.1.1. Vectors in the plane

Given the familiar $x-$ and $y-$axes, a vector is represented geometrically by a directed line segment from the origin, O to a terminal point A. The vector is then fully defined by the ordered pair (a_1, a_2) that represent the terminal point, A. The coordinates, a_1 and a_2 are also called the components of the vector.

We often use bold face letters to denote vectors,

$$\mathbf{a} = [a_1, a_2] \text{ or in column form } \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}.$$

Example 1.1.1 *The point* $\boldsymbol{A}$ *has coordinates* $(\mathbf{2}, \mathbf{3})$ *the vector from the origin* $\boldsymbol{O}$ *to* $\boldsymbol{A}$ *is given by* $\mathbf{a} = \vec{\boldsymbol{O}}\boldsymbol{A} = [\mathbf{2}, \mathbf{3}]$*. We may also write the vector in column form* $\mathbf{a} = \begin{bmatrix} \mathbf{2} \\ \mathbf{3} \end{bmatrix}$*.*

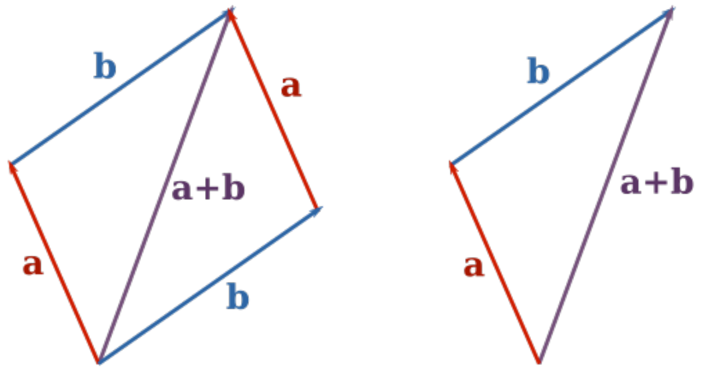


Suppose we have two vectors $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$. Two vectors, $\mathbf{a}$ and $\mathbf{b}$ are equal if and only if $a_1 = b_1$ and $a_2 = b_2$.
The first basic vector operation is vector addition defined as

$$\mathbf{a} + \mathbf{b} = [a_1, a_2] + [b_1, b_2] = [a_1 + b_1, a_2 + b_2].$$

This addition describes firstly moving the distance and direction of $\mathbf{a}$ and then moving the distance and direction of $\mathbf{b}$. Note

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}.$$



The second basic vector operation is scalar multiplication. Given a vector $\mathbf{a} = [a_1, a_2]$, and a scalar, $\alpha \in \mathbb{R}$, multiplying vector $\mathbf{a}$ with scalar α gives

$$\alpha\mathbf{a} = \alpha[a_1, a_2] = [\alpha a_1, \alpha a_2].$$

Using these operations we can also calculate the difference of $\mathbf{a}$ and $\mathbf{b}$ as

$$\mathbf{a} - \mathbf{b} = \mathbf{a} + (-1)\mathbf{b} = \mathbf{a} + (-\mathbf{b}),$$

where $-\mathbf{b} = (-1)\mathbf{b}$ is the negative of $\mathbf{b}$.

Example 1.1.2 *Given the points* $\boldsymbol{A} = (\mathbf{1}, \mathbf{-4})$ *and* $\boldsymbol{B} = (\mathbf{2}, \mathbf{3})$ *with corresponding vectors* $\vec{\boldsymbol{OA}} = \mathbf{a} = [\mathbf{1}, \mathbf{-4}]$ *and* $\vec{\boldsymbol{OB}} = \mathbf{b} = [\mathbf{2}, \mathbf{3}]$ *the vector describing the path from* $\boldsymbol{A}$ *to* $\boldsymbol{B}$, $\vec{\boldsymbol{AB}}$, *is given by* $[\mathbf{2}, \mathbf{3}] - [\mathbf{1}, \mathbf{-4}] = [\mathbf{2}, \mathbf{3}] + (\mathbf{-1})[\mathbf{1}, \mathbf{-4}] = [\mathbf{1}, \mathbf{7}]$. *We have described the path by going from* $\boldsymbol{A}$ *to* $\boldsymbol{O}$, $-\mathbf{a}$ *and then to* $\boldsymbol{B}$.

We already saw a geometric demonstration on the commutativity of vector addition. Note, that this also follows from the properties of vector addition and addition on real numbers. More generally, purely derived from the properties of vector addition and scalar multiplication and the properties of addition and multiplication of real numbers, the following properties hold.

Theorem 1.1.3 *Let* $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^2$ *be vectors in the plane and* $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}$ *scalars.*

1. $\mathbf{a} + \mathbf{b} \in \mathbb{R}^2$ *Closure under summation.*
2. $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ *Commutativity*
3. $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$ *Associativity*
4. $\exists \ \mathbf{0}$, *such that* $\mathbf{a} + \mathbf{0} = \mathbf{a}$ *Additive identity, zero vector*
5. $\forall \mathbf{a} \ \exists - \mathbf{a}$ *such that* $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$ *Additive inverse*
6. $\boldsymbol{\alpha}\mathbf{a} \in \mathbb{R}^2$ *Closure under scalar multiplication*
7. $\boldsymbol{\alpha}(\mathbf{a} + \mathbf{b}) = \boldsymbol{\alpha}\mathbf{a} + \boldsymbol{\alpha}\mathbf{b}$ *Distributivity*
8. $(\boldsymbol{\alpha} + \boldsymbol{\beta})\mathbf{a} = \boldsymbol{\alpha}\mathbf{a} + \boldsymbol{\beta}\mathbf{a}$ *Distributivity*
9. $\boldsymbol{\alpha}(\boldsymbol{\beta}\mathbf{a}) = (\boldsymbol{\alpha}\boldsymbol{\beta})\mathbf{a}$ *Associativity of multiplication*
10. $\mathbf{1a} = \mathbf{a}$ *Multiplicative identity*

