Answers of Exercises
in
Linear Algebra with Python
–Theory and Applications–

May 31, 2023

Preface

This document contains answers to all the exercises in ”Linear Algebra with Python –Theory and Applications–”. It is natural for beginners to have a hard time figuring out what they need to show to get the correct answer that the authors are looking for, and even if they do figure out what they need to show, where to start. For such beginners, we have given solutions for each exercise with as few gaps as possible. However, the solutions listed here are not necessarily the only path to the correct answer for the exercise. The reader with a deeper understanding should try to create original solutions.

Contents

Chapter 1

Mathematics and Python

Exercise 1.1	Sect. 1.1. p.2	Computing truth tables
Exercise 1.2	Sect. 1.1. p.2	The assosiative and distributibe lows of logics
Exercise 1.3	Sect. 1.2. p.3	Properties of the four arithmetic operations
Exercise 1.4	Sect. 1.2. p.3	Properties of complex numbers
Exercise 1.5	Sect. 1.2. p.4	de Moivre's theorem
Exercise 1.6	Sect. 1.2. p.4	Graphs of Maclaurin's expansions*
Exercise 1.7	Sect. 1.3. p.7	Power sets
Exercise 1.8	Sect. 1.3. p.8	The distributibe lows of sets
Exercise 1.9	Sect. 1.5. p.12	Enumeration of mappings

* Using Python

Exercise 1.1. Show the equivalences

$$P \to Q \;\Leftrightarrow\; \neg\, P \vee Q \quad \text{and} \quad P \leftrightarrow Q \;\Leftrightarrow\; (P \to Q) \wedge (Q \to P)$$

by making the truth value tables for them. Moreover, confirm the equivalences by Python.

P	Q	$\neg\, P$	$P \to Q$	$\neg P \vee Q$
true	true	false	true	true
true	false	false	false	false
false	true	true	true	true
false	false	true	true	true

Since the truth values (of shadowed cells) in the columns of $P \to Q$ and $\neg P \vee Q$ coincide through every rows, it follows that $P \to Q \;\Leftrightarrow\; \neg\, P \vee Q$ holds. If you also add the column of truth values of $(P \to Q) \leftrightarrow (\neg P \vee Q)$ to the above table, you find that all truth values of the logical formula $(P \to Q) \leftrightarrow (\neg P \vee Q)$ are true. That is, the formula is a *tautology*. In general, a formula $f\,(P_1, P_2, \ldots, P_n)$ created from the propositions $P_1, P_2, \ldots, P_n$ and logilal operators is called a tautology if and only if the truth value of $f\,(P_1, P_2, \ldots, P_n)$ is always true, regardless of the truth values assigned to each $P_1, P_2, \ldots, P_n$.

P	Q	$P \to Q$	$Q \to P$	$P \leftrightarrow Q$	$(P \to Q) \wedge (Q \to P)$
true	true	true	true	true	true
true	false	false	true	false	false
false	true	true	false	false	false
false	false	true	true	true	true

The truth values in the columns of $P \leftrightarrow Q$ and $(P \to Q) \wedge (Q \to P)$ coincide through every rows. Then, the logical formula $(P \leftrightarrow Q) \leftrightarrow \big((P \to Q) \wedge (Q \to P)\big)$ becomes a tautology.

Exercise 1.2. Prove the associative laws

$$(P \wedge Q) \wedge R \;\Leftrightarrow\; P \wedge (Q \wedge R), \quad (P \vee Q) \vee R \;\Leftrightarrow P \vee (Q \vee R)$$

and the distributive laws

$$P \wedge (Q \vee R) \;\Leftrightarrow\; (P \wedge Q) \vee (P \wedge R), \quad P \vee (Q \wedge R) \;\Leftrightarrow (P \vee Q) \wedge (P \vee R).$$

In the following truth value tables, these can be verified by looking at the shaded cells.

the associative laws:

P	Q	R	$P \wedge Q$	$Q \wedge R$	$(P \wedge Q) \wedge R$	$P \wedge (Q \wedge R)$
true	true	true	true	true	true	true
true	true	false	true	false	false	false
true	false	true	false	false	false	false
true	false	false	false	false	false	false
false	true	true	false	true	false	false
false	true	false	false	false	false	false
false	false	true	false	false	false	false
false	false	false	false	false	false	false

P	Q	R	$P \vee Q$	$Q \vee R$	$(P \vee Q) \vee R$	$P \vee (Q \vee R)$
true	true	true	true	true	true	true
true	true	false	true	true	true	true
true	false	true	true	true	true	true
true	false	false	true	false	true	true
false	true	true	true	true	true	true
false	true	false	true	true	true	true
false	false	true	false	true	true	true
false	false	false	false	false	false	false

the distributive laws:

P	Q	R	$Q \vee R$	$P \wedge Q$	$P \wedge R$	$P \wedge (Q \vee R)$	$(P \wedge Q) \vee (P \wedge R)$
true	true	true	true	true	true	true	true
true	true	false	true	true	false	true	true
true	false	true	true	false	true	true	true
true	false	false	false	false	false	false	false
false	true	true	true	false	false	false	false
false	true	false	true	false	false	false	false
false	false	true	true	false	false	false	false
false	false	false	false	false	false	false	false

P	Q	R	$Q \wedge R$	$P \vee Q$	$P \vee R$	$P \vee (Q \wedge R)$	$(P \vee Q) \wedge (P \vee R)$
true	true	true	true	true	true	true	true
true	true	false	false	true	true	true	true
true	false	true	false	true	true	true	true
true	false	false	false	true	true	true	true
false	true	true	true	true	true	true	true
false	true	false	false	true	false	false	false
false	false	true	false	false	true	false	false
false	false	false	false	false	false	false	false

Supplementary problem. Prove De Morgan's law for propositional logic:

$$\neg\,(P \wedge Q) \;\Leftrightarrow \neg P \vee \neg Q, \qquad \neg\,(P \vee Q) \;\Leftrightarrow \neg P \wedge \neg Q.$$

Exercise 1.3. Deduce the equalities $0x = x0 = 0$ from properties R1–R9.

This follows from the following series of equational transformations using R1–R9:

$$
\begin{aligned}
0x &= x0 && \text{(by R5)} \\
&= x0 + 0 && \text{(by R3)} \\
&= x0 + \big(x0 + (-x0)\big) && \text{(by R4)} \\
&= (x0 + x0) + (-x0) && \text{(by R2)} \\
&= x\,(0 + 0) + (-x0) && \text{(by R9)} \\
&= x0 + (-x0) && \text{(by R3)} \\
&= 0 && \text{(by R4)}
\end{aligned}
$$

Supplementary problem. Deduce the equalities $(-1) \times (-1) = 1$ from properties R1–R9.

Exercise 1.4. Prove properties Z1 – Z6 above from properties R1 – R9 of real numbers.

Let $z = x + iy, z_1 = x_1 + iy_1, z_2 = x_2 + iy_2$ (x, y, x_1, y_1, x_2, y_2 are real numbers, i is the imaginary unit $i^2 = -1$).

1.
$$z\overline{z} = (x + iy)(x - iy) = x^2 - xiy + iyx - iyiy = x^2 + y^2 = |z|^2$$

2.
$$\begin{aligned} & |z_1z_2|^2 = \left|(x_1 + iy_1)(x_2 + iy_2)\right|^2 = \left|(x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)\right|^2 \\ = & (x_1x_2 - y_1y_2)^2 + (x_1y_2 + x_2y_1)^2 = x_1^2x_2^2 + y_1^2y_2^2 + x_1^2y_2^2 + x_2^2y_1^2 \\ = & \left(x_1^2 + y_1^2\right)\left(x_2^2 + y_2^2\right) = |z_1|^2 |z_2|^2 \end{aligned}$$

If 5 below is shown already, 2 can be shown as follows:

$$|z_1z_2|^2 = z_1z_2\overline{z_1z_2} = z_1z_2\overline{z_1}\,\overline{z_2} = z_1\overline{z_1}z_2\overline{z_2} = |z_1|^2 |z_2|^2 .$$

3. If z is a real number, then $y = 0$ and $z = x = x - iy = \overline{z}$. Conversely, if $z = \overline{z}$, then $x + iy = x - iy$. Hence, we get $2iy = 0$ and thus $y = 0$. Therefore, $z = x$ is real.

4.
$$\text{LHS} = \overline{(x_1 + x_2) + i(y_1 + y_2)} = (x_1 + x_2) - i(y_1 + y_2) = \text{RHS}$$

5.
$$\begin{aligned} & \text{LHS} = \overline{(x_1 + iy_1)(x_2 + iy_2)} = \overline{(x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)} \\ = & (x_1x_2 - y_1y_2) - i(x_1y_2 + x_2y_1) = (x_1 - iy_1)(x_2 - iy_2) = \text{RHS} \end{aligned}$$

6. Let $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$ (x_1, y_1, x_2, y_2 are real numbers). Then,

$$\begin{aligned} & |z_i + z_2| \leqq |z_1| + |z_2| \\ \Leftrightarrow\ & |z_i + z_2|^2 \leqq \left(|z_1| + |z_2|\right)^2 \\ \Leftrightarrow\ & |z_i + z_2|^2 \leqq |z_1|^2 + 2|z_1z_2| + |z_2|^2 \\ \Leftrightarrow\ & (x_1 + x_2)^2 + (y_1 + y_2)^2 \leqq x_1^2 + y_1^2 + 2\left|(x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)\right| + x_2^2 + y_2^2 \\ \Leftrightarrow\ & 2x_1x_2 + 2y_1y_2 \leqq 2\sqrt{(x_1x_2 - y_1y_2)^2 + (x_1y_2 + x_2y_1)^2} \\ \Leftrightarrow\ & (x_1x_2 + y_1y_2)^2 \leqq (x_1x_2 - y_1y_2)^2 + (x_1y_2 + x_2y_1)^2 \\ \Leftrightarrow\ & 2x_1x_2y_1y_2 \leqq x_1^2y_2^2 + x_2^2y_1^2 \\ \Leftrightarrow\ & 0 \leqq (x_1y_2 - x_2y_1)^2 \end{aligned}$$

Alternatively, we can prove it as follows:

$$\begin{aligned}
|z_1 + z_2|^2 &= (z_1 + z_2)(\overline{z_1} + \overline{z_2}) = z_1\overline{z_1} + z_1\overline{z_2} + z_2\overline{z_1} + z_2\overline{z_2} = z_1\overline{z_1} + z_1\overline{z_2} + \overline{z_1\overline{z_2}} + z_2\overline{z_2} \\
&= |z_1|^2 + 2\mathrm{Re}\,(z_1\overline{z_2}) + |z_1|^2 \leqq |z_1|^2 + 2\,|z_1\overline{z_2}| + |z_1|^2 = |z_1|^2 + 2\,|z_1|\,|z_2| + |z_1|^2 \\
&= (|z_1| + |z_2|)^2
\end{aligned}$$

Here we use the facts $z + \overline{z} = 2\mathrm{Re}\,z$ and that $\mathrm{Re}\,z \leqq |z|$ for complex number z.

Exercise 1.5. Prove the following formulas.

(1) $e^{i(\theta_1+\theta_2)} = e^{i\theta_1}e^{i\theta_2}$ for any $\theta_1, \theta_2 \in \mathbb{R}$ (Hint: use the trigonometric addition theorem).

(2) $\left(e^{i\theta}\right)^n = e^{in\theta}$ (**de Moivre's theorem**) holds for any natural number n (Hint: use mathematical induction).

(1) By Euler's formula and the addition theorem of cos, we have

$$\begin{aligned} \text{LHS} &= \cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2) \\ &= \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 + i(\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2) \\ &= (\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2) \;=\; \text{RHS} \end{aligned}$$

(2) If $n = 1$, the theorem holds. Suppose that the theorem is true for $n = k$. Then we have $\left(e^{i\theta}\right)^k = e^{ik\theta}$ and also

$$\left(e^{i\theta}\right)^{k+1} = \left(e^{i\theta}\right)^k e^{i\theta} = e^{ik\theta}e^{i\theta} = e^{ik\theta+i\theta} = e^{i(k+1)\theta}$$

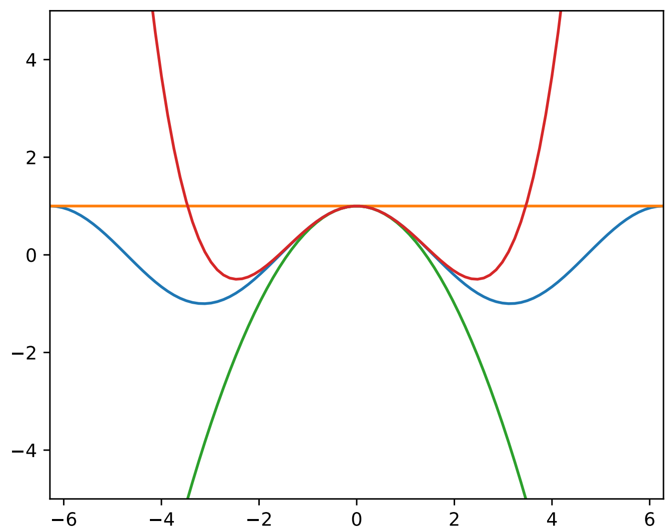
is saticefied. The first equality is followed by the power low, the socond is the assumption of the mathematiical induction, and the third equality follows 1. Therefore the theorem is true for $n = k + 1$.

Exercise 1.6. Using Python and its library of Matplotlib, draw the graphs in Figure 1.1.

Program: `ps_sin.py` (the Maclaurin expansion of $\sin x$)

In [1]:

```
from numpy import linspace, pi, sin
import matplotlib.pyplot as plt

x0, x1, y0, y1 = -2*pi, 2*pi, -5, 5
X = linspace(x0, x1, 100)

plt.axis('scaled'), plt.xlim(x0, x1), plt.ylim(y0, y1)
plt.plot(X, sin(X))
plt.plot(X, X)
plt.plot(X, X - X**3/6)
plt.plot(X, X - X**3/6 + X**5/120)
plt.show()
```

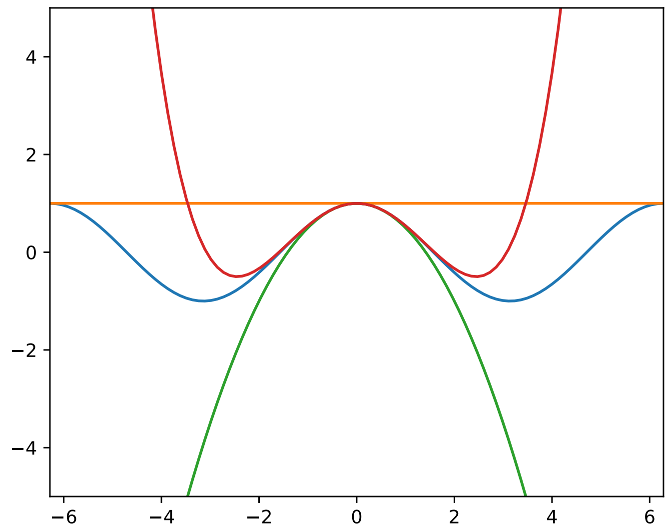


Program: `ps_cos.py` (the Maclaurin expansion of $\cos x$)

In [1]:

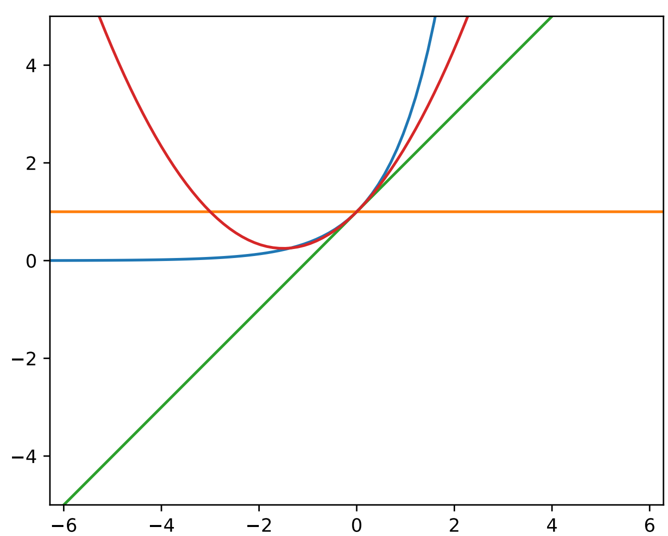
```
from numpy import linspace, pi, cos
import matplotlib.pyplot as plt

x0, x1, y0, y1 = -2*pi, 2*pi, -5, 5
X = linspace(x0, x1, 100)

plt.axis('scaled'), plt.xlim(x0, x1), plt.ylim(y0, y1)
plt.plot(X, cos(X))
plt.plot(X, X**0)
plt.plot(X, X**0 - X**2/2)
plt.plot(X, X**0 - X**2/2 + X**4/24)
```

In [1]:
```
plt.show()
```



Program: `ps_exp.py` (the Maclaurin expansion of e^x)

In [1]:
```
from numpy import linspace, pi, exp
import matplotlib.pyplot as plt

x0, x1, y0, y1 = -2*pi, 2*pi, -5, 5
X = linspace(x0, x1, 100)

plt.axis('scaled'), plt.xlim(x0, x1), plt.ylim(y0, y1)
plt.plot(X, exp(X))
plt.plot(X, X**0)
plt.plot(X, X**0 + X)
plt.plot(X, X**0 + X + X**2/3)
plt.show()
```



Exercise 1.7. Write down the extensional description of the power set $2^{\{1,2,3\}}$. Also, prove that the number of elements of the power set of a set with n elements is 2^n.

$$2^{\{1,2,3\}} \;=\; \left\{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\right\}.$$

Let A be a set consisting of n elements. One subset B of A is determined by whether each element of A is included or not. Hence, the total number of these cases is 2^n. This number is also equal to the number of integers that an n-bit binary number can represent. The following table maps a subset of $2^{\{1,2,3\}}$ to one of 2^3 integers.

	1	2	3	
$\emptyset$	0	0	0	0
$\{3\}$	0	0	1	1
$\{2\}$	0	1	0	2
$\{2,3\}$	0	1	1	3
$\{1\}$	1	0	0	4
$\{1,3\}$	1	0	1	5
$\{1,2\}$	1	1	0	6
$\{1,2,3\}$	1	1	1	7

Another idea is that the number of combinations that extracts k piaces from n pieces is as ${}_n\mathrm{C}_k$, so it is calculated as follows using the binomial theorem.

$${}_n\mathrm{C}_0 + {}_n\mathrm{C}_1 + {}_n\mathrm{C}_2 + \cdots + {}_n\mathrm{C}_n \;=\; (1+1)^2 \;=\; 2^n.$$

Supplementary problem:

1. Write down the extensional description of $\left(2^{\{1,2\}}\right)^2$.

2. Write down the extensional description of $2^{\{1,2\}^2}$.

Exercise 1.8. Prove the equalities $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ and $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

The former follows from the following series of logical equivalence transformations:

$$
\begin{aligned}
& x \in A \cap (B \cup C) & \\
\Leftrightarrow\ & (x \in A) \wedge (x \in B \cup C) & (\because x \in X \cap Y \Leftrightarrow (x \in X) \wedge (x \in Y)) \\
\Leftrightarrow\ & (x \in A) \wedge \big((x \in B) \vee (x \in C)\big) & (\because x \in X \cup Y \Leftrightarrow (x \in X) \vee (x \in Y)) \\
\Leftrightarrow\ & \big((x \in A) \wedge (x \in B)\big) \vee \big((x \in A) \vee (x \in C)\big) & (\because P \wedge (Q \vee R) \Leftrightarrow (P \wedge Q) \vee (P \wedge R)) \\
\Leftrightarrow\ & (x \in A \cap B) \vee (x \in A \cap C) & (\because (x \in X) \wedge (x \in Y) \Leftrightarrow x \in X \cap Y) \\
\Leftrightarrow\ & x \in (A \cap B) \cup (A \cap C) & (\because (x \in X) \vee (x \in Y) \Leftrightarrow x \in X \cup Y).
\end{aligned}
$$

We can show the latter similarly.

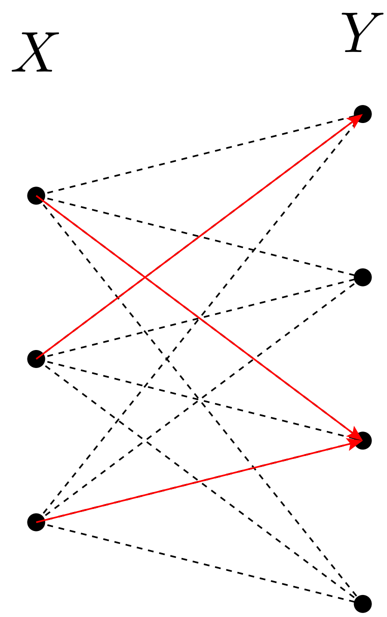
$$
\begin{aligned}
& x \in A \cup (B \cap C) & \\
\Leftrightarrow\ & (x \in A) \vee (x \in B \cap C) & (\because x \in X \cup Y \Leftrightarrow (x \in X) \vee (x \in Y)) \\
\Leftrightarrow\ & (x \in A) \vee \big((x \in B) \wedge (x \in C)\big) & (\because x \in X \cap Y \Leftrightarrow (x \in X) \wedge (x \in Y)) \\
\Leftrightarrow\ & \big((x \in A) \vee (x \in B)\big) \wedge \big((x \in A) \vee (x \in C)\big) & (\because P \vee (Q \wedge R) \Leftrightarrow (P \vee Q) \wedge (P \vee R)) \\
\Leftrightarrow\ & (x \in A \cup B) \wedge (x \in A \cup C) & (\because (x \in X) \vee (x \in Y) \Leftrightarrow x \in X \cup Y) \\
\Leftrightarrow\ & x \in (A \cup B) \cap (A \cup C) & (\because (x \in X) \wedge (x \in Y) \Leftrightarrow x \in X \cap Y).
\end{aligned}
$$

Supplementary problem. Prove De Morgan's law for set theory:

$$(A \cap B)^{\complement} = A^{\complement} \cup B^{\complement}, \qquad (A \cup B)^{\complement} = A^{\complement} \cap B^{\complement}.$$

Exercise 1.9. Prove the above fact on the size of the power set Y^X.

Let consider a mapping $f \in Y^X$. Since $f(x)$ can be assigned to any element of Y, there are n different ways of assigning for each $x \in X$ where n is the number of elements of Y. Since x moves over a total of m elements of X, we can consider n^m possible different mappings f.



The set of red arrows in the above graph represents an example of mappings from X to Y. Each point of X must have only one arrow emitted from it. On the other hand, there may be a point in Y that none of the arrows reach, or two or more may reach a single point in Y. Therefore, one arrow emitted from the point of X can reach one of any points of Y independently of the other arrows.

Supplementary problem:

1. Suppose $m > n$, count the number of surjections from X onto Y.
2. Suppose $m = n$, count the number of bictions from X onto Y.
3. Suppose $m < n$, count the number of injections from X into Y.

