Selected ESQC mathematics exercises

Student handout

Simen Kvaal

Chapter 1: Sets and functions

Exercises

Exercise 1.1: Days of the week as a set

The days of the week form a simple example of a finite set. Recall that we write the elements of a set inside braces, such as

$$A = \{a, b, c\}.$$

a. Write the days of the week as a set D using braces $\{\ \}$.

b. How many elements does D contain? Write your answer using the notation $|D|$ for the number of elements in a set.

c. Is Monday an element of D? Write your answer using the symbol $\in$.

d. Is January an element of D? Write your answer using the symbol $\notin$.

e. Let W be the set consisting of Saturday and Sunday. Write W using braces. Is W a subset of D? Express your answer using the symbol $\subseteq$.

f. Let

$$M = \{\text{Monday}, \text{Tuesday}, \text{Wednesday}, \text{Thursday}, \text{Friday}\}.$$

What is $M \cup W$? What is $M \cap W$?

g. Suppose another student writes the days in a different order,

$$D' = \{\text{Sunday}, \text{Saturday}, \text{Friday}, \text{Thursday}, \text{Wednesday}, \text{Tuesday}, \text{Monday}\}.$$

Is $D' = D$?

Exercise 1.9: Classically allowed region for Morse potential

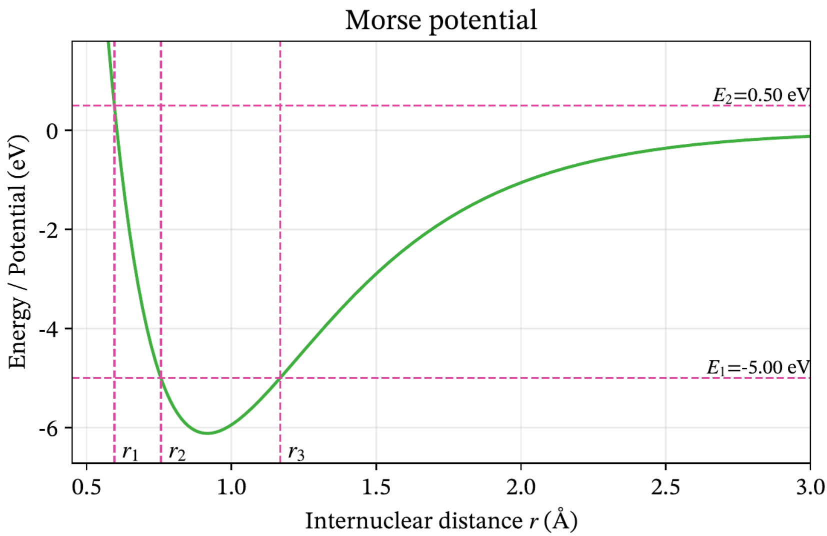

Figure 1: Morse potential

Figure 1 shows the *Morse potential* - a common model potential for vibrational motion of a diatomic molecule:

$$V(r) = D_e \left(1 - e^{-a(r-r_e)}\right)^2 - D_e,$$

where $D_e > 0$ is the dissociation energy, $r_e > 0$ is the equilibrium bond length, and $a > 0$ is a parameter that determines the width of the potential well. The total energy of the molecule is denoted by $E \in \mathbb{R}$. The classically allowed region for the total energy E is the set S_E of all bond lengths r such that $V(r) \leq E$:

$$S_E = \{r \in \mathbb{R} \mid V(r) \leq E\}$$

a. Explain the notation $S_E = \{r \in \mathbb{R} \mid V(r) \leq E\}$ in words. In particular, what does the vertical bar $\mid$ mean?

b. Can you find a condition on E such that S_E is the empty set?

c. Can you find a condition on E such that S_E is the whole real line?

d. Based on the figure, which of the following statements are true?

- $S_{E_1} = (r_2, r_3)$
- $S_{E_1} = [r_2, r_3]$
- $S_{E_1} = \{r_2, r_3\}$

e. Based on the figure, which of the following statements are true?

