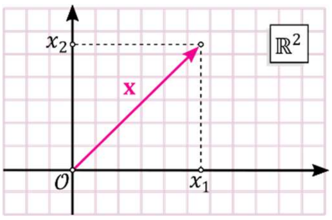
LECTURE 00

Vector Spaces and Linear Algebra

Brief Theory

The Euclidean vector space $\mathbb{R}^m$



- Let m be a non-vanishing natural number, that is, $m \in \mathbb{N} \setminus \{0\} = \mathbb{N}^*$. Then, the members of the real **Euclidean vector space** $\mathbb{R}^m$ are called **vectors** and are denoted by lowercase bold latin letters, such as $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$. Provided a coordinate system $\mathcal{O}$, a vector $\mathbf{x} \in \mathbb{R}^m$ is written as a column of m ordered real numbers x_1, x_2, ..., x_m that are called the **components** of $\mathbf{x}$ in $\mathcal{O}$, that is,

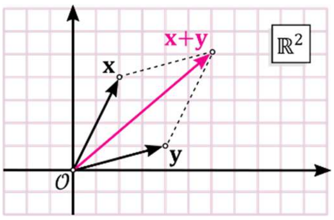
$$\mathbf{x} =_{\mathcal{O}} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}.$$

The row vector whose columns are the rows of $\mathbf{x}$ is called the **transpose** of $\mathbf{x}$ and is denoted by

$$\mathbf{x}^\top =_{\mathcal{O}} [x_1, x_2, \dots, x_m].$$

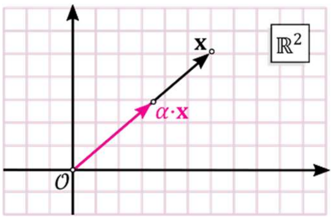
Further, $\mathbf{x} =_{\mathcal{O}} [x_1, x_2 \dots, x_m]^\top$.

Unless otherwise marked, the components of all vectors are given in a common coordinate system $\mathcal{O}$, and hence, the symbol $\mathcal{O}$ is omitted from the equality sign, when writing the components of a vector.



- In $\mathbb{R}^m$, **vector addition** $(\star + \star)\colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$ and **scalar multiplication** $(\star \cdot \star)\colon \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$are induced by the corresponding operations in $\mathbb{R}$. In particular, if $\alpha \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, then

$$\mathbf{x} + \mathbf{y} = [x_1 + y_1, \dots x_m + y_m]^\top, \qquad \alpha \cdot \mathbf{x} = [\alpha x_1, \dots, \alpha x_m]^\top.$$



- Let $\emptyset \neq Q \subset \mathbb{R}^m$. If $\mathbf{x} + \mathbf{y} \in Q$ and $\alpha \cdot \mathbf{x} \in Q$ for all $\alpha \in \mathbb{R}$ and all $\mathbf{x}, \mathbf{y} \in Q$, then Q is said to be a **subspace** of $\mathbb{R}^m$.

- Consider $n \in \mathbb{N}^*$ vectors $\mathbf{x}^1$, $\mathbf{x}^2$, ..., $\mathbf{x}^n \in \mathbb{R}^m$. Any vector

$$\mathbf{x} = \alpha_1 \mathbf{x}^1 + \alpha_2 \mathbf{x}^2 + \dots + \alpha_n \mathbf{x}^n,$$

where α_1, α_2, ..., $\alpha_n \in \mathbb{R}$, is said to be a **linear combination** of the vectors $\mathbf{x}^1$, $\mathbf{x}^2$, ..., $\mathbf{x}^n$.

- Let $\mathbf{x} = \alpha_1\mathbf{x}^1 + \alpha_2\mathbf{x}^2 + \cdots + \alpha_n\mathbf{x}^n$ be a linear combinations of the vectors $\mathbf{x}^1$, $\mathbf{x}^2$, ..., $\mathbf{x}^n \in \mathbb{R}^m$. If $\mathbf{x} = \mathbf{0}$ entails that $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$, then the vectors $\mathbf{x}^1$, $\mathbf{x}^2$, ..., $\mathbf{x}^n$ are said to be **linearly independent**. Otherwise, if there exists a vanishing linear combination, while not all α_1, α_2, ..., α_n are equal to zero, then the vectors $\mathbf{x}^1$, $\mathbf{x}^2$, ..., $\mathbf{x}^n$ are called **linearly dependent**.

- A set of vectors $\mathbf{x}^1$, $\mathbf{x}^2$, ..., $\mathbf{x}^m \in \mathbb{R}^m$ is said to **span** $\mathbb{R}^m$, if every vector in $\mathbb{R}^m$ can be written as a linear combination of the vectors $\mathbf{x}^1$, $\mathbf{x}^2$, ..., $\mathbf{x}^m$. The set of all linear combinations of the vectors $\mathbf{x}^1$, $\mathbf{x}^2$, ..., $\mathbf{x}^m$ is denoted by

$$\langle \mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^m \rangle = \{\mathbf{x} \in \mathbb{R}^m \,|\, \mathbf{x} = \alpha_1\mathbf{x}^1 + \alpha_2\mathbf{x}^2 + \cdots + \alpha_m\mathbf{x}^m\}.$$

If $\mathbf{x}^1$, $\mathbf{x}^2$, ..., $\mathbf{x}^m \in \mathbb{R}^m$ span $\mathbb{R}^m$, then $\langle \mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^m \rangle = \mathbb{R}^m$.

- If the vectors $\mathbf{x}^1$, $\mathbf{x}^2$, ..., $\mathbf{x}^m \in \mathbb{R}^m$ are
  (a) linearly independent and
  (b) span $\mathbb{R}^m$,

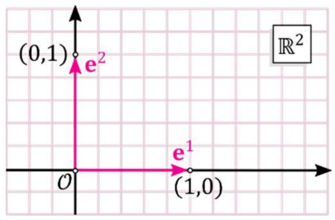
then $\mathbf{x}^1$, $\mathbf{x}^2$, ..., $\mathbf{x}^m$ constitute a **basis** of $\mathbb{R}^m$ and are called **basis vectors**. The largest number of linearly independent vectors in $\mathbb{R}^m$ is $\dim(\mathbb{R}^m) = m$ and is called the **dimension** of $\mathbb{R}^m$.