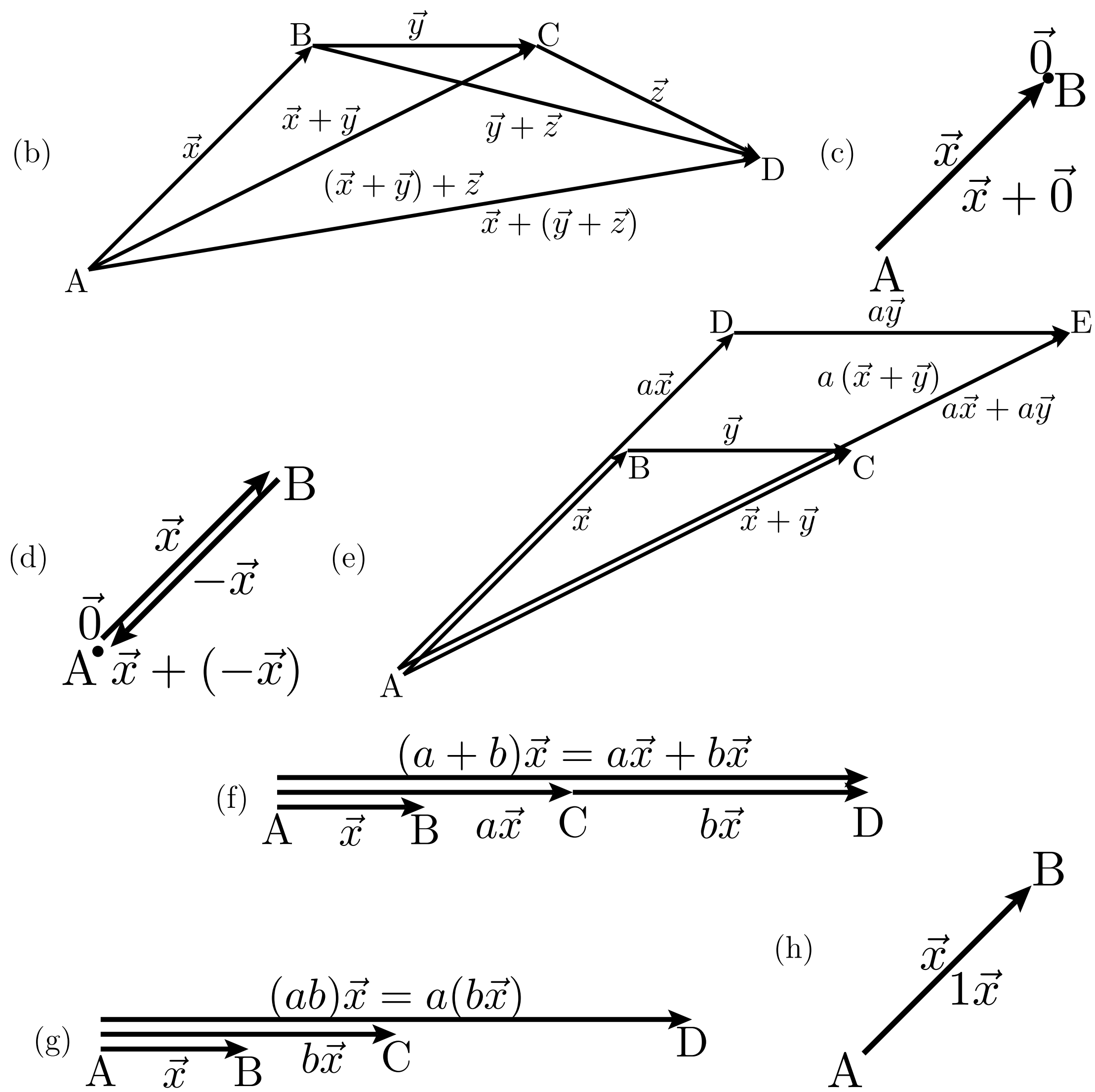
Chapter 2

Linear Spaces and Linear Mappings

Exercise 2.1	Sect. 2.1.	p.31	Geometrical sense of axioms of linear space
Exercise 2.2	Sect. 2.1.	p.32	Axioms of linear space in 3D*
Exercise 2.3	Sect. 2.1.	p.33	Properties derived from the axioms of linear space
Exercise 2.4	Sect. 2.1.	p.34	Axioms of linear space in $\mathbb{K}^n$
Exercise 2.5	Sect. 2.1.	p.35	Axioms of linear space in $\mathbb{K}^X$
Exercise 2.6	Sect. 2.1.	p.37	Graphs of complex-valued functions*
Exercise 2.7	Sect. 2.2.	p.37	Subspaces and linear combination
Exercise 2.8	Sect. 2.2.	p.38	Two subspaces of $\mathbb{K}^2$
Exercise 2.9	Sect. 2.2.	p.38	All subspaces of the 2D plane and 3D space
Exercise 2.10	Sect. 2.2.	p.39	Intersection of subspaces
Exercise 2.11	Sect. 2.2.	p.39	Set operations of subspaces
Exercise 2.12	Sect. 2.3.	p.39	Linear mappings and linear combination
Exercise 2.13	Sect. 2.2.	p.40	Derivative of polynomials
Exercise 2.14	Sect. 2.3.	p.40	Linear spaces of linear mappings
Exercise 2.15	Sect. 2.2.	p.42	Some linear mappings on $\mathbb{R}^2$
Exercise 2.16	Sect. 2.3.	p.42	Derivative of polynomials

* Using Python

Exercise 2.1. Explain the geometrical meaning of Axioms (b) − (h) in the linear space in Example 2.1.



Exercise 2.2. Using `vec3d.py`, check the other axioms of linear space.

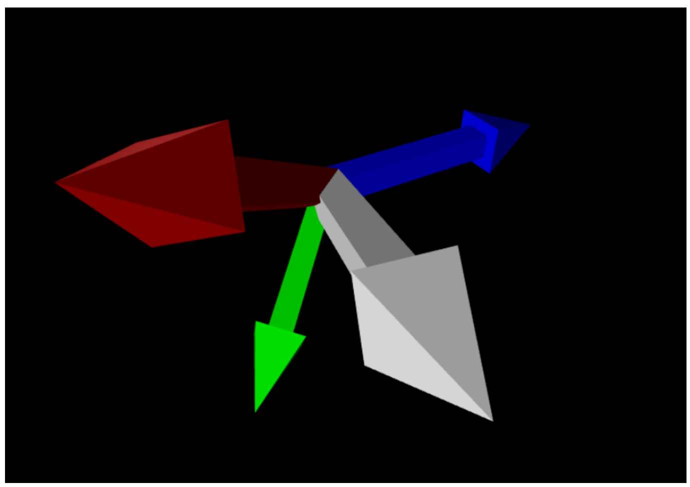
Program: `axioms3d.py`

In [1]:
```
from vpython import *

o = vec(0, 0, 0)

def axiom_a(x, y):
    arrow(axis=x, color=color.red)
    arrow(axis=y, color=color.green)
    arrow(axis=x+y, color=color.yellow)
    arrow(axis=y+x, color=color.yellow)
    print(x+y, y+x)

def axiom_b(x, y, z):
    arrow(axis=x, color=color.red)
    arrow(axis=y, color=color.green)
    arrow(axis=z, color=color.blue)
    arrow(axis=(x+y)+z, color=color.white)
    arrow(axis=x+(y+z), color=color.white)
    print((x+y)+z, x+(y+z))

def axiom_c(x):
    arrow(axis=x, color=color.red)
    arrow(axis=x+o, color=color.green)
    print(x, x+o)

def axiom_d(x):
    arrow(axis=x, color=color.red)
    arrow(axis=-x, color=color.green)
    print(o, x+(-x))

def axiom_e(a, x, y):
    arrow(axis=x, color=color.red)
    arrow(axis=y, color=color.green)
    arrow(axis=a*(x+y), color=color.yellow)
    arrow(axis=a*x+a*y, color=color.yellow)
    print(a*(x+y), a*x+a*y)

def axiom_f(a, b, x):
    arrow(axis=x, color=color.red)
    arrow(axis=(a+b)*x, color=color.yellow)
    arrow(axis=a*x+b*x, color=color.yellow)
    print((a+b)*x, a*x+b*x)

def axiom_g(a, b, x):
    arrow(axis=x, color=color.red)
    arrow(axis=(a*b)*x, color=color.yellow)
```

In [1]:

```
46     arrow(axis=a*(b*x), color=color.yellow)
47     print((a*b)*x, a*(b*x))
48
49
50 def axiom_h(x):
51     arrow(axis=x, color=color.red)
52     arrow(axis=1*x, color=color.yellow)
53     print(x, 1*x)
54
55 x = vec(1, 2, 3)
56 y = vec(1, -1, 2)
57 z = vec(1, 0, -3)
58 a = 2
59 b = 3
60
61 axiom_b(x, y, z)
```

Lines 55 – 61: Check Axiom (b).



Exercise 2.3. Prove the following:

(1) $\overbrace{\boldsymbol{x}+\boldsymbol{x}+\cdots+\boldsymbol{x}}^{n} = n\boldsymbol{x}$, (hint: use Axioms (f), (h) and mathematical induction)

(2) $ab \neq 0 \Rightarrow \dfrac{\boldsymbol{x}}{a} + \dfrac{\boldsymbol{y}}{b} = \dfrac{b\boldsymbol{x}+a\boldsymbol{y}}{ab}$,

(3) If $\boldsymbol{x} \neq \boldsymbol{0}$, then $a\boldsymbol{x} = \boldsymbol{x} \Rightarrow a = 1$,

(4) If $\boldsymbol{x} \neq \boldsymbol{0}$, then $a\boldsymbol{x} = b\boldsymbol{x} \Rightarrow a = b$.

(1) $\boldsymbol{x}+\boldsymbol{x} = 1\boldsymbol{x}+1\boldsymbol{x} = (1+1)\,\boldsymbol{x} = 2\boldsymbol{x}$

(2) This assertion is true for $n = 1$. Suppose it is true for $n = k$. Because

$$\overbrace{\boldsymbol{x}+\boldsymbol{x}+\cdots+\boldsymbol{x}}^{k+1} = \overbrace{\boldsymbol{x}+\boldsymbol{x}+\cdots+\boldsymbol{x}}^{k}+\boldsymbol{x} = k\boldsymbol{x}+\boldsymbol{x} = k\boldsymbol{x}+1\boldsymbol{x} = (k+1)\,\boldsymbol{x},$$

the assertion is true for $n = k+1$. Therefore the assertion is true for any natural number n.

(3) Suppose $ab \neq 0$. Then

$$\begin{aligned} \text{LHS} &= \frac{1}{a}\boldsymbol{x} + \frac{1}{b}\boldsymbol{b} = \left(\frac{1}{a}+\frac{1}{b}\right)\boldsymbol{x} = \frac{b+a}{ab}\boldsymbol{x} \\ &= \left(\frac{1}{ab}(b+a)\right)\boldsymbol{x} = \frac{1}{ab}\big((b+a)\,\boldsymbol{x}\big) = \text{RHS}. \end{aligned}$$

(4) If $a\boldsymbol{x} = \boldsymbol{x}$, we have $a\boldsymbol{x} - \boldsymbol{x} = \boldsymbol{0}$ and then $(a-1)\,\boldsymbol{x} = \boldsymbol{0}$. Since $\boldsymbol{x} \neq \boldsymbol{0}$, $a - 1 = 0$ must hold and therefore wehave $a = 1$.

(5) If $a\boldsymbol{x} = b\boldsymbol{x}$, we have $a\boldsymbol{x} - b\boldsymbol{x} = \boldsymbol{0}$ and then $(a-b)\,\boldsymbol{x} = \boldsymbol{0}$. Because $\boldsymbol{x} \neq \boldsymbol{0}$, $a - b = 0$ must hold and then we have $a = b$.

Exercise 2.4. Prove that $\mathbb{K}^n$ satisfies Axioms (b)–(h).

Let denote

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \ \vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \ \vec{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}.$$

(b)

$$\begin{aligned} (\vec{x}+\vec{y})+\vec{z} &= \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} (x_1+y_1)+z_1 \\ (x_2+y_2)+z_2 \\ \vdots \\ (x_n+y_n)+z_n \end{bmatrix} = \begin{bmatrix} x_1+(y_1+z_1) \\ x_2+(y_2+z_2) \\ \vdots \\ x_n+(y_n+z_n) \end{bmatrix} \\ &= \vec{x} + \begin{bmatrix} y_1+z_1 \\ y_2+z_2 \\ \vdots \\ y_n+z_n \end{bmatrix} = \vec{x}+(\vec{y}+\vec{z}) \end{aligned}$$

(c)

$$\vec{x}+\vec{0} = \begin{bmatrix} x_1+0 \\ x_2+0 \\ \vdots \\ x_n+0 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \vec{x}$$

(d)

$$\vec{x}+(-\vec{x}) = \vec{x} + \begin{bmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{bmatrix} = \begin{bmatrix} x_1+(-x_1) \\ x_2+(-x_2) \\ \vdots \\ x_n+(-x_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \vec{0}$$

(e)

$$\begin{aligned} a(\vec{x}+\vec{y}) &= a\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix} = \begin{bmatrix} a(x_1+y_1) \\ a(x_2+y_2) \\ \vdots \\ a(x_n+y_n) \end{bmatrix} = \begin{bmatrix} ax_1+ay_1 \\ ax_2+ay_2 \\ \vdots \\ ax_n+ay_n \end{bmatrix} = \begin{bmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{bmatrix} + \begin{bmatrix} ay_1 \\ ay_2 \\ \vdots \\ ay_n \end{bmatrix} \\ &= a\vec{x}+a\vec{y} \end{aligned}$$

(f)

$$(a+b)\,\vec{x} \;=\; \begin{bmatrix} (a+b)\,x_1 \\ (a+b)\,x_2 \\ \vdots \\ (a+b)\,x_n \end{bmatrix} \;=\; \begin{bmatrix} ax_1+bx_1 \\ ax_2+bx_2 \\ \vdots \\ ax_n+bx_n \end{bmatrix} \;=\; \begin{bmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{bmatrix} + \begin{bmatrix} bx_1 \\ bx_2 \\ \vdots \\ bx_n \end{bmatrix} \;=\; a\vec{x}+b\vec{x}$$

(g)

$$(ab)\,\vec{x} \;=\; \begin{bmatrix} (ab)\,x_1 \\ (ab)\,x_2 \\ \vdots \\ (ab)\,x_n \end{bmatrix} \;=\; \begin{bmatrix} a\,(bx_1) \\ a\,(bx_2) \\ \vdots \\ a\,(bx_n) \end{bmatrix} \;=\; a\begin{bmatrix} bx_1 \\ bx_2 \\ \vdots \\ bx_n \end{bmatrix} \;=\; a\,(b\vec{x})$$

(h)

$$1\vec{x} \;=\; \begin{bmatrix} 1x_1 \\ 1x_2 \\ \vdots \\ 1x_n \end{bmatrix} \;=\; \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \;=\; \vec{x}$$

Exercise 2.5. Check that $\mathbb{K}^X$ satisfies Axioms (b) – (h).

Suppose $f, g, h \in \mathbb{K}^X$.

(b) Since

$$\begin{aligned}\big((f+g)+h\big)(x) &= (f+g)(x)+h(x) = \big(f(x)+g(x)\big)+f(x)\\ &= f(x)+\big(g(x)+f(x)\big) = f(x)+(g+h)(x)\\ &= \big(f+(g+h)\big)(x)\end{aligned}$$

for all $x \in X$, we have $(f+g)+h = f+(g+h)$.

(c) Since

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

for all $x \in X$, we have $f+0=f$.

(d) Since

$$\big(f+(-f)\big)(x) = f(x)+(-f)\big)(x) = f(x)+\big(-f(x)\big) = 0 = 0(x)$$

for all $x \in X$, we have $f+(-f)=0$.

(e) Since

$$\begin{aligned}\big(a(f+g)\big)(x) &= a(f+g)(x) = a\big(f(x)+g(x)\big) = af(x)+ag(x)\\ &= (af)(x)+(ag)(x) = (af+ag)(x)\end{aligned}$$

for all $x \in X$, we have $a(f+g) = af+ag$.

(f) Since

$$\big((a+b)f\big)(x) = (a+b)f(x) = af(x)+bf(x) = (af)(x)+(bf)(x) = (af+ag)(x)$$

for all $x \in X$, we have $(a+b)f = af+bf$.

(g) Since

$$\big((ab)f\big)(x) = (ab)f(x) = a\big(bf(x)\big) = a(bf)(x) = \big(a(bf)\big)(x)$$

for all $x \in X$, we have $(ab)f = a(bf)$.

(h) Since

$$(1f)(x) = 1f(x) = f(x)$$

for all $x \in X$, we have $1f = f$.

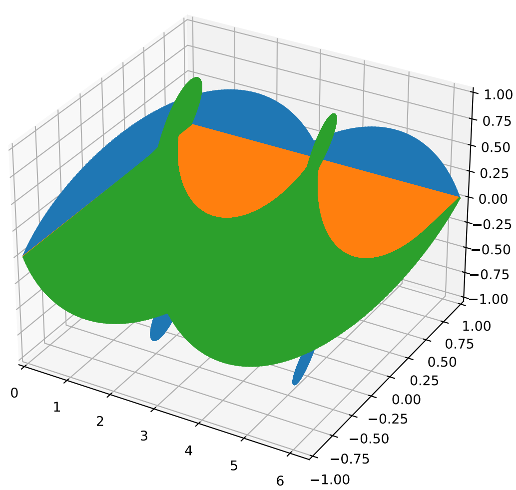
Exercise 2.6. Using the above program, observe the shapes of graphs of the functions for different n. In particular, choose n around 500 and around 1000.

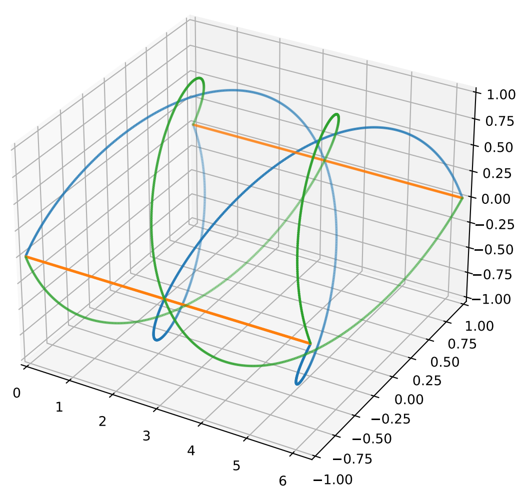
For example, the following program checks the shape of the graph when $n = 499, 500, 501$.

Program: `cfunc.py`

In [1]:
```
from numpy import exp, pi, linspace
import matplotlib.pyplot as plt

f = lambda n, t: exp(1j * n * t)
t = linspace(0, 2 * pi, 1001)

fig = plt.figure(figsize=(7, 7))
ax = fig.add_subplot(111, projection='3d')
ax.set_xlim(0, 2*pi), ax.set_ylim(-1, 1), ax.set_zlim(-1, 1)
for n in [100, 101]:
    z = f(n, t)
    ax.plot(t, z.real, z.imag)
    #ax.scatter(t, z.real, z.imag,s=1)
plt.show()
```

When viewed as a line graph (using `plot`-method), the large, violently oscillating curve looks almost like a curved surface (the left figure). If we don't connect adjacent points with lines (using `scatter`-method), the plotted points will appear to be continuous curves (the right figure).





Exercise 2.7. Prove that S1 and S2 are equivalent to the following condition.

S5. $a\boldsymbol{x} + b\boldsymbol{y} \in W$ for any $a, b \in \mathbb{K}$ and for any $\boldsymbol{x}, \boldsymbol{y} \in W$.

Furthermore, for a subspace W, prove that

$$a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_n\boldsymbol{x}_n \in W$$

for any $a_1, a_2, \ldots, a_n \in \mathbb{K}$ and $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in W$.

Suppose that W satisfies S1 and S2. Let $a, b \in \mathbb{K}$ and $\boldsymbol{x}, \boldsymbol{y} \in W$. From S2, we have $a\boldsymbol{x}, b\boldsymbol{y} \in W$. and then $a\boldsymbol{x} + b\boldsymbol{y} \in W$ follows S1. Conversely, suppose that $W \neq \emptyset$ satisfies S5. When $a = b = 1$, S1 is derived and when $b = 0$, S2 is deraived.

The second half is based on mathematical induction. Let W be a subspace. W satisefies the condition for $n = 1$ by S2. Suppose that W satisefies it for $n \leqq k$. By the assumption of mathematical induction $a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_k\boldsymbol{x}_k \in W$ and also $a_{k+1}\boldsymbol{x}_{k+1} \in W$. By using S1 we get

$$a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_k\boldsymbol{x}_k + a_{k+1}\boldsymbol{x}_{k+1} \;=\; (a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_k\boldsymbol{x}_k) + a_{k+1}\boldsymbol{x}_{k+1} \;\in\; W.$$

Hence, W satisfies the condition for $n = k + 1$.

Exercise 2.8. Prove that $\{(x,0) \mid x \in \mathbb{K}\}$ and $\{(x,x) \mid x \in \mathbb{K}\}$ are subspaces of $\mathbb{K}^2$.

$W_1 \stackrel{\text{def}}{=} \{(x,0) \mid x \in \mathbb{K}\}$ is a non-empty subset of $\mathbb{K}^2$ because $(0,0)$ belongs to W_1. Since

$$a\,(x,0) + b\,(y,0) \;=\; (ax+by,0) \;\in\; W_1$$

for $a,b,x,y \in \mathbb{K}$, it follows that W_1 is closed under the linear combination. Thus, W_1 is a subspace of $\mathbb{K}^2$.

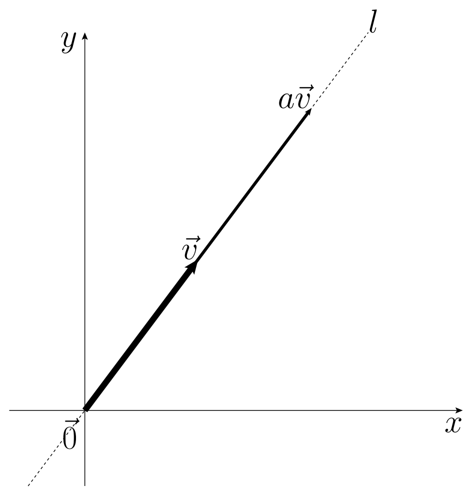
$W_2 \stackrel{\text{def}}{=} \{(x,x) \mid x \in \mathbb{K}\}$ is a non-empty subset of $\mathbb{K}^2$ because $(0,0)$ belongs to W_2. Since

$$a\,(x,x) + b\,(y,y) \;=\; (ax+by, ax+by) \;\in\; W_2,$$

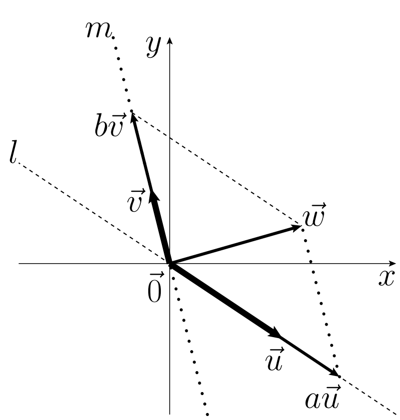
for $a,b,x,y \in \mathbb{K}$, it follows that W_2 is closed under the linear combination. Thus, W_2 is a subspace of $\mathbb{K}^2$.

Exercise 2.9. Explain geometrically why subspaces of a plane and space are given as above.

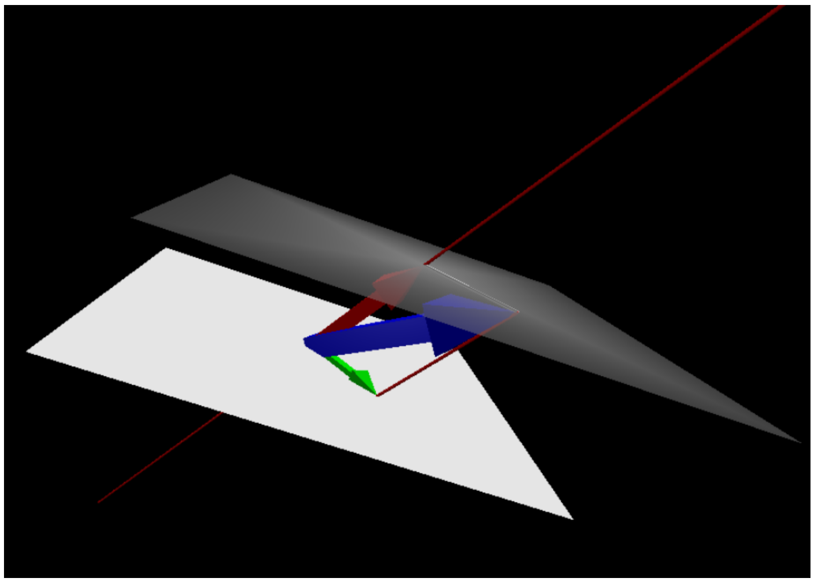
We first consider a subspace of the palne $\mathbb{R}^2$. $\left\{\vec{0}\right\}$ is a trivial subspace of $\mathbb{R}^2$. Suppose a subspace W of $\mathbb{R}^2$ contains a vector $\vec{v}$ other than the zero vector. Then, all vectors on the straight line l passing through the origin on which $\vec{v}$ lies are scalar multiples of $\vec{v}$, and l includes the origin and is closed with respect to a linear combination. Thus, l is a subspace.



Suppose a subspace contains two vectors $\vec{u}$ and $\vec{v}$ that are not collinear through the origin. Let l and m be the straight lines through the origin on which $\vec{u}$ and $\vec{v}$ lie, respectively. Consider a vector $\vec{w}$ that does not lie on either l or m. The point where a line passing through $\vec{w}$ and parallel to m intersects l can be written as $a\vec{u}$, and the point where a line passing through $\vec{w}$ and parallel to l intersects m can be written as $b\vec{v}$. Since $\vec{w} = a\vec{u} + b\vec{v}$, any vector on $\mathbb{R}^2$ can be written as a linear combination of $\vec{u}$ and $\vec{v}$. So if the subspace is neither trivial nor a straight line through the origin, it is $\mathbb{R}^2$.



Next we consider a subspace of the sapce $\mathbb{R}^3$. It follows from consideration of the case of $\mathbb{R}^2$ that the set consisting of only the origin, lines through the origin, and planes through the origin are all subspaces. Let the subspace W contain the plane C through the origin and the vector $\vec{v}$ not on C. All points on the line l through the origin containing $\vec{v}$ are vectors in the subspace W. Consider the vector $\vec{w}$ in $\mathbb{R}^3$ that belongs to neither C nor l. The point where the plane parallel to C and passing through $\vec{w}$ intersects l is the vector of W. The point where the line parallel to l and passing through $\vec{w}$ intersects l is also the vector of W. Since $\vec{w}$ can be written as the sum of these two vectors, $\vec{w}$ is the vector of W. So W is equal to the whole of $\mathbb{R}^3$. So if the subspace is neither a line through the origin nor a plane through the origin, it is the whole space $\mathbb{R}^3$.



Exercise 2.10. Prove that $\bigcap_{i \in I} W_i$ is a subspace of V.

For any $i \in I$, W_i is a subspace and $\mathbf{0} \in W_i$, so $\mathbf{0} \in \bigcap_{i \in I} W_i$ and $\bigcap_{i \in I} W_i$ is not an empty set. It suffices to show that $\bigcap_{i \in I} W_i$ is closed under the linear combination. Let $a, b \in \mathbb{K}$ and $\boldsymbol{x}, \boldsymbol{y} \in \bigcap_{i \in I} W_i$. Fix an arbitrary $i \in I$. Since W_i is a subspace and $\boldsymbol{x}, \boldsymbol{y} \in W_i$, we have $a\boldsymbol{x} + b\boldsymbol{y} \in W_i$. Here, $i \in I$ was arbitrary, so $a\boldsymbol{x} + b\boldsymbol{y} \in \bigcap_{i \in I} W_i$ holds.