Each property can be proven using the properties for vector addition and scalar multiplication. For example,

$$\begin{aligned}
\boldsymbol{\alpha}(\mathbf{a}+\mathbf{b}) &= \boldsymbol{\alpha}\left([\boldsymbol{a}_1, \boldsymbol{a}_2] + [\boldsymbol{b}_1, \boldsymbol{b}_2]\right) \\
&= \boldsymbol{\alpha}[\boldsymbol{a}_1 + \boldsymbol{b}_1, \boldsymbol{a}_2 + \boldsymbol{b}_2] \\
&= [\boldsymbol{\alpha}(\boldsymbol{a}_1 + \boldsymbol{b}_1), \boldsymbol{\alpha}(\boldsymbol{a}_2 + \boldsymbol{b}_2)] \\
&= [\boldsymbol{\alpha}\boldsymbol{a}_1 + \boldsymbol{\alpha}\boldsymbol{b}_1, \boldsymbol{\alpha}\boldsymbol{a}_2 + \boldsymbol{\alpha}\boldsymbol{b}_2] \\
&= [\boldsymbol{\alpha}\boldsymbol{a}_1, \boldsymbol{\alpha}\boldsymbol{a}_2] + [\boldsymbol{\alpha}\boldsymbol{b}_1, \boldsymbol{\alpha}\boldsymbol{b}_2] \\
&= \boldsymbol{\alpha}[\boldsymbol{a}_1, \boldsymbol{a}_2] + \boldsymbol{\alpha}[\boldsymbol{b}_1, \boldsymbol{b}_2] \\
&= \boldsymbol{\alpha}\mathbf{a} + \boldsymbol{\alpha}\mathbf{b}.
\end{aligned}$$

Exercise 1.1.4 *Let* $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^2$ *then prove that* $(\mathbf{a}+\mathbf{b})+\mathbf{c} = \mathbf{a}+(\mathbf{b}+\mathbf{c})$*.*

1.1.2. Vectors in $\mathbb{R}^n$

Everything in two dimensions extends to three dimensions. A point $\boldsymbol{A}$ in three dimensions has coordinates $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ with corresponding vector

$$\boldsymbol{OA} = [\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}], \text{ or column vector } \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \\ \boldsymbol{z} \end{bmatrix} \in \mathbb{R}^3.$$

For any $n > 0$ integer we can define $\mathbb{R}^n$ also known as $n-$ space:

- $\mathbb{R}^1 =$ set of all real numbers
- $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(v_1, v_2) \mid v_1, v_2 \in \mathbb{R}\}$
- $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R} = \{(v_1, v_2, v_3) \mid v_1, v_2, v_3 \in \mathbb{R}\}$

  .

  .

- $\mathbb{R}^n = \{(v_1, v_2, \ldots, v_n) \mid v_1, v_2, \ldots v_n \in \mathbb{R}\}$

where $(v_1, v_2, \ldots, v_n)$ is a point in $n-$dimensional space with corresponding vector of the form,

$$\mathbf{v} = [v_1, v_2, \ldots, v_n] \text{ or } \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_n \end{bmatrix} \in \mathbb{R}^n.$$

The vector addition and scalar multiplication properties, which we introduced for vectors in the plane $n = 2$, apply in the general case and are called the standard operations in $\mathbb{R}^n$:

$$[u_1, \dots, u_n] + [v_1, \dots, v_n] = [u_1 + v_1, \dots, u_n + v_n],$$
$$\alpha[u_1, \dots, u_n] = [\alpha u_1, \dots, \alpha u_n].$$

Remark 1.1.5 *Note that when we use notation* $(u_1, \dots, u_n)$ *or* $(v_1, \dots, v_n)$ *to represent the coordinates then the analogous expressions are*

$$(u_1, \dots, u_n) + (v_1, \dots, v_n) = (u_1 + v_1, \dots, u_n + v_n),$$
$$\alpha(u_1, \dots, u_n) = (\alpha u_1, \dots, \alpha u_n).$$

Implied by the properties of the standard operations all the properties described in 1.1.3 also apply for vectors in $\mathbb{R}^n$.

Therefore, problems like

$$\boldsymbol{\alpha}(\mathbf{x} + 3\mathbf{y}) = \mathbf{u} + \boldsymbol{\beta}\mathbf{x} + 2(\mathbf{v} - \mathbf{x}) \quad \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}, \boldsymbol{\beta} \neq \boldsymbol{\alpha} + 2$$

for vectors $\mathbf{u}, \mathbf{v}, \mathbf{x}, \mathbf{y}$ in $\mathbb{R}^n$ can be solved for $\mathbf{x}$ using standard algebraic operations.

$$\boldsymbol{\alpha}\mathbf{x} + 3\boldsymbol{\alpha}\mathbf{y} = \mathbf{u} + \mathbf{x}(\boldsymbol{\beta} - 2) + 2\mathbf{v}$$

$$\mathbf{x}(\boldsymbol{\alpha} - \boldsymbol{\beta} + 2) = \mathbf{u} + 2\mathbf{v} - 3\boldsymbol{\alpha}\mathbf{y}$$

$$\mathbf{x} = \frac{\mathbf{u} + 2\mathbf{v} - 3\boldsymbol{\alpha}\mathbf{y}}{\boldsymbol{\alpha} - \boldsymbol{\beta} + 2}$$

1.1.3. The Dot Product and Length

Let

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_n \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ u_n \end{bmatrix},$$

then the dot product, $\mathbf{u} \cdot \mathbf{v}$, is defined by

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots u_n v_n = \sum_{i=1}^{n} u_i v_i.$$

Theorem 1.1.6 *Given* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ *and scalar* $c \in \mathbb{R}$, *the following properties hold*

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ *Commutativity*
2. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ *Distributivity*
3. $c(\mathbf{u} \cdot \mathbf{v}) = (c\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (c\mathbf{v}) = c\mathbf{u} \cdot \mathbf{v}$.
4. $\mathbf{u} \cdot \mathbf{u} \geq \mathbf{0}$ *and* $\mathbf{u} \cdot \mathbf{u} = \mathbf{0}$ *if and only if* $\mathbf{u} = \mathbf{0}$.

The length of a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_n \end{bmatrix}$, is the scalar

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

A unit vector is any vector of length 1.

Example 1.1.7 *For any vector* $\mathbf{v} \neq \mathbf{0}$ *we have the corresponding unit vector* $\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}$.

Finding such a unit vector is known as normalizing.