Let $\mathbf{x}^1$, $\mathbf{x}^2$, ..., $\mathbf{x}^m \in \mathbb{R}^m$ be a basis of $\mathbb{R}^m$. Then a $\mathbf{x} \in \mathbb{R}^m$ can be written as $\mathbf{x} = \alpha_1\mathbf{x}^1 + \alpha_2\mathbf{x}^2 + \cdots + \alpha_m\mathbf{x}^m$ and the real scalar weights α_1, α_2, ..., α_n are called the **components** of $\mathbf{x}$ in that basis.

- The m-dimensional Euclidean vector space $\mathbb{R}^m$ is equipped with the so-called **standard basis vectors** $\mathbf{e}^1$, $\mathbf{e}^2$, ..., $\mathbf{e}^m$, where

$$\mathbf{e}^i = [0, \dots, 0, 1, 0, \dots, 0]^\top \in \mathbb{R}^m, \qquad i \in \{1, 2, \dots, m\},$$

and the unique non-zero element is located at position with index i.



Unless otherwise marked, the components of all vectors in $\mathbb{R}^m$ are given in the standard basis.

- The **ℓ^2-scalar product** $(\star, \star)_2 \colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is the symmetric and bilinear mapping whose values are

$$(\mathbf{x}, \mathbf{y})_2 = \mathbf{x}^\top\mathbf{y} = \mathbf{y}^\top\mathbf{x} = x_1y_1 + x_2y_2 + \cdots + x_my_m,$$

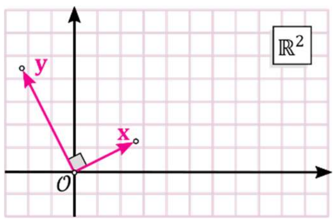
with $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^m$. Mark that $(\mathbf{x}, \mathbf{x})_2 \geq 0$ for all $\mathbf{x} \in \mathbb{R}^m$.

Recall that the ℓ^2-scalar product is given by

$$(\mathbf{x}, \mathbf{y})_2 = \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \cos\theta,$$

where θ is the angle between the vectors $\mathbf{x}$, $\mathbf{y}$.

Two vectors $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^m$ are orthogonal (perpendicular) if, and only if, $(\mathbf{x}, \mathbf{y})_2 = 0$; that is, $\mathbf{x} \perp \mathbf{y} \Leftrightarrow (\mathbf{x}, \mathbf{y})_2 = 0$.

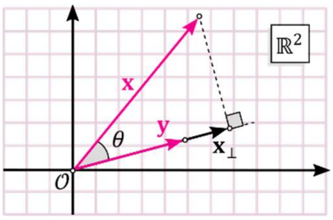


- The $\boldsymbol{\ell^2}$**-norm** $\|\star\|_2\colon \mathbb{R}^m \to \mathbb{R}$ is defined by $\|\mathbf{x}\|_2 = \sqrt{(\mathbf{x},\mathbf{x})_2}$, with $\mathbf{x} \in \mathbb{R}^m$. The ℓ^2-norm satisfies the following properties;
  (a) $\|\mathbf{x}\|_2 \geq 0$ for all $\mathbf{x} \in \mathbb{R}^m$;
  (b) $\|\mathbf{x}\|_2 = 0$ if, and only if, $\mathbf{x} = \mathbf{0}$;
  (c) $\|\alpha\mathbf{x}\|_2 = |\alpha| \cdot \|\mathbf{x}\|_2$ for all $\alpha \in \mathbb{R}$ and all $\mathbf{x} \in \mathbb{R}^m$;
  (d) $\|\mathbf{x}+\mathbf{y}\|_2 \leq \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2$ for all $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^m$.

A vector $\mathbf{x} \in \mathbb{R}^m$ whose ℓ^2-norm is unity, $\|\mathbf{x}\|_2 = 1$, is called **unit vector**. Provided a non-vanishing vector $\mathbf{x} \in \mathbb{R}^m$, the vector $\mathbf{x}/\|\mathbf{x}\|_2$ is a unit vector in the direction of $\mathbf{x}$. The ℓ^2-norm of a vector is also called the Euclidean norm or the magnitude of that vector,

$$\|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_m^2}.$$

The standard basis vectors $\mathbf{e}^1$, $\mathbf{e}^2$, ..., $\mathbf{e}^m \in \mathbb{R}^m$ *are orthonormal, that is,* $(\mathbf{e}^i,\mathbf{e}^j)_2 = 0$ *for all distinct* i, $j \in \{1,2,\dots,m\}$, *while if* $i = j$, *then* $(\mathbf{e}^i,\mathbf{e}^i)_2 = 1$ *and* $\left\|\mathbf{e}^i\right\|_2 = 1$ *for all* $i \in \{1,2,\dots,m\}$. *Further, the* i*-th component* x_i *of a vector* $\mathbf{x} \in \mathbb{R}^m$ *can be written as* $x_i = (\mathbf{x},\mathbf{e}^i)_2$.



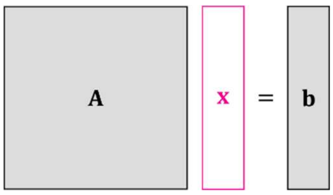
- If $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^m$ and $\mathbf{y}$ is a unit vector, that is $\|\mathbf{y}\|_2 = 1$, the absolute scalar product $|(\mathbf{x},\mathbf{y})_2|$ is the magnitude of the vector $\mathbf{x}_\perp$ that is formed by projecting $\mathbf{x}$ on $\mathbf{y}$.

Square linear systems

- Let a_i^j, x_i, $b_i \in \mathbb{R}$ for all i, $j \in \{1,2,\dots,m\}$, where all a_{ij} and b_i are known, while all x_i are uknown. Then, the algebraic system

$$a_i^1 x_1 + a_i^2 x_2 + \cdots + a_i^m x_m = b_i, \qquad i \in \{1,2,\dots,m\},$$

can be written as $\mathbf{A}\mathbf{x} = \mathbf{b}$, with $\mathbf{A} \in \mathbb{R}^{m\times m}$ being a square matrix and $\mathbf{x}$, $\mathbf{b} \in \mathbb{R}^m$ being vectors defined by



$$\underbrace{\begin{bmatrix} a_1^1 & \cdots & a_1^m \\ \vdots & \ddots & \vdots \\ a_m^1 & \cdots & a_m^m \end{bmatrix}}_{=\mathbf{A}} \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}}_{=\mathbf{x}} = \underbrace{\begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}}_{=\mathbf{b}}.$$

- A square matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ is said to be **symmetric**, if it is equal to its transpose; that is, $\mathbf{A} = \mathbf{A}^\top$, meaning that $a_i^j = a_j^i$ for all i, $j \in \{1,2,\dots,m\}$.