Exercise 2.11. Prove that for a subspace W of V, the complement $W^{\complement}$ can never be a subspace of V. Moreover, for subspaces W_1 and W_2 of V, does $W_1 \cup W_2$ or $W_1 \setminus W_2$ become a subspace?

Neither $W^{\complement}$ nor $W_1 \setminus W_2$ can be a subspace because they never have the zero vector.

$W_1 \cup W_2$ cannot be a subspace unless W_1 and W_2 are contained in the other. This can be shown as follows. Suppose that $W_1 \cup W_2$ were a subspace. Chose $w_1, w_2 ¥inV$ such that $\boldsymbol{w}_1 \in W_1, \boldsymbol{w}_1 \notin W_2, \boldsymbol{w}_2 \in W_2, \boldsymbol{w}_2 \notin W_1$, then $\boldsymbol{w}_1 + \boldsymbol{w}_2 \in W_1 \cup W_2$. So here we don't lose generality as $\boldsymbol{w}_1 + \boldsymbol{w}_2 \in W_1$. However, $\boldsymbol{w}_2 = (\boldsymbol{w}_1 + \boldsymbol{w}_2) + (-\boldsymbol{w}_1) \in W_1$, so a contradiction is.

Exercise 2.12. Show that M1 and M2 are equivalent to the following condition.

M5. $\boldsymbol{f}(a\boldsymbol{x} + b\boldsymbol{y}) = a\boldsymbol{f}(\boldsymbol{x}) + b\boldsymbol{f}(\boldsymbol{y})$ for any $a, b \in \mathbb{K}$ and $\boldsymbol{x}, \boldsymbol{y} \in V$.

Moreover, prove that a linear mapping $\boldsymbol{f}$ satisfies

$$\boldsymbol{f}(a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_n\boldsymbol{x}_n) = a_1\boldsymbol{f}(\boldsymbol{x}_1) + a_2\boldsymbol{f}(\boldsymbol{x}_2) + \cdots + a_n\boldsymbol{f}(\boldsymbol{x}_n)$$

for any $a_1, a_2, \ldots, a_n \in \mathbb{K}$ and $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in V$.

Suppose that$\boldsymbol{f}$ saticefies M1 and M2. For any $a, b \in \mathbb{K}$ and $\boldsymbol{x}, \boldsymbol{y} \in V$,

$$\boldsymbol{f}(a\boldsymbol{x} + b\boldsymbol{y}) = \boldsymbol{f}(a\boldsymbol{x}) + \boldsymbol{f}(b\boldsymbol{y}) = a\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{y}).$$

Hence $\boldsymbol{f}$ saticefies M5. Conversely, suppose that $\boldsymbol{f}$ saticefies M5. For any $\boldsymbol{x}, \boldsymbol{y} \in V$ taking $a = b = 1$

$$\boldsymbol{f}(a\boldsymbol{x} + b\boldsymbol{y}) = \boldsymbol{f}(1\boldsymbol{x} + 1\boldsymbol{y}) = 1\boldsymbol{f}(\boldsymbol{x}) + 1\boldsymbol{f}(\boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{y}).$$

Thus M1 is derived.

For any $a \in \mathbb{K}$ and $\boldsymbol{x} \in V$

$$\boldsymbol{f}(a\boldsymbol{x}) = \boldsymbol{f}(a\boldsymbol{x} + 0\boldsymbol{x}) = a\boldsymbol{f}(\boldsymbol{x}) + 0\boldsymbol{f}(\boldsymbol{x}) = a\boldsymbol{f}(\boldsymbol{x}).$$

Hence M1 holds.

The latter half follows from mathematical induction. When $n = 1$, the assertion is true by M2. Suppose that it is true for $n = k$. Then

$$\begin{aligned} & \boldsymbol{f}(a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_k\boldsymbol{x}_k + a_{k+1}\boldsymbol{x}_{k+1}) \\ = \; & \boldsymbol{f}(a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_k\boldsymbol{x}_k) + \boldsymbol{f}(a_{k+1}\boldsymbol{x}_{k+1}) \\ = \; & a_1\boldsymbol{f}(\boldsymbol{x}_1) + a_2\boldsymbol{f}(\boldsymbol{x}_2) + \cdots + a_n\boldsymbol{f}(\boldsymbol{x}_n) + a_{k+1}\boldsymbol{f}(\boldsymbol{x}_{k+1}). \end{aligned}$$

The first equality follows from M1. The second one follows from the assumption of mathematical induction and M2. Hence the assertion is true for $n = k + 1$.

Exercise 2.13. Prove that the scalar multiple of a linear mapping and the composition of linear mappings are linear mappings.

Let $c \in \mathbb{K}$ and also $\boldsymbol{f} : V \to W$ be a linear mapping. Since

$$\begin{aligned} & (c\boldsymbol{f})(a\boldsymbol{x} + b\boldsymbol{y}) = c\boldsymbol{f}(a\boldsymbol{x} + b\boldsymbol{y}) = c\big(a\boldsymbol{f}(\boldsymbol{x}) + b\boldsymbol{f}(\boldsymbol{y})\big) \\ = {} & a\big(c\boldsymbol{f}(\boldsymbol{x})\big) + b\big(c\boldsymbol{f}(\boldsymbol{y})\big) = a(c\boldsymbol{f})(\boldsymbol{x}) + b(c\boldsymbol{f})(\boldsymbol{y}). \end{aligned}$$

for any $a, b \in \mathbb{K}$ and $\boldsymbol{x}, \boldsymbol{y} \in V$, the scalar multiple $c\boldsymbol{f}$ of $\boldsymbol{f}$ is a linear mapping.

Let both $\boldsymbol{f} : V \to W$ and $\boldsymbol{g} : U \to V$ be linear mapping. Since

$$\begin{aligned} & (\boldsymbol{f} \circ \boldsymbol{g})(a\boldsymbol{x} + b\boldsymbol{y}) = \boldsymbol{f}\big(\boldsymbol{g}(a\boldsymbol{x} + b\boldsymbol{y})\big) = \boldsymbol{f}\big(a\boldsymbol{g}(\boldsymbol{x}) + b\boldsymbol{g}(\boldsymbol{y})\big) \\ = {} & a\boldsymbol{f}\big(\boldsymbol{g}(\boldsymbol{x})\big) + b\boldsymbol{f}\big(\boldsymbol{g}(\boldsymbol{y})\big) = a(\boldsymbol{f} \circ \boldsymbol{g})(\boldsymbol{x}) + b(\boldsymbol{f} \circ \boldsymbol{g})(\boldsymbol{y}). \end{aligned}$$

for any $a, b \in \mathbb{K}$ and $\boldsymbol{x}, \boldsymbol{y} \in U$, the composition $\boldsymbol{f} \circ \boldsymbol{g} : U \to W$ of $\boldsymbol{f}$ and $\boldsymbol{g}$ is a linear mapping.

Exercise 2.14. Prove Axioms (a)–(h) of linear space for $L(V,W)$.

Let $\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h} \in L(V,W)$.

(a) Since

$$\begin{aligned} (\boldsymbol{f}+\boldsymbol{g})(\boldsymbol{x}) &= \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x}) \\ &= \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{x}) \\ &= (\boldsymbol{g}+\boldsymbol{f})(\boldsymbol{x}) \end{aligned}$$

for all $\boldsymbol{x} \in V$, we have $\boldsymbol{f}+\boldsymbol{g} = \boldsymbol{g}+\boldsymbol{f}$.

Note that $L(V,W) \subseteq W^V$. Hence, (b)–(h) follow from Exercise 2.13 and Exercise 2.5.

Exercise 2.15. For given constants $a, b, c, d \in \mathbb{R}$ define the mapping $\boldsymbol{f} : \mathbb{R}^2 \to \mathbb{R}^2$ by sending $(x, y) \in \mathbb{R}^2$ to $(u, v) \in \mathbb{R}^2$, where (u, v) is determined by

$$\begin{cases} u &= ax + by \\ v &= cx + dy. \end{cases}$$

Prove that $\boldsymbol{f}$ is a linear mapping. In particular, in each of the cases $(a, b, c, d) = (1, 2, 2, 3)$ and $(a, b, c, d) = (1, 2, 2, 4)$ determine the kernel and the range of $\boldsymbol{f}$.

Since

$$\boldsymbol{f}\big((x, y)\big) = (ax + by, cx + dy),$$

we have

$$\begin{aligned} \boldsymbol{f}\big((x_1, y_1) + (x_2, y_2)\big) &= \boldsymbol{f}\big((x_1 + x_2, y_1 + y_2)\big) \\ &= \big(a(x_1 + x_2) + b(y_1 + y_2), c(x_1 + x_2) + d(y_1 + y_2)\big) \\ &= \big((ax_1 + by_1) + (ax_2 + by_2), (cx_1 + dy_1) + (cx_2 + dy_2)\big) \\ &= (ax_1 + by_1, cx_1 + dy_1) + (ax_2 + by_2, cx_2 + dy_2) \\ &= \boldsymbol{f}\big((x_1, y_1)\big) + \boldsymbol{f}\big((x_2, y_2)\big). \end{aligned}$$

This says that $\boldsymbol{f}$ preserves the vector addition. Preserving the scalor multiplication follows from

$$\begin{aligned} \boldsymbol{f}\big(\alpha(x, y)\big) &= \boldsymbol{f}\big((\alpha x, \alpha y)\big) = (a\alpha x + b\alpha y, c\alpha x + d\alpha y) \\ &= \alpha(ax + by, cx + dy) = \alpha \boldsymbol{f}\big((x, y)\big) \end{aligned}$$

for any $\alpha \in \mathbb{R}$.

Let $(a, b, c, d) = (1, 2, 2, 3)$. Since $f\big((x, y)\big) = (x + 2y, 2x + 3y)$ and

$$\operatorname{kernel}(\boldsymbol{f}) = \big\{(x, y) \;\big|\; x + 2y = 0, 2x + 3y = 0\big\},$$

by solving the simultaneous equations

$$\begin{cases} x + 2y &= 0 \\ 2x + 3y &= 0 \end{cases},$$

we have $x = y = 0$. Thus $\operatorname{kernel}(\boldsymbol{f})$ consists of only the origin. On the other hand,

$$\operatorname{range}(\boldsymbol{f}) = \big\{(u, v) \;\big|\; \text{there exist } x, y \in \mathbb{R} \text{ such that } u = x + 2y, v = 2x + 3y\big\},$$

and for any $(u, v) \in \mathbb{R}^2$ solving the simaltaneous equations

$$\begin{cases} x + 2y &= u \\ 2x + 3y &= v \end{cases}$$

of x, y, we get $x = -3u + 2v, y = 2u - v$. This says that range $(\boldsymbol{f})$ is the whole space $\mathbb{R}^2$.

Let $(a, b, c, d) = (1, 2, 2, 4)$. Since $f\big((x, y)\big) = (x + 2y, 2x + 4y)$,

$$\text{kernel}\,(\boldsymbol{f}) \;=\; \big\{(x, y) \;\big|\; x + 2y = 0, 2x + 4y = 0\big\},$$

solving simaltaneous equation

$$\begin{cases} x + 2y &= 0 \\ 2x + 4y &= 0 \end{cases},$$

of x and y, every (x, y) with $x+2y = 0$ is the solution. Hence, kernel $(\boldsymbol{f})$ is the line $x+2y = 0$. On the other hand,

$$\text{range}\,(\boldsymbol{f}) \;=\; \big\{(u, v) \;\big|\; \text{there exist } x, y \in \mathbb{R} \text{ such that } u = x + 2y, v = 2x + 4y\big\}.$$

Let x and y move arbitrarily on $\mathbb{R}$ and consider u and v satisfying the simultaneous equation

$$\begin{cases} x + 2y &= u \\ 2x + 4y &= v \end{cases}.$$

Then u moves whole $\mathbb{R}$ and v always takes the value of twice u. Hence range $(\boldsymbol{f})$ is the line $y = 2x$.

Exercise 2.16. Consider the set $V = \left\{ax^2 + bx + c \mid a, b, c \in \mathbb{R}\right\}$ of all constant, linear and quadratic functions.

(1) Prove that V is a linear space over $\mathbb{R}$.

(2) Prove that $W = \left\{ax + b \mid a, b \in \mathbb{R}\right\}$ is a subspace of V.

(3) For a function $f \in V$, $D(f)$ denotes the derivative f' of f. For example, $D(1) = 0$, $D(x) = 1$, $D(x^2) = 2x$. Prove that $D : V \to V$ is a linear mapping.

(4) Determine range (D) and kernel (D).

(1) Vector sum:

$$\left(a_1x^2 + b_1x + c_1\right) + \left(a_2x^2 + b_2x + c_2\right) = (a_1 + a_2)x^2 + (b_1 + b_2)x + (c_1 + c_2)$$

Scalor multiple:

$$\alpha\left(ax^2 + bx + c\right) + \left(a_2x^2 + b_2x + c_2\right) = \alpha ax^2 + \alpha bx + \alpha c$$

The zero vector:

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

The inverse vector:

$$-\left(ax^2 + bx + c\right) = (-a)x^2 + (-b)x + (-c)$$

According to the laws concerning the four arithmetic operations of symbolic expressions, we have axioms (a) – (h) easily.

(2) Since $ax + b = 0x^2 + ax + b \in V$, we have $W \subseteq V$. Furthermore, $0 \in W$ and

$$\alpha(a_1x + b_1) + \beta(a_2x + b_2) = (\alpha a_1 + \beta a_2)x + (\alpha b_1 + \alpha b_2) \in W,$$

so W is a subspace of V.

(3) If $f = a_1x^2 + b_1x + c_1$ and $g = a_2x^2 + b_2x + c_2$, then from the polynomial differentiation formula we get

$$\begin{aligned}
D(\alpha f + \beta g) &= \left(\alpha\left(a_1x^2 + b_1x + c_1\right) + \beta\left(a_2x^2 + b_2x + c_2\right)\right)' \\
&= \left((\alpha a_1 + \alpha a_2)x^2 + (\alpha b_1 + \alpha b_2)x + (\alpha c_1 + \alpha c_2)\right)' \\
&= 2(\alpha a_1 + \alpha a_2)x + (\alpha b_1 + \alpha b_2) \\
&= \alpha(2a_1x + b_1) + \beta(2a_2x + b_2) \\
&= \alpha\left(a_1x^2 + b_1x + c_1\right)' + \beta\left(a_2x^2 + b_2x + c_2\right)' \\
&= \alpha D(f) + \beta D(g)
\end{aligned}$$

and the linearity is shown.

(4) range (D) is the whole W of polynomials of maximum degree 2 or less. kernel (D) is a set of all constant functions.

Chapter 3

Basis and Dimension

Exercise 3.1	Sect. 3.1. p.52	Linear combinations*
Exercise 3.2	Sect. 3.2. p.53	Linear independence
Exercise 3.3	Sect. 3.3. p.55	Properties of a basis
Exercise 3.4	Sect. 3.3. p.56	Linear isomorphism
Exercise 3.5	Sect. 3.3. p.57	A basis of the set of polynomials
Exercise 3.6	Sect. 3.5. p.62	Properties of direct sum subspaces
Exercise 3.7	Sect. 3.5. p.63	Indepencence of subspaces
Exercise 3.8	Sect. 3.5. p.63	Direct sum of two subspaces

* Using Python

Exercise 3.1. Express $\vec{x}$ as a linear combination of the vectors in A in each case below, and check your solution by solving it by SymPy. Moreover, draw $\vec{x}$ and the vectors in A using Matplotlib in case 1 and using VPython in case 2.

(1) $\vec{x} = (17, -10)\,,\ A = \{(5, -4)\,,\ (4, -5)\}$.

(2) $\vec{x} = (-16, 1, 10)\,,\ A = \{(1, 0, 0)\,, (1, 1, 0)\,, (1, 1, 1)\}$.

(1) Let

$$(17, -10) \ = \ a\,(5, -4) + b\,(4, -5)\,.$$

Since

$$\text{LHS} \ = \ (5a + 4b, -4a - 5b)\,,$$

we have the following simultaneous equations

$$\begin{cases} 5a + 4b = 17 & \cdots (1) \\ -4a - 5b = -10 & \cdots (2) \end{cases}$$

of unknown variables a and b. In order to eliminate a, from $(1) \times 4 + (2) \times 5$ we have

$$\begin{array}{rrcr} & 20a + 16b & = & 68 \\ +) & -20a - 25b & = & -50 \\ \hline & -9b & = & 18 \end{array}$$

and get $b = -2$. Substituting it in (1), from

$$5a - 8 \ = \ 17$$

we get $a = 5$. Hence

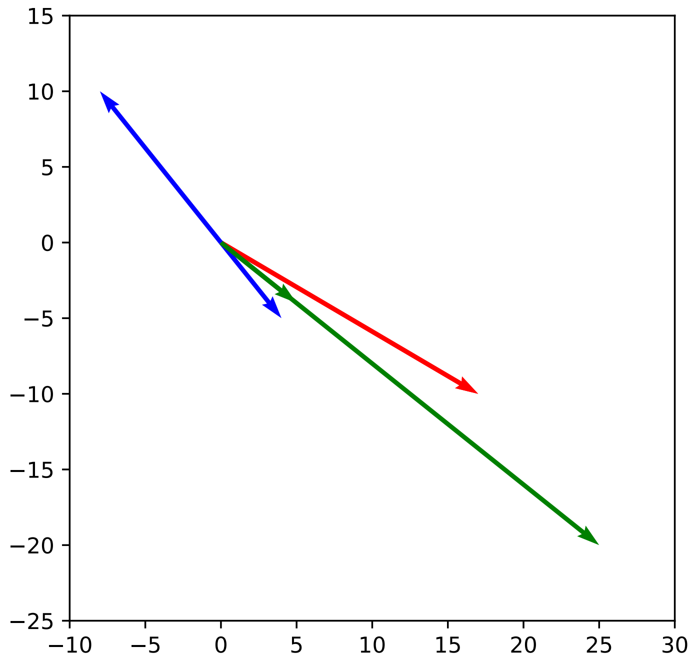
$$(17, -10) \ = \ 5\,(5, -4) - 2\,(4, -5)\,.$$

Program: `vec2d.py`

In [1]:
```
from numpy import array
import matplotlib.pyplot as plt

o = array([0, 0])
x = array([17, -10])
A = [array([5, -4]), array([4, -5])]

def arrow(v, c):
    plt.quiver(0, 0, v[0], v[1], color=c, units='xy', scale=1)

fig = plt.figure(figsize=(5, 5))
plt.axis('scaled'), plt.xlim(-10, 30), plt.ylim(-25, 15)

```

In [1]:
```
14 arrow(x, 'r')
15 arrow(A[0], 'g')
16 arrow(A[1], 'b')
17
18 arrow(5*A[0] - 2*A[1], 'r')
19 arrow(5*A[0], 'g')
20 arrow(-2*A[1], 'b')
21
22 plt.show()
23 #plt.savefig('vec2d.pdf')
```

Lines 22, 23: If you comment out line 22 and remove the comment out in line 23, the image will be saved with the specified file name. As in this example, if the file extension is `pdf`(resp. `jpg` or `png`), the image will be saved as a pdf (resp. jpg or png) file.



(2) Let

$$(-16, 1, 10) \ = \ a\,(1, 0, 0) + b\,(1, 1, 0) + c\,(1, 1, 1)\,.$$

Since

$$\text{LHS} \ = \ (a + b + c,\ b + c,\ c)\,,$$

we have simultaneous equations

$$\begin{cases} a + b + c \ = \ -16 & \cdots\ (1) \\ \qquad b + c \ = \ 1 & \cdots\ (2) \\ \qquad\quad c \ = \ 10 & \cdots\ (3) \end{cases}.$$

of a, b and c. From (3), $c = 10$. Substituting this to (2), $b + 10 = 1$ and $b = -9$. Further substituting this to (1), $a - 9 + 10 = -16$ and then $a = -17$. Therefore

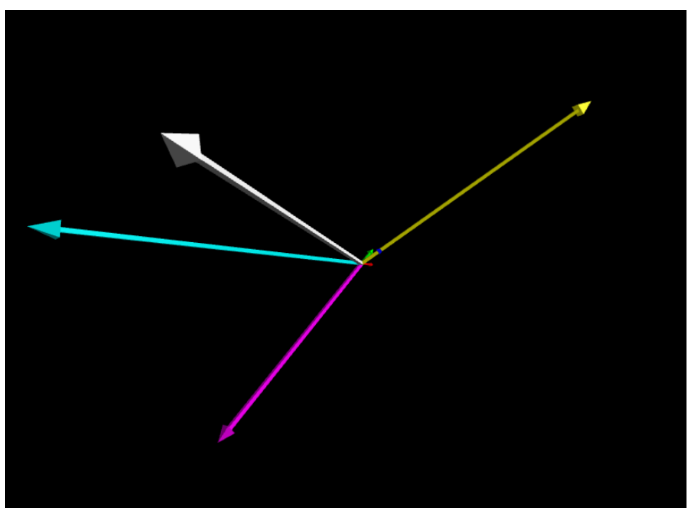
$$(-16, 1, 10) \ = \ -17\,(1, 0, 0) - 9\,(1, 1, 0) + 10\,(1, 1, 1)\,.$$

Program: `vec3d.py`

In [1]:

```
1  from vpython import *
2
3  x = vec(-16,1,10)
4  A = [vec(1,0,0), vec(1,1,0), vec(1,1,1)]
5  arrow(axis=x, shaftwidth=0.2)
6  arrow(axis=A[0], color=color.red)
7  arrow(axis=A[1], color=color.green)
8  arrow(axis=A[2], color=color.blue)
9
10 arrow(axis=-17*A[0]-9*A[1]+10*A[2], shaftwidth=0.2)
11 arrow(axis=-17*A[0], color=color.cyan, shaftwidth=0.2)
12 arrow(axis=-9*A[1], color=color.magenta, shaftwidth=0.2)
13 arrow(axis=10*A[2], color=color.yellow, shaftwidth=0.2)
14
15 #input('Hit Enter key')
16 #scene.capture('vec3d')
```

Lines 15, 16: If you uncomment this line, the image displayed in the browser will be saved in the specified file in png format when you press the `Enter` key in the shell window after changing the viewpoint or scaling with the mouse. It is saved with the name `vec3d.png` (the extension is automatically added). The file is saved in the download file save folder specified by the browser.



Exercise 3.2. Solve the equations (a) and (b) by hand, and decide if they are linearly independent.

$$\text{(a)}\begin{cases} x + 2y + 3z = 0 & \cdots (1) \\ 2x + 3y + 4z = 0 & \cdots (2) \\ 3x + 4y + 5z = 0 & \cdots (3) \end{cases}$$

(1) × 2 − (2):

$$\begin{array}{rrrrrrrr} & 2x & + & 4y & + & 6z & = & 0 \\ -) & 2x & + & 3y & + & 4z & = & 0 \\ \hline & & & y & + & 2z & = & 0 \quad \cdots (4) \end{array}$$

(1) × 3 − (3):

$$\begin{array}{rrrrrrrr} & 3x & + & 6y & + & 9z & = & 0 \\ -) & 3x & + & 4y & + & 5z & = & 0 \\ \hline & & & 2y & + & 4z & = & 0 \quad \cdots (5) \end{array}$$

Here (4) and (5) are equivalent. For example, by substituting the pair $y = 2, z = -1$ which satisfis them to (1) (or (2) or (3)) we have $x = 1$. Hence we get

$$1\,(1,2,3) + 2\,(2,3,4) + (-1)\,(3,4,5) \ = \ (0,0,0)$$

and we can conclude that $\big\{(1,2,3)\,,\ (2,3,4)\,,\ (3,4,5)\big\}$ is linearly independent.

$$\text{(b)}\begin{cases} x + 2y + 3z = 0 & \cdots (1) \\ 2x + 3y + z = 0 & \cdots (2) \\ 3x + y + 2z = 0 & \cdots (3) \end{cases}$$

(1) × 2 − (2):

$$\begin{array}{rrrrrrrr} & 2x & + & 4y & + & 6z & = & 0 \\ -) & 2x & + & 3y & + & z & = & 0 \\ \hline & & & y & + & 5z & = & 0 \quad \cdots (4) \end{array}$$

(1) × 3 − (3):

$$\begin{array}{rrrrrrrr} & 3x & + & 6y & + & 9z & = & 0 \\ -) & 3x & + & y & + & 2z & = & 0 \\ \hline & & & 5y & + & 7z & = & 0 \quad \cdots (5) \end{array}$$

(4) × 5 − (5):

$$\begin{array}{rrrrrr} & 5y & + & 25z & = & 0 \\ -) & 5y & + & 7z & = & 0 \\ \hline & & & 18z & = & 0 \end{array}$$

Henve $z = 0$. Substituting it to (4) we have $y = 0$ and the doing it to (1) we have $x = 0$. Therefore,

$$0\,(1,2,3) + 0\,(2,3,1) + 0\,(3,1,2) \ = \ (0,0,0)$$

is a unique expression, so that $\{(1,2,3)\,,\,(2,3,1)\,,\,(3,1,2)\}$ is linearly independent.