- $S_{E_2} = (r_1, +\infty)$
- $S_{E_2} = [r_1, +\infty)$
- $S_{E_2} = \{r_1, r_2, r_3\}$

Exercise 1.10: Molecules as sets

In this exercise, a *molecule* is informally defined only by its chemical composition: the number of atoms of each element occurring in its molecular formula. Thus, H_2 and CO are two distinct molecules, while HCOOH and CO_2H_2 represent the same molecule, since both contain one C atom, two H atoms, and two O atoms.

We ignore molecular geometry, bonding, isotopes, charge, and all questions of chemical stability. In particular, any formal combination of atoms is deemed a molecule, whether or not such a species could exist physically.

We restrict our attention to the four elements C, H, O, and N. Let M denote the set of all molecules that can be constructed from these elements. A molecule must contain at least one atom.

One useful way of describing a molecule is by the ordered quadruple

$$(n_{\mathrm{C}}, n_{\mathrm{H}}, n_{\mathrm{O}}, n_{\mathrm{N}}), \tag{0.1}$$

where each n_X is the number of atoms of element X in the molecule. For example,

$$\mathrm{CH_4} \longleftrightarrow (1, 4, 0, 0),$$

and

$$\mathrm{N_2O} \longleftrightarrow (0, 0, 1, 2).$$

a. Explain why Equation 0.1 uniquely determines an element of M. Give 4-5 additional examples of molecules in terms of this representation.

b. Express M in terms of $\mathbb{N}_0 = \{0, 1, 2, \dots\}$. What is the cardinality of M? Is it finite, countably infinite, or uncountably infinite? Justify your answer.

c. Let

$$M_2 = \{m \in M \mid m \text{ contains exactly two atoms}\}.$$

Find $|M_2|$ and list all elements of M_2.

d. Let $S \subset M$ be the set

$$S = \{m \in M \mid m \text{ contains at least one H atom}\}.$$

Describe S in terms of the quadruple representation above. Is S finite, countably infinite, or uncountably infinite? Justify your answer.

d. Describe the complement $M \setminus S$. What is its cardinality?

e. Define

$$C = \{m \in M \mid m \text{ contains at least one C atom}\}.$$

Describe the following sets both in words and in terms of the quadruple representation:

$$C \cap S, \qquad C \cup S, \qquad C \setminus S.$$

Determine the cardinality of each set.

f. For $n \geq 1$, define

$$M_n = \{m \in M \mid m \text{ contains exactly } n \text{ atoms}\}.$$

Thus, a molecule in M_n satisfies

$$n_{\mathrm{C}} + n_{\mathrm{H}} + n_{\mathrm{O}} + n_{\mathrm{N}} = n.$$

Show that

$$|M_n| = \binom{n+3}{3}.$$

Check that your formula agrees with your answer to part **b**.

g. Explain why

$$M = \bigcup_{n=1}^{\infty} M_n.$$

Are the sets M_n pairwise disjoint? Use this decomposition to give another argument that M is countably infinite.

h. Show that M and S have the same cardinality. Explain why this is possible even though S is a proper subset of M.

Chapter 3: Summing over indices

Exercises

Exercise 3.1: Very basic sums

Expand the following sums and calculate their values:

a) $\sum_{i=1}^{4} i$

b) $\sum_{i=1}^{4} i^2$

c) $\sum_{k=0}^{3} 2^k$

Exercise 3.2: Simplifying as a sum

Write each expression using summation notation:

a) $x_1 + x_2 + x_3 + x_4 + x_5$

b) $1 + q + q^2 + q^3 + q^4$

c) $a_1 b_1 + a_2 b_2 + a_3 b_3$

Chapter 6: Vectors and matrices

The vector space $\mathbb{R}^n$ consists of all column vectors

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad x_i \in \mathbb{R},$$

with the usual vector addition and scalar multiplication. Similarly, $\mathbb{C}^n$ consists of column vectors with complex components. More generally, a vector space V over a field $\mathbb{F}$ is a set equipped with vector addition and scalar multiplication such that, for all $x, y, z \in V$ and $\alpha, \beta \in \mathbb{F}$,