- A symmetric matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ is said to be **positive definite**, if $\mathbf{x}^\top\mathbf{A}\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$. When this is the case, all eigenvalues of $\mathbf{A}$ are greater than zero.

- A square matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ is said to be **sparse**, if most of its entries are vanishing.

- A matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ can be written as an ordered collection of its columns, that is, $\mathbf{A} = [\,\mathbf{a}^1 \mid \mathbf{a}^2 \mid \cdots \mid \mathbf{a}^m]$, where $\mathbf{a}^j \in \mathbb{R}^m$ for all $j \in \{1,2,\dots,m\}$. Hence, the left-hand side of $\mathbf{Ax} = \mathbf{b}$ can be written as a linear combination of the columns of $\mathbf{A}$,

$$\mathbf{Ax} = x_1\mathbf{a}^1 + x_2\mathbf{a}^2 + \cdots + x_m\mathbf{a}^m.$$

- The **column space** $\mathcal{C}(\mathbf{A})$ of a matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ is the space spanned by the columns of $\mathbf{A}$, that is, $\mathcal{C}(\mathbf{A}) = \langle \mathbf{a}^1, \mathbf{a}^2, \dots, \mathbf{a}^m \rangle$, where $\mathbf{a}^j \in \mathbb{R}^m$, with $j \in \{1,2,\dots,m\}$, are the columns of $\mathbf{A}$. The dimension of $\mathcal{C}(\mathbf{A})$ is called the (column) **rank** of $\mathbf{A}$ and is denoted by $\operatorname{rank}(\mathbf{A}) = \dim(\mathcal{C}(\mathbf{A}))$.

- If the columns of a matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ are linearly independent, $\mathbf{A}$ is said to have **full-rank** or to be **non-singular**. When this is the case, $\operatorname{rank}(\mathbf{A}) = m$.

- When the rank of a matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ is smaller than m, then $\mathbf{A}$ is said to be **rank-deficient** or **singular** and $\det(\mathbf{A}) = 0$.

- Let $\mathbf{A} \in \mathbb{R}^{m\times m}$. The set $\mathcal{N}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{Ax} = \mathbf{0}\}$ is called the **kernel** or the **nullspace** of $\mathbf{A}$. The dimension of the kernel of $\mathbf{A}$ is called its **nullity** and is denoted by $\operatorname{nullity}(\mathbf{A}) = \dim(\mathcal{N}(\mathbf{A}))$. The rank and the nullity of $\mathbf{A}$ satisfy the so-called **rank-nullity theorem**

$$\operatorname{rank}(\mathbf{A}) + \operatorname{nullity}(\mathbf{A}) = m.$$

Since the members $\mathbf{x}$ of the kernel of a matrix $\mathbf{A}$ satisfy the equation $\mathbf{Ax} = \mathbf{0}$, each vector in the kernel of $\mathbf{A}$ is orthogonal to the rows of that matrix.

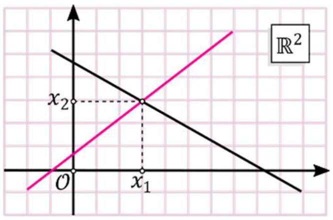
- Let $\mathbf{Ax} = \mathbf{b}$, with $\mathbf{A} \in \mathbb{R}^{m\times m}$ and $\mathbf{x}$, $\mathbf{b} \in \mathbb{R}^m$. If the matrix $\mathbf{A}$ has full-rank, the inverse matrix $\mathbf{A}^{-1} \in \mathbb{R}^{m\times m}$ exists and is defined by

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_m,$$

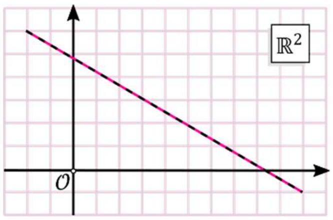
where $\mathbf{I}_m$ is the $m \times m$ identity matrix. Hence, the solution $\mathbf{x}$ is uniquely determined by $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. If $\mathbf{A}$ is singular, then

  (a) $\mathbf{b} \in \mathcal{C}(\mathbf{A})$ entails that the system $\mathbf{Ax} = \mathbf{b}$ has infinitely many solutions, while

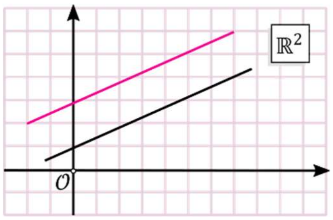
  (b) $\mathbf{b} \notin \mathcal{C}(\mathbf{A})$ entails that the system $\mathbf{Ax} = \mathbf{b}$ has no solutions.



Unique solution



Infinitely many solutions



No solutions

Eigenvalue problems

- Let $\mathbf{A} \in \mathbb{R}^{m\times m}$, $\mathbf{x} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$, and $\alpha \in \mathbb{C}$. Then, equation

$$\mathbf{Ax} = \alpha\mathbf{x}$$

is said to be the **eigenvalue problem** of $\mathbf{A}$, $\mathbf{x}$ is called an **eigenvector** of $\mathbf{A}$, and α is called the **eigenvalue** of $\mathbf{A}$ that is associated with $\mathbf{x}$.

- If $\mathbf{x}$ is an eigenvector of $\mathbf{A} \in \mathbb{R}^{m\times m}$, then all $\beta\mathbf{x}$, with $\beta \in \mathbb{C} \setminus \{0\}$, are also eigenvectos of $\mathbf{A}$, and hence, eigenvalue problems are not uniquely solvable.