1.1.4. Important Inequalities

Given two vectors, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, the Cauchy-Schwarz inequality states:

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

Claim: $|\mathbf{u}\cdot\mathbf{v}| \leq \|\mathbf{u}\|\|\mathbf{v}\|$

proof 1.1.8 *Suppose, neither* $\mathbf{u}$ *nor* $\mathbf{v}$ *is zero otherwise we have* $\mathbf{0} \leq \mathbf{0}$ *and we are done. Then,*

$$\begin{aligned} \mathbf{0} &\leq \left\| \frac{\mathbf{u}}{\|\mathbf{u}\|} - \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\|^{\mathbf{2}} \\ &= \left(\frac{\mathbf{u}}{\|\mathbf{u}\|} - \frac{\mathbf{v}}{\|\mathbf{v}\|}\right)\cdot\left(\frac{\mathbf{u}}{\|\mathbf{u}\|} - \frac{\mathbf{v}}{\|\mathbf{v}\|}\right) = \mathbf{1} - \frac{\mathbf{2u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} + \mathbf{1} \end{aligned}$$

Therefore,

$$\frac{\mathbf{2u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} \leq \mathbf{2},$$

and

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

Note that, if $\mathbf{u}\cdot\mathbf{v} \geq \mathbf{0}$ *then we have*

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

If $\mathbf{u} \cdot \mathbf{v} < \mathbf{0}$ *then consider*

$$\begin{aligned} \mathbf{0} &\le \left\| \frac{\mathbf{u}}{\|\mathbf{u}\|} + \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\|^{\mathbf{2}} \\ &= \left(\frac{\mathbf{u}}{\|\mathbf{u}\|} + \frac{\mathbf{v}}{\|\mathbf{v}\|} \right) \cdot \left(\frac{\mathbf{u}}{\|\mathbf{u}\|} + \frac{\mathbf{v}}{\|\mathbf{v}\|} \right) = \mathbf{1} + \frac{\mathbf{2u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} + \mathbf{1}. \end{aligned}$$

This leads to

$$-\frac{\mathbf{2u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \le \mathbf{2},$$

and

$$-\mathbf{u} \cdot \mathbf{v} \le \|\mathbf{u}\| \|\mathbf{v}\|.$$

Since $-\mathbf{u} \cdot \mathbf{v} = |\mathbf{u} \cdot \mathbf{v}|$, *we again have*

$$|\mathbf{u} \cdot \mathbf{v}| \le \|\mathbf{u}\| \|\mathbf{v}\|.$$

This completes the proof. $\square$

For any two vectors, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, the Triangle inequality states:

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

Claim: $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

proof 1.1.9

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) = \mathbf{u}\cdot\mathbf{u} + 2\mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{v} \\
&\leq \mathbf{u}\cdot\mathbf{u} + 2|\mathbf{u}\cdot\mathbf{v}| + \mathbf{v}\cdot\mathbf{v} \\
&\leq \mathbf{u}\cdot\mathbf{u} + 2\|\mathbf{u}\|\|\mathbf{v}\| + \mathbf{v}\cdot\mathbf{v} \quad \textit{C.S.I.} \\
&= \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2
\end{aligned}$$

Therefore,

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

This leads to

$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|. \qquad \square$$

In summary, for vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ we have the two important inequalities:

1. Cauchy-Schwarz

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

2. Triangle Inequality

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

Exercise 1.1.10 *Suppose that $x_i \in \mathbb{R}$, $i = 1, \ldots, n$ and*

$$\sum_{i=1}^{n} x_i = 1.$$

Prove that

$$\sum_{i=1}^{n} x_i^2 \geq \frac{1}{n}.$$

1.1.5. Angle between vectors

For two nonzero vectors in $\mathbb{R}^n$, the Cauchy-Schwartz inequality implies that

$$-1 \leq \frac{\mathbf{u} \cdot \mathbf{v}}{||\mathbf{u}|| \, ||\mathbf{v}||} \leq 1.$$

Note, that any number in the interval $[-1, 1]$ is the cosine of a unique angle θ. If $n = 2, 3$ this angle is the angle between the corresponding planar or spatial vectors. This concept can be generalized for vectors in $\mathbb{R}^n$.

Therefore, we define the angle between vectors $\mathbf{u}$ and $\mathbf{v}$ as

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{||\mathbf{u}|| \, ||\mathbf{v}||}.$$

Two vectors in $\mathbb{R}^n$ are said to be orthogonal if $\mathbf{u} \cdot \mathbf{v} = \mathbf{0}$.

Exercise 1.1.11 *Find all vectors orthogonal to* $\mathbf{u} = [4, -1, 0]$.

Theorem 1.1.12 *Pythagoras in* $\mathbb{R}^n$*:*

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

if and only if $\mathbf{u}\cdot\mathbf{v} = \mathbf{0}$ *(orthogonal).*

Claim: $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$

proof 1.1.13

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

Clearly, if $\mathbf{u}\cdot\mathbf{v} = \mathbf{0}$ *then*

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2. \qquad \square$$

Exercise 1.1.14 *Show that the Pythagorean theorem is true for* $\mathbf{u} = [\mathbf{1},\mathbf{1},\mathbf{0}]$ *and* $\mathbf{v} = [\mathbf{0},\mathbf{0},\mathbf{1}]$.

1.2. Applications

1.2.1. Modular Arithmetic

Firstly, we need to define the space $\mathbb{Z}_m$ integers modular m e.g. binary consists of integers modular 2, the set $\{0, 1\}$. In general, $\mathbb{Z}_m$ is the set of scalars $\{0, 1, 2, \dots, m-1\}$ and for any elements $x, y \in \mathbb{Z}_m$ we have the rules

$$(x+y)|_m = \text{remainder}\left(\frac{(x+y)}{m}\right),$$

$$xy|_m = \text{remainder}\left(\frac{xy}{m}\right).$$

Example 1.2.1 $\mathbb{Z}_3 = \{0, 1, 2\}$ *and*

$$(1+2)|_3 = 0, \quad (2+2)|_3 = 1,$$

$$2 \times 2|_3 = 1.$$

We denote by $\mathbb{Z}_m^n$ the $n-$dimensional vector modular m.

Exercise 1.2.2 *Fill in the table using* $\mathbb{Z}_4$

$+$	$\mathbf{0}$	$\mathbf{1}$	$\mathbf{2}$	$\mathbf{3}$
$\mathbf{0}$				
$\mathbf{1}$				
$\mathbf{2}$				
$\mathbf{3}$				

and

$\times$	$\mathbf{0}$	$\mathbf{1}$	$\mathbf{2}$	$\mathbf{3}$
$\mathbf{0}$				
$\mathbf{1}$				
$\mathbf{2}$				
$\mathbf{3}$				

1.2.2. Error Detecting Codes

A hiker is using a GPS system that receives information in binary. He wants to know whether to go North, East, South, or West. Assume the GPS received the following binary code in $\mathbb{Z}_2^2$

$$[0,0] = \boldsymbol{N}, \quad [0,1] = \boldsymbol{E}, \quad [1,1] = \boldsymbol{S}, \quad [1,0] = \boldsymbol{W}.$$

If in sending the message one error at most may occur then this could be disastrous. Since changing one entry (modular 2), would give a different direction.