Exercise 3.3. Let X be a basis of V. Prove the following facts.

(1) A set obtained by adding a new vector to X or removing any vector of X is no longer a basis of V.

(2) A set obtained by replacing any vector of X with its nonzero scalar multiple remains a basis.

(3) A set obtained by replacing any vector of X with a sum of it and a scalar multiple of another vector of X remains a basis.

(1) For any $\boldsymbol{v} \in V$, consider $Y \stackrel{\text{def}}{=} X \cup \{\boldsymbol{v}\}$. Since X generates V, $\boldsymbol{v}$ is writen as a linear combination of vectors in Y. Hence Y is dependent and can not be a basis of V. On the other hand, for and $\boldsymbol{v} \in X$, consider $Y \stackrel{\text{def}}{=} X \setminus \{\boldsymbol{v}\}$. Since X was independent, $\boldsymbol{v}$ can not be written as a linear combination of vectors in Y. Hence Y does not generate V and can not be a basis of V.

(2) Let $X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ be a basis and $c \neq 0$. If

$$a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_i c\boldsymbol{x}_i + \cdots + a_n\boldsymbol{x}_n \ = \ \mathbf{0},$$

then

$$a_1 \ = \ a_2 \ = \ \cdots \ = \ a_i c \ = \ \cdots \ = \ a_n \ = \ 0.$$

Because $c \neq 0$, we have $a_i = 0$. Therefore, $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, c\boldsymbol{x}_i, \ldots, \boldsymbol{x}_n\}$ is linearly dependdent. Let $\boldsymbol{y} \in V$ be arbitrary. Then,

$$\boldsymbol{y} \ = \ a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_i\boldsymbol{x}_i + \cdots + a_n\boldsymbol{x}_n$$

for some $a_1, a_2, \ldots, a_n \in \mathbb{K}$. Since $c \neq 0$, we have

$$\boldsymbol{y} \ = \ a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + \frac{a_i}{c} c\boldsymbol{x}_i + \cdots + a_n\boldsymbol{x}_n.$$

Thus, $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, c\boldsymbol{x}_i, \ldots, \boldsymbol{x}_n\}$ spans V.

(3) Consider any pair of different vectors $\boldsymbol{v}, \boldsymbol{w} \in X$ and put $\boldsymbol{u} \stackrel{\text{def}}{=} \boldsymbol{v} + a\boldsymbol{w}$ for some $a \in \mathbb{K}$. Let Y be the result of removing $\boldsymbol{v}$ from X and adding $\boldsymbol{u}$ in it, that is

$$Y \stackrel{\text{def}}{=} (X \setminus \{\boldsymbol{v}\}) \cup \{\boldsymbol{u}\} \ = \ \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n\},$$

where $\boldsymbol{y}_1 = \boldsymbol{u}$ and $\boldsymbol{y}_2 = \boldsymbol{w}$. First we will show that Y is linear independent. Let

$$b_1\boldsymbol{y}_1 + b_2\boldsymbol{y}_2 + \cdots + b_n\boldsymbol{y}_n \ = \ \mathbf{0}$$

Then

$$b_1(\boldsymbol{v} + a\boldsymbol{w}) + b_2\boldsymbol{w} + \cdots + b_n\boldsymbol{y}_n \ = \ b_1\boldsymbol{v} + (ab_1 + b_2)\boldsymbol{w} + \cdots + b_n\boldsymbol{y}_n \ = \ \mathbf{0}.$$

Since X is independent, we have

$$b_1 \;=\; ab_1 + b_2 \;=\; b_3 \;=\; \cdots \;=\; b_n \;=\; 0$$

and also $b_2 = 0$. Thus, Y is independent. Next,we will show that Y generates V. Since X generates V, for any $\boldsymbol{x} \in V$ there extst $c_1, c_2, \ldots, c_n \in \mathbb{K}$ such that

$$\boldsymbol{x} \;=\; c_1\boldsymbol{v} + c_2\boldsymbol{w} + c_3\boldsymbol{y}_3 + \cdots + c_n\boldsymbol{y}_n.$$

Recalling $\boldsymbol{v} = \boldsymbol{u} - a\boldsymbol{w}$, we have

$$\boldsymbol{x} \;=\; c_1\boldsymbol{u} + (c_2 - ac_1)\,\boldsymbol{w} + c_3\boldsymbol{y}_3 + \cdots + c_n\boldsymbol{y}_n.$$

Since $\boldsymbol{x} \in V$ was arbitrary, this says that Y generates V.

Exercise 3.4. Let V and W be linear spaces and let $\boldsymbol{f} : V \to W$ be a linear isomorphism. Let $X = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\} \subset V$. Prove the following facts.

(1) If X generates V, then $\boldsymbol{f}(X) = \{\boldsymbol{f}(\boldsymbol{v}_1), \boldsymbol{f}(\boldsymbol{v}_2), \ldots, \boldsymbol{f}(\boldsymbol{v}_k)\}$ generates W.

(2) If X is linearly independent, then so is $\boldsymbol{f}(X)$.

(3) If X is a basis of V, then $\boldsymbol{f}(X)$ is a basis of W.

(1) Let $\boldsymbol{w} \in W$ be arbitrary. Because $\boldsymbol{f}$ is surjective, there exists $\boldsymbol{v} \in V$ such that $\boldsymbol{w} = \boldsymbol{f}(\boldsymbol{v})$, Since $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ generates V, we have

$$\boldsymbol{v} = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_k\boldsymbol{v}_k.$$

hence

$$\boldsymbol{w} = \boldsymbol{f}(\boldsymbol{v}) = \boldsymbol{f}(a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_k\boldsymbol{v}_k) = a_1\boldsymbol{f}(\boldsymbol{v}_1) + a_2\boldsymbol{f}(\boldsymbol{v}_2) + \cdots + a_k\boldsymbol{f}(\boldsymbol{v}_k).$$

That is, $\boldsymbol{w} \in W$ can be written as a linear combination of $\{\boldsymbol{f}(\boldsymbol{v}_1), \boldsymbol{f}(\boldsymbol{v}_2), \ldots, \boldsymbol{f}(\boldsymbol{v}_k)\}$. Hence, $\boldsymbol{w} \in W$was arbitrary, $\{\boldsymbol{f}(\boldsymbol{v}_1), \boldsymbol{f}(\boldsymbol{v}_2), \ldots, \boldsymbol{f}(\boldsymbol{v}_k)\}$ generate W.

(2) Let

$$a_1\boldsymbol{f}(\boldsymbol{v}_1) + a_2\boldsymbol{f}(\boldsymbol{v}_2) + \cdots + a_k\boldsymbol{f}(\boldsymbol{v}_k) = \boldsymbol{0}_W.$$

Then

$$\boldsymbol{f}(a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_k\boldsymbol{v}_k) = \boldsymbol{f}(\boldsymbol{0}_V).$$

Bacause $\boldsymbol{f}$ is injective,we have

$$a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_k\boldsymbol{v}_k = \boldsymbol{0}_V.$$

Since $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ is linear independent, we get $a_1 = a_2 = \cdots = a_k = 0$. Thus, $\{\boldsymbol{f}(\boldsymbol{v}_1), \boldsymbol{f}(\boldsymbol{v}_2), \ldots, \boldsymbol{f}(\boldsymbol{v}_k)\}$ is independent.

(3) This follows from (1) and (2)..

Exercise 3.5. Let V be the linear space of all polynomials in x of degree up to 2 with real coefficients. Prove that $\{\, x^2+2x+3,\, 2x^2+3x+1,\, 3x^2+x+2 \,\}$ becomes a basis of V. Moreover, represent x^2-2x+1 as a vector in $\mathbb{R}^3$ on this basis.

Because $\{x^2, x, 1\}$ is a base of V, the dimension of V is 3. Hence if

$$\left\{\, x^2+2x+3,\, 2x^2+3x+1,\, 3x^2+x+2 \,\right\}$$

which has three elements of V is linear independent, it must be a basis of V. Let

$$a\left(x^2+2x+3\right)+b\left(2x^2+3x+1\right)+c\left(3x^2+x+2\right) \;=\; 0.$$

Since

$$(a+2b+3c)\,x^2+(2a+3b+c)\,x+(3a+b+2c) \;=\; 0,$$

we have simaltaneous equations

$$\begin{cases} a + 2b + 3c = 0 & \cdots (1) \\ 2a + 3b + c = 0 & \cdots (2) \\ 3a + b + 2c = 0 & \cdots (3) \end{cases}$$

of unknown variable a, b and c.
(1) $\times$ 2 $-$ (2):

$$\begin{array}{rrrrrrrrl} & 2a & + & 4b & + & 6c & = & 0 & \\ -) & 2a & + & 3b & + & c & = & 0 & \\ \hline & & & b & + & 5c & = & 0 & \cdots (4) \end{array}$$

(1) $\times$ 3 $-$ (3):

$$\begin{array}{rrrrrrrrl} & 3a & + & 6b & + & 9c & = & 0 & \\ -) & 3a & + & b & + & 2c & = & 0 & \\ \hline & & & 5b & + & 7c & = & 0 & \cdots (5) \end{array}$$

(4) $\times$ 5 $-$ (5):

$$\begin{array}{rrrrrr} & 5b & + & 25c & = & 0 \\ -) & 5b & + & 7c & = & 0 \\ \hline & & & 18c & = & 0 \end{array}$$

We first get $c=0$. Secondly, by substituting it to (5), we have.$b=0$. Finally, by substituting them to (1), we have $a=0$. Thus the linear independency has been shown.

We show the latter. Put

$$\begin{aligned} x^2-2x+1 &= a\left(x^2+2x+3\right)+b\left(2x^2+3x+1\right)+c\left(3x^2+x+2\right) \\ &= (a+2b+3c)\,x^2+(2a+3b+c)\,x+(3a+b+2c)\,. \end{aligned}$$

Hence we have simaltaneous equations

$$\begin{cases} a + 2b + 3c = 1 & \cdots (1) \\ 2a + 3b + c = -2 & \cdots (2) \\ 3a + b + 2c = 1 & \cdots (3) \end{cases}$$

of unknown variables a, b and c.
$(1) \times 2 - (2)$:

$$\begin{array}{rrrrrrrrl} & 2a & + & 4b & + & 6c & = & 2 & \\ -) & 2a & + & 3b & + & c & = & -2 & \\ \hline & & & b & + & 5c & = & 4 & \cdots (4) \end{array}$$

$(1) \times 3 - (3)$:

$$\begin{array}{rrrrrrrrl} & 3a & + & 6b & + & 9c & = & 3 & \\ -) & 3a & + & b & + & 2c & = & 1 & \\ \hline & & & 5b & + & 7c & = & 2 & \cdots (5) \end{array}$$

$(4) \times 5 - (5)$:

$$\begin{array}{rrrrrr} & 5b & + & 25c & = & 20 \\ -) & 5b & + & 7c & = & 2 \\ \hline & & & 18c & = & 18 \end{array}$$

We first get $c = 1$. Secondly, by substituting it to (4), we have $b = -1$. Secondly, by substituting them to (1), we have $a = 0$. Therefore the vector $\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \in \mathbb{R}^3$ is the representation of $x^2 - 2x + 1$.

Exercise 3.6. (1) Prove that W above is a subspace of V.
(2) Prove that W is equal to the subspace generated by $W_1 \cup W_2 \cup \cdots \cup W_k$.
(3) For any i with $1 \leqq i < k$, prove

$$W_1 + W_2 + \cdots + W_k = (W_1 + W_2 + \cdots + W_i) + (W_{i+1} + W_2 + \cdots + W_k).$$

(1) It is easily seen that W is a non-empty subset of V. Consider $a, b \in \mathbb{K}$ and

$$\boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_k,\ \boldsymbol{y}_1 + \boldsymbol{y}_2 + \cdots + \boldsymbol{y}_k \in W$$

with $\boldsymbol{x}_i, \boldsymbol{y}_i \in W_i$ for every $i = 1, 2, \ldots, k$. Then

$$\begin{aligned} & a(\boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_k) + b(\boldsymbol{y}_1 + \boldsymbol{y}_2 + \cdots + \boldsymbol{y}_k) \\ = \ & (a_1\boldsymbol{x}_1 + b\boldsymbol{y}_1) + (a_2\boldsymbol{x}_2 + b_2\boldsymbol{y}_2) + \cdots + (a_k\boldsymbol{x}_k + b_k\boldsymbol{y}_k) \end{aligned}$$

and $a_i\boldsymbol{x}_i + b_i\boldsymbol{y}_i \in W_i$ for every $i = 1, 2, \ldots, k$. Hence the right hand side of the above equality is the vector of W. Since W is a non-empty subset of V and closed under linear combination, W is a subspace of V.

(2) Since for any $x_i \in W_i$ $(i = 1, 2, \ldots, k)$

$$\boldsymbol{x}_i \ = \ \mathbf{0} + \cdots + \mathbf{0} + \boldsymbol{x}_i + \mathbf{0} + \cdots + \mathbf{0} \ \in \ W,$$

we have $W_i \subseteq W$ for every $i = 1, 2, \ldots, n$.and slso $W_1 \cup W_2 \cup \cdots \cup W_k \subset W$. Let W' be a subspace of V containing $W_1 \cup W_2 \cup \cdots \cup W_k$. Because W' is closed under vector addition, we have

$$\boldsymbol{x}_1 + \boldsymbol{x}_2 \cdots + \boldsymbol{x}_k \ \in \ W'$$

for any $\boldsymbol{x}_1 \in W_1, \boldsymbol{x}_2 \in W_2, \ldots, \boldsymbol{x}_k \in W_k$. This says that $W \subseteq W'$. Therefore, W is the smallest subspace of V containing $W_1 \cup W_2 \cup \cdots \cup W_k$, that is, W is the subspace generated by $W_1 \cup W_2 \cup \cdots \cup W_k$.

(3) Let $\boldsymbol{y} \in W_1 + W_2 + \cdots + W_i$ and $\boldsymbol{z} \in W_{i+1} + W_{i+2} + \cdots + W_k$. Because $\boldsymbol{y}$ and $\boldsymbol{z}$ can be written such as

$$\boldsymbol{y} \ = \ \boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_i, \quad \boldsymbol{z} \ = \ \boldsymbol{x}_{i+1} + \boldsymbol{x}_{i+2} + \cdots + \boldsymbol{x}_k$$

for some $\boldsymbol{x}_1 \in W_1, \boldsymbol{x}_2 \in W_2, \ldots, \boldsymbol{x}_k \in W_k$. Then

$$\boldsymbol{y} + \boldsymbol{z} \ = \ \boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_k \ \in W,$$

so that the RHS is included in W. Conversely, any $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_k \ \in W$ is in the RHS, because

$$\boldsymbol{x} \ = \ (\boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_i) + (\boldsymbol{x}_{i+1} + \cdots + \boldsymbol{x}_k).$$

Exercise 3.7. For nonzero subspaces W_1, W_2 of V, prove that the following statements are equivalent.
(1) $\{W_1, W_2\}$ is linearly independent,
(2) Any element $\boldsymbol{x}$ in $W_1 + W_2$ is uniquely written as $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$ with $\boldsymbol{x}_1 \in W_1$ and $\boldsymbol{x}_2 \in W_2$,
(3) $W_1 \cap W_2 = \{\mathbf{0}\}$.

(1) implies (2): We assume (1). Suppose that $\boldsymbol{x} \in W_1 + W_2$ satisfies

$$\boldsymbol{x} \ = \ \boldsymbol{x}_1 + \boldsymbol{x}_2 \ = \ \boldsymbol{x}_1' + \boldsymbol{x}_2'$$

for some $\boldsymbol{x}_1, \boldsymbol{x}_1' \in W_1$ and $\boldsymbol{x}_2, \boldsymbol{x}_2' \in W_2$. Then

$$(\boldsymbol{x}_1 - \boldsymbol{x}_1') + (\boldsymbol{x}_2 - \boldsymbol{x}_2') \ = \ \mathbf{0}.$$

Since $\{W_1, W_2\}$ is independent, both terms on the left hand side or either of them cannot be non-zero vector(s), that is. both must be the zero vector. Hence $\boldsymbol{x}_1 = \boldsymbol{x}_1'$ and $\boldsymbol{x}_2 = \boldsymbol{x}_2'$.

(2) implies (3): Consider $\boldsymbol{x} \in W_1 \cap W_2$. Then

$$\boldsymbol{x} + \mathbf{0} \ = \ \mathbf{0} + \boldsymbol{x} \in W_1 + W_2.$$

Then, (2) derives $\boldsymbol{x} = \mathbf{0}$.

(3) implies (1): Let $0 \neq \boldsymbol{x}_1 \in W_1$ and $\mathbf{0} \neq \boldsymbol{x}_2 \in W_2$. Suppose

$$a\boldsymbol{x}_1 + b\boldsymbol{x}_2 \ = \ \mathbf{0}.$$

We will sow that $a = b = 0$. If not so, let $a \neq 0$ without of loss of generarity. Then

$$\boldsymbol{x}_1 \ = \ \frac{b}{a}\boldsymbol{x}_2 \ \in W_2.$$

Hence $\boldsymbol{x}_1 = \mathbf{0}$, and this is contradicton.

Exercise 3.8. Prove the results stated just above.

Check the following:

1. Vector addition and scalor multiplication on $V \times W$ defined above satisfy axioms of linear space in Chapter 2.
2. Both V' and W' are closed under vector addition and scalor multiplication.
3. $V' \cup W'$ generates $V \times W$.
4. $\{V', W'\}$ is linearly independent.

1. (a)

$$\begin{aligned}(\boldsymbol{v}_1, \boldsymbol{w}_1) + (\boldsymbol{v}_2, \boldsymbol{w}_2) &= (\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}_1 + \boldsymbol{w}_2) \\ &= (\boldsymbol{v}_2 + \boldsymbol{v}_1, \boldsymbol{w}_2 + \boldsymbol{w}_1) \\ &= (\boldsymbol{v}_2, \boldsymbol{w}_2) + (\boldsymbol{v}_1, \boldsymbol{w}_1)\end{aligned}$$

(b)

$$\begin{aligned}\big((\boldsymbol{v}_1, \boldsymbol{w}_1) + (\boldsymbol{v}_2, \boldsymbol{w}_2)\big) + (\boldsymbol{v}_3, \boldsymbol{w}_3) &= (\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}_1 + \boldsymbol{w}_2) + (\boldsymbol{v}_3, \boldsymbol{w}_3) \\ &= \big((\boldsymbol{v}_1 + \boldsymbol{v}_2) + \boldsymbol{v}_3, (\boldsymbol{w}_1 + \boldsymbol{w}_2) + \boldsymbol{w}_3\big) \\ &= \big(\boldsymbol{v}_1 + (\boldsymbol{v}_2 + \boldsymbol{v}_3), \boldsymbol{w}_1 + (\boldsymbol{w}_2 + \boldsymbol{w}_3)\big) \\ &= (\boldsymbol{v}_1, \boldsymbol{w}_1) + (\boldsymbol{v}_2 + \boldsymbol{v}_3, \boldsymbol{w}_2 + \boldsymbol{w}_3) \\ &= (\boldsymbol{v}_1, \boldsymbol{w}_1) + \big((\boldsymbol{v}_2, \boldsymbol{w}_2) + (\boldsymbol{v}_3, \boldsymbol{w}_3)\big)\end{aligned}$$

(c)

$$\begin{aligned}(\boldsymbol{v}, \boldsymbol{w}) + (\mathbf{0}_V, \mathbf{0}_W) &= (\boldsymbol{v} + \mathbf{0}_V, \boldsymbol{w} + \mathbf{0}_W) \\ &= (\boldsymbol{v}, \boldsymbol{w})\end{aligned}$$

(d)

$$\begin{aligned}(\boldsymbol{v}, \boldsymbol{w}) + \big(-(\boldsymbol{v}, \boldsymbol{w})\big) &= (\boldsymbol{v}, \boldsymbol{w}) + (-\boldsymbol{v}, -\boldsymbol{w}) \\ &= \big(\boldsymbol{v} + (-\boldsymbol{v}), \boldsymbol{w} + (-\boldsymbol{v})\big) \\ &= (\mathbf{0}_V, \mathbf{0}_W)\end{aligned}$$

(e)

$$\begin{aligned}a\big((\boldsymbol{v}_1, \boldsymbol{w}_1) + (\boldsymbol{v}_2, \boldsymbol{w}_2)\big) &= a(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}_1 + \boldsymbol{w}_2) \\ &= a(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}_1 + \boldsymbol{w}_2) \\ &= \big(a(\boldsymbol{v}_1 + \boldsymbol{v}_2), a(\boldsymbol{w}_1 + \boldsymbol{w}_2)\big) \\ &= (a\boldsymbol{v}_1 + a\boldsymbol{v}_2, a\boldsymbol{w}_1 + a\boldsymbol{w}_2) \\ &= (a\boldsymbol{v}_2, a\boldsymbol{w}_2) + (a\boldsymbol{v}_1, a\boldsymbol{w}_1) \\ &= a(\boldsymbol{v}_2, \boldsymbol{w}_2) + a(\boldsymbol{v}_1, \boldsymbol{w}_1)\end{aligned}$$

(f)

$$\begin{aligned}(a+b)(\boldsymbol{v},\boldsymbol{w}) &= \big((a+b)\boldsymbol{v},(a+b)\boldsymbol{w}\big)\\ &= (a\boldsymbol{v}+b\boldsymbol{v},a\boldsymbol{w}+b\boldsymbol{w})\\ &= (a\boldsymbol{v},a\boldsymbol{w})+(b\boldsymbol{v},b\boldsymbol{w})\\ &= a(\boldsymbol{v},\boldsymbol{w})+b(\boldsymbol{v},\boldsymbol{w})\end{aligned}$$

(g)

$$\begin{aligned}(ab)(\boldsymbol{v},\boldsymbol{w}) &= \big((ab)\boldsymbol{v},(ab)\boldsymbol{w}\big)\\ &= \big(a(b\boldsymbol{v}),a(b\boldsymbol{w})\big)\\ &= a(b\boldsymbol{v},b\boldsymbol{w})\\ &= a\big(b(\boldsymbol{v},\boldsymbol{w})\big)\end{aligned}$$

(h)

$$\begin{aligned}1(\boldsymbol{v},\boldsymbol{w}) &= (1\boldsymbol{v},1\boldsymbol{w})\\ &= (\boldsymbol{v},\boldsymbol{w})\end{aligned}$$

2. It is clear that $V' \subseteq V\times W$ and $(\mathbf{0}_V,\mathbf{0}_W)\in V'$. Hence, V' is a nonempty subset of $V\times W$. For $a,b\in\mathbb{K}$ and $(\boldsymbol{v}_1,\mathbf{0}_W),(\boldsymbol{v}_2,\mathbf{0}_W)\in V'$

   $$a(\boldsymbol{v}_1,\mathbf{0}_W)+b(\boldsymbol{v}_2,\mathbf{0}_W) = (a\boldsymbol{v}_1+b\boldsymbol{v}_2,\mathbf{0}_W) \in V'.$$

   Hence, V' is closed under the linear combination. Therefore, V' is a linear subspace of $V\times W$.

   It is clear that $W' \subseteq V\times W$ and $(\mathbf{0}_V,\mathbf{0}_W)\in W'$. Hence, W' is a nonempty subset of $V\times W$. For $a,b\in\mathbb{K}$ and $(\mathbf{0}_V,\boldsymbol{w}_1),(\mathbf{0}_V,\boldsymbol{w}_2)\in W'$

   $$a(\mathbf{0}_V,\boldsymbol{w}_1)+b(\mathbf{0}_V,\boldsymbol{w}_1) = (\mathbf{0}_V,a\boldsymbol{w}_1+b\boldsymbol{w}_2,) \in W'.$$

   Hence, W' is closed under the linear combination. Therefore, W' is a linear subspace of $V\times W$.

3. Since any $(\boldsymbol{v},\boldsymbol{w})\in V\times W$ can be writted as

   $$(\boldsymbol{v},\boldsymbol{w}) = (\boldsymbol{v},\mathbf{0}_W)+(\mathbf{0}_V,\boldsymbol{w}),$$

   and $(\boldsymbol{v},\mathbf{0}_W),(\mathbf{0}_V,\boldsymbol{w})\in V'\cup W'$, $V'\cup W'$ generates $V\times W$.

4. Suppose $\boldsymbol{v}\neq\mathbf{0}_V$, $\boldsymbol{w}\neq\mathbf{0}_W$ and

   $$a(\boldsymbol{v},\mathbf{0}_W)+b(\mathbf{0}_V,\boldsymbol{w}) = (\mathbf{0}_V,\mathbf{0}_W).$$

   Then, $(a\boldsymbol{v},b\boldsymbol{w})=(\mathbf{0}_V,\mathbf{0}_W)$ and we have $a=b=0$. Thus, $\{V',W'\}$ is independent.