- The eigenvalue problem of a matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ can be written as

$$(\mathbf{A} - \alpha\mathbf{I}_m)\mathbf{x} = \mathbf{0}.$$

If $\mathbf{A} - \alpha\mathbf{I}_m$ is non-singular, then $\mathbf{x} = \mathbf{0}$ and is not a valid solution, since it contradicts the assumption $\mathbf{x} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$. Hence, $\mathbf{A} - \alpha\mathbf{I}_m$ has to be singular, and as such, its determinant must be vanishing, that is,

$$\det(\mathbf{A} - \alpha\mathbf{I}_m) = 0.$$

The polynomial whose values are given by $\phi_{\mathbf{A}}(\alpha) = \det(\mathbf{A} - \alpha\mathbf{I}_m)$ is called the **characteristic polynomial** of $\mathbf{A}$ and its roots are the eigenvalues of $\mathbf{A}$.

- The set $\sigma(\mathbf{A})$ of the eigenvalues of a matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ is called the **spectrum** of $\mathbf{A}$, that is,

$$\sigma(\mathbf{A}) = \{\alpha \in \mathbb{C} \mid \phi_{\mathbf{A}}(\alpha) = 0\}.$$

The absolutely largest eigenvalue of $\mathbf{A}$ is said to be the **dominant eigenvalue** and the corresponding eigenvector is called a **dominant eigenvector**. The number $\rho(\mathbf{A}) = \max\{|\alpha| \colon \alpha \in \sigma(\mathbf{A})\}$ is called the **spectral radius** of $\mathbf{A}$.

Worked Examples

0.1 Let $\boldsymbol{Q}$ be the set of vectors in $\mathbb{R}^3$ that are orthogonal to $\mathbf{a} = [3, 2, 1]^\top$. (a) Prove that $\boldsymbol{Q}$ is a subspace of $\mathbb{R}^3$, (b) find a basis for $\boldsymbol{Q}$, and (c) determine its dimension.

SOLUTION

Let $\mathbf{x} = [x_1, x_2, x_3]^\top \in \mathbb{R}^3$ be an arbitrary vector that is orthogonal to $\mathbf{a}$, that is,

$$\mathbf{a} \perp \mathbf{x} \Leftrightarrow \mathbf{a}^\top\mathbf{x} = 0 \Leftrightarrow 3x_1 + 2x_2 + x_3 = 0,$$

and hence, $Q = \{\mathbf{x} \in \mathbb{R}^3 \mid 3x_1 + 2x_2 + x_3 = 0\}$.

(a) For $\mathbf{x} = \mathbf{0} \in \mathbb{R}^3$, $\mathbf{a}^\top\mathbf{x} = 0$, and hence $\mathbf{0} \in Q$. Further, consider $\mathbf{x}, \mathbf{y} \in Q$, that is,

$$\mathbf{a}^\top\mathbf{x} = 0, \qquad \mathbf{a}^\top\mathbf{y} = 0. \text{ (I)}$$

Hence,

$$\mathbf{a}^\top(\mathbf{x}+\mathbf{y}) = \mathbf{a}^\top\mathbf{x} + \mathbf{a}^\top\mathbf{y} \overset{\text{(I)}}{=} 0,$$

meaning that $\mathbf{x}+\mathbf{y} \in Q$. Similarly, $\alpha\mathbf{x} \in Q$ for all $\mathbf{x} \in Q$ and all $\alpha \in \mathbb{R}$. Thus, Q is a subspace of $\mathbb{R}^3$.

(b) By employing the equation that is satisfied by the members of Q, the third component x_3 can be eliminated, that is, $x_3 = -3x_1 - 2x_2$. Hence, every $\mathbf{x} \in Q$ can be written as

$$\begin{aligned}\mathbf{x} = [x_1, x_2, x_3]^\top &= [x_1, x_2, -3x_1 - 2x_2]^\top \\ &= x_1 \underbrace{[1,0,-3]^\top}_{\mathbf{q}^1} + x_2 \underbrace{[0,1,-2]^\top}_{\mathbf{q}^2} = x_1\mathbf{q}^1 + x_2\mathbf{q}^2\end{aligned}$$

and the vectors $\mathbf{q}^1$, $\mathbf{q}^2$ span Q, that is, $Q = \langle \mathbf{q}^1, \mathbf{q}^2 \rangle$. Further, since $[\,\mathbf{q}^1 \mid \mathbf{q}^2\,][x_1, x_2]^\top = \mathbf{0}$ entails $x_1 = x_2 = 0$, the vectors $\mathbf{q}^1$, $\mathbf{q}^2$ are linearly independent. Thus, $\mathbf{q}^1$ and $\mathbf{q}^2$ are basis vectors.

The components of the members of Q satisfy $3x_1 + 2x_2 + x_3 = 0$, which means that they belong to a plane, and hence, $\dim(Q) = 2$.

(c) The dimension of a (sub)space coincides with the number of basis vectors; hence, $\dim(Q) = 2$.

0.2 The so-called ℓ^p-norms, with $p \geq 1$, are defined by

$$\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_m|^p)^{1/p}$$

for all $\mathbf{x} \in \mathbb{R}^m$. Prove that

$$\lim_{p\to\infty} \|\mathbf{x}\|_p = \max\{|x_i| \mid i \in \{1, 2, \dots, m\}\}.$$

SOLUTION

Let $\mathbf{x} \in \mathbb{R}^m$ and $M = \max\{|x_i| \mid i \in \{1,2,\dots,m\}\}$. Then,

$$\begin{aligned}\|\mathbf{x}\|_p &= (|x_1|^p + |x_2|^p + \cdots + |x_m|^p)^{1/p} \\ &\leq (M^p + M^p + \cdots + M^p)^{1/p} = Mm^{1/p}.\end{aligned}$$

Further, $\|\mathbf{x}\|_p \geq M$. Thus, $M \leq \|\mathbf{x}\|_p \leq Mm^{1/p}$ and an application of the squeeze theorem,

$$M = \lim_{p\to\infty} M \leq \lim_{p\to\infty} \|\mathbf{x}\|_p \leq \lim_{p\to\infty} Mm^{1/p} = M,$$

results in the sought equation.

Since $M = 0$ implies $\mathbf{x} = \mathbf{0}$, it is easy to check that the function with values

$$\|\mathbf{x}\|_\infty = \lim_{p\to\infty} \|\mathbf{x}\|_p$$

defines a norm, namely the ℓ^∞-norm.

0.3 Prove the so-called Cauchy-Schwarz inequality

$$|(\mathbf{x}, \mathbf{y})_2| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, with $m > 1$.