Using $\mathbb{Z}_2^3$ we can use

$$[0,0,0] = \boldsymbol{N}, \quad [0,1,1] = \boldsymbol{E}, \quad [1,1,0] = \boldsymbol{S}, \quad [1,0,1] = \boldsymbol{W}.$$

Now, if at most one error occurs we immediately detect the error, since the resulting vector will not be recognised as any of $\boldsymbol{N}, \boldsymbol{E}, \boldsymbol{S}$, or $\boldsymbol{W}$.

1.2.3. Bar Codes

A bar code is of the form

$$\mathbf{u} = \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \cdot \\ \cdot \\ \boldsymbol{u}_{11} \\ \boldsymbol{d} \end{bmatrix} \in \mathbb{Z}_{10}^{12}.$$

The first eleven entries give the information concerning the product and the manufacturer. The last digit, $\boldsymbol{d}$, is the check digit and is such that $\mathbf{c} \cdot \mathbf{u} = \mathbf{0}$ in $\mathbb{Z}_{10}$, where $\mathbf{c}$ is check vector.

In this case

$$\mathbf{c} = \begin{bmatrix} \mathbf{3} \\ \mathbf{1} \\ \mathbf{3} \\ \mathbf{1} \\ \mathbf{3} \\ \mathbf{1} \\ \mathbf{3} \\ \mathbf{1} \\ \mathbf{3} \\ \mathbf{1} \\ \mathbf{3} \\ \mathbf{1} \end{bmatrix}.$$

Hence,

$$\mathbf{c} \cdot \mathbf{u} =$$

$$(\mathbf{3}(\boldsymbol{u}_1 + \boldsymbol{u}_3 + \boldsymbol{u}_5 + \boldsymbol{u}_7 + \boldsymbol{u}_9 + \boldsymbol{u}_{11}) +$$

$$(\boldsymbol{u}_2 + \boldsymbol{u}_4 + \boldsymbol{u}_6 + \boldsymbol{u}_8 + \boldsymbol{u}_{10}) + \boldsymbol{d})|_{10} = \mathbf{0}.$$

Consider the case

$$\mathbf{u} = \begin{bmatrix} \mathbf{0} \\ \mathbf{7} \\ \mathbf{4} \\ \mathbf{9} \\ \mathbf{2} \\ \mathbf{7} \\ \mathbf{0} \\ \mathbf{2} \\ \mathbf{0} \\ \mathbf{9} \\ \mathbf{4} \\ \mathbf{6} \end{bmatrix}.$$

When scanned the scanner picked up the number

$$\mathbf{v} = \begin{bmatrix} \mathbf{0} \\ \mathbf{7} \\ \mathbf{4} \\ \mathbf{2} \\ \mathbf{9} \\ \mathbf{7} \\ \mathbf{0} \\ \mathbf{2} \\ \mathbf{0} \\ \mathbf{9} \\ \mathbf{4} \\ \mathbf{6} \end{bmatrix}.$$

Giving $\mathbf{c} \cdot \mathbf{v} = \mathbf{4} \neq \mathbf{0}$ in $\mathbb{Z}_{10}$. Thus an error has been detected.

Exercise 1.2.3 *Will a single error or transposition of adjacent digits always be detected?*

1.2.4. International Standard Book Number

The international standard book number (ISBN) takes the form

$$\mathbf{a} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \\ \boldsymbol{x}_4 \\ \boldsymbol{x}_5 \\ \boldsymbol{x}_6 \\ \boldsymbol{x}_7 \\ \boldsymbol{x}_8 \\ \boldsymbol{x}_9 \\ \boldsymbol{X} \end{bmatrix} \quad \text{and check vector} \quad \mathbf{c} = \begin{bmatrix} \mathbf{10} \\ \mathbf{9} \\ \mathbf{8} \\ \mathbf{7} \\ \mathbf{6} \\ \mathbf{5} \\ \mathbf{4} \\ \mathbf{3} \\ \mathbf{2} \\ \mathbf{1} \end{bmatrix},$$

and it is required that $\mathbf{c} \cdot \mathbf{a} = \mathbf{0}$ in $\mathbb{Z}_{11}$. Note that,

$$\mathbf{c} \cdot \mathbf{a} = \sum_{i=1}^{10} (11 - i) a_i.$$

Exercise 1.2.4 *Will the ISBN check sum always detect single errors? How about the transposition of adjacent elements? Will they be detected?*

1.3. Vectors and Matrices

Matrices and matrix properties will be discussed in detail during Lectures 7-9. Here however we give a preliminary introduction to matrices in order to clearly distinguish row and column vectors and exploit their use to give an alternative representation of the sum associated with the dot product.

An m by n matrix is a rectangular list of elements with m rows and n columns. Using this definition, we can provide a new description for row and column vectors.

1.3.1. Column and row vectors

A row vector or row matrix is a $\mathbf{1} \times \boldsymbol{n}$ matrix,

$$[\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n],$$

which can be denoted by

$$[\qquad \mathbf{a} \qquad],$$

where $\mathbf{a} \in \mathbb{R}^{1 \times n}$.

A column vector or column matrix is a $\boldsymbol{n} \times 1$ matrix,

$$\begin{bmatrix} \boldsymbol{a}_1 \\ \cdot \\ \cdot \\ \boldsymbol{a}_n \end{bmatrix},$$

which can be denoted by

$$\begin{bmatrix} \\ \mathbf{a} \\ \\ \end{bmatrix},$$

where $\mathbf{a} \in \mathbb{R}^{n \times 1}$.