1. **Commutativity of addition:** $x + y = y + x$.
2. **Associativity of addition:** $x + (y + z) = (x + y) + z$.
3. **Additive identity:** There exists a zero vector $0 \in V$ such that $x + 0 = x$.
4. **Additive inverse:** For every $x \in V$, there exists $-x \in V$ such that $x + (-x) = 0$.
5. **Distributivity over vector addition:** $\alpha(x + y) = \alpha x + \alpha y$.
6. **Distributivity over scalar addition:** $(\alpha + \beta)x = \alpha x + \beta x$.
7. **Associativity of scalar multiplication:** $(\alpha\beta)x = \alpha(\beta x)$.
8. **Scalar identity:** $1x = x$.

For a matrix $A \in \mathbb{R}^{m \times n}$ and a column vector $x \in \mathbb{R}^n$, the matrix-vector product $y = Ax \in \mathbb{R}^m$ is defined componentwise by

$$y_i = \sum_{j=1}^{n} A_{ij} x_j.$$

For matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times p}$, the matrix product $C = AB \in \mathbb{R}^{m\times p}$ is defined by

$$C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}.$$

The same definitions of matrix-vector and matrix-matrix multiplication apply over $\mathbb{C}$.

Exercises: Basic vector and matrix calculations

Exercise 6.1: Basic vector operations

Consider the vectors

$$\mathbf{x} = (1, 2, -1), \qquad \mathbf{y} = (3, -1, 2), \qquad \mathbf{z} = (0, 2, 1)$$

in the vector space $\mathbb{R}^3$. (We here write the vectors as an ordered tuple instead of a column.)

a. Compute $\mathbf{x} + \mathbf{y}$.

b. Compute $\mathbf{x} - \mathbf{y}$.

c. Compute $2\mathbf{x}$ and $-3\mathbf{y}$.

d. Compute the linear combination $2\mathbf{x} + 3\mathbf{y}$.

e. Compute $\mathbf{x} + 2\mathbf{y} - \mathbf{z}$.

f. Find numbers α and β such that $\alpha\mathbf{x} + \beta\mathbf{y} = (7, 0, 3)$.

g. Verify by direct calculation that $2(\mathbf{x} + \mathbf{y}) = 2\mathbf{x} + 2\mathbf{y}$. Which vector-space axiom does this illustrate?

h. Is $(1, 2)$ an element of this vector space? Is $(0, 0, 0)$ an element? Explain briefly.

Exercise 6.2: Basic matrix-vector products

Consider the matrices and vectors

$$A = \begin{pmatrix} 2 & -1 \\ 3 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 3 & 1 \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \qquad \mathbf{y} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}.$$

a. Compute $A\mathbf{x}$.

b. Compute $B\mathbf{y}$.

c. Compute $A(2\mathbf{x})$ and $2(A\mathbf{x})$. Verify that they are equal.

d. Can the products $A\mathbf{y}$ and $B\mathbf{x}$ be formed? Explain your answer in terms of the dimensions of the matrices and vectors.

Exercise 6.3: Matrix-matrix products

Consider the matrices

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 0 \\ -1 & 4 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 3 & 2 \end{pmatrix}.$$

a. Compute AB.

b. Compute BA. Is $AB = BA$?

c. Compute AC. What are the dimensions of the resulting matrix?

d. Can the product CA be formed? Can the product BC be formed? Explain your answers using the dimensions of the matrices.

Exercise 6.4: Linear independence and basis

We work in the space $V = \mathbb{R}^2$, plane vectors. We define two vectors,

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

a. Show by elementary means that the two vectors are linearly independent, that is: If $x_1\mathbf{b}_1 + x_2\mathbf{b}_2 = 0$, then $x_1 = x_2 = 0$ must hold.

b. The vectors form a basis since they are linearly independent. Explain in your own words what a basis is.
c. Make a 2D drawing or plot of them as arrows from the origin.