SOLUTION

The left-hand side of the given inequality is

Given the definition of the component-wise or Hadamard product,

$$\mathbf{x} \odot \mathbf{y} = [x_1 y_1, \ldots, x_m y_m]^\top,$$

the quantity A is equal to the ℓ^1-norm of $\mathbf{x} \odot \mathbf{y}$.

$$\begin{aligned} |(\mathbf{x}, \mathbf{y})_2| &= |x_1 y_1 + x_2 y_2 + \cdots + x_m y_m| \\ &\le |x_1 y_1| + |x_2 y_2| + \cdots + |x_m y_m| = A. \text{ (I)} \end{aligned}$$

Let $\not{f}\colon \mathbb{R} \to \mathbb{R}$ be the function with values

$$\begin{aligned} \not{f}(\alpha) &= (\alpha|x_1| + |y_1|)^2 + (\alpha|x_2| + |y_2|)^2 + \cdots + (\alpha|x_m| + |y_m|)^2 \\ &= \|\mathbf{x}\|_2^2 \alpha^2 + 2A\alpha + \|\mathbf{y}\|_2^2 \ge 0. \end{aligned}$$

Since $\not{f}(\alpha)$ is a second-degree polynomial and $\not{f}(\alpha) \ge 0$ for all $\alpha \in \mathbb{R}$, the equation $\not{f}(\alpha) = 0$ has a single or no roots in $\mathbb{R}$. Thus,

$$4A^2 - 4\|\mathbf{x}\|_2^2 \|\mathbf{y}\|_2^2 \le 0 \Leftrightarrow A \le \|\mathbf{x}\|_2 \|\mathbf{y}\|_2. \text{ (II)}$$

The sought inequality follows from (I) and (II).

0.4 Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times m}$ satisfy the equation $\mathbf{A} + \mathbf{B} + 2020\mathbf{AB} = \mathbf{0}$. Prove that (a) $\mathbf{I}_m + 2020\mathbf{A}$ is invertible and that (b) $\mathbf{AB} = \mathbf{BA}$.

SOLUTION

(a) The given equation can be recast as follows;

$$\begin{aligned} \mathbf{A} + \mathbf{B} + 2020\mathbf{AB} = \mathbf{0} &\Leftrightarrow \mathbf{A} + (\mathbf{I}_m + 2020\mathbf{A})\mathbf{B} = \mathbf{0} \\ &\Leftrightarrow 2020\mathbf{A} + 2020(\mathbf{I}_m + 2020\mathbf{A})\mathbf{B} = \mathbf{0} \\ &\Leftrightarrow \mathbf{I}_m + 2020\mathbf{A} + 2020(\mathbf{I}_m + 2020\mathbf{A})\mathbf{B} = \mathbf{I}_m \\ &\Leftrightarrow (\mathbf{I}_m + 2020\mathbf{A})(\mathbf{I}_m + 2020\mathbf{B}) = \mathbf{I}_m, \text{ (I)} \end{aligned}$$

and hence, $\mathbf{I}_m + 2020\mathbf{A}$ is non-singular with inverse

$$(\mathbf{I}_m + 2020\mathbf{A})^{-1} = \mathbf{I}_m + 2020\mathbf{B}.$$

(b) From equation (I) follows that

An invertible matrix $\mathbf{A}$ commutes with its inverse, that is,

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A}.$$

$$\begin{aligned} (\mathbf{I}_m + 2020\mathbf{B})(\mathbf{I}_m + 2020\mathbf{A}) &= \mathbf{I}_m \\ &\Leftrightarrow \mathbf{I}_m + 2020\mathbf{B} + (\mathbf{I}_m + 2020\mathbf{B})2020\mathbf{A} = \mathbf{I}_m \\ &\Leftrightarrow \mathbf{B} + (\mathbf{I}_m + 2020\mathbf{B})\mathbf{A} = \mathbf{0}, \end{aligned}$$

and hence, $\mathbf{B} + \mathbf{A} + 2020\mathbf{BA} = \mathbf{0}$. (II). The sought equation follows from (II) and the given equation.

Two matrices $\mathbf{A}$, $\mathbf{B}$ are said to be similar, if there exists a non-singular matrix $\mathbf{P}$, such that

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}.$$

0.5 If $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times m}$ satisfy the equation $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$ for some $\mathbf{P} \in \mathbb{R}^{m\times m}$, then prove that $\mathbf{A}$ and $\mathbf{B}$ have the same eigenvalues.

SOLUTION

Let $\phi_{\mathbf{A}}(x) = \det(\mathbf{A} - x\mathbf{I}_m)$ and $\phi_{\mathbf{B}}(x) = \det(\mathbf{B} - x\mathbf{I}_m)$ be the characteristic polynomials of $\mathbf{A}$ and $\mathbf{B}$, respectively. Then,

$\det(\mathbf{AB}) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$

$$\begin{aligned} \phi_{\mathbf{B}}(x) = \det(\mathbf{B} - x\mathbf{I}_m) &= \det(\mathbf{P}^{-1}\mathbf{AP} - x\mathbf{P}^{-1}\mathbf{P}) \\ &= \det[\mathbf{P}^{-1}(\mathbf{A} - x\mathbf{I}_m)\mathbf{P}] \\ &= \det(\mathbf{P}^{-1}) \cdot \det(\mathbf{A} - x\mathbf{I}_m) \cdot \det(\mathbf{P}) \\ &= \det(\mathbf{P}^{-1}\mathbf{P}) \cdot \det(\mathbf{A} - x\mathbf{I}_m) = \det(\mathbf{A} - x\mathbf{I}_m) \\ &= \phi_{\mathbf{A}}(x) \end{aligned}$$

for all $x \in \mathbb{C}$. Hence, the characteristic polynomials of $\mathbf{A}$ and $\mathbf{B}$ are equal, and as such, they have the same roots, which are the common eigenvalues of $\mathbf{A}$ and $\mathbf{B}$.

0.6 **Let $\mathbf{A} \in \mathbb{R}^{m\times m}$ be non-singular, symmetric, and positive definite and consider the function $f\colon \mathbb{R}^m \to \mathbb{R}$ with values**

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x} + \mathbf{c}.$$

Prove that f has a unique minimum at $\mathbf{x}_* \in \mathbb{R}^m$ that satisfies the equation $\mathbf{A}\mathbf{x}_* = \mathbf{b}$.