We use operation transpose to convert a row matrix into a column matrix.

e.g.

$$[\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n]^{\boldsymbol{T}} = \begin{bmatrix} \boldsymbol{a}_1 \\ \cdot \\ \cdot \\ \boldsymbol{a}_n \end{bmatrix} \quad \text{or} \quad [\quad \mathbf{a} \quad]^{\boldsymbol{T}} = \begin{bmatrix} \\ \mathbf{a} \\ \\ \end{bmatrix}$$

Similarly, operation transpose converts a column matrix into a row matrix.

$$\begin{bmatrix} \boldsymbol{a}_1 \\ \cdot \\ \cdot \\ \boldsymbol{a}_n \end{bmatrix}^{\boldsymbol{T}} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n] \text{ or } \begin{bmatrix} \\ \mathbf{a} \\ \\ \end{bmatrix}^{\boldsymbol{T}} = [\qquad \mathbf{a} \qquad]$$

1.3.2. Dot Product revisited

Let $\mathbf{u}$ and $\mathbf{v}$ be two column vectors or column matrices of $n \times 1$,

$$\mathbf{v} = \begin{bmatrix} u_1 \\ \cdot \\ \cdot \\ u_n \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} v_1 \\ \cdot \\ \cdot \\ v_n \end{bmatrix}.$$

Then the multiplication of a row matrix with a column matrix is defined by

$$\mathbf{u}^T \mathbf{v} = [u_1, \ldots, u_n] \begin{bmatrix} v_1 \\ \cdot \\ \cdot \\ v_n \end{bmatrix} = \sum_{i=1}^n u_i v_i.$$

In fact this forms the building block for general matrix multiplication.
For now notice that the dot product $\mathbf{u} \cdot \mathbf{v}$, which was defined as $\sum_{i=1}^n u_i v_i$, represents the same sum.

$$\mathbf{u}^T\mathbf{v} = [u_1, \ldots, u_n]\begin{bmatrix} v_1 \\ \cdot \\ \cdot \\ v_n \end{bmatrix} = \sum_{i=1}^n u_i v_i = \mathbf{u}\cdot\mathbf{v}.$$

Therefore,

$$\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} = \mathbf{v}^T\mathbf{u} = \mathbf{v}\cdot\mathbf{u}.$$

In later chapters $\mathbf{u}^T\mathbf{v}$ and $\mathbf{u}\cdot\mathbf{v}$ will be used interchangeably to represent the same sum.

1.3.3. Vectors and Matrices

A set of column vectors $\mathbf{a}_i \in \mathbb{R}^{m \times 1}, i = 1, ..., n$ are contained in an $m \times n$ matrix:

$$A_c = \left[\begin{array}{c|c|c} & & \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ & & \end{array} \right] \in \mathbb{R}^{m \times n}$$

Similarly an $n \times m$ matrix contains row vectors, $\mathbf{a}_i^T \in \mathbb{R}^{1 \times m}$, $i = 1, \ldots, n$.

$$A_r = \left[\begin{array}{c} \mathbf{a}_1^T \\ \hline \cdot \\ \hline \mathbf{a}_n^T \end{array} \right] \in \mathbb{R}^{n \times m}$$

1.4. Vector Spaces and Subspaces

1.4.1. Vector Spaces

The properties of vector addition and scalar multiplication, the standard operations in $\mathbb{R}^n$ give rise to a series of ten properties (rules), which can be derived from the properties of standard operations and the properties of addition and multiplication of real numbers. These rules, which we explicitly listed (see Theorem 1.1.3) for vectors in the plane $(n = 2)$, apply for all vectors in $\mathbb{R}^n$. By having a proper definition of addition and scalar multiplication other mathematical quantities can share these properties.

Definition 1.4.1 *Let $\boldsymbol{V}$ be a set on which two operations, addition and scalar multiplication have been defined. If $\mathbf{u}, \mathbf{v} \in \boldsymbol{V}$, the addition is denoted by $\mathbf{u} + \mathbf{v}$ and if $\boldsymbol{\alpha}$ is a scalar then scalar multiplication is denoted by $\boldsymbol{\alpha}\mathbf{u}$. If the following rules hold for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \boldsymbol{V}$ and for all scalars $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, then $\boldsymbol{V}$ is called a vector space and its elements are called vectors.*

1. $\mathbf{u}+\mathbf{v} \in \boldsymbol{V}$ *Closure under addition*
2. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$ *Commutativity*
3. $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$ *Associativity*
4. *There exists an element (zero vector), denoted* $\mathbf{0}$, *such that* $\mathbf{u}+\mathbf{0}=\mathbf{u}$ *Additive identity*
5. *For every* $\mathbf{u} \in \boldsymbol{V}$ *there exists an element* $-\mathbf{u} \in \boldsymbol{V}$ *such that* $\mathbf{u}+(-\mathbf{u})=\mathbf{0}.$ *Additive inverse*

$\uparrow$ *Addition* $\downarrow$ *Scalar Multiplication*

6. $\boldsymbol{\alpha}\mathbf{u} \in \boldsymbol{V}$ *Closure under scalar multiplication*
7. $\boldsymbol{\alpha}(\mathbf{u}+\mathbf{v})=\boldsymbol{\alpha}\mathbf{v}+\boldsymbol{\alpha}\mathbf{u}$ *Distributivity*
8. $(\boldsymbol{\alpha}+\boldsymbol{\beta})\mathbf{u}=\boldsymbol{\alpha}\mathbf{u}+\boldsymbol{\beta}\mathbf{u}$ *Distributivity*
9. $\boldsymbol{\alpha}(\boldsymbol{\beta}\mathbf{u})=(\boldsymbol{\alpha}\boldsymbol{\beta})\mathbf{u}$ *Associativity*
10. $\mathbf{1}\mathbf{u}=\mathbf{u}$ *Scalar identity*

When specifying a vector space, the set of vectors, set of scalars and the two operations need to be clearly described. In the following examples, the set of scalars is the set of real numbers.

Remark 1.4.2 *The definition of a vector space does not specify what the set $\boldsymbol{V}$ consists of.*

The following are examples of vector spaces.

Example 1.4.3 *For any* $n \geq 1$ *and integer,* $\mathbb{R}^n$ *is a vector space. In this course its also the most important. For* $n = 2$, *planar vectors satisfy the ten properties (see Theorem 1.1.3), which can be derived from the standard operations of vector addition and scalar multiplication.* $\mathbb{R}^1$ *with the familiar operations of addition and multiplication is a vector space.*

Example 1.4.4 *For any $n \geq 1$ and integer, let $\mathcal{P}^n$ denote the set of all polynomials with maximum degree n or less with real coefficients. Polynomials in $\mathcal{P}^n$ can be written as*

$$p(x) = p_0 + p_1 x + p_2 x^2 + \ldots + p_n x^n,$$

where $p_0, p_1, p_2, \ldots, p_n \in \mathbb{R}$. Given another polynomial

$$q(x) = q_0 + q_1 x + q_2 x^2 + \ldots + q_n x^n,$$

the sum (addition) of two polynomials are defined as

$$p(x)+q(x) = (p_0+q_0)+(p_1+q_1)x+(p_2+q_2)x^2+\cdots+(p_n+q_n)x^n.$$

In addition, scalar multiplication by scalar $c \in \mathbb{R}$ is defined as

$$cp(x) = cp_0 + cp_1 x + cp_2 x^2 + \ldots + cp_n x^n.$$

Exercise 1.4.5 *Show that $\mathcal{P}^2$ is a vector space.*

Example 1.4.6 *Consider the set*

$$C[a,b] = \{f \mid f : [a,b] \to \mathbb{R}, f \text{ continuous}\}.$$

of continuous functions defined on $[a,b]$. *Furthermore, we can define addition as*