Chapter 4

Matrices

Exercise 4.1	Sect. 4.1.	p.71	Handling matrices in NumPy*
Exercise 4.2	Sect. 4.1.	p.73	Handling matrices in SymPy and LaTeX-outputs*
Exercise 4.3	Sect. 4.2.	p.75	Matrix representation and a basis
Exercise 4.4	Sect. 4.2.	p.75	Scattering of matrix ranges*
Exercise 4.5	Sect. 4.2.	p.76	Matrix representation of differantiation
Exercise 4.6	Sect. 4.2.	p.78	Matrix representation of mappings on 2- or 3-d spaces
Exercise 4.7	Sect. 4.3.	p.79	Computing matrix products*
Exercise 4.8	Sect. 4.3.	p.86	Generation of matrix product problems*
Exercise 4.9	Sect. 4.4.	p.86	Check of regularity of matrices
Exercise 4.10	Sect. 4.4.	p.88	$\boldsymbol{BA} = \boldsymbol{I}$ for non-square matrices*
Exercise 4.11	Sect. 4.4.	p.89	The power low of square matrices
Exercise 4.12	Sect. 4.5.	p.90	Properties of adjoint matrices
Exercise 4.13	Sect. 4.6.	p.93	Culculation time of matrix products*

* Using Python

Exercise 4.1. Observe what happens, if we replace `A.tolist` in `In[4]` with `list(A)` in the above dialog. Moreover, observe how the results change if we replace `B = A.copy()` in `In[5]` with `B = A`.

In [2]:
```
A = array(A); A
```

Out[2]:
```
array([[1, 2, 3],
       [4, 5, 6]])
```

In [3]:
```
print(A)
```

```
[[1 2 3]
 [4 5 6]]
```

In [4]:
```
L = list(A); L
```

Out[4]:
```
[array([1, 2, 3]), array([4, 5, 6])]
```

If an array `A` is converted to a list `L` by `list(A)`, `L` is a list, but the elements of `L` are arrays.

In [5]:
```
B = A; B
```

If you change `B = A.copy()` to `B = A`, `A` and `B` point to the same object in memory.

Out[5]:
```
array([[1, 2, 3],
       [4, 5, 6]])
```

In [6]:
```
A == B
```

Out[6]:
```
array([[ True,  True,  True],
       [ True,  True,  True]])
```

In [7]:
```
(A == B).all()
```

Out[7]:
```
True
```

In [8]:
```
B[0, 1] = 1
```

If you change an element of `B`, the same element of `A` also changes.

In [9]:
```
(A == B).all()
```

Out[9]:
```
True
```

In [10]:
```
A
```

```
Out[10]: array([[1, 1, 3],
                [4, 5, 6]])
```

```
In [11]: B
```

```
Out[11]: array([[1, 1, 3],
                [4, 5, 6]])
```

It's not just that `A` and `B` have the same content, but that `A` and `B` look at the same thing.

Exercise 4.2. The following program outputs a calculation problem of matrices in a form of a code in math mode of LaTeX. Typesetting the obtained tex code, we will have Figure4.1 Get another problem by changing the seed `2021` and solve it.

By changing the seed of the random number in Line 4, a different problem is generated. By changing Line 9 of the program as follows, the LaTeX formula with the answer will be displayed.

Program: `latex1.py`

In [1]:

```
1 from numpy.random import seed, randint, choice
2 from sympy import Matrix, latex
3 
4 seed(2021)
5 m, n = randint(2, 4, 2)
6 X = [-3, -2, -1, 1, 2, 3, 4, 5]
7 A = Matrix(choice(X, (m, n)))
8 B = Matrix(choice(X, (m, n)))
9 print(f'{latex(A)} + {latex(B)} = {latex(A+B)}')
```

```
\left[\begin{matrix}-2 & -3 & 3\\4 & 4 & 2\end{matrix}\right] + \left[\
    begin{matrix}5 & 4 & 1\\3 & 3 & 3\end{matrix}\right] = \left[\begin{
    matrix}3 & 1 & 4\\7 & 7 & 5\end{matrix}\right]
```

Type setting this expressin of LaTeX we have

$$\left[\begin{matrix}-2 & -3 & 3\\4 & 4 & 2\end{matrix}\right] + \left[\begin{matrix}5 & 4 & 1\\3 & 3 & 3\end{matrix}\right] = \left[\begin{matrix}3 & 1 & 4\\7 & 7 & 5\end{matrix}\right].$$

Exercise 4.3. Let $V = \mathbb{K}^n$ and $W = \mathbb{K}^m$, and let $\boldsymbol{f} : V \to W$ be a linear mapping and $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{bmatrix}$ be the representation matrix of $\boldsymbol{f}$ on the standard bases. Prove the following statements.

1. $\boldsymbol{f}$ is surjective if and only if $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}$ generates W.
2. $\boldsymbol{f}$ is injective if and only if $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}$ is linearly independent.
3. $\boldsymbol{f}$ is bijective if and only if $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}$ is a basis of W.

Recall that for any $\boldsymbol{v} = (v_1, v_2, \ldots, v_n) \in V$

$$\boldsymbol{f}(\boldsymbol{v}) \;=\; \boldsymbol{A}\boldsymbol{v} \;=\; v_1\boldsymbol{a}_1 + v_2\boldsymbol{a}_2 + \cdots + v_n\boldsymbol{a}_n.$$

1.

$$\begin{aligned}
& \boldsymbol{f} \text{ is surjective} \\
\Leftrightarrow\;& \operatorname{range}(\boldsymbol{f}) \\
\Leftrightarrow\;& \left\{\boldsymbol{f}(\boldsymbol{v}) \;\middle|\; \boldsymbol{v} \in V\right\} = W \\
\Leftrightarrow\;& \left\{\boldsymbol{A}\boldsymbol{v} \;\middle|\; \boldsymbol{v} \in V\right\} = W \\
\Leftrightarrow\;& \left\{v_1\boldsymbol{a}_1 + v_2\boldsymbol{a}_2 + \cdots + v_n\boldsymbol{a}_n \;\middle|\; v_1, v_2, \ldots, v_n \in \mathbb{K}\right\} = W \\
\Leftrightarrow\;& \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\} \text{ is linearly independent.}
\end{aligned}$$

2.

$$\begin{aligned}
& \boldsymbol{f} \text{ is injective} \\
\Leftrightarrow\;& \operatorname{kernel}(\boldsymbol{f}) = \{\boldsymbol{0}_V\} \\
\Leftrightarrow\;& \left\{\boldsymbol{v} \;\middle|\; \boldsymbol{f}(\boldsymbol{v}) = \boldsymbol{0}_W\right\} = \{\boldsymbol{0}_V\} \\
\Leftrightarrow\;& \left\{\boldsymbol{v} \;\middle|\; \boldsymbol{A}\boldsymbol{v} = \boldsymbol{0}_W\right\} = \{\boldsymbol{0}_V\} \\
\Leftrightarrow\;& \left\{(v_1, v_2, \ldots, v_n) \;\middle|\; v_1\boldsymbol{a}_1 + v_2\boldsymbol{a}_2 + \cdots + v_n\boldsymbol{a}_n = \boldsymbol{0}_W\right\} = \left\{(0, 0, \ldots, 0)\right\} \\
\Leftrightarrow\;& v_1\boldsymbol{a}_1 + v_2\boldsymbol{a}_2 + \cdots + v_n\boldsymbol{a}_n = \boldsymbol{0}_W \text{ implies } v_1 = v_2 = \cdots = v_n = 0 \\
\Leftrightarrow\;& \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\} \text{ is linearly independent.}
\end{aligned}$$

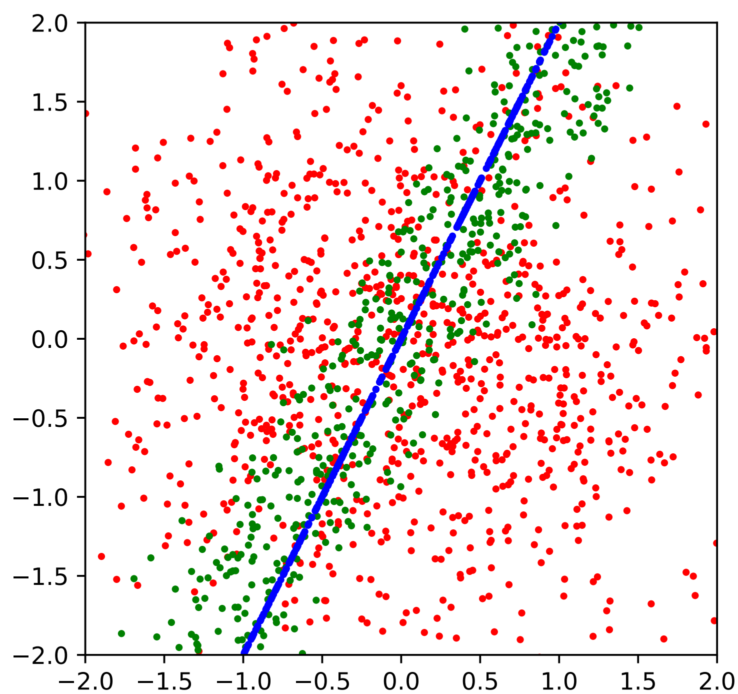
3. It is the consequence of 1 and 2.

Exercise 4.4. Generate 1000 vectors subject to the 2-dimensional standard normal distribution and multiply them by the following matrix (1) or (2), and display the vectors as points on a plane. Do a similar experiment using matrices (3) or (4) for the 3-dimensional standard normal distribution.

$$(1)\ \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \qquad (2)\ \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \qquad (3)\ \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix} \qquad (4)\ \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$$

Program: `prob4_3_2d.py`

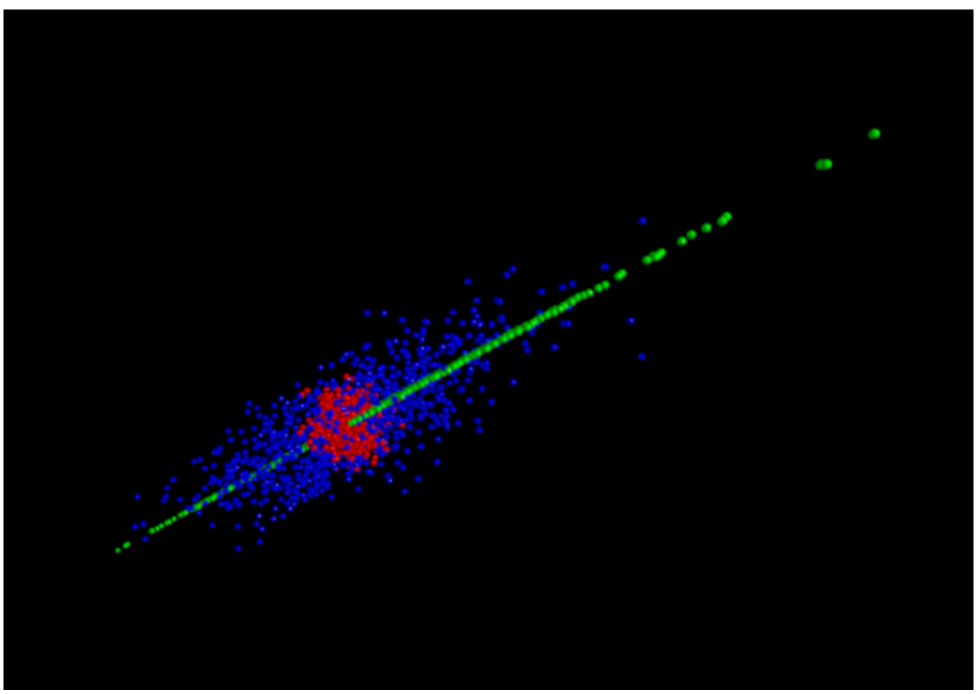
In [1]:
```
import matplotlib.pyplot as plt
from numpy import array, random

A = array([[1, 2], [2, 3]])
B = array([[1, 2], [2, 4]])
P = random.normal(0, 1, (1000, 2))
Q = array([A.dot(p) for p in P])
R = array([B.dot(p) for p in P])
plt.axis("scaled"), plt.xlim(-2, 2), plt.ylim(-2, 2)
plt.scatter(P[:, 0], P[:, 1], s=4, color='r')
plt.scatter(Q[:, 0], Q[:, 1], s=4, color='g')
plt.scatter(R[:, 0], R[:, 1], s=4, color='b')
plt.show()
```



The red dots represent 1000 vectors sampled from a 2-dimensional standard normal distribution. The green dots are the images of these 1000 red dots under the linear transformation represented by matrix (1). The distribution of these dots shows a spread with an area. On the other hand, the blue dots are the images of the 1000 red dots under the linear transformation represented by matrix (2), and these dots are distributed along a single line. If the number of samples is increased or the variance of the normal distribution is increased, both the red and green dots will spread over a wider range. However, the blue dots will never spread outside the line.

Program: `prob4_3_3d.py`

In [1]:
```
from vpython import *
from numpy import array, random

A = array([[1, 2, 3], [2, 3, 4], [3, 4, 5]])
B = array([[1, 2, 3], [2, 3, 1], [3, 1, 2]])
X = B
P = random.normal(0, 1, (1000, 3))
Q = array([A.dot(p) for p in P])
R = array([B.dot(p) for p in P])
points(pos=[vector(*p) for p in P], radius=3, color=color.red)
points(pos=[vector(*q) for q in Q], radius=3, color=color.green)
points(pos=[vector(*r) for r in R], radius=3, color=color.blue)
```



The red dots represent 1000 vectors sampled from a 3-dimensional standard normal distribution. The green dots are the images of these 1000 red dots under the linear transformation represented by matrix (3). These dots are distributed along a single plane. On the other hand, the blue dots are the images of the 1000 red dots under the linear transformation represented by matrix (2), and the distribution of these dots shows a spread with a volume. If the number of samples is increased or the variance of the normal distribution is increased, both the red and blue dots will spread over a wider range. in the space. However, the green dots will never spread outside the plane.

Exercise 4.5. Let V be the linear space consisting of all polynomials in x of degree at most 4 and W be the linear space consisting of all polynomials of degree at most 2. Find the matrix representation of the second order differentiation $\dfrac{d^2}{dx^2}$, which is a linear mapping from V to W, on the bases $\{1, x, x^2, x^3, x^4\}$ and $\{1, x, x^2\}$ of V and W respectively.

Let $\{x^5, x^4, x^3, x^2, x, 1\}$ be a basis of V and $\{x^3, x^2, x, 1\}$ a basis of W. Since the linear mapping $\dfrac{d^2}{dx^2}$ transform each element of the basis of V such as

$$\begin{array}{rcl} x^5 & \mapsto & 20x^3 + 0x^2 + 0x + 0 \\ x^4 & \mapsto & 0x^3 + 12x^2 + 0x + 0 \\ x^3 & \mapsto & 0x^3 + 0x^2 + 6x + 0 \\ x^2 & \mapsto & 0x^3 + 0x^2 + 0x + 2 \\ x & \mapsto & 0x^3 + 0x^2 + 0x + 0 \\ 1 & \mapsto & 0x^3 + 0x^2 + 0x + 0, \end{array}$$

the matrix representation of this linear mappimg is

$$\begin{bmatrix} 20 & 0 & 0 & 0 & 0 & 0 \\ 0 & 12 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \end{bmatrix}.$$

Indeed, the formula

$$\frac{d^2}{dx^2}\left(a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0\right) = 20a_5x^3 + 12a_4x^2 + 6a_3x + 2a_2$$

is represented by

$$\begin{bmatrix} 20 & 0 & 0 & 0 & 0 & 0 \\ 0 & 12 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_5 \\ a_4 \\ a_3 \\ a_2 \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} 20a_5 \\ 12a_4 \\ 6a_3 \\ 2a_2 \end{bmatrix}.$$

Here, the polynomials are expressed in order of descending powers, but if they are of ascending powers, the matrix representation will change as follows. From

$$\begin{aligned} 1 &\mapsto 0 + 0x + 0x^2 + 0x^3 \\ x &\mapsto 0 + 0x + 0x^2 + 0x^3 \\ x^2 &\mapsto 2 + 0x + 0x^2 + 0x^3 \\ x^3 &\mapsto 0 + 6x + 0x^2 + 0x^3 \\ x^4 &\mapsto 0 + 0x + 12x^2 + 0x^3 \\ x^5 &\mapsto 0 + 0x + 0x^2 + 20x^3, \end{aligned}$$

the representation matrix is

$$\begin{bmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 0 & 0 & 20 \end{bmatrix}.$$

Then,

$$\frac{d^2}{dx^2}\left(a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + a_5 x^5\right) = 2a_2 + 6a_3 x + 12a_4 x^2 + 20a_5 x^3$$

is represented by

$$\begin{bmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 0 & 0 & 20 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} = \begin{bmatrix} 2a_2 \\ 6a_3 \\ 12a_4 \\ 20a_5 \end{bmatrix}.$$

Exercise 4.6. Find the representation matrix of each of the following linear mappings on the standard bases.

1. Linear mapping $(x, y) \mapsto (x + y, x, y)$ from $\mathbb{K}^2$ to $\mathbb{K}^3$,
2. Linear mapping $(x, y, z) \mapsto (x + y, y + z, x + z)$ from $\mathbb{K}^3$ to $\mathbb{K}^3$.
3. Linear mapping $(x, y, z) \mapsto (x + y, y + z)$ from $\mathbb{K}^3$ to $\mathbb{K}^2$.

1. From
$$\begin{aligned} \begin{bmatrix}1\\0\end{bmatrix} &\mapsto \begin{bmatrix}0\\1\end{bmatrix} = 0\begin{bmatrix}1\\0\end{bmatrix} + 1\begin{bmatrix}0\\1\end{bmatrix} \\ \begin{bmatrix}0\\1\end{bmatrix} &\mapsto \begin{bmatrix}1\\0\end{bmatrix} = 1\begin{bmatrix}1\\0\end{bmatrix} + 0\begin{bmatrix}0\\1\end{bmatrix} \end{aligned}$$
we have the matrix representation $\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$.

2. From
$$\begin{aligned} \begin{bmatrix}1\\0\end{bmatrix} &\mapsto \begin{bmatrix}1\\1\\0\end{bmatrix} = 1\begin{bmatrix}1\\0\\0\end{bmatrix} + 1\begin{bmatrix}0\\1\\0\end{bmatrix} + 0\begin{bmatrix}0\\0\\1\end{bmatrix} \\ \begin{bmatrix}0\\1\end{bmatrix} &\mapsto \begin{bmatrix}1\\0\\1\end{bmatrix} = 1\begin{bmatrix}1\\0\\0\end{bmatrix} + 0\begin{bmatrix}0\\1\\0\end{bmatrix} + 1\begin{bmatrix}0\\0\\1\end{bmatrix} \end{aligned}$$
we have the matrix representation $\begin{bmatrix}1 & 1\\1 & 0\\0 & 1\end{bmatrix}$.

3. From
$$\begin{aligned} \begin{bmatrix}1\\0\\0\end{bmatrix} &\mapsto \begin{bmatrix}1\\0\end{bmatrix} = 1\begin{bmatrix}1\\0\end{bmatrix} + 0\begin{bmatrix}0\\1\end{bmatrix} \\ \begin{bmatrix}0\\1\\0\end{bmatrix} &\mapsto \begin{bmatrix}1\\1\end{bmatrix} = 1\begin{bmatrix}1\\0\end{bmatrix} + 1\begin{bmatrix}0\\1\end{bmatrix} \\ \begin{bmatrix}0\\0\\1\end{bmatrix} &\mapsto \begin{bmatrix}0\\1\end{bmatrix} = 0\begin{bmatrix}1\\0\end{bmatrix} + 1\begin{bmatrix}0\\1\end{bmatrix} \end{aligned}$$
we have the matrix representation $\begin{bmatrix}1 & 1 & 0\\0 & 1 & 1\end{bmatrix}$.

Exercise 4.7. Select two, allowing duplicates, from the next matrices, and compute the product of them if it is definable.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \quad \boldsymbol{D} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

The following program is a slightly modified version of the program `problems.py` on page 110 of this book. SymPy is used instead of NumPy for matrix calculations. If an error occurs when the matrix product cannot be defined, we will continue the program execution by exception handling.

Program: `problems2.py`

In [1]:
```
1  from sympy import *
2
3  A = Matrix([[1, 2], [3, 4]])
4  B = Matrix([[1, 2, 3], [4, 5, 6]])
5  C = Matrix([[1, 2], [3, 4], [5, 6]])
6  D = Matrix([[1, 2, 3], [4, 5, 6], [7, 8, 9]])
7
8  print(r'\begin{eqnarray*}')
9  for X in ('A', 'B', 'C', 'D'):
10     for Y in ('A', 'B', 'C', 'D'):
11         try:
12             U, V = eval(X), eval(Y)
13             XY = r'\boldsymbol{%s%s}' % (X,Y)
14             print(fr'{XY}&=&{latex(U)}{latex(V)}\ =\ {latex(U*V)}\\')
15         except ShapeError:
16             continue
17 print(r'\end{eqnarray*}')
```

Lines 11 .–.16: Line 14 raises an error named `ShapeError` if the matrix product of `U*V` is not defined. If this error occurs, execute the exception handling on line 16 and continue the `for` loop on lines 11-16. If no error occurs, execute the `print` function on line 14.

This program will output the product of two combinations that allow duplication of the given matrices and the solution in LaTeX format. Typesetting it gives the following result.

$$\boldsymbol{AA} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix}$$

$$\boldsymbol{AB} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \end{bmatrix}$$

$$\boldsymbol{BC} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 22 & 28 \\ 49 & 64 \end{bmatrix}$$

$$\boldsymbol{BD} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 30 & 36 & 42 \\ 66 & 81 & 96 \end{bmatrix}$$

$$\boldsymbol{CA} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 7 & 10 \\ 15 & 22 \\ 23 & 34 \end{bmatrix}$$

$$\boldsymbol{CB} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \\ 29 & 40 & 51 \end{bmatrix}$$

$$\boldsymbol{DC} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 22 & 28 \\ 49 & 64 \\ 76 & 100 \end{bmatrix}$$

$$\boldsymbol{DD} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 30 & 36 & 42 \\ 66 & 81 & 96 \\ 102 & 126 & 150 \end{bmatrix}$$

Exercise 4.8. Typeset the LaTeX code output by Program `latex2.py` to obtain 5 calculation problems (Figure 4.4). Solve the problems.

By modifying the program as follows, the correct answer will be output at the same time, when the value of the variable `ans` is `True`.

Program: `latex2.py`

In [1]:
```
from numpy.random import seed, choice
from sympy import Matrix, latex

seed(2021)
ans = False
template = r'''
\begin{enumerate}
\item $%s%s = %s$
\item $%s%s = %s$
\item $%s%s = %s$
\item $%s%s = %s$
\item $%s%s = %s$
\end{enumerate}
'''

matrices= ()
for no in range(5):
    m, el, n = choice([2, 3], 3)
    X = [-3, -2, -1, 1, 2, 3, 4, 5]
    A = Matrix(choice(X, (m, el)))
    B = Matrix(choice(X, (el, n)))
    matrices += (latex(A), latex(B), latex(A*B) if ans else '')
print(template % matrices)
```

1. $\begin{bmatrix}-3 & 3 & 4\\ 4 & 2 & 5\end{bmatrix}\begin{bmatrix}4 & 1 & 3\\ 3 & 3 & -3\\ 2 & 4 & 4\end{bmatrix} = \begin{bmatrix}5 & 22 & -2\\ 32 & 30 & 26\end{bmatrix}$

2. $\begin{bmatrix}3 & 5\\ -2 & 4\end{bmatrix}\begin{bmatrix}-2 & 2 & 3\\ -1 & -1 & -3\end{bmatrix} = \begin{bmatrix}-11 & 1 & -6\\ 0 & -8 & -18\end{bmatrix}$

3. $\begin{bmatrix}-3 & -1 & 4\\ 1 & 2 & 3\\ -3 & 3 & -2\end{bmatrix}\begin{bmatrix}4 & -1 & 4\\ 5 & 3 & 4\\ -2 & 3 & 4\end{bmatrix} = \begin{bmatrix}-25 & 12 & 0\\ 8 & 14 & 24\\ 7 & 6 & -8\end{bmatrix}$

4. $\begin{bmatrix}2 & 3\\ -1 & 1\\ -2 & -1\end{bmatrix}\begin{bmatrix}-1 & 5\\ -3 & -1\end{bmatrix} = \begin{bmatrix}-11 & 7\\ -2 & -6\\ 5 & -9\end{bmatrix}$

5. $\begin{bmatrix} 1 & -2 & 4 \\ -2 & -2 & -1 \end{bmatrix} \begin{bmatrix} 5 & 3 & 1 \\ 1 & 2 & 5 \\ 5 & 1 & -2 \end{bmatrix} = \begin{bmatrix} 23 & 3 & -17 \\ -17 & -11 & -10 \end{bmatrix}$

Exercise 4.9. For the following matrices, see if they are regular matrices by checking the linear independence of their column vectors.

$$(1)\ \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \qquad (2)\ \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \qquad (3)\ \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix} \qquad (4)\ \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$$

It is possible to construct simultaneous equations according to the definition and check whether the solutions are unique or not, but it is relatively easy to see whether or not they are linearly independent by the following considerations.

(1) Since the scalar multiple of $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is never $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$, the column vectors are linearly independent and the matrix is a regular matrix.

(2) Since $2\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, the column vectors are linearly dependent and this matrix is not a regular matrix.

(3) Since $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and then $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$, the column vectors are linearly dependent and the matrix is not a regular matrix.

(4) The first and second columns are linearly independent. Check if the 3rd column can be represented by a linear combination of the 1st and 2nd columns. Putting

$$\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = a \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + b \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix},$$

we get

$$\begin{cases} a + 2b = 3 & \cdots\ (1) \\ 2a + 3b = 1 & \cdots\ (2) \\ 3a + b = 2 & \cdots\ (3). \end{cases}$$

$(1) \times 2 - (2)$ derives $b = 5$, and $(1) \times 3 - (3)$ does $b = \frac{7}{5}$, so the simaltaneous equations has no solution. Therefore, the three column vectors are linearly independent and the matrix is regular.