SOLUTION

The i-th component of the gradient

$$\nabla f = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_m}\right]^\top$$

of $f\colon \mathbb{R}^m \to \mathbb{R}$ is

$$\begin{aligned}\frac{\partial f}{\partial x_i} &\overset{\text{(I)}}{=} \frac{1}{2}\frac{\partial \mathbf{x}^\top}{\partial x_i}\mathbf{A}\mathbf{x} + \frac{1}{2}\mathbf{x}^\top\mathbf{A}\frac{\partial \mathbf{x}}{\partial x_i} - \mathbf{b}^\top\frac{\partial \mathbf{x}}{\partial x_i} \overset{\text{(II)}}{=} \frac{1}{2}(\mathbf{e}^i)^\top\mathbf{A}\mathbf{x} + \frac{1}{2}\mathbf{x}^\top\mathbf{A}\mathbf{e}^i \\ &\quad - \mathbf{b}^\top\mathbf{e}^i = \frac{1}{2}(\mathbf{A}\mathbf{x})_i + \frac{1}{2}(\mathbf{A}^\top\mathbf{x})_i - b_i \overset{\text{(III)}}{=} (\mathbf{A}\mathbf{x})_i - b_i,\end{aligned}$$

$$[0,\cdots,0,1,0,\cdots,0]\begin{bmatrix}\vdots\\ \mathbf{A}\mathbf{x}\\ \vdots\end{bmatrix} = (\mathbf{A}\mathbf{x})_i$$

that is, $\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$. Equality (I) follows from the differentiation rule for the product of two functions, in equality (II) the standard basis vectors are used, while equality (III) follows from the symmetry of $\mathbf{A}$. The location of a candidate extremum can be found by setting

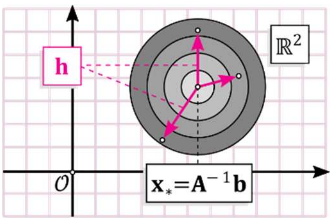
$$\nabla f(\mathbf{x}_*) = \mathbf{0} \Leftrightarrow \mathbf{A}\mathbf{x}_* = \mathbf{b}.$$

Since $\mathbf{A}$ is non-singular, the system $\mathbf{A}\mathbf{x}_* = \mathbf{b}$ has a unique solution $\mathbf{x}_* = \mathbf{A}^{-1}\mathbf{b}$. To characterize the extremum at $\mathbf{x}_*$, the function f is expanded around $\mathbf{x}_*$. The Taylor expansion of f stops at second order terms, since f is a polynomial of degree two. Hence, for all directions $\mathbf{h} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$,

$$\begin{aligned}f(\mathbf{x}_* + \mathbf{h}) &= f(\mathbf{x}_*) + \mathbf{h}^\top\nabla f(\mathbf{x}_*) + \frac{1}{2}\mathbf{h}^\top\mathbf{H}_f(\mathbf{x}_*)\mathbf{h} \\ &= f(\mathbf{x}_*) + \frac{1}{2}\mathbf{h}^\top\mathbf{A}\mathbf{h} > f(\mathbf{x}_*),\end{aligned}$$

since $\mathbf{h}^\top\mathbf{A}\mathbf{h} > 0$ because $\mathbf{A}$ is positive definite. Here, $\mathbf{H}_f$ is the Hessian matrix, that is the matrix whose entries are the second-order partial derivatives of f. Since $\nabla f(\mathbf{x})$ is linear in $\mathbf{x}$, the Hessian matrix is $\mathbf{H}_f(\mathbf{x}) = \mathbf{A}$. Inequality $f(\mathbf{x}_* + \mathbf{h}) > f(\mathbf{x}_*)$ implies that the

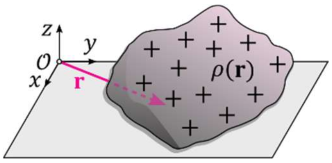
$$\mathbf{H}_f = \begin{bmatrix}\frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m}\\ \vdots & \ddots & \vdots\\ \frac{\partial^2 f}{\partial x_m \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_m^2}\end{bmatrix}$$



function values in any direction are greater than the function value at $\mathbf{x}_*$, meaning that the function $\oint$ attains a unique global minimum value $\oint(\mathbf{x}_*)$ that is located at $\mathbf{x}_* = \mathbf{A}^{-1}\mathbf{b}$.

Computing Lab

Problem description



- Let $\mathbf{r} = [x, y, z]^\top$ be the position vector in $\mathbb{R}^3$ with respect to the origin $\mathcal{O}(0,0,0)$ and consider an electric charge density $\rho(\mathbf{r})$ that is located in vacuum and is at rest. Then, the differential form of the **Gauss law** is

$$\nabla \cdot \mathbf{E}(\mathbf{r}) = \frac{\rho(\mathbf{r})}{\epsilon_0},$$

where

$$\nabla = \left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right]^\top$$

is the **nabla operator** in the considered Cartesian coordinate system, $\nabla \cdot \mathbf{E}(\mathbf{r}) = \mathrm{div}\big(\mathbf{E}(\mathbf{r})\big)$ is the **divergence** of the electric field $\mathbf{E}(\mathbf{r}) = \big[E_x(\mathbf{r}), E_y(\mathbf{r}), E_z(\mathbf{r})\big]^\top$, and ϵ_0 is the vacuum permittivity. If the field $\mathbf{E}(\mathbf{r})$ is **irrotational** or **curl-free**, $\nabla \times \mathbf{E}(\mathbf{r}) = \mathrm{rot}\big(\mathbf{E}(\mathbf{r})\big) = \mathbf{0}$, then there exists a **scalar potential** function $u\colon \mathbb{R}^3 \to \mathbb{R}$, such that

$$\mathbf{E}(\mathbf{r}) = -\nabla u(\mathbf{r}) = -\left[\frac{\partial u(\mathbf{r})}{\partial x}, \frac{\partial u(\mathbf{r})}{\partial y}, \frac{\partial u(\mathbf{r})}{\partial z}\right]^\top.$$

Hence, the Gauss law is recast in the form

$$\nabla \cdot \nabla u(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon_0} \Leftrightarrow \Delta u(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon_0}, \text{(I)}$$

where

$$\Delta = \nabla \cdot \nabla = \frac{\partial}{\partial x}\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\frac{\partial}{\partial y} + \frac{\partial}{\partial z}\frac{\partial}{\partial z} = \frac{\partial}{\partial x} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

is the **Laplace operator**.