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

and scalar multiplication as

$$(cf)(x) = cf(x).$$

One can also define the additive identity as

$$i_a : [a,b] \to \mathbb{R}, i_a(x) = 0.$$

Then, for any $f \in C[a,b]$

$$(f + i_a)(x) = f(x) + 0 = f(x)$$

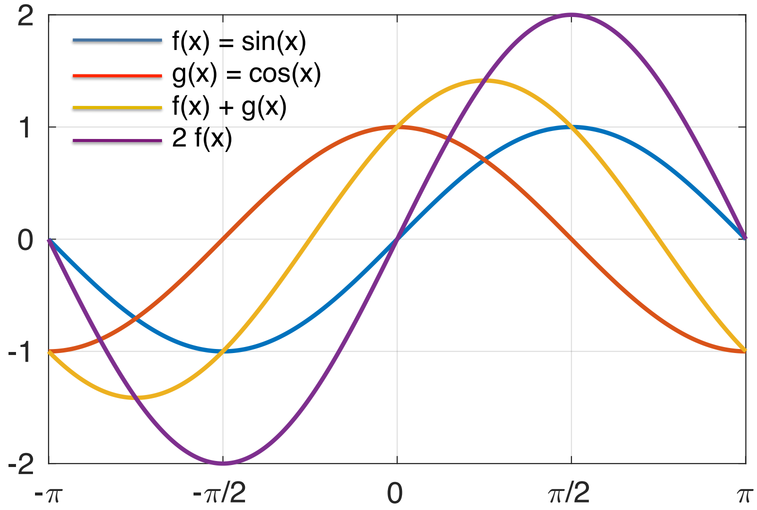
Note that $C[a,b]$ *(with the defined addition and scalar multiplication) forms a vector space, but we will not prove all axioms here. We note, however, that the following statement is true.*

If $f, g \in C[a,b]$ *then* $f + g \in C[a,b]$ *and* $cf \in C[a,b]$ *holds for any* $c \in \mathbb{R}$.

For example,

$$f : [-\pi, \pi] \to \mathbb{R}, f(x) = sin(x)$$

$$g : [-\pi, \pi] \to \mathbb{R}, g(x) = cos(x).$$



Exercise 1.4.7 *Does* $\mathbb{N}$*, the set of natural numbers (with standard addition and scalar multiplication) form a vector space?*

Exercise 1.4.8 *Consider the set of all third degree polynomials with addition and scalar multiplication properties defined above. Does it form a vector space?*

1.4.2. Vector subspaces

Definition 1.4.9 *A subspace, $\boldsymbol{W}$, of a vector space, $\boldsymbol{V}$, is a non-empty subset satisfying all the requirements of a vector space under the operations of addition and scalar multiplication defined in $\boldsymbol{V}$.*

Example 1.4.10 $V = \mathbb{R}^3$, $\boldsymbol{W} = \{(\boldsymbol{x_1}, \mathbf{2}, \boldsymbol{x_3}) \mid \boldsymbol{x_1}, \boldsymbol{x_3} \in \mathbb{R}\}$. $\boldsymbol{W}$ *is not a subspace* $\boldsymbol{V}$ *because* $(\mathbf{0}, \mathbf{0}, \mathbf{0}) \notin \boldsymbol{W}$, *additive identity is not in* $\boldsymbol{W}$.

Given that $\boldsymbol{W} \subset \boldsymbol{V}$ and have the same operations of addition and scalar multiplications we may expect that some of the vector space properties are inherited so we do not have to investigate all the ten axioms in Definition 1.4.1. Fortunately, there is a related theorem.

Theorem 1.4.11 *If* $\boldsymbol{W} \subseteq \boldsymbol{V}, \boldsymbol{V} \neq \emptyset$ *then* $\boldsymbol{W}$ *is a subspace if and only if* $\mathbf{u} + \mathbf{v} \in \boldsymbol{W}$ *for all* $\mathbf{u}, \mathbf{v} \in \boldsymbol{W}$ *and given any scalar* $\boldsymbol{c}$ *then* $\boldsymbol{c}\mathbf{w} \in \boldsymbol{W}$ *for all* $\mathbf{w} \in \boldsymbol{W}$*.*

proof 1.4.12 $\Rightarrow$ *If W is a subspace then it satisfies all the vector space axioms so it is closed under addition and scalar multiplication.*

$\Leftarrow$ *If W is closed under addition and scalar multiplication then axioms 1 and 6 are already satisfied. In addition, vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in W$ are in V. Axioms 1 and 6 together with the fact that $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ imply that axioms 2-3 and axioms 7-10 are satisfied. Furthermore, we can multiply any $\mathbf{u} \in W$ with $c = 0$ or $c = -1$ to generate the additive identity (axiom 4) or inverse (axiom 5).* $\square$

In other words, the theorem states that if $\boldsymbol{W}$ is a nonempty subset of $\boldsymbol{V}$ then it is enough to investigate closure under addition and closure under scalar multiplication (items 1. and 6. in Definition 1.4.1) in order to conclude whether or not $\boldsymbol{W}$ is a subspace.

Example 1.4.13 *The space $\mathcal{P}^{n-1}$ is a subspace of $\mathcal{P}^n$. In fact, $\mathcal{P}^n$ is a subspace of $\mathcal{P}^m$ if $n \leq m$ and the subspaces satisfy*

$$\mathcal{P}^0 \subset \mathcal{P}^1 \subset \mathcal{P}^2 \subset \cdots \subset \mathcal{P}^n.$$

Exercise 1.4.14 $W = \{(x_1, x_2, x_3) \mid x_1, x_2, x_3 \in \mathbb{R}_+\}$. $V = \mathbb{R}^3$
Is W a subspace of V?

Exercise 1.4.15 $W = \{(x, y) \mid x^2 + y^2 = 1\}$. $V = \mathbb{R}^2$
Is W a subspace of V?

1.5. Complex Numbers and Complex Vector Spaces (optional)

1.5.1. Complex Numbers

The general solution of the quadratic equation $ax^2 + bx + c = 0$ can be written as

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

where $b^2 - 4ac$ is called the discriminant. If $b^2 - 4ac \geq 0$ then the solution(s) are real numbers but what can we conclude in cases when the discriminant is negative?