Exercise 6.5: Gaussian elimination and vector decomposition

Let

$$\mathbf{v} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Using Gaussian elimination, compute the coefficients x_1 and x_2 such that

$$\mathbf{v} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2.$$

Illustrate this decomposition in the graph from the previous exercise, using the parallelogram rule: The sum of two vectors is obtained by placing the origin of one at the end of the second, or vice versa.

Exercise 6.6: Office coordinates

A person walks from home. First, they walk to the subway station. This is 1 km north and 500 m east as the crow flies. The subway takes the person to the city centre, which 3 km east and 2 km south of the subway station, as the crow flies. The person then walks 200 m south and 400 m west to their office.

What are the coordinates of the office, relative to the person's home?

What is the distance to home, as the crow flies?

Exercise 6.7: Boat sailing on the ocean

A boat is sailing on the ocean, at a constant speed 20 km/h relative to the water surface, in the northeast direction. However, the water surface is moving too, at a constant velocity. After two hours, the boat is located 20 km north and 10 km east of the starting point. What is the speed and direction of ocean drift?

Exercise 6.8: An exercise by Robert Beezer

A three-digit number has two properties. The tens-digit and the ones-digit add up to 5. If the number is written with the digits in the reverse order, and then subtracted from the original number, the result is 792. Use a system of equations to find all of the three-digit numbers with these properties.

Exercises: More vectors and matrices

Exercise 6.9: Compute the action of a matrix on vectors

Let A be the $n \times n$ matrix defined by $A_{i,i+1} = i^2$ for $1 \leq i < n$, and zero otherwise. Compute the action of A and A^T on the standard basis for $\mathbb{R}^n$. Compute the action of A and A^T on an arbitrary vector $\mathbf{x} \in \mathbb{R}^n$.

Exercise 6.10: Matrix-matrix products, rows and columns

A useful way to understand matrix-matrix multiplication is to think of a matrix as a collection of vectors. This gives two complementary interpretations of the product AB:

- A acts separately on each **column** of B.
- B acts separately on each **row** of A, when the rows are regarded as row vectors.

These viewpoints are useful both for understanding matrix multiplication and for computations. For example, if the columns of B represent several vectors, then AB applies the linear transformation A to *all of these vectors at once*.

a. Consider

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

Write B as a collection of its three column vectors,

$$B = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3].$$

Write down $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_3$ explicitly.

b. Compute $A\mathbf{b}_1$, $A\mathbf{b}_2$, and $A\mathbf{b}_3$ separately. Put the three resulting vectors next to each other as the columns of a new matrix:

$$[A\mathbf{b}_1, A\mathbf{b}_2, A\mathbf{b}_3].$$

c. Compute AB directly using ordinary matrix multiplication. Compare your answer with part b.

d. Now consider the general case. Let

$$A \in M(n, m, \mathbb{F}), \qquad B \in M(m, o, \mathbb{F}),$$

and write the columns of B as

$$B = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_o].$$

Explain why

$$AB = [A\mathbf{b}_1, A\mathbf{b}_2, \dots, A\mathbf{b}_o].$$

In other words, explain why the jth column of AB is $A\mathbf{b}_j$.

e. There is a similar interpretation in terms of rows. Write A using its row vectors,

$$A = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix}.$$

Show that

$$AB = \begin{bmatrix} \mathbf{a}_1^T B \\ \mathbf{a}_2^T B \\ \vdots \\ \mathbf{a}_n^T B \end{bmatrix}.$$

Thus, the ith row of AB is the row vector $\mathbf{a}_i^T B$.

f. Suppose the columns of B are several vectors that we want to transform using the same linear transformation A. Explain why the identity from part d means that we can collect all the vectors into one matrix B and perform a single matrix-matrix multiplication AB.