- By assuming that the scalar potential is constant in the y and z directions, problem (I) is constrained along the x-axis. Further, the charge density is assumed to be $\rho(x) = \epsilon_0$ throughout the interval $\mathrm{X} = (0,1)$ and that the scalar potential $u(x)$ vanishes at the boundary of X, that is, u satisfies the so-called **Dirichlet boundary conditions** $u(0) = u(1) = 0$ (II). The Gauss law (I) together with the boundary

Let u be a sufficiently smooth function. Similar to scalar(-vector) multiplication, if $u(\mathbf{x}) \in \mathbb{R}$, then the gradient of u is

$$\nabla u = \left[\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}\right]^\top.$$

Similar to the ℓ^2-scalar product, if $u(\mathbf{x}) \in \mathbb{R}^3$, then the divergence of u is

$$\nabla \cdot u = \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} + \frac{\partial u_z}{\partial z}.$$

Similar to the cross product, if $u(\mathbf{x}) \in \mathbb{R}^3$, then the rotation of u is

$$\nabla \times u = \det\left(\begin{bmatrix} \mathbf{e}^1 & \mathbf{e}^2 & \mathbf{e}^3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ u_x & u_y & u_z \end{bmatrix}\right).$$

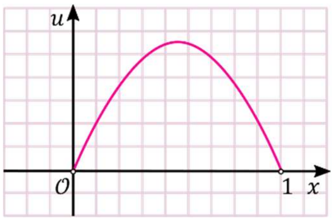
conditions (II) constitute a **boundary value problem** that needs to be solved for the scalar potential u. In particular, the

$$u''(x) = -1 \quad \forall x \in \mathrm{X}, \qquad u(0) = u(1) = 0, \quad \text{(III)}$$

where $u'' = \mathrm{d}^2/\mathrm{d}x^2$ is the second-order derivative of u with respect to the space variable x.

The exact solution of boundary value problem (III) is

$$u(x) = -\frac{1}{2}x(x-1).$$

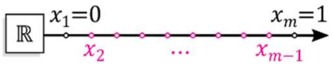


Discretization

- Let $x_1, x_2, \ldots, x_m \in \mathrm{X} \cup \{0,1\}$ be m equidistant points, with $x_1 = 0$ and $x_m = 1$. Hence,

$$0 = x_1 < x_2 < \cdots < x_{m-1} < x_m = 1$$

and $x_{i+1} = x_i + h$ for all $i \in \{1,2,\ldots,m-1\}$. Here, $0 < h \ll 1$ is a measure of the fineness of the resulting partition X_h of X. Given the function values $u(x_i)$, **Taylor's theorem** provides the approximations



$$u(x_{i-1}) = u(x_i - h) \cong u(x_i) - hu'(x_i) + \frac{1}{2}h^2u''(x_i),$$
$$u(x_{i+1}) = u(x_i + h) \cong u(x_i) + hu'(x_i) + \frac{1}{2}h^2u''(x_i).$$

By truncating the infinite Taylor expansion of a smooth function u around some point x, an approximation of u around x is obtained. The most commonly used approximation is the one obtained by preserving only linear in h terms and neglecting second and higher-order terms.

By adding these two approximations and setting $u_i = u(x_i)$, the derivative $u''(x_i)$ is approximated by

$$u''(x_i) \cong \frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} \quad \text{(IV)}$$

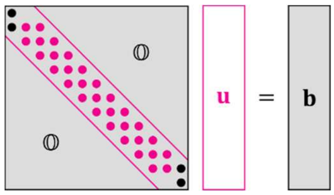
for all $i \in \{2,3,\ldots,m-1\}$, that is, for all points in X.

- The points with indices $i = 1$ and $i = m$ require some extra attention, since the boundary conditions $u_1 = u_m = 0$ must be considered. For that purpose, the equations

$$u_1 = 0 \Leftrightarrow 1 \cdot u_1 + 0 \cdot u_2 + \cdots + 0 \cdot u_{m-1} + 0 \cdot u_m = 0,$$
$$u_m = 0 \Leftrightarrow 0 \cdot u_1 + 0 \cdot u_2 + \cdots + 0 \cdot u_{m-1} + 1 \cdot u_m = 0$$

are introduced. By substituting the approximations (IV) of $u''(x_i)$ into BVP (III), the following system of equations is obtained;

$$\underbrace{\frac{1}{h^2}\begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_{m-1} \\ u_m \end{bmatrix}}_{\mathbf{u}} = \underbrace{\begin{bmatrix} 0 \\ -1 \\ -1 \\ \vdots \\ -1 \\ 0 \end{bmatrix}}_{\mathbf{b}},$$



where the first and last rows of the coefficient matrix **A** ensure that the Dirichlet boundary conditions are satisfied by the computed solution.

For symmetric problems in more than one space dimensions, the approach of «additional equations» for imposing Dirichlet boundary conditions breaks the symmetry of the matrix **A**. *This artifact can be avoided by using one of the following methods.*
- *The boundary nodes are not included as unknowns, while their contribution is considered by modifying the right-hand side of the resulting linear system.*
- *The diagonal entries of* **A** *that are associated with boundary nodes and the corresponding* **b** *components are multiplied with a large number, in the so-called penalty methods.*

Brief python implementation

- The following code lines introduce the required tools.

```
In [1]: import numpy as np
In [2]: import scipy.sparse as sp
In [3]: from scipy.sparse.linalg import spsolve
```

- There are two ways to generate the points x_i; either by choosing the number m of these points, or by choosing the spacing h. Here, these points are generated with the following code snippet.

```
In [4]: h = 0.001
In [5]: m = int(1/h+1)
In [6]: Xh = np.linspace(0, 1, m)
```

The right-hand side vector **b** is computed by

```
In [7]: b = -np.ones((m, 1))
```