For example, consider the equation,

$$x^2 + 2 = 0.$$

In this case the discriminant is $b^2 - 4ac = -8$ but there is no real number we could square to obtain -8. To overcome this problem, we introduce the imaginary unit,

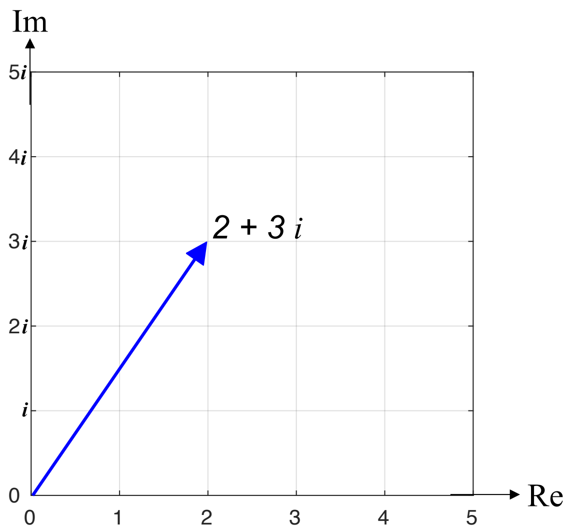
$$i = \sqrt{-1},$$

where $i^2 = -1$. Using the imaginary unit, we have $\sqrt{-8} = 2\sqrt{2}i$ and $(2\sqrt{2}i)(2\sqrt{2}i) = -8$.

Definition 1.5.1 *If $a, b \in \mathbb{R}$ then the number $a + bi$ is a complex number, where a is the real part and bi is the imaginary part of the complex number. (A complex number is uniquely determined by its imaginary and real parts.)*

1.5.2. Complex Plane

If we associate the complex number $a + bi$ with the ordered pair (a, b) then complex numbers can be represented in a coordinate plane, which is called the complex plane. For example, the complex number $2 + 3i$ can be represented in the complex plane.



Note that the horizontal axis is the real axis and the vertical axis is the imaginary axis.

1.5.3. Addition and Scalar Multiplication

The sum of two complex numbers $\boldsymbol{a+bi}$ and $\boldsymbol{c+di}$ are defined as

$$\boldsymbol{(a+bi)+(c+di)=(a+c)+(b+d)i}.$$

If $\boldsymbol{c} \in \mathbb{R}$ and $\boldsymbol{a+bi}$ is a complex number then the scalar multiple of $\boldsymbol{c}$ and $\boldsymbol{a+bi}$ is defined as:

$$\boldsymbol{c(a+bi)=ca+cbi}.$$

Exercise 1.5.2 *(optional) Show that with addition and scalar multiplication, the set of complex numbers forms a vector space (over $\mathbb{R}$).*

1.5.4. Multiplication, Conjugates and Division

The multiplication of complex numbers $a + bi$ and $c + di$ is performed by applying the distributive property so that

$$(a + bi)(c + di) = ac - bd + (ad + bc)i.$$

To facilitate introducing the concept of complex conjugates consider using the quadratic formula to solve the equation $x^2 + 4x + 5 = 0$ and get

$$x = \frac{-4 \pm \sqrt{4^2 - 20}}{2} = \frac{-4 \pm 2i}{2} = -2 \pm i,$$

where $-2 + i$ and $-2 - i$ are called complex conjugates.

Definition 1.5.3 *The conjugate of a the complex number* $z = a + bi$ *is* $\overline{z} = a - bi$.

Some properties of complex conjugates are outlined here:

Theorem 1.5.4 *Let* $z = a + bi$ *be a complex number and* $\overline{z}$ *its conjugate. Then the following properties are true:*

1. $z\overline{z} = a^2 + b^2$
2. $z\overline{z} = 0$ *iff* $z = 0$
3. $z\overline{z} \geq 0$
4. $\overline{\overline{z}} = z$

Given that a complex number can be represented as a vector in the complex plane, the length of this vector can characterise the complex number. This length is also called the modulus of the complex number.

Definition 1.5.5 *The modulus of the complex number* $z = a + bi$ *is* $|z| = \sqrt{a^2 + b^2}$.

Exercise 1.5.6 *(optional) Show that* $|z|^2 = z\overline{z}$.

In order to demonstrate division using complex numbers, consider $s = a + bi$ and $q = c + di$ and assume that $s/q = x + yi$ and $c^2 + d^2 \neq 0$. Furthermore, we could write

$$a + bi = s = (x + yi)(c + di) = (cx - dy) + (cy + dx)i$$

Next, the system of linear equations,

$$\begin{aligned} a &= cx - dy \\ b &= cy + dx, \end{aligned}$$

can be solved for x and y, leading to

$$\begin{aligned} x &= \frac{ca + bd}{c^2 + d^2} = \frac{ca + bd}{q\bar{q}} \\ y &= \frac{cb - ad}{c^2 + d^2} = \frac{cb - ad}{q\bar{q}}. \end{aligned}$$

The quotient of the complex numbers $s = a + bi$ and $q = c + di$ can also be written as

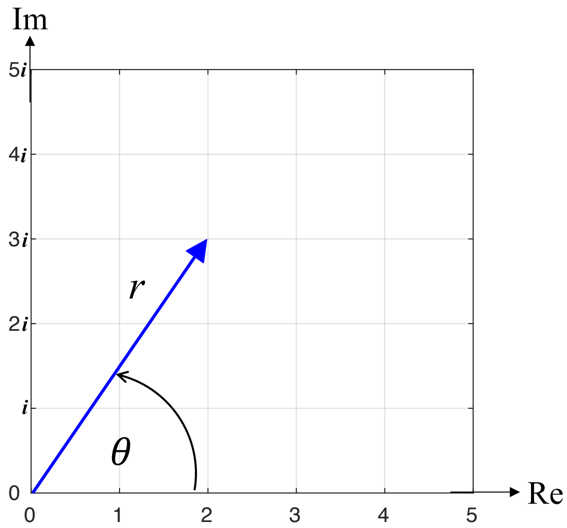
$$\frac{s}{q} = \frac{a + bi}{c + di} = \frac{ca + bd}{q\bar{q}} + \frac{cb - ad}{q\bar{q}} i = \frac{s\bar{q}}{\mid q \mid^2}.$$

A very practical representation for complex numbers is the polar form.