Exercise 6.11

Show that each column of $C = AB$ is a linear combination of the columns of A with coefficients determined by B,

$$\mathbf{c}_i = \sum_j \mathbf{a}_j B_{ji}.$$

Show a similar result for the rows of C.

Chapter 9: Eigenvalues and eigenvectors

Eigenvalue problems are among the most important problems in linear algebra, and they occur throughout quantum chemistry. In an eigenvalue problem, we are given a square matrix $A \in \mathbb{C}^{n\times n}$ and seek a number $\lambda \in \mathbb{C}$ and a nonzero vector $\mathbf{x} \in \mathbb{C}^n$ such that

$$A\mathbf{x} = \lambda\mathbf{x}.$$

The number λ is called an *eigenvalue* of A, and the nonzero vector $\mathbf{x}$ is called an *eigenvector* corresponding to λ. In column-vector notation,

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

The equation $A\mathbf{x} = \lambda\mathbf{x}$ says that multiplication by A has a particularly simple effect on an eigenvector: it only multiplies the vector by a number. Thus, the direction represented by $\mathbf{x}$ is left unchanged, apart from multiplication by the scalar λ. Notice that $\mathbf{x} = \mathbf{0}$ is excluded from the definition, since $A\mathbf{0} = \lambda\mathbf{0}$ holds for every λ and therefore tells us nothing about the matrix.

The eigenvalue equation can equivalently be written as

$$(A - \lambda I)\mathbf{x} = \mathbf{0},$$

where I is the identity matrix. Since we require $\mathbf{x} \neq \mathbf{0}$, the matrix $A - \lambda I$ must have a nontrivial null space and therefore cannot be invertible. For a finite-dimensional square matrix, this is equivalent to

$$\det(A - \lambda I) = 0.$$

This equation is called the **characteristic equation** or **secular equation** of A. Its solutions are the eigenvalues of A. Once an eigenvalue λ has been found, the corresponding eigenvectors are obtained by solving the homogeneous linear system $(A - \lambda I)\mathbf{x} = \mathbf{0}$.

An especially important situation occurs when A has enough linearly independent eigenvectors to form a basis of $\mathbb{C}^n$. Suppose these eigenvectors are

$$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n,$$

where the subscript labels the different vectors. Thus,

$$\mathbf{x}_j = \begin{bmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{bmatrix}$$

denotes the j'th eigenvector, while x_{ij} denotes its ith component. If $\mathbf{x}_j$ has eigenvalue λ_j, then

$$A\mathbf{x}_j = \lambda_j \mathbf{x}_j, \qquad j = 1, \ldots, n.$$

We can collect these eigenvectors as the columns of a matrix

$$X = [\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_n]$$

and collect their corresponding eigenvalues in the diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

The n separate eigenvalue equations can then be written as the single matrix equation

$$AX = XD.$$

Because the eigenvectors form a basis, X is invertible. Consequently,

$$A = XDX^{-1}.$$

A matrix for which such an eigenvector basis exists is called *diagonalizable*. Diagonalization is useful because the action of a diagonal matrix is particularly simple: it acts independently on each component. In a basis of eigenvectors, the matrix A therefore has its simplest possible form. This idea is fundamental in quantum mechanics, where many important matrices, in particular Hermitian matrices, possess an orthonormal basis of eigenvectors.

In the exercises below, we study how eigenvalues and eigenvectors are found, when a matrix can be diagonalized, and why Hermitian and, more generally, normal matrices have especially useful eigenvalue decompositions.

Exercises

Exercise 9.1: Checking candidate eigenvectors

Consider the real symmetric matrix

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

Recall that a nonzero vector $\mathbf{x}$ is an eigenvector of A if

$$A\mathbf{x} = \lambda\mathbf{x}$$

for some number λ. In other words, after multiplication by A, the vector must point in the same or exactly opposite direction: it may be stretched, compressed, or multiplied by a negative number, but its direction cannot otherwise change.