Exercise 4.10. Is it possible to make $\boldsymbol{BA} = \boldsymbol{I}$ hold for the same $\boldsymbol{A}$ and $\boldsymbol{B}$ as above ?

In [1]:
```
from sympy import Matrix, solve, eye
from sympy.abc import a, b, c, d, e, f

A = Matrix([[1, 2, 3], [2, 3, 4]])
B = Matrix([[a, b], [c, d], [e, f]])
ans = solve(A*B - eye(2), [a, b, c, d, e, f])
print(ans)
```

```
{a: e - 3, c: 2 - 2*e, b: f + 2, d: -2*f - 1}
```

Solve the equation $\boldsymbol{BA} = \boldsymbol{I}$ for a, b, c, d, e, f.

In [2]:
```
solve(B * A - eye(3), [a, b, c, d, e, f])
```

Out[2]:
```
[]
```

`[]` means no solution. Try to actually calculate $\boldsymbol{BA}$.

In [3]:
```
B * A
```

Out[3]:
```
Matrix([
[a + 2*b, 2*a + 3*b, 3*a + 4*b],
[c + 2*d, 2*c + 3*d, 3*c + 4*d],
[e + 2*f, 2*e + 3*f, 3*e + 4*f]])
```

For $\boldsymbol{BA} = \boldsymbol{I}$,

$$\begin{bmatrix} a+2b & 2a+3b & 3a+4b \\ c+2d & 2c+3d & 3c+4d \\ e+2f & 2e+3f & 3e+4f \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

must hold. There do not exist a and b which simultaneously satisfy $a + 2b = 1$, $2a + 3b = 0$ and $3a + 4b = 0$ by comparing the first row of the matrices on both sides. Therefore, $\boldsymbol{BA}$ never becomes the identity matrix.

Exercise 4.11. For regular matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ prove the following.

(1) $(\boldsymbol{A}^p)^{-1} = \left(\boldsymbol{A}^{-1}\right)^p$ holds for a nonnegative integer p.

(2) Define $\boldsymbol{A}^{-p} \stackrel{\text{def}}{=} (\boldsymbol{A}^p)^{-1}$ for positive integer p. Then, the *exponent law* $\boldsymbol{A}^p\boldsymbol{A}^q = \boldsymbol{A}^{p+q}$ holds for any integers p and q.

(3) If $\boldsymbol{A}$ and $\boldsymbol{B}$ are similar, so are $\boldsymbol{A}^p$ and $\boldsymbol{B}^p$ for any integer p.

(1) Since

$$\begin{aligned}
\left(\boldsymbol{A}^{-1}\right)^p \boldsymbol{A}^p &= \overbrace{\boldsymbol{A}^{-1}\boldsymbol{A}^{-1}\cdots\boldsymbol{A}^{-1}}^{p}\overbrace{\boldsymbol{A}\boldsymbol{A}\cdots\boldsymbol{A}}^{p} \\
&= \overbrace{\boldsymbol{A}^{-1}\boldsymbol{A}^{-1}\cdots\boldsymbol{A}^{-1}}^{p-1}\boldsymbol{A}^{-1}\boldsymbol{A}\overbrace{\boldsymbol{A}\boldsymbol{A}\cdots\boldsymbol{A}}^{p-1} \\
&= \overbrace{\boldsymbol{A}^{-1}\boldsymbol{A}^{-1}\cdots\boldsymbol{A}^{-1}}^{p-1}\overbrace{\boldsymbol{A}\boldsymbol{A}\cdots\boldsymbol{A}}^{p-1} \\
&\vdots \\
&= \boldsymbol{A}^{-1}\boldsymbol{A} \\
&= \boldsymbol{I},
\end{aligned}$$

by the uniqueness of the matrix inverse we have $(\boldsymbol{A}^p)^{-1} = \left(\boldsymbol{A}^{-1}\right)^p$.

The proof by mathematical induction is as follows. When $p = 0$, $\boldsymbol{I}^{-1} = \boldsymbol{I}$, so the equality holds. Suppose that the equality holds when $p = k$. Then

$$\left(\boldsymbol{A}^{k+1}\right)^{-1} = \left(\boldsymbol{A}\boldsymbol{A}^k\right)^{-1} = \left(\boldsymbol{A}^k\right)^{-1}\boldsymbol{A}^{-1} = \left(\boldsymbol{A}^{-1}\right)^k\boldsymbol{A}^{-1} = \left(\boldsymbol{A}^{-1}\right)^{k+1}.$$

Thus the equality hold when $p = k + 1$.

(2) When both p and q are non-negative integers, $\boldsymbol{A}^p\boldsymbol{A}^q = \boldsymbol{A}^{p+q}$ hold. Then

$$\boldsymbol{A}^{-p}\boldsymbol{A}^{-q} = (\boldsymbol{A}^p)^{-1}(\boldsymbol{A}^q)^{-1} = (\boldsymbol{A}^q\boldsymbol{A}^p)^{-1} = \left(\boldsymbol{A}^{q+p}\right)^{-1} = \boldsymbol{A}^{-(q+p)} = \boldsymbol{A}^{(-p)+(-q)}$$

holds. When $p \geqq q$,

$$\boldsymbol{A}^p\boldsymbol{A}^{-q} = \boldsymbol{A}^{p-q}\boldsymbol{A}^q\left(\boldsymbol{A}^{-1}\right)^q = \boldsymbol{A}^{p-q} = \boldsymbol{A}^{p+(-q)}.$$

When $p \leqq q$,

$$\boldsymbol{A}^p\boldsymbol{A}^{-q} = \boldsymbol{A}^p\left(\boldsymbol{A}^{-1}\right)^q = \boldsymbol{A}^p\left(\boldsymbol{A}^{-1}\right)^p\left(\boldsymbol{A}^{-1}\right)^{q-p} = \left(\boldsymbol{A}^{-1}\right)^{q-p} = \boldsymbol{A}^{-(q-p)} = \boldsymbol{A}^{p+(-q)}.$$

When $p \geqq q$,

$$\boldsymbol{A}^{-p}\boldsymbol{A}^q = \left(\boldsymbol{A}^{-1}\right)^p\boldsymbol{A}^q = \left(\boldsymbol{A}^{-1}\right)^{p-q}\left(\boldsymbol{A}^{-1}\right)^q\boldsymbol{A}^q = \left(\boldsymbol{A}^{-1}\right)^{p-q} = \boldsymbol{A}^{-(p-q)} = \boldsymbol{A}^{-p+q}$$

When $p \leqq q$,

$$\boldsymbol{A}^{-p}\boldsymbol{A}^q = \left(\boldsymbol{A}^{-1}\right)^p\boldsymbol{A}^p\boldsymbol{A}^{-p+q} = \boldsymbol{A}^{-p+q}.$$

From the above, the exponential law holds for any combination of signs of p and q.

(3) If $\boldsymbol{A}$ and $\boldsymbol{B}$ are similar, there exists a regular matrix $\boldsymbol{V}$ such that $\boldsymbol{A} = \boldsymbol{V}^{-1}\boldsymbol{B}\boldsymbol{V}$. Then, since

$$\boldsymbol{A}^{-1} = \left(\boldsymbol{V}^{-1}\boldsymbol{B}\boldsymbol{V}\right)^{-1} = \boldsymbol{V}^{-1}\boldsymbol{B}^{-1}\left(\boldsymbol{V}^{-1}\right)^{-1} = \boldsymbol{V}^{-1}\boldsymbol{B}^{-1}\boldsymbol{V},$$

$\boldsymbol{A}^{-1}$ and $\boldsymbol{B}^{-1}$ are similar too. When $p = 0, 1$, $\boldsymbol{A}^p$ and $\boldsymbol{B}^p$ are clearly similar. Let p be an integer greater than 1. Then, since

$$\boldsymbol{A}^p = \left(\boldsymbol{V}^{-1}\boldsymbol{B}\boldsymbol{V}\right)^p = \overbrace{\boldsymbol{V}^{-1}\boldsymbol{B}\boldsymbol{V}\cdot\boldsymbol{V}^{-1}\boldsymbol{B}\boldsymbol{V}\cdot\cdots\cdot\boldsymbol{V}^{-1}\boldsymbol{B}\boldsymbol{V}}^{p} = \boldsymbol{V}^{-1}\boldsymbol{B}^p\boldsymbol{V},$$

$\boldsymbol{A}^p$ and $\boldsymbol{B}^p$ are similar. Thus, $\boldsymbol{A}^{-p} = (\boldsymbol{A}^p)^{-1}$ and $\boldsymbol{B}^{-p} = (\boldsymbol{B}^p)^{-1}$ are also similar.

Exercise 4.12. Prove equalities 1 − 3 above.

Here, we give a proof of the equality $(\boldsymbol{AB})^* = \boldsymbol{B}^*\boldsymbol{A}^*$. Let

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{l1} & a_{l2} & \cdots & a_{lm} \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}.$$

Then the (i, j)-th component of the product matrix $\boldsymbol{AB}$ is $\sum_{k=1}^{m} a_{ik}b_{kj}$. The (i, j)-th component of $(\boldsymbol{AB})^*$ is $\sum_{k=1}^{m} \overline{a_{jk}}\overline{b_{ki}}$. On the other hand, since

$$\boldsymbol{B}^* = \begin{bmatrix} \overline{b_{11}} & \overline{b_{21}} & \cdots & \overline{b_{m1}} \\ \overline{b_{12}} & \overline{b_{22}} & \cdots & \overline{b_{m2}} \\ \vdots & \vdots & & \vdots \\ \overline{b_{1n}} & \overline{b_{2n}} & \cdots & \overline{b_{mn}} \end{bmatrix}, \quad \boldsymbol{A}^* = \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \cdots & \overline{a_{l1}} \\ \overline{a_{21}} & \overline{a_{22}} & \cdots & \overline{a_{l2}} \\ \vdots & \vdots & & \vdots \\ \overline{a_{1m}} & \overline{a_{2m}} & \cdots & \overline{a_{lm}} \end{bmatrix},$$

the (i, j)-th component of the product matrix $\boldsymbol{B}^*\boldsymbol{A}^*$ is $\sum_{k=1}^{m} \overline{b_{ki}}\overline{a_{jk}}$. Hence we have the equality. The remaining equalities of 1 and 2 are left to the reader. Next, we prove the socond equality of 3. If $\boldsymbol{A}$ is a regular matrix, by 3 we have

$$\boldsymbol{I} = \boldsymbol{I}^* = \left(\boldsymbol{A}^{-1}\boldsymbol{A}\right)^* = \boldsymbol{A}^*\left(\boldsymbol{A}^{-1}\right)^*$$

and also

$$\boldsymbol{I} = \boldsymbol{I}^* = \left(\boldsymbol{A}\boldsymbol{A}^{-1}\right)^* = \left(\boldsymbol{A}^{-1}\right)^*\boldsymbol{A}^*.$$

By the uniqueness of the inverse matrix, the inverse of $\boldsymbol{A}^*$ is $\left(\boldsymbol{A}^{-1}\right)^*$, so we have the second equality in 3. The first one is similarly seen.

Exercise 4.13. Draw the graphs of $\dfrac{g(n)}{f(n)}$ and $\dfrac{h(n)}{g(n)}$, and observe them. It is expected that the former is a linear function and the latter is a constant function. Is that true? If necessary, try to calculate up to $n = 3000$.

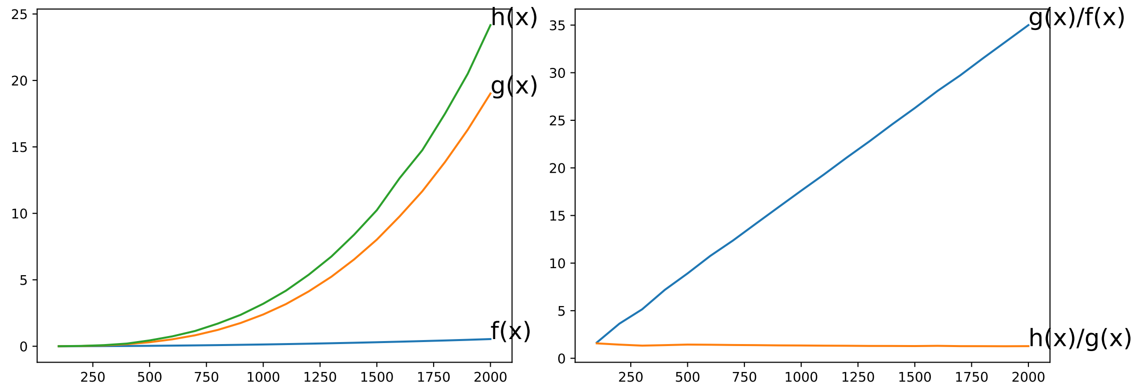
Program: `prob4_11.py`

In [1]:

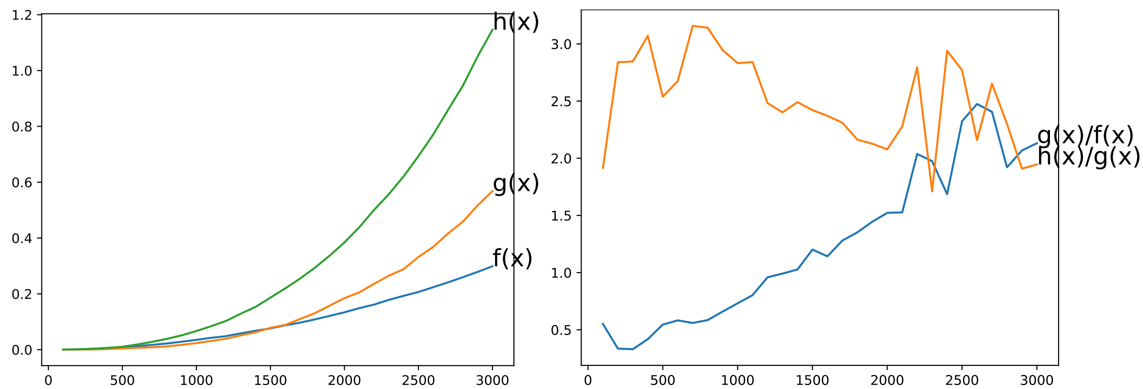
```
1  from numpy.random import normal
2  from numpy.linalg import inv
3  import matplotlib.pyplot as plt
4  from time import time
5
6  N = range(100, 2100, 100)
7  T = [[], [], []]
8
9  for n in N:
10     t0 = time()
11     A = normal(0, 1, (n, n))
12     t1 = time()
13     A.dot(A)
14     t2 = time()
15     inv(A)
16     t3 = time()
17     print(n, end=', ')
18     t = (t0, t1, t2, t3)
19     for i in range(3):
20         T[i].append(t[i + 1] - t[i])
21
22 T1 = [t1/t0 for t0, t1 in zip(T[0],T[1])]
23 T2 = [t2/t1 for t1, t2 in zip(T[1],T[2])]
24 plt.plot(N, T1)
25 plt.plot(N, T2)
26 plt.text(N[-1], T1[-1], 'g(x)/f(x)', fontsize=18)
27 plt.text(N[-1], T2[-1], 'h(x)/g(x)', fontsize=18)
28
29 plt.show()
```

This program is rewritten from line 22 onwards of `mat_product5.py` on pages 124 and 125 of this book.

Note that the function `time` defined in the library `time` on line 4 (used in lines 10, 12, 14, and 16) measures time with a precision higher than 1 second depending on the system. It's not always possible, so fractions on lines 22 and 23, for example, can cause divide-by-zero errors. In such cases, try the function `perf_counter`, which is defined in the same library `time`, instead of the function `time`.



The left is the experimental result of `mat_product5.py`, and the right is the experimental result of the above program. Both are done with Raspberry Pi 4 (1.5GHz, Cortex-A72)



The left is the experimental result of `mat_product5.py`, and the right is the experimental result of the above program (using `perf_counter`). Both were measured in increments of 100 up to $n = 3000$ with iMac4 (3.2GHz intel Core i5). Mac and Windows have faster processing speed than Raspberry Pi, but the OS also performs other processing in the background, so the graph on the left shows its effect.

Chapter 5

Elementary Operations and Matrix Invariants

Exercise 5.1	Sect. 5.1.	p.96	Inverse of ementary matrices
Exercise 5.2	Sect. 5.1.	p.96	Elementary matrices and operations*
Exercise 5.3	Sect. 5.2.	p.102	Properties of the rank of matrices
Exercise 5.4	Sect. 5.2.	p.102	Generating problems for ranks*
Exercise 5.5	Sect. 5.3.	p.104	Computing signatures of permutations*
Exercise 5.6	Sect. 5.3.	pp.104,105	Determnants of 2- or 3×3-matrix*
Exercise 5.7	Sect. 5.3.	p.111	Regular matrices in randomly genarated matrices*
Exercise 5.8	Sect. 5.3.	p.111	Generating problems for detarminants*
Exercise 5.9	Sect. 5.4.	p.113	Properties of the trace
Exercise 5.10	Sect. 5.5.	p.117	Generating problems of linear systems*
Exercise 5.11	Sect. 5.6.	p.120	Generating problems of inverse matices
Exercise 5.12	Sect. 5.6.	p.120	Culculation time of inverse matrices*
Exercise 5.13	Sect. 5.6.	p.120	Cofactor matrices
Exercise 5.14	Sect. 5.6.	p.121	Cramer's formula*

* Using Python

Exercise 5.1. Confirm the above results by calculating the products $\boldsymbol{E}_1^{(i,j,c)}\boldsymbol{E}_1^{(i,j,-c)}$, $\boldsymbol{E}_2^{(i,j)}\boldsymbol{E}_2^{(i,j)}$ and $\boldsymbol{E}_3^{(i,c)}\boldsymbol{E}_3^{(i,1/c)}$.

$$\begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & c & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & -c & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} = \boldsymbol{I}$$

$$\begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & c & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & -c & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

The inverse matrix of $\boldsymbol{E}_1^{(i,j,c)}$ is $\boldsymbol{E}_1^{(i,j,-c)}$ and also $\det\left(\boldsymbol{E}_1^{(i,j,c)}\right) = 1$.

$\boldsymbol{E}_2^{(i,j)}\boldsymbol{E}_2^{(i,j)} = \boldsymbol{I}$ and $\boldsymbol{E}_3^{(i,c)}\boldsymbol{E}_3^{(i,1/c)} = \boldsymbol{I}$ are similarly shown.

$$\begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

The inverse matrix of $\boldsymbol{E}_2^{(i,j)}$ is itself and also $\det\left(\boldsymbol{E}_2^{(i,j)}\right) = -1$.

$$\begin{bmatrix} 1 & \cdots & 0 & \cdots & \cdot & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & c & \cdots & \cdot & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ \cdot & \cdots & \cdot & \cdots & 1 & \cdots & \cdot \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \cdot & \cdots & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & \cdots & 0 & \cdots & \cdot & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & \dfrac{1}{c} & \cdots & \cdot & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ \cdot & \cdots & \cdot & \cdots & 1 & \cdots & \cdot \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \cdot & \cdots & 1 \end{bmatrix}$$

The inverse matrix of $\boldsymbol{E}_3^{(i,c)}$ is $\boldsymbol{E}_3^{(i,1/c)}$ and also $\det\left(\boldsymbol{E}_3^{(i,c)}\right) = c$.

Exercise 5.2. Check what happens in $\mathbb{R}^2$ if we apply elementary matrices as linear mappings to vector (x, y). Also, check how matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ changes if we multiply it by elementary matrices from the left or right.

$$\begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + \alpha y \\ y \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ \alpha x + y \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}$$

$$\begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \alpha x \\ y \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ \alpha y \end{bmatrix}$$

$$\begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a + \alpha c & b + \alpha d \\ c & d \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & \alpha a + b \\ c & \alpha c + d \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ \alpha a + c & \alpha b + d \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} = \begin{bmatrix} a + \alpha b & b \\ c + \alpha d & d \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} b & a \\ d & c \end{bmatrix}$$

$$\begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \alpha a & \alpha b \\ c & d \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \alpha a & b \\ \alpha c & d \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ \alpha c & \alpha d \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} = \begin{bmatrix} a & \alpha b \\ c & \alpha d \end{bmatrix}$$

The following is a program that outputs the calculation of this question in LaTeX format.

Program: `elementary_latex.py`

In [1]:

```
from sympy import *
from sympy.abc import a, b, c, d, x, y, alpha

E11 = Matrix([[1, alpha], [0, 1]])
E12 = Matrix([[1, 0], [alpha, 1]])
E2 = Matrix([[0, 1], [1, 0]])
E31 = Matrix([[alpha, 0], [0, 1]])
E32 = Matrix([[1, 0], [0, alpha]])

v = Matrix([[x], [y]])
A = Matrix([[a, b], [c, d]])

print(r'\begin{eqnarray*}')
print(f'{latex(E11)}{latex(v)}&=&{latex(E11*v)}\\\\')
print(f'{latex(E12)}{latex(v)}&=&{latex(E12*v)}\\\\')
print(f'{latex(E2)}{latex(v)}&=&{latex(E2*v)}\\\\')
print(f'{latex(E31)}{latex(v)}&=&{latex(E31*v)}\\\\')
print(f'{latex(E32)}{latex(v)}&=&{latex(E32*v)}')
print(r'\end{eqnarray*}')

print()

print(r'\begin{eqnarray*}')
print(f'{latex(E11)}{latex(A)}&=&{latex(E11*A)}\\\\')
print(f'{latex(A)}{latex(E11)}&=&{latex(A*E11)}\\\\')
print(f'{latex(E12)}{latex(A)}&=&{latex(E12*A)}\\\\')
print(f'{latex(A)}{latex(E12)}&=&{latex(A*E12)}\\\\')
print(f'{latex(E2)}{latex(A)}&=&{latex(E2*A)}\\\\')
print(f'{latex(A)}{latex(E2)}&=&{latex(A*E2)}\\\\')
print(f'{latex(E31)}{latex(A)}&=&{latex(E31*A)}\\\\')
print(f'{latex(A)}{latex(E31)}&=&{latex(A*E31)}\\\\')
print(f'{latex(E32)}{latex(A)}&=&{latex(E32*A)}\\\\')
print(f'{latex(A)}{latex(E32)}&=&{latex(A*E32)}')
print(r'\end{eqnarray*}')
```

The expression of the output formula may appear different. For example, $x+\alpha y$ may be output as $\alpha y+x$. The appearance of the matrix also differs depending on whether the array environment is used or the matrix environment is used.

$$\begin{bmatrix}1 & \alpha\\0 & 1\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}\alpha y + x\\y\end{bmatrix}$$

$$\begin{bmatrix}1 & 0\\\alpha & 1\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x\\\alpha x + y\end{bmatrix}$$

$$\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}y\\x\end{bmatrix}$$

$$\begin{bmatrix}\alpha & 0\\0 & 1\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}\alpha x\\y\end{bmatrix}$$

$$\begin{bmatrix}1 & 0\\ 0 & \alpha\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} = \begin{bmatrix}x\\ \alpha y\end{bmatrix}$$

$$\begin{bmatrix}1 & \alpha\\ 0 & 1\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix} = \begin{bmatrix}a+\alpha c & \alpha d+b\\ c & d\end{bmatrix}$$

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}1 & \alpha\\ 0 & 1\end{bmatrix} = \begin{bmatrix}a & a\alpha+b\\ c & \alpha c+d\end{bmatrix}$$

$$\begin{bmatrix}1 & 0\\ \alpha & 1\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix} = \begin{bmatrix}a & b\\ a\alpha+c & \alpha b+d\end{bmatrix}$$

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}1 & 0\\ \alpha & 1\end{bmatrix} = \begin{bmatrix}a+\alpha b & b\\ \alpha d+c & d\end{bmatrix}$$

$$\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix} = \begin{bmatrix}c & d\\ a & b\end{bmatrix}$$

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} = \begin{bmatrix}b & a\\ d & c\end{bmatrix}$$

$$\begin{bmatrix}\alpha & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix} = \begin{bmatrix}a\alpha & \alpha b\\ c & d\end{bmatrix}$$

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}\alpha & 0\\ 0 & 1\end{bmatrix} = \begin{bmatrix}a\alpha & b\\ \alpha c & d\end{bmatrix}$$

$$\begin{bmatrix}1 & 0\\ 0 & \alpha\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix} = \begin{bmatrix}a & b\\ \alpha c & \alpha d\end{bmatrix}$$

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & \alpha\end{bmatrix} = \begin{bmatrix}a & \alpha b\\ c & \alpha d\end{bmatrix}$$

Exercise 5.3. Prove the following facts about the rank of a matrix.

(1) $\operatorname{rank}(\boldsymbol{A}) \leqq \min\{m, n\}$ for $m \times n$ matrix $\boldsymbol{A}$.

(2) $\operatorname{rank}(\boldsymbol{AB}) \leqq \min\{\operatorname{rank}(\boldsymbol{A}), \operatorname{rank}(\boldsymbol{B})\}$ for $l \times m$ matrix $\boldsymbol{A}$ and for $m \times n$ matrix $\boldsymbol{B}$.