Definition 1.5.7 *The polar form of a nonzero complex number* $z = a + bi$ *is*

$$z = r(\cos(\theta) + i\sin(\theta)),$$

where $a = r\cos(\theta), b = r\sin(\theta), r = |z| = \sqrt{a^2 + b^2}$ *and* $\tan(\theta) = b/a$. *We call* r *the modulus and* θ *the argument of the complex number.*



Note, that there are infinitely many choices for the argument and therefore it is often preferred to use the value of θ such that $-\pi < \theta \leq \pi$. This value is called the principal argument and it is denoted by $Arg(z)$. Two complex numbers in polar form are equal if and only if they have the same modulus and principal argument.

By using the polar form representation of complex numbers $\boldsymbol{s = r_1(\cos(\theta_1) + i\sin(\theta_1))}$ and $\boldsymbol{q = r_2(\cos(\theta_2) + i\sin(\theta_2))}$ their product and quotient (if $\boldsymbol{q \neq 0}$) becomes:

$$\boldsymbol{sq = r_1 r_2(\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2))}$$

$$\boldsymbol{\frac{s}{q} = \frac{r_1}{r_2}(\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2))}$$

1.5.5. Complex Vector Spaces

Analogously to $\mathbb{R}^n$, for any $n > 0$ integer we can define $\mathbb{C}^n$, which consists of ordered n-tuples of complex numbers:

- $\mathbb{C}^1 =$ set of all complex numbers
- $\mathbb{C}^2 = \mathbb{C} \times \mathbb{C} = \{(v_1, v_2) \mid v_1, v_2 \in \mathbb{C}\}$
- $\mathbb{C}^3 = \mathbb{C} \times \mathbb{C} \times \mathbb{C} = \{(v_1, v_2, v_3) \mid v_1, v_2, v_3 \in \mathbb{C}\}$

$\vdots$

- $\mathbb{C}^n = \{(v_1, v_2, \ldots, v_n) \mid v_1, v_2, \ldots v_n \in \mathbb{C}\}$

where $(v_1, v_2, \ldots, v_n)$ is a point in $n-$dimensional complex space with corresponding vector of the form,

$$\mathbf{v} = [v_1, v_2, \ldots, v_n] \text{ or } \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_n \end{bmatrix} \in \mathbb{C}^n.$$

Using $v_k = a_k + b_k i$ for $k = 1, \ldots n$, we can also write

$$\mathbf{v} = [a_1 + b_1 i, a_2 + b_2 i, \ldots, a_n + b_n i] \text{ or } \begin{bmatrix} a_1 + b_1 i \\ a_2 + b_2 i \\ \cdot \\ \cdot \\ a_n + b_n i \end{bmatrix} \in \mathbb{C}^n.$$

The vector addition and scalar multiplication with complex numbers in $\mathbb{C}^n$ are performed componentwise analogously to the vector addition and scalar multiplication in $\mathbb{R}^n$. Therefore, for any $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ and $\boldsymbol{\alpha} \in \mathbb{C}$, we have

$$[\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n] + [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n] = [\boldsymbol{u}_1 + \boldsymbol{v}_1, \ldots, \boldsymbol{u}_n + \boldsymbol{v}_n],$$
$$\boldsymbol{\alpha}[\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n] = [\boldsymbol{\alpha u}_1, \ldots, \boldsymbol{\alpha u}_n].$$

Therefore, $\mathbb{C}^n$ is a complex vector space under vector addition and scalar multiplication.

1.5.6. Applications

Many applications exploit the use of Euler's formula, which states that for any real number $\varphi \in \mathbb{R}$, we have

$$e^{i\varphi} = \cos(\varphi) + i\ \sin(\varphi).$$

This formula establishes a relationship between trigonometric functions and the the exponential function. In the special case of $\varphi = \pi$, the formula also leads to the famous Euler's identity,

$$e^{i\pi} + 1 = 0.$$

One of the beauty of Euler's identity is that it unites three constants, Euler's number, e, the imaginary unit, i and π.

The application of complex numbers is substantial in mathematics, engineering and sciences. Here we only mention a few applications of complex numbers.

- One simple practical application of complex numbers is to derive trigonometric identities. For example, for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}$ we have:

$$\begin{aligned} \cos(x+y) + i\sin(x+y) &= e^{i(x+y)} \\ &= e^{ix}\, e^{iy} \\ &= (\cos(x) + i\ \sin(x))(\cos(y) + i\ \sin(y)) \\ &= (\cos(x)\cos(y) - \sin(x)\sin(y)) \\ &\quad + i\ (\cos(x)\sin(y) + \sin(x)\cos(y)). \end{aligned}$$

Therefore,

$$\begin{aligned} \cos(x+y) &= (\cos(x)\cos(y) - \sin(x)\sin(y)), \\ \sin(x+y) &= (\cos(x)\sin(y) + \sin(x)\cos(y)). \end{aligned}$$

For another example, note that

$$\begin{aligned}
1 = e^0 &= e^{ix}e^{-ix} \\
&= (\cos(x) + i\sin(x))(\cos(-x) + i\sin(-x)) \\
&= (\cos(x) + i\sin(x))(\cos(x) - i\sin(x)) \\
&= \cos(x)^2 + \sin(x)^2.
\end{aligned}$$

- In signal processing, the Fourier transform of a signal is informative about what frequencies are present in your signal. It decomposes a function of time into frequencies that make it up. The Fourier transform of an integrable function $f : \mathbb{R} \longrightarrow \mathbb{C}$ is defined as

$$\hat{f}(y) = \int_{-\infty}^{\infty} f(x) e^{-i(2\pi xy)} dx.$$

- Iterations with complex numbers can be used to build certain fractals, which are objects that are self-similar across different scales. Fractals has applications from medicine, biology, astronomy and computer science. For example, in computer graphics it is often essential to be able to generate realistic and not repeating landscapes. This can be achieved by fractal terrain generations methods that are designed to produce fractal behaviour that mimics the appearance of natural terrains.

  One of the most famous fractal is the Mandelbrot Set named after Benoit Mandelbrot and is based on a sequence of complex numbers:

$$z_n = (z_{n-1})^2 + c, \qquad z_1 = c.$$

For some $c \in \mathbb{C}$, the sequence is bounded so that $|z_i| < N$ for all i. In this case c is in the Mandelbrot Set. For some other $c \in \mathbb{C}$ the sequence is unbounded; it diverges and becomes infinitely large and c is not in the Mandelbrot Set. Given this behaviour it is possible to create a diagram by assigning black to all points in the Mandelbrot Set. For all other points a lighter colour is assigned based on how quickly the corresponding series diverges.