For each of the following candidate vectors, compute $A\mathbf{x}$ and decide whether $\mathbf{x}$ is an eigenvector of A. If it is, find the corresponding eigenvalue λ.

a. Consider

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Compute $A\mathbf{x}$. Can the result be written as $\lambda\mathbf{x}$? If so, what is λ?

b. Consider

$$\mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Compute $A\mathbf{x}$ and determine whether $\mathbf{x}$ is an eigenvector. If it is, find its eigenvalue.

c. Consider

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Compute $A\mathbf{x}$. Is the result a scalar multiple of $\mathbf{x}$? Is $\mathbf{x}$ an eigenvector?

d. Consider

$$\mathbf{x} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

Without doing the full matrix-vector multiplication at first, compare this vector with the vector in part a. Do you expect it to be an eigenvector? Check your answer by computing $A\mathbf{x}$.

e. The vectors in parts a and d are different vectors, but they point in the same direction. What eigenvalue do they have? What does this tell you about scalar multiples of an eigenvector?

f. Compare the two eigenvectors found in parts a and b. Compute their ordinary dot product. What do you observe?

Exercise 9.2: Eigenvalues in chemical kinetics

Consider the reversible reaction

$$\mathrm{A} \underset{k}{\overset{k}{\rightleftharpoons}} \mathrm{B}.$$

The positive number k is called the *rate constant*. It tells us how rapidly the reaction takes place. Here we use the same rate constant for the forward reaction $\mathrm{A} \to \mathrm{B}$ and the backward reaction $\mathrm{B} \to \mathrm{A}$, to keep the example simple.

Let $a(t)$ and $b(t)$ denote the concentrations of A and B. For a first-order reaction, the reaction rate is proportional to the concentration of the reacting species. Thus, A is converted into B at a rate ka, while B is converted into A at a rate kb. The larger k is, the faster the reaction proceeds. If concentration is measured in $\mathrm{mol\,L^{-1}}$ and time in seconds, k has units of $\mathrm{s^{-1}}$.

The concentration of A therefore decreases due to the forward reaction at a rate ka and increases due to the backward reaction at a rate kb. Similarly, the concentration of B increases at a rate ka and decreases at a rate kb. This gives the rate equations

$$\frac{da}{dt} = -ka + kb, \qquad \frac{db}{dt} = ka - kb.$$

If we collect the concentrations in the vector

$$\mathbf{x}(t) = \begin{bmatrix} a(t) \\ b(t) \end{bmatrix},$$

the rate equations can be written as

$$\frac{d\mathbf{x}}{dt} = K\mathbf{x}, \qquad K = \begin{bmatrix} -k & k \\ k & -k \end{bmatrix}.$$

We will see that the eigenvectors of K describe two particularly simple ways in which the concentrations can change.

a. Consider the vector

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Calculate $K\mathbf{x}_1$ and show that $\mathbf{x}_1$ is an eigenvector of K. What is its eigenvalue? Explain chemically why nothing changes when $a = b$.

b. Now consider

$$\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Calculate $K\mathbf{x}_2$ and show that it is also an eigenvector. What is its eigenvalue?

c. Suppose that the concentration vector has the form

$$\mathbf{x}(t) = c(t)\mathbf{x}_2.$$

Insert this into the rate equation and use your result from part b to show that

$$\frac{dc}{dt} = -2kc.$$

Thus $c(t) = c(0)e^{-2kt}$. What happens to this contribution as time increases? What happens more rapidly if the rate constant k is increased? Explain why the two eigenvectors $\mathbf{x}_1$ and $\mathbf{x}_2$ can be interpreted as an *equilibrium mode* and a *relaxation mode*, respectively.