(1) Because $\operatorname{rank}(\boldsymbol{A})$ is the dimension of $\operatorname{range}(\boldsymbol{A})$ and $\operatorname{range}(\boldsymbol{A}) \subseteq \mathbb{K}^m$, the inequality $\operatorname{rank}(\boldsymbol{A}) \leqq m$ is clear. On the other hand, $\operatorname{range}(\boldsymbol{A})$ is generated by all column vectors of $\boldsymbol{A}$. If necessary, we can get a basis for $\operatorname{range}(\boldsymbol{A})$ by removing the appropriate column vectors. Therefore, $\operatorname{rank}(\boldsymbol{A}) \leqq n$ is shown. The latter half can be also shown using the dimension theoem as follows: $\boldsymbol{A}$ is a linear mapping from $\mathbb{K}^n$ to $\mathbb{K}^m$. Let k be the dimension of $\operatorname{kernel}(\boldsymbol{A})$. The the dimension theorem says that

$$n = \operatorname{rank}(\boldsymbol{A}) + k.$$

hence, we have $n \geqq \operatorname{rank}(\boldsymbol{A})$.

(2) Let the shape of $\boldsymbol{A}$ be (l, m) and the one of $\boldsymbol{B}$ be (m, n). Since

$$\operatorname{range}(\boldsymbol{AB}) = \left\{\boldsymbol{ABx} \;\middle|\; \boldsymbol{x} \in \mathbb{K}^n\right\} = \left\{\boldsymbol{Ay} \;\middle|\; \boldsymbol{y} \in \operatorname{range}(\boldsymbol{B})\right\} \subseteq \operatorname{range}(\boldsymbol{A}),$$

we have $\operatorname{rank}(\boldsymbol{AB}) \leqq \operatorname{rank}(\boldsymbol{A})$. On the other hand, let $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k\}$ be a basis of $\operatorname{range}(\boldsymbol{B})$, then, $\{\boldsymbol{Ay}_1, \boldsymbol{Ay}_2, \ldots, \boldsymbol{Ay}_k\}$ generates $\operatorname{range}(\boldsymbol{AB})$ (it's not necessarily linearly independent). Because $k = \operatorname{rank}(\boldsymbol{B})$, we get $\operatorname{rank}(\boldsymbol{AB}) \leqq \operatorname{rank}(\boldsymbol{B})$.

Exercise 5.4. The following program produces calculation problems of matrix rank. Find the rank of each matrix generated by this program.

The following program is a modification of the program `prob_rank.py` in this book. It outputs the questions in LaTeX format and also outputs the answers.

Program: `prob_rank2.py`

In [1]:
```
from numpy.random import seed, choice, permutation
from sympy import Matrix, latex


def f(P, m1, m2, n):
    if n > min(m1, m2):
        return Matrix(choice(P, (m1, m2)))
    else:
        while True:
            X, Y = choice(P, (m1, n)), choice(P, (n, m2))
            A = Matrix(X.dot(Y))
            if A.rank() == n:
                return A


m1, m2 = 3, 4
seed(2020)
ans = []
print(r'\begin{enumerate}')
for i in permutation(max(m1, m2)):
    A = f([-3, -2, -1, 1, 2, 3], m1, m2, i + 1)
    print(rf'\item ${latex(A)}$')
    ans.append(A.rank())
print(r'\end{enumerate}')
print(f'Answer: {ans}')
```

Problems:

1. $\begin{bmatrix} 4 & 5 & -5 & 4 \\ 8 & 1 & -17 & 16 \\ 0 & -3 & -3 & 0 \end{bmatrix}$

2. $\begin{bmatrix} -10 & 4 & -8 & -8 \\ 7 & -2 & 8 & 4 \\ 0 & -1 & -3 & 2 \end{bmatrix}$

3. $\begin{bmatrix} -2 & 2 & -2 & 1 \\ -1 & -3 & -2 & -2 \\ -1 & -1 & -2 & -3 \end{bmatrix}$

4. $\begin{bmatrix} 3 & -1 & -3 & -1 \\ 9 & -3 & -9 & -3 \\ -3 & 1 & 3 & 1 \end{bmatrix}$

Answer: [3, 2, 3, 1]

Below is an example calculated by hand. Other elementary transformation procedures are possible.

1.

$$\begin{bmatrix} 4 & 5 & -5 & 4 \\ 8 & 1 & -17 & 16 \\ 0 & -3 & -3 & 0 \end{bmatrix}$$

(Line 2 − Line 1 × 2)

$$\begin{bmatrix} 4 & 5 & -5 & 4 \\ 0 & -9 & -7 & 8 \\ 0 & -3 & -3 & 0 \end{bmatrix}$$

(Line 3 × 3)

$$\begin{bmatrix} 4 & 5 & -5 & 4 \\ 0 & -9 & -7 & 8 \\ 0 & -9 & -9 & 0 \end{bmatrix}$$

(Line 3 − Line 2)

$$\begin{bmatrix} 4 & 5 & -5 & 4 \\ 0 & -9 & -7 & 8 \\ 0 & 0 & -2 & 0 \end{bmatrix}$$

Thus, the rank is 3.

2.

$$\begin{bmatrix} -10 & 4 & -8 & -8 \\ 7 & -2 & 8 & 4 \\ 0 & -1 & -3 & 2 \end{bmatrix}$$

(Line 1 ÷ 2) → (Line 2 × 5)

$$\begin{bmatrix} -5 & 2 & -4 & -4 \\ 35 & -10 & 40 & 20 \\ 0 & -1 & -3 & 2 \end{bmatrix}$$

(Row 2 + Row 1 × 7)

$$\begin{bmatrix} -5 & 2 & -4 & -4 \\ 0 & 4 & 12 & -8 \\ 0 & -1 & -3 & 2 \end{bmatrix}$$

(Row 3 × 4 + Row 2)

$$\begin{bmatrix} -5 & 2 & -4 & -4 \\ 0 & 4 & 12 & -8 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus, the rank is 2.

3.

$$\begin{bmatrix} -2 & 2 & -2 & 1 \\ -1 & -3 & -2 & -2 \\ -1 & -1 & -2 & -3 \end{bmatrix}$$

(Row 2 × 2 – Row 1) $\rightarrow$ (Row 3 × 2 – Row 1)

$$\begin{bmatrix} -2 & 2 & -2 & 1 \\ 0 & -8 & -2 & -5 \\ 0 & -4 & -2 & -7 \end{bmatrix}$$

(Row 3 × 2) $\rightarrow$ (Row 3 – Row 2)

$$\begin{bmatrix} -2 & 2 & -2 & 1 \\ 0 & -8 & -2 & -5 \\ 0 & 0 & -2 & -9 \end{bmatrix}$$

Thus, the rank is 3.

4.

$$\begin{bmatrix} 3 & -1 & -3 & -1 \\ 9 & -3 & -9 & -3 \\ -3 & 1 & 3 & 1 \end{bmatrix}$$

(Row 2 – Row 1 × 3) $\rightarrow$ (Row 3 – Row 1)

$$\begin{bmatrix} 3 & -1 & -3 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus, the rank is 1.

Exercise 5.5. Enumerate all permutations of orders 2 and 3, and calculate their signatures.

Here is a program that finds all permutations of order n and their signs.

Program: `permutation.py`

In [1]:
```
def perms(seq):
    if len(seq) <= 1:
        yield seq, 1
    else:
        for p, s in perms(seq[1:]):
            a = seq[0]
            for i in range(len(p) + 1):
                yield p[:i] + [a] + p[i:], s * (-1)**i
```

This program utilizes the mechanism of generators in Python. The permutations of order 2 and their signs are:

In [2]:
```
for p, s in perms([1, 2]):
    print(p, s)
```

Out[2]:
```
[1, 2] 1
[2, 1] -1
```

The permutations of order 3 and their signs are:

In [3]:
```
for p, s in perms([1, 2, 3]):
    print(p, s)
```

Out[3]:
```
[1, 2, 3] 1
[2, 1, 3] -1
[2, 3, 1] 1
[1, 3, 2] -1
[3, 1, 2] 1
[3, 2, 1] -1
```

Exercise 5.6. Prove the following formulas for the determinants of matrices of orders 2 and 3.

(1)
$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{22}.$$

(2)
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{matrix} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \end{matrix}$$

The reader should try both calculation according to the definition of determinats and calculation using elementary transformations.

Next is a program that uses SymPy to calculate the determinant in accordance with the definition and to compare it to the result computed by the `det` method of the `Matrix` class.

Program: `determinant.py`

In [1]:

```
def perms(seq):
    if len(seq) <= 1:
        yield seq, 1
    else:
        for p, s in perms(seq[1:]):
            a = seq[0]
            for i in range(len(p) + 1):
                yield p[:i] + [a] + p[i:], s * (-1)**i

if __name__ in '__main__':
    from sympy import *

    A = [[S('a11'), S('a12'), S('a13'), S('a14')],
         [S('a21'), S('a22'), S('a23'), S('a24')],
         [S('a31'), S('a32'), S('a33'), S('a34')],
         [S('a41'), S('a42'), S('a43'), S('a44')]]

    for N in[2, 3, 4]:
        P = perms(list(range(N)))
        D = 0
        for p, s in P:
            term = s
            for i in range(N):
                term *= A[i][p[i]]
            D += term
        print(D)

        B = Matrix(A)[:N,:N]
        print(B.det() - D)
```

```
a11*a22 - a12*a21
0
a11*a22*a33 - a11*a23*a32 - a12*a21*a33 + a12*a23*a31 + a13*a21*a32 - a13*
    a22*a31
0
a11*a22*a33*a44 - a11*a22*a34*a43 - a11*a23*a32*a44 + a11*a23*a34*a42 + a11
    *a24*a32*a43 - a11*a24*a33*a42 - a12*a21*a33*a44 + a12*a21*a34*a43 + a12
    *a23*a31*a44 - a12*a23*a34*a41 - a12*a24*a31*a43 + a12*a24*a33*a41 + a13
    *a21*a32*a44 - a13*a21*a34*a42 - a13*a22*a31*a44 + a13*a22*a34*a41 + a13
    *a24*a31*a42 - a13*a24*a32*a41 - a14*a21*a32*a43 + a14*a21*a33*a42 + a14
    *a22*a31*a43 - a14*a22*a33*a41 - a14*a23*a31*a42 + a14*a23*a32*a41
0
```

Exercise 5.7. Write a program that randomly creates 10000 square matrices of order n whose elements are integers between 0 and 10. Count the number of regular matrices for several n.

Below is a program that performs 10 experiments with n changed from 1 to 10.

Program: `random_matrix.py`

In [1]:

```
1  from numpy.random import seed, randint
2  from numpy.linalg import matrix_rank
3  
4  def f(n):
5      S = 0
6      for i in range(10000):
7          A = randint(0, 10, (n, n))
8          if matrix_rank(A) < n:
9              S += 1
10     return S
11 
12 seed(2021)
13 for m in range(10):
14     print([f(n) for n in range(1, 11)])
```

Lines 8, 9: Whether or not a square matrix of order n is regular is determined by whether the rank is equal to n.

```
\begin{verbatim}
[1005, 577, 206, 51, 4, 0, 0, 0, 0, 0]
[1044, 536, 228, 41, 11, 1, 0, 0, 0, 0]
[957, 582, 211, 54, 12, 0, 1, 0, 0, 0]
[960, 545, 190, 50, 5, 1, 0, 0, 0, 0]
[1052, 567, 194, 50, 10, 1, 0, 0, 0, 0]
[963, 543, 190, 48, 8, 3, 0, 0, 0, 0]
[984, 587, 204, 43, 17, 2, 1, 0, 0, 0]
[1010, 557, 210, 62, 9, 2, 0, 0, 0, 0]
[1022, 556, 217, 59, 9, 0, 0, 0, 0, 0]
[985, 542, 214, 45, 15, 1, 0, 0, 0, 0]
```

Each line lists the number of occurrences of singular matrices from $n = 1$ to $n = 10$. There are 10 lines because we ran the experiment 10 times. It can be seen that the occurrences of singular matrices are reduced as n increases.

Exercise 5.8. The following program generates two matrices. Calculate the determinant of generated matrices by hand using elementary operations.

The following program is a modified version of `prob_det.py` that prints out the questions in LaTeX format and also these answers.

Program: `prob_det2.py`

In [1]:

```
1  from numpy.random import seed, choice, permutation
2  from sympy import Matrix, latex
3  
4  def f(P, m, p):
5      while True:
6          A = Matrix(choice(P, (m, m)))
7          if p == 0:
8              if A.det() == 0:
9                  return A
10         elif A.det() != 0:
11             return A
12 
13 m = 3
14 seed(2020)
15 ans = []
16 print(f'Problem:')
17 print(r'\begin{enumerate}')
18 for p in permutation(2):
19     A = f([-3, -2, -1, 1, 2, 3], m, p)
20     prob = rf'\item ${latex(A)}$'.replace('[', '|').replace(']', '|')
21     print(prob)
22     ans.append(A.det())
23 print(r'\end{enumerate}')
24 print(f' Answer{ans}')
```

```
Problem:
\begin{enumerate}
\item $\left|\begin{matrix}-3 & 1 & 1\\1 & 3 & 1\\-3 & 3 & -3\end{matrix}\right|$
\item $\left|\begin{matrix}-2 & 1 & -1\\-3 & 2 & 3\\3 & -2 & -3\end{matrix}\right|$
\end{enumerate}
Answer[48, 0]
```

Line 20: Replace both `[` and `]` in the LaTeX string `latex(A)` with `|` to display the matrix as the determinant.

Problem:

1. $\begin{vmatrix} -3 & 1 & 1 \\ 1 & 3 & 1 \\ -3 & 3 & -3 \end{vmatrix}$

2. $\begin{vmatrix} -2 & 1 & -1 \\ -3 & 2 & 3 \\ 3 & -2 & -3 \end{vmatrix}$

Answer: [48, 0]

1.

$$\begin{aligned}
\begin{vmatrix} -3 & 1 & 1 \\ 1 & 3 & 1 \\ -3 & 3 & -3 \end{vmatrix} &= -\begin{vmatrix} 1 & 3 & 1 \\ -3 & 1 & 1 \\ -3 & 3 & -3 \end{vmatrix} && \text{(Swap Row 1 and Row 2)} \\
&= -\begin{vmatrix} 1 & 3 & 1 \\ 0 & 10 & 4 \\ 0 & 12 & 0 \end{vmatrix} && \begin{matrix}\text{(Row 2 + Row 1} \times 3) \\ \text{(Row 3 + Row 1} \times 3)\end{matrix} \\
&= \begin{vmatrix} 1 & 1 & 3 \\ 0 & 4 & 10 \\ 0 & 0 & 12 \end{vmatrix} && \text{(Swap Row 2 and Row 3)} \\
&= 1 \times 4 \times 12 = 48
\end{aligned}$$

2.

$$\begin{bmatrix} -2 & 1 & -1 \\ -3 & 2 & 3 \\ 3 & -2 & -3 \end{bmatrix} = \begin{bmatrix} -2 & 1 & -1 \\ -3 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix} \qquad \text{(Row 2 + Row 2)}$$

$$= 0 \qquad \text{(If a row is all 0s, its determinant value is 0)}$$

Typeset output for `m=4`

Problem:

1. $\begin{vmatrix} -3 & 1 & 1 & 1 \\ 3 & 1 & -3 & 3 \\ -3 & -3 & -3 & -1 \\ -2 & 1 & 1 & -1 \end{vmatrix}$

2. $\begin{vmatrix} -3 & -2 & -3 & 2 \\ 2 & 1 & 1 & -2 \\ 2 & 1 & -1 & -2 \\ 1 & 2 & -3 & 2 \end{vmatrix}$

Answer: [-88, 0]

Before solving this problem, we introduce one formula that simplifies the calculation of the determinant. This formula is also a special case of a formula called cofactor expansion (see answer to question 5.11).

$$\left|\begin{array}{c|cccc} a_{11} & * & * & \cdots & * \\ \hline 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & a_{n3} & \cdots & a_{nn} \end{array}\right| = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}$$

The horizontal line below the first row on the left side and the vertical line on the right side of the first column are inserted to make the formulas easier to see. The number of parts of $*$ can be anything. This is not difficult to derive from the definition of determinant.

1.

$$\begin{aligned}
\begin{vmatrix} -3 & 1 & 1 & 1 \\ 3 & 1 & -3 & 3 \\ -3 & -3 & -3 & -1 \\ -2 & 1 & 1 & -1 \end{vmatrix} &= -\begin{vmatrix} 1 & -3 & 1 & 1 \\ 1 & 3 & -3 & 3 \\ -3 & -3 & -3 & -1 \\ 1 & -2 & 1 & -1 \end{vmatrix} && \text{(Swap Row 1 and Row 2)} \\
&= -\begin{vmatrix} 1 & -3 & 1 & 1 \\ 0 & 6 & -4 & 2 \\ 0 & -12 & 0 & 2 \\ 0 & 1 & 0 & -2 \end{vmatrix} && \begin{array}{l} \text{(Row 2} - \text{Row 1)} \\ \text{(Row 3} + \text{Row 1} \times 3) \\ \text{(Row 3} - \text{Row 1)} \end{array} \\
&= -1 \times \begin{vmatrix} 6 & -4 & 2 \\ -12 & 0 & 2 \\ 1 & 0 & -2 \end{vmatrix} \\
&= \begin{vmatrix} 1 & 0 & -2 \\ -12 & 0 & 2 \\ 6 & -4 & 2 \end{vmatrix} && \text{(Swap Row 1 and Row 3)} \\
&= \begin{vmatrix} 1 & 0 & -2 \\ 0 & 0 & -22 \\ 0 & -4 & 14 \end{vmatrix} && \begin{array}{l} \text{(Row 2} + \text{Row 1} \times 12) \\ \text{(Row 3} - \text{Row 1} \times 6) \end{array} \\
&= -1 \times (-4) \times (-22) \ = \ -88
\end{aligned}$$

2,

$$
\begin{aligned}
\begin{vmatrix} -3 & -2 & -3 & 2 \\ 2 & 1 & 1 & -2 \\ 2 & 1 & -1 & -2 \\ 1 & 2 & -3 & 2 \end{vmatrix}
&= -\begin{vmatrix} 1 & 2 & -3 & 2 \\ 2 & 1 & 1 & -2 \\ 2 & 1 & -1 & -2 \\ -3 & -2 & -3 & 2 \end{vmatrix} \qquad \text{(Swap Row 1 and Row 4)} \\
&= -\begin{vmatrix} 1 & 2 & -3 & 2 \\ 0 & -3 & 7 & -6 \\ 0 & -3 & 5 & -6 \\ 0 & 4 & -12 & 8 \end{vmatrix} \qquad \begin{matrix} \text{(Row 2} - \text{Row } 1 \times 2) \\ \text{(Row 3} - \text{Row } 1 \times 2) \\ \text{(Row 2} + \text{Row } 1 \times 3) \end{matrix} \\
&= -1 \times \begin{vmatrix} -3 & 7 & -6 \\ -3 & 5 & -6 \\ 4 & -12 & 8 \end{vmatrix} \\
&= -4 \times \begin{vmatrix} -3 & 7 & -6 \\ -3 & 5 & -6 \\ 1 & -3 & 2 \end{vmatrix} \qquad \text{(Row } 3 \div 4) \\
&= 4 \times \begin{vmatrix} 1 & -3 & 2 \\ -3 & 5 & -6 \\ -3 & 7 & -6 \end{vmatrix} \qquad \text{(Swap Row 1 and Row 3)} \\
&= 4 \times \begin{vmatrix} 1 & -3 & 2 \\ 0 & -4 & 0 \\ 0 & -2 & 0 \end{vmatrix} \\
&= 4 \times (-4) \times (-2) \times \begin{vmatrix} 1 & -3 & 2 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{vmatrix} \qquad \begin{matrix} \text{(Row } 2 \div (-4)) \\ \text{(Row } 3 \div (-4)) \end{matrix} \\
&= 0 \qquad \text{(If there are the same two rows, the determinant value is 0)}
\end{aligned}
$$

Exercise 5.9. There are some small gaps and omissions in the arguments in this section. In the first part we say "are immediate from", "a simple calculation" and "using this fact repeatedly", which contain some gaps. In the second half, we omit the proof of existence of $\boldsymbol{X}$. Fill these gaps one by one.

- "The equalities 1 to 3 are immediate from the definition.": Let

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \ \boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}.$$

1. Since $\boldsymbol{A}+\boldsymbol{B} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \cdots & a_{nn}+b_{nn} \end{bmatrix}$, it follows that

$$\text{LHS} = \sum_{i=1}^{n} (a_{ii}+b_{ii}) = \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = \text{RHS}$$

2. Since $c\boldsymbol{A} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{n1} & ca_{n2} & \cdots & ca_{nn} \end{bmatrix}$, it follows that

$$\text{LHS} = \sum_{i=1}^{n} ca_{ii} = c\sum_{i=1}^{n} a_{ii} = \text{RHS}$$

3. Since $\boldsymbol{A}^{*} = \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \cdots & \overline{a_{n1}} \\ \overline{a_{12}} & \overline{a_{22}} & \cdots & \overline{a_{n2}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \cdots & \overline{a_{nn}} \end{bmatrix}$, it follows that

$$\text{LHS} = \sum_{i=1}^{n} \overline{a_{ii}} = \overline{\sum_{i=1}^{n} a_{ii}} = \text{RHS}$$

- "A simple calculation leads to $\operatorname{Tr}(\boldsymbol{AE}) = \operatorname{Tr}(\boldsymbol{EA})$.": When $\boldsymbol{E} = \boldsymbol{E}_1^{(i,j,c)}$, since both $\boldsymbol{AE}$ and $\boldsymbol{EA}$ have the same ordering of the diagonal components as $\boldsymbol{A}$, we have $\operatorname{Tr}(\boldsymbol{AE}) = \operatorname{Tr}(\boldsymbol{EA}) = \operatorname{Tr}(\boldsymbol{A})$.

When $\boldsymbol{E} = \boldsymbol{E}_2^{(i,j)}$, since the ordering of the diagonal components of $\boldsymbol{AE}$ and $\boldsymbol{EA}$ are just the (i,i) component and the (j,j) component of the diagonal component of $\boldsymbol{A}$ that are exchanged respectively, we have $\mathrm{Tr}\,(\boldsymbol{AE}) = \mathrm{Tr}\,(\boldsymbol{EA}) = \mathrm{Tr}\,(\boldsymbol{A})$.

When $\boldsymbol{E} = \boldsymbol{E}_3^{(i,c)}$, since the ordering of the diagonal components of $\boldsymbol{AE}$ and $\boldsymbol{EA}$ are just c-times of the (i,i) component of the diagonal component of $\boldsymbol{A}$, we have $\mathrm{Tr}\,(\boldsymbol{AE}) = \mathrm{Tr}\,(\boldsymbol{EA}) = c\mathrm{Tr}\,(\boldsymbol{A})$.

- " using this fact repeatedly, we get $\mathrm{Tr}\,(\boldsymbol{AB}) = \mathrm{Tr}\,(\boldsymbol{BA})$.": First, if $\boldsymbol{B}$ is a diagonal matrix, since

$$\boldsymbol{AB} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{22} & \cdots & a_{1n}b_{nn} \\ a_{21}b_{11} & a_{22}b_{22} & \cdots & a_{2n}b_{nn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}b_{11} & a_{n2}b_{22} & \cdots & a_{nn}b_{nn} \end{bmatrix}, \quad \boldsymbol{BA} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{11} & \cdots & a_{1n}b_{11} \\ a_{21}b_{22} & a_{22}b_{22} & \cdots & a_{2n}b_{22} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}b_{nn} & a_{n2}b_{nn} & \cdots & a_{nn}b_{nn} \end{bmatrix},$$

note that $\mathrm{Tr}\,(\boldsymbol{AB}) = \mathrm{Tr}\,(\boldsymbol{BA}) = \sum_{i=1}^{n} a_{ii}b_{ii}$. $\boldsymbol{B}$ can be represented by

$$\boldsymbol{B} = \boldsymbol{B}_1\boldsymbol{B}_2\cdots\boldsymbol{B}_k,$$

where each of $\boldsymbol{B}_1, \boldsymbol{B}_2, \ldots, \boldsymbol{B}_k$ is either a elementary matrix or a diagonal matrix. Hence

$$\begin{aligned} \mathrm{Tr}\,(\boldsymbol{AB}) &= \mathrm{Tr}\,(\boldsymbol{AB}_1\boldsymbol{B}_2\cdots\boldsymbol{B}_k) \\ &= \mathrm{Tr}\,(\boldsymbol{B}_k\boldsymbol{AB}_1\boldsymbol{B}_2\cdots\boldsymbol{B}_{k-1}) \\ &= \mathrm{Tr}\,(\boldsymbol{B}_{k-1}\boldsymbol{B}_k\boldsymbol{AB}_1\boldsymbol{B}_2\cdots\boldsymbol{B}_{k-2}) \\ &\vdots \\ &= \mathrm{Tr}\,(\boldsymbol{B}_1\boldsymbol{B}_2\cdots\boldsymbol{B}_k\boldsymbol{A}) \\ &= \mathrm{Tr}\,(\boldsymbol{BA}) \end{aligned}$$

is derived.