- To construct the coefficient matrix **A**, observe that **A** is sparse with $3m - 4 = 2999$ nonvanishing entries, that is, $\cong 0.3\%$ of the total number m^2 of its entries. Matrices with mostly vanishing elements are called sparse and numerical software stores them by squeezing out the zero elements; in python, functions from the `scipy.sparse` package are used for sparse matrices.

```
In [8]: d0 = -2*np.ones((1,m))
In [9]: d0[0,0], d0[0,-1] = 1, 1 # BC entries
In[10]: d1 = np.ones((1, m-1))
In[11]: d1[0,0] = 0              # BC entries
In[12]: d2 = np.flip(d1, 1) #
In[13]: A = (1/h**2)*(sp.diags(d0, [0]) + \
            sp.diags(d1, [1]) + sp.diags(d2,[-1]))
```

Further, for sparse linear systems, specialized solution algorithms are available. Given $\mathbf{A}$ and $\mathbf{b}$, the numerical solution $\mathbf{u}$ is computed as the solution to the linear system $\mathbf{Au} = \mathbf{b}$ by typing

```
In[14]: u = spsolve(A, b)
```

Evaluation Quiz

Double-choice questions

Instructions. *Characterize the following statements as true (T) or false (F) and justify your choice. Unless otherwise noted, m is a natural number greater than unity.*

0.1 $\mathbb{R}^{m-1}$ is a subspace of $\mathbb{R}^m$.

0.2 A single vector in $\mathbb{R}^m$ is linearly independent.

0.3 If $\mathbf{A} \in \mathbb{R}^{m\times m}$, then $\det(\mathbf{A}) = 0$ implies $\operatorname{rank}(\mathbf{A}) = m$.

0.4 Let $\mathbf{A} \in \mathbb{R}^{m\times m}$ be a matrix. If $\mathbf{Ax} = \mathbf{0}$, then $\mathbf{x} = \mathbf{0}$.

0.5 If $\mathbf{A} \in \mathbb{R}^{m\times m}$ is a non-singular matrix, then the dimension of the kernel of $\mathbf{A}$ is unity.

0.6 If the columns of $\mathbf{A} \in \mathbb{R}^{m\times m}$ are linearly independent, then $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$.

0.7 Let $\mathbf{A} \in \mathbb{R}^{m\times m}$ be a non-singular matrix. If $\mathbf{A}^\top = \mathbf{A}^{-1}$, then $\|\mathbf{Ax}\|_2 = \|\mathbf{x}\|_2$ for all $\mathbf{x} \in \mathbb{R}^m$.

0.8 If $\mathbf{A} \in \mathbb{R}^{m\times m}$ is a matrix and the system $\mathbf{Ax} = \mathbf{b}$ has a solution, then $\det(\mathbf{A}) \neq 0$.

0.9 If $\mathbf{A} \in \mathbb{R}^{2\times 2}$ is a matrix with strictly positive determinant and vanishing sum of diagonal entries, then all eigenvalues of $\mathbf{A}$ are real numbers.

0.10 If the matrices $\mathbf{A}$, $\mathbf{B} \in \mathbb{R}^{m\times m}$ are similar, then they have the same set of eigenvectors.

Exercises and problems

Instructions. *Solve the following tasks and justify each step of the solution process, in detail.*

0.11 Prove that $A = \{[\alpha, 2\alpha]^\top \in \mathbb{R}^2 \mid \alpha \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^2$.

0.12 Prove that (a) $A = \{[w, x, y, z] \in \mathbb{R}^4 \mid w + x + y + z = 0\}$ is a subspace of $\mathbb{R}^4$. (b) Find a basis for A and (c) determine the dimension of A.

0.13 Prove that $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x} - 2\mathbf{y}, 2\mathbf{x} + \mathbf{y} \rangle$ for all $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^m$.

0.14 Let $\mathscr{f}\colon \mathbb{R}^3 \to \mathbb{R}^3$ be the function with values $\mathscr{f}(x_1, x_2, x_3) = [x_1 + x_2 + x_3, 2x_1 - x_2, -x_1 + 2y_2 + x_3]^\top$. Prove that the nullity of the matrix that can be associated with $\mathscr{f}$ is unity.

0.15 Prove that if $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times m}$ such that $\mathbf{AB} = \mathbf{A}$ and $\mathbf{B}^2 = 0$, then $\mathbf{A} = \mathbf{0}$.

0.16 Let $\mathbf{A} \in \mathbb{R}^{m \times m}$ such that $\mathbf{A}^2 - 2\alpha\mathbf{A} + \mathbf{I}_m = \mathbf{0}$, where $\alpha \in \mathbb{R}$. (a) For which values of α is the matrix $\mathbf{A} + \mathbf{I}_m$ invertible and (b) for which values of α is the matrix $\mathbf{A} - \mathbf{I}_m$ invertible?

0.17 Prove that the matrices $\mathbf{AB}$ and $\mathbf{BA}$, where $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times m}$, have the same eigenvalues.

0.18 Prove that $\|\mathbf{x} + \mathbf{y}\|_2^2 = \|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2$ if, and only if, the vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ are orthogonal.

0.19 Prove that $\|\mathbf{x} + \mathbf{y}\|_2^2 + \|\mathbf{x} - \mathbf{y}\|_2^2 = 2\|\mathbf{x}\|_2^2 + 2\|\mathbf{y}\|_2^2$ for all vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$.

0.20 Prove that any two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ satisfy the triangular inequality $\|\mathbf{x} + \mathbf{y}\|_2 \leq \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2$.

Mini project

Instructions. *Work the following task and report by providing detailed procedures, results, code, and graphical representations as part of your resolution.*

0.21 Let $\Omega = (0,1) \times (0,1)$ and consider the boundary value problem (BVP) $\Delta u = -1$ in Ω, $u = 0$ at $x = 0$, and $u = 1$ at $x = 1$.

(a) Derive and describe an electrostatics problem that is modelled by the given BVP.

(b) Approximate Δu using Taylor's theorem and the same grid spacing $h_x = h_y = h$ in both x and y directions.

(c) Use python to generate random data throughout Ω and apply the Laplace operator twice. Justify why the Laplace operator is sometimes called a smoothing operator.

(d) Assemble the discrete system that corresponds to the given BVP and impose the Dirichlet boundary conditions in two distinct ways. Compare the sparsity pattern of the coefficient matrix to that obtained in the one-dimensional case.

(e) Solve the discrete system that corresponds to the given BVP for $h \in \{0.001, 0.01, 0.05, 0.1\}$, compute the solution times, and make a plot of the most accurate solution throughout Ω.