Exercise 9.3: Eigenvalues and diagonalization of a Hermitian matrix

Consider the matrix

$$A = \begin{bmatrix} 2 & -\mathrm{i} \\ \mathrm{i} & 2 \end{bmatrix}.$$

a. Show that A is Hermitian, that is, verify that

$$A^H = A.$$

b. The eigenvalues of A are the values of λ for which

$$\det(A - \lambda I) = 0.$$

Set up the matrix $A - \lambda I$, calculate its determinant, and find the two eigenvalues λ_1 and λ_2. Verify that both eigenvalues are real. (The determinant of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $ad - bc$.)

c. For each eigenvalue λ_j, solve

$$(A - \lambda_j I)\mathbf{x}_j = \mathbf{0}$$

to find a corresponding eigenvector $\mathbf{x}_j$. Normalize the eigenvectors so that

$$\mathbf{x}_j^H \mathbf{x}_j = 1.$$

d. Verify explicitly that your two normalized eigenvectors are orthogonal,

$$\mathbf{x}_1^H \mathbf{x}_2 = 0.$$

Thus, the eigenvectors form an orthonormal basis of $\mathbb{C}^2$.

e. Form the matrix U whose columns are the normalized eigenvectors,

$$U = [\mathbf{x}_1 \ \mathbf{x}_2].$$

Show explicitly that U is unitary by calculating $U^H U$.

f. Form the diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$

where the ordering of the eigenvalues must agree with the ordering of the eigenvectors in U.

g. Verify by direct matrix multiplication that

$$A = UDU^H.$$

Equivalently, verify that $U^H AU = D$. Explain briefly why $U^{-1} = U^H$ in this diagonalization.

Exercise 9.4: A matrix that cannot be diagonalized

Consider the matrix

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

a. Find the eigenvalues of A by solving the characteristic equation $\det(A - \lambda I) = 0$. How many distinct eigenvalues does A have?

b. Find all eigenvectors corresponding to the eigenvalue you found. How many linearly independent eigenvectors does A have?

c. To diagonalize a 2×2 matrix, we need two linearly independent eigenvectors, which can be used as the columns of an invertible matrix U. Explain why this is impossible for A, and conclude that A is not diagonalizable.

Chapter 10: General vector spaces

A vector space needs vector addition and scalar multiplication, but it need not resemble Euclidean coordinate space.

You may be very familiar with Euclidean space $\mathbb{R}^n$ as a vector space, and quite used to the notion that we have an inner product and a corresponding distance measure on this space. However, linear spaces are much more general. An inner product or norm is not part of the definition of a general vector space. Only the a "linear structure" matters: Consider for example a general vector space over $\mathbb{R}$: An abstract set V with operations such that one can *add vectors*,

$$z = x + y \in V,$$

and *multiply vectors by constants* $c \in \mathbb{R}$,

$$z = cx \in V$$

where $c \in \mathbb{R}$. The full set of axioms is as follows. We focus on vector spaces over the field $\mathbb{R}$, but the definition can be generalized to vector spaces over any field such as $\mathbb{C}$.

> **Definition: Vector space.** A *vector space over the field* $\mathbb{R}$ is a set V together with a binary *vector addition* $+ : V \times V \to V$ and *scalar multiplication* $\cdot : \mathbb{R} \times V \to V$ such that, for all $x, y, z \in V$ and all $\alpha, \beta \in \mathbb{R}$, the following axioms are true:
>
> 1. There exists a $0 \in V$ such that $0 + x = x$ for all $x \in V$
> 2. $x + (y + z) = (x + y) + z$
> 3. $x + y = y + x$
> 4. For every $x \in V$, there exists $x' \in V$ such that $x + x' = 0$
> 5. $(\alpha\beta) \cdot x = \alpha \cdot (\beta \cdot x)$
> 6. $1 \cdot x = x$
> 7. $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$

8. $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$

In this exercise, we will train our ability to think abstractly, and also try to get acquainted with general vector spaces.

Example: Somerset Apple Cake. Ingredients:

- 170 g butter
- 170 g light brown sugar
- 3 eggs
- 1 tablespoon liquid honey
- 1.5 teaspoons cinnamon
- 0.5 teaspoon cloves
- 0.5 teaspoon nutmeg
- 1 teaspoon baking powder
- 240 g wheat flour
- 700 g apples, diced
- 100 ml apple cider, apple juice, or milk
- 100 g light sultana raisins (optional)

Instructions:

- Cream the room-temperature butter with the sugar until light and fluffy.
- Beat in the eggs and honey.
- Sift the flour, baking powder, and spices. Add them to the mixture and stir until the batter is smooth and lump-free.
- Stir in the milk or apple cider (and the raisins, if you choose to use them; they can be soaked in cider for a couple of hours beforehand, if desired).
- Finally, fold the apple pieces into the batter.
- Pour the batter into a greased round tin with parchment paper at the bottom, or use a deep, round ovenproof dish (24 cm in diameter).

- Bake the cake in the middle of the oven at 160°C for 1.5 hours (check with a cake tester to ensure it's fully baked).
- Serve warm or cold with a dollop of whipped cream, clotted cream, vanilla ice cream, or whatever you like.

Exercises

Exercise 10.1: Cooking with vectors 1

We can think of an ingredient list as a vector. For example, suppose that our universe of ingredients consists only of

$$(\text{butter}, \text{sugar}, \text{eggs}, \text{flour}, \text{apples}).$$

After choosing suitable units for each ingredient, an ingredient list can then be represented by a vector

$$x = (x_1, x_2, x_3, x_4, x_5),$$

where each component gives the amount of one ingredient.

a. Using grams for butter, sugar, flour, and apples, and number of eggs for eggs, write the corresponding five-component vector for the Somerset Apple Cake recipe.
b. What would the vector $(100, 0, 2, 300, 0)$ mean? What does a zero component mean?
c. What is the dimension of this vector space? Give a natural set of basis vectors and explain what each basis vector represents.
d. Suppose that x and y are the ingredient vectors for two recipes. What does the vector $x + y$ represent?
e. What does the vector $2x$ represent? What about $\frac{1}{2}x$? Does the scalar 2 have a physical unit?

f. Our choice of units is part of the way we represent the recipe. Suppose milk were another component. Explain why we should not represent 1 litre of milk as 1 in one recipe and 1000 ml of milk as 1000 in another without taking account of the different units.

g. Typical recipes contain ingredients that are not among our five chosen ingredients. How could we enlarge our vector space to accommodate more recipes? What happens to the dimension of the space?

Exercise 10.2: Cooking with vectors 2

In Exercise 10.1, we represented recipes using a fixed list of possible ingredients. If you did not do this exercise already, please do so first. In this exercise, we make this construction more general.

Let S be the set of all possible ingredients. An ingredient list can be represented by a function

$$f : S \to \mathbb{R},$$

where $f(s)$ gives the amount of ingredient s. We assume that $f(s) = 0$ for all but finitely many ingredients. The set L of all such functions is a vector space, with addition and scalar multiplication defined in the usual way.

a. Explain how this description generalizes the finite-dimensional ingredient vectors used in Exercise 10.1. If S contains infinitely many ingredients, what can you say about the dimension of L?

b. Mathematically, L contains vectors with negative components. Can such vectors be interpreted as ordinary ingredient lists? What about the zero vector?

c. Define the subset

$$L_+ = \{f \in L \mid f(s) \geq 0 \text{ for every } s \in S\}.$$

Is L_+ a vector space? Which vector-space properties fail? Is L_+ closed under addition? Is it closed under multiplication by *nonnegative* scalars?

d. Even if $f, g \in L_+$ are ingredient lists for two perfectly good recipes, $f + g$ need not describe a sensible recipe. What information is missing from an ingredient vector?

e. We might try to describe the steps of a recipe by operators acting on L. Consider, for example, the instruction "cream the butter and sugar". Can this operation be represented naturally as a linear map $T : L \to L$ on our present ingredient space? Explain the challenges one encounters.

f. What additional information might have to be included in our mathematical description before operations such as mixing, creaming, heating, and baking could be represented?

g. In quantum mechanics, we similarly choose a mathematical space to represent physical states and operators to represent physical quantities and transformations. Why is the choice of the space important for determining what the mathematical model can describe?