- "to prove our assertion, it is enough to show that $\varphi(\boldsymbol{U}_{ij}) = \alpha\,\delta_{ij}$ for some constant $\alpha \in \mathbb{K}$.": Since

$$\boldsymbol{A} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\boldsymbol{U}_{ij},$$

if the above condition holds,

$$\varphi(\boldsymbol{A}) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\varphi(\boldsymbol{U}_{ij}) = \sum_{i=1}^{n} a_{ii}\varphi(\boldsymbol{U}_{ii}) = \alpha\sum_{i=1}^{n} a_{ii} = \alpha\mathrm{Tr}\,(\boldsymbol{A})$$

is derived.

- “Since $\boldsymbol{U}_{ij} = \boldsymbol{U}_{ij}\boldsymbol{U}_{jj}$ and $\boldsymbol{U}_{jj}\boldsymbol{U}_{ij} = \boldsymbol{O}$, ‘’: Multiplying a matrix by $\boldsymbol{U}_{jj}$ from the right sets all other elements of the matrix to 0 except for the jth column. Therefore, when $\boldsymbol{U}_{ij}$ is multiplied from the right, jth column remains as it is, so $\boldsymbol{U}_{ij}$ does not change as it is . On the other hand, multiplying a matrix by $\boldsymbol{U}_{jj}$ from the left will set all other elements to 0 except for the jth row of the matrix. Therefore, if $\boldsymbol{U}_{ij}$ is multiplied from the left, the jth row (all components of this row are 0) remains unchanged, and all other components are 0, so it becomes $\boldsymbol{O}$.
- “Since $\boldsymbol{U}_{ii} = \boldsymbol{E}_2^{(i,j)}\boldsymbol{U}_{jj}\boldsymbol{E}_2^{(i,j)}$ and $\boldsymbol{E}_2^{(i,j)}\boldsymbol{E}_2^{(i,j)} = \boldsymbol{I}$,” Multiplying a matrix from the left by $\boldsymbol{E}_2^{(i,j)}$ swaps the ith and jth rows of the matrix, and multiplying from the left The ith column and the jth column are exchanged. Therefore, by multiplying from both sides of $\boldsymbol{U}_{jj}$, the (i,i) component and the (j,j) component are only exchanged. On the other hand, $\boldsymbol{E}_2^{(i,j)}\boldsymbol{E}_2^{(i,j)} = \boldsymbol{I}$ was already checked in question 5.1.
- The last paragraph: For

$$\boldsymbol{X} = \begin{bmatrix} \varphi(\boldsymbol{U}_{11}) & \varphi(\boldsymbol{U}_{21}) & \cdots & \varphi(\boldsymbol{U}_{n1}) \\ \varphi(\boldsymbol{U}_{12}) & \varphi(\boldsymbol{U}_{22}) & \cdots & \varphi(\boldsymbol{U}_{n2}) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi(\boldsymbol{U}_{1n}) & \varphi(\boldsymbol{U}_{2n}) & \cdots & \varphi(\boldsymbol{U}_{nn}) \end{bmatrix},$$

we have $\varphi(\boldsymbol{A}) = \mathrm{Tr}(\boldsymbol{AX})$, because

$$\mathrm{Tr}(\boldsymbol{AX}) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\varphi(\boldsymbol{U}_{ij}) = \varphi\left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\boldsymbol{U}_{ij}\right) = \varphi(\boldsymbol{A}).$$

Exercise 5.10. The following program randomly generates a problem of linear system with unknown variables x, y, z. By changing the parameters, it generates a problem with or without solution, or with a unique or infinitely many solutions. It prints out the created problem in LaTeX math mode code. It also output the solution. Solve the output problem by hand.

The program `Prob_eqn.py` given in this book outputs expressions using matrices and vecors, but has been improved to the following program which outputs problems and answers in the form of simultaneous equations.

Program: `prob_eqn2.py`

In [1]:

```
from numpy.random import seed, choice, shuffle
from sympy import Matrix, latex, solve, zeros
from sympy.abc import x, y, z

template = r''Problem:'
$\left\{
\begin{array}{ccccccccc}
%s&x&+&%s&y&+&%s&z&=&%s\\
%s&x&+&%s&y&+&%s&z&=&%s\\
%s&x&+&%s&y&+&%s&z&=&%s\\
\end{array}
\right.$\qquad'''

def f(P, m, n):
    while True:
        A0 = choice(P, (3, 4))
        A = Matrix(A0)
        if A[:, :3].rank() == m and A.rank() == n:
            break
    A, b = A[:, :3], A[:, 3]
    u = Matrix([[x], [y], [z]])
    problem = template % tuple(A0.flatten())
    print(problem)
    answer = solve(A * u - b, [x, y, z])
    if answer == []:
        print('Answer:  no solution')
    else:
        print(r'Answer: $\left\{\begin{array}{ccc}')
        print(rf'x&=&{latex(answer[x])}\\')
        print(rf'x&=&{latex(answer[y])}\\')
        if z in answer:
            print(rf'z&=&{latex(answer[z])}\\')
        else:
            print(r'z&:&\text{arbitrary constant}\\')
        print(r'\end{array}\right.$')
        print()
```

In [1]:
```

seed(2020)
print(r'\begin{enumerate}')
print(r'\item ')
m, n = 3, 3
f(range(2, 10), m, n)
print(r'\item ')
m, n = 2, 2
f(range(2, 10), m, n)
print(r'\item ')
m, n = 2, 3
f(range(2, 10), m, n)
print(r'\end{enumerate}')
```

1. Problem: $\begin{cases} 2x + 2y + 5z = 8 \\ 5x + 5y + 7z = 5 \\ 9x + 2y + 7z = 2 \end{cases}$ Answer: $\begin{cases} x = -\frac{18}{11} \\ x = -\frac{13}{11} \\ z = \frac{30}{11} \end{cases}$

2. Problem: $\begin{cases} 7x + 8y + 8z = 3 \\ 5x + 8y + 6z = 5 \\ 8x + 8y + 9z = 2 \end{cases}$ Answer: $\begin{cases} x = -z - 1 \\ x = \frac{5}{4} - \frac{z}{8} \\ z : \text{arbitrary constant} \end{cases}$

3. Problem: $\begin{cases} 3x + 4y + 4z = 8 \\ 2x + 3y + 3z = 6 \\ 5x + 9y + 9z = 5 \end{cases}$ Answer: no solution

How to solve problem 1:

x	y	z	
2	2	5	8
5	5	7	5
9	2	7	2

x	y	z		
2	2	5	8	
10	10	14	10	(Row 2 × 2)
18	4	14	4	(Row 3 × 2)

x	y	z		
2	2	5	8	
0	0	−11	−30	(Row 2 − Row 1 × 5)
0	−14	−31	−68	(Row 3 − Row 1 × 9)

x	y	z		
2	2	5	8	
0	−14	−31	−68	(Swap Row 2 and 3)
0	0	−11	−30	

$$\begin{cases} 2x + 2y + 5z = 8 & \cdots(1) \\ \quad -14y - 31z = -68 & \cdots(2) \\ \qquad -11z = 30 & \cdots(3) \end{cases}$$

By (3) we get $z = \dfrac{30}{11}$ Substituting this to (2) we have

$$\begin{aligned} -14y - 31 \times \frac{30}{11} &= -68 \\ -154y - 930 &= -748 \\ -154y &= 182 \\ y &= -\frac{13}{11} \end{aligned}$$

Substituting the to (1) we get

$$\begin{aligned} 2x + 2 \times \left(-\frac{13}{11}\right) + 5 \times \frac{30}{11} &= 8 \\ 22x - 26 + 150 &= 88 \\ 22x &= -36 \\ x &= -\frac{18}{11}. \end{aligned}$$

$$\text{Answer} \begin{cases} x = -\frac{18}{11} \\ y = -\frac{13}{11} \\ z = \frac{30}{11} \end{cases}$$

How to solve problem 2:

x	y	z	
7	8	8	3
5	8	6	5
8	8	9	2

x	y	z		
−1	0	−1	1	(Row 1 − Row 3)
5	8	6	5	
8	8	9	2	

x	y	z		
−1	0	−1	1	
0	8	1	10	(Row 2 − Row 1 × 5)
0	8	1	10	(Row 3 − Row 1 × 8)

x	y	z		
−1	0	−1	1	
0	8	1	10	
0	0	0	0	(Row 3 − Row 2)

$$\begin{cases} -\ x \quad - \quad z = 1 & \cdots (1) \\ \quad 8\ y + z = 10 & \cdots (2) \\ \quad\quad 0\ z = 0 & \cdots (3) \end{cases}$$

By (3), z is an arbitrary constant. Substituting this to(2), we get $y = \dfrac{5}{4} - \dfrac{z}{8}$. Substituting this to (1), we have$x = -1 - z$.

$$\text{Answer} \begin{cases} x = -1 - z \\ y = \frac{5}{4} - \frac{z}{8} \\ z : \text{arbitrary constant} \end{cases}$$

How to solve problem 3:

x	y	z	
3	4	4	8
2	3	3	6
5	9	9	5

x	y	z		
1	1	1	2	(Row 1 − Row 2)
2	3	3	6	
5	9	9	5	

x	y	z		
1	1	1	2	
0	1	1	2	(Row 2 − Row 1 × 2)
0	4	4	−5	(Row 3 − Row 1 × 5)

x	y	z		
1	1	1	2	
0	1	1	2	
0	0	0	−13	(Row 3 − Row 2)

$$\begin{cases} x + y + z = 2 & \cdots (1) \\ \quad y + z = 2 & \cdots (2) \\ \quad\quad 0\ z = -13 & \cdots (3) \end{cases}$$

(3) will never hold, so there is no solution in these simultaneous equations.

Exercise 5.11. Generate matrices using program `prob_det.py` in Exercise 6.8 and calculate their inverse matrices by the sweeping method.

1. $\begin{bmatrix} -3 & 1 & 1 \\ 1 & 3 & 1 \\ -3 & 3 & -3 \end{bmatrix}$:

$$\left[\begin{array}{ccc|ccc} -3 & 1 & 1 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 & 1 & 0 \\ -3 & 3 & -3 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 1 & 0 & 1 & 0 \\ -3 & 1 & 1 & 1 & 0 & 0 \\ -3 & 3 & -3 & 0 & 0 & 1 \end{array}\right] \quad \text{(Swap rows 1 and 2)}$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 1 & 0 & 1 & 0 \\ 0 & 10 & 4 & 1 & 3 & 0 \\ 0 & 12 & 0 & 0 & 3 & 1 \end{array}\right] \quad \begin{array}{l} (\text{Row } 2 + \text{Row } 1 \times 3) \\ (\text{Row } 3 + \text{Row } 1 \times 3) \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 1 & 0 & 1 & 0 \\ 0 & 1 & \frac{2}{5} & \frac{1}{10} & \frac{3}{10} & 0 \\ 0 & 12 & 0 & 0 & 3 & 1 \end{array}\right] \quad (\text{Row } 2 \div 10)$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -\frac{1}{5} & -\frac{3}{10} & \frac{1}{10} & 0 \\ 0 & 1 & \frac{2}{5} & \frac{1}{10} & \frac{3}{10} & 0 \\ 0 & 0 & -\frac{24}{5} & -\frac{6}{5} & -\frac{3}{5} & 1 \end{array}\right] \quad \begin{array}{l} (\text{Row } 1 - \text{Row } 2 \times 3) \\ \\ (\text{Row } 3 - \text{Row } 2 \times 12) \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -\frac{1}{5} & -\frac{3}{10} & \frac{1}{10} & 0 \\ 0 & 1 & \frac{2}{5} & \frac{1}{10} & \frac{3}{10} & 0 \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{8} & -\frac{5}{24} \end{array}\right] \quad (\text{Row } 3 \div -\tfrac{24}{5})$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{4} & \frac{1}{8} & -\frac{1}{24} \\ 0 & 1 & 0 & 0 & \frac{1}{4} & \frac{1}{12} \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{8} & -\frac{5}{24} \end{array}\right] \quad \begin{array}{l} (\text{Row } 1 + \text{Row } 3 \times \frac{1}{5}) \\ (\text{Row } 2 - \text{Row } 3 \times \frac{2}{5}) \end{array}$$

Therefore the inverse matrix is $\begin{bmatrix} -\frac{1}{4} & \frac{1}{8} & -\frac{1}{24} \\ 0 & \frac{1}{4} & \frac{1}{12} \\ \frac{1}{4} & \frac{1}{8} & -\frac{5}{24} \end{bmatrix}$.

2. $\begin{bmatrix} -2 & 1 & -1 \\ -3 & 2 & 3 \\ 3 & -2 & -3 \end{bmatrix}$

$$\left[\begin{array}{rrr|rrr} -2 & 1 & -1 & 1 & 0 & 0 \\ -3 & 2 & 3 & 0 & 1 & 0 \\ 3 & -2 & -3 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{rrr|lrr} 1 & -1 & -4 & 1 & -1 & 0 \\ -3 & 2 & 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right] \qquad \begin{array}{l} (\text{Row } 1 - \text{Row } 2) \\ \\ (\text{Row } 2 + \text{Row } 1) \end{array}$$

$$\left[\begin{array}{lrr|lrr} 1 & -1 & -4 & 1 & -1 & 0 \\ 0 & -1 & -9 & 3 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right] \qquad (\text{Row } 2 + \text{Row } 1 \times 3)$$

$$\left[\begin{array}{lrr|rrr} 1 & -1 & -4 & 1 & -1 & 0 \\ 0 & 1 & 9 & -3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right] \qquad (\text{Row } 2 \times -1)$$

$$\left[\begin{array}{lll|rll} 1 & 0 & 5 & -2 & 1 & 0 \\ 0 & 1 & 9 & -3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right] \qquad (\text{Row } 1 + \text{Row } 2)$$

There is no inverse matrix because the sweeping method cannot be continued anymore.

Exercise 5.12. Write a program using NumPy to calculate the inverse of a matrix of any order by the sweeping method. Compare its calculation time with the of function `inv` provided in NumPy for randomly generated matrices of large order.

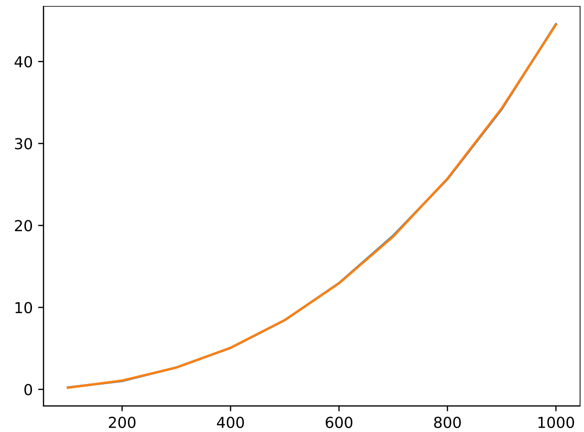
Program: `sweep.py`

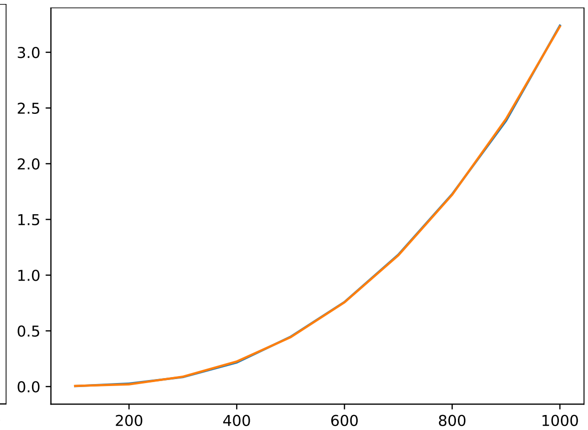
In [1]:

```
1  from numpy import *
2  from time import time
3  import matplotlib.pyplot as plt
4  
5  def inv(A):
6      m = len(A)
7      C = concatenate((A, eye(m)), axis=1)
8      for i in range(m):
9          C[i, :] /= C[i, i]
10         for k in range(m):
11             if k != i:
12                 C[k, :] -= C[k, i] * C[i, :]
13     return C[:, m:]
14 
15 X = array(range(10, 510, 10))
16 Y = []
17 for n in X:
18     A = random.normal(0, 1, (n, n))
19     t0 = time()
20     inv(A)
21     #linalg.inv(A)
22     t1 = time()
23     Y.append(t1 - t0)
24     print(n, t1 - t0)
25 
26 S = linalg.pinv(array([X**0, X**1, X**2, X**3]))
27 T = S.transpose()
28 c = T.dot(Y)
29 print(c)
30 
31 plt.plot(X, Y)
32 plt.plot(X, c[0] + c[1]*X + c[2]*X**2 + c[3]*X**3)
33 plt.show()
```

Line 2: For the function texttt time, see the answer of Question 4.11.
Lines 5 − 13: A function to calculate the inverse matrix by the sweeping method.
Line 21: Comment out the 20th line and remove the comment out of this line, which allows you to measure the calculation time using the function that calculates the inverse matrix in Numpy's module `linalg`.
Lines 26 − 28: The calculation time of the inverse matrix of the square matrix order n is of the n^3th order. Using the generalized reverse matrix described in Chapter 10, the coefficient of a cubic function that approximates the graph is calculated.

Line 29: Displllay coeficients of $1, x, x^2, x^3$ respectively.





The left is calculated using `inv`. The approximate function is as follows.

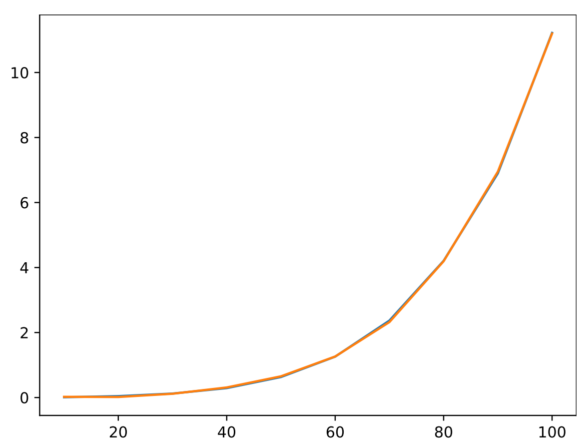
$y = -1.15176487 \times 10^{-1} + 7.41374215 \times 10^{-4} x + 2.17633872 \times 10^{-5} x^2 + 2.20985648 \times 10^{-8} x^3.$

The right is a calculation using Numpy's `linalg.inv`. More than 10 times faster than above. The approximate function is as follows.

$y = 3.17283176 \times 10^{-2} - 4.19975415 \times 10^{-4} x + 1.34962027 \times 10^{-6} x^2 + 2.27055604 \times 10^{-9} x^3.$

Both are Raspberry Pi 4.

By the way, if you use the formula of the reverse matrix (texttt inv.Py) using cofactor matrices at the end of the chapter, it will take 10 seconds or more in the 100th regular matrix as follows. This is quite slow compared to the above. `inv.py` uses the numpy function `linalg.det` for the calculation of matrices.



$y = 2.07779648 \times 10^{-1} - 3.17384809 \times 10^{-2} x + 1.57720787 \times 10^{-3} x^2 - 2.88631266 \times 10^{-5} x^3 + 2.72963110 \times 10^{-7} x^4 - 2.82740676 \times 10^{-11} x^5$

Exercise 5.13. Prove the equalities above for the cofactor matrix.

First, we show the following formula which is called the cofactor expansion of determinant.

$$\begin{aligned}|\boldsymbol{A}| &= a_{1j}\Delta_{1j} + a_{2j}\Delta_{2j} + \cdots + a_{nj}\Delta_{nj} \qquad (j = 1, 2, \ldots, n) \\ &= a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \cdots + a_{in}\Delta_{in} \qquad (i = 1, 2, \ldots, n)\end{aligned}$$

We show the first equation for $j = 1$, which is called the cofactor expansion of Column 1; to show

$$|\boldsymbol{A}| \;=\; a_{11}\Delta_{11} + a_{21}\Delta_{21} + \cdots + a_{n1}\Delta_{n1} \qquad \cdots (1).$$

By a property of determinant, we have

$$\begin{aligned}&\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \\ = &\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \\ = &\; a_{11}\begin{vmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{vmatrix} + a_{21}\begin{vmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 1 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + a_{n1}\begin{vmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_{n2} & \cdots & a_{nn} \end{vmatrix}.\end{aligned}$$

By repeatedly exchanging rows (or columns), it can be deformed as follows.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & 0 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \\ a_{i1} & a_{i2} & \cdots & 1 & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \\ a_{n1} & a_{n2} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = (-1)^{i+j}\begin{vmatrix} 1 & a_{i1} & \cdots & a_{i,j-1} & \cdots & a_{in} \\ 0 & a_{11} & \cdots & a_{1,j-1} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \\ 0 & a_{i-1,1} & \cdots & a_{i-1,j-1} & \cdots & a_{i-1,n} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \\ 0 & a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = \Delta_{ij}$$

This shows (1) when $j = 1$. The proof in which $j = 2, \ldots, n$ (the cofactor expansion of the jth column)

$$|\boldsymbol{A}| \;=\; a_{1j}\Delta_{1j} + a_{2j}\Delta_{2j} + \cdots + a_{nj}\Delta_{nj}$$

can be shown similarly. The second equality which is called the cofactor expansion with respect to rows is due to the one with respect to columns, because the value of determinant is invariant under transposition.

Next, we show$\boldsymbol{A}\boldsymbol{A}' = |\boldsymbol{A}|\,\boldsymbol{I}$ for

$$\boldsymbol{A}' \stackrel{\text{def}}{=} \begin{bmatrix} \Delta_{11} & \Delta_{21} & \cdots & \Delta_{n1} \\ \Delta_{12} & \Delta_{22} & \cdots & \Delta_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{1n} & \Delta_{2n} & \cdots & \Delta_{nn} \end{bmatrix}.$$

The (i, j)th component of $\boldsymbol{A}\boldsymbol{A}'$ is

$$a_{i1}\Delta_{j1} + a_{i2}\Delta_{j2} + \cdots + a_{in}\Delta_{jn}.$$

When $i = j$, it is the cofactor expansion of $|\boldsymbol{A}|$ with respect to rows and equals $|\boldsymbol{A}|$. When $i \neq j$, let $\boldsymbol{B}$ be the matrix obtained from $\boldsymbol{A}$ by replacing the j row by the i row. Then the determinant of $\boldsymbol{B}$ is 0 because two different rows are equal. Hence by considering the cofactor expansion of $\boldsymbol{B}$ with respect to the jth row, we have

$$a_{i1}\Delta_{j1} + a_{i2}\Delta_{j2} + \cdots + a_{in}\Delta_{jn} \;=\; 0.$$

Thus $\boldsymbol{A}\boldsymbol{A}' = |\boldsymbol{A}|\,\boldsymbol{I}$ is proved.

If the linear equation $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ has a solution, then $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$ and we have

$$\boldsymbol{x} \;=\; \frac{\boldsymbol{A}'\boldsymbol{b}}{|\boldsymbol{A}|}.$$

Here, the ith row of the numerator $\boldsymbol{A}'\boldsymbol{b}$ on the right side equals to the determinant of the matrix $\boldsymbol{A}$ with the jth column replaced by $\boldsymbol{b}$ (consider the cofactor expansion at the jth column of that matrix). This is called Cramer's formula.

Exercise 5.14. Give the formula for the solution of the system of equations $\boldsymbol{Ax} = \boldsymbol{b}$ with

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

where A is regular. That is, express explicitly the solutions for unknowns x_1, x_2, x_3 in terms of coefficients a_{ij} and b_i ($1 \leqq i, j \leqq 3$).

Program: `linearsystem.py`

In [1]:
```
from sympy import *
from sympy.abc import x, y, z

var('a11 a12 a13 a21 a22 a23 a31 a32 a33 b1 b2 b3')
ans = solve([a11*x + a12*y + a13*z - b1,
             a21*x + a22*y + a23*z - b2,
             a31*x + a32*y + a33*z - b3], [x, y, z])

print(r'\[\left\{\begin{array}{ccc}')
print(rf'x&=&{latex(ans[x])}\\[24pt]'.replace('frac', 'dfrac'))
print(rf'y&=&{latex(ans[y])}\\[24pt]'.replace('frac', 'dfrac'))
print(rf'z&=&{latex(ans[z])}\\'.replace('frac', 'dfrac'))
print(r'\end{array}\right.\]')
```

$$\left\{\begin{array}{ccc}
x&=&\dfrac{a_{12}a_{23}b_3 - a_{12}a_{33}b_2 - a_{13}a_{22}b_3 + a_{13}a_{32}b_2 + a_{22}a_{33}b_1 - a_{23}a_{32}b_1}{a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}}\\[24pt]
y&=&\dfrac{-a_{11}a_{23}b_3 + a_{11}a_{33}b_2 + a_{13}a_{21}b_3 - a_{13}a_{31}b_2 - a_{21}a_{33}b_1 + a_{23}a_{31}b_1}{a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}}\\[24pt]
z&=&\dfrac{a_{11}a_{22}b_3 - a_{11}a_{32}b_2 - a_{12}a_{21}b_3 + a_{12}a_{31}b_2 + a_{21}a_{32}b_1 - a_{22}a_{31}b_1}{a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}}
\end{array}\right.$$

To obtain the above results by hand, expand the following Cramer's formula:

$$x = \dfrac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \qquad y = \dfrac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \qquad z = \dfrac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$$

Chapter 6

Inner Product and Fourier Expansion

Exercise 6.1	Sect. 6.1.	p.124	$\|\cdot\|_1$ and $\|\cdot\|_\infty$
Exercise 6.2	Sect. 6.1.	p.124	Properties of the standard inner product
Exercise 6.3	Sect. 6.1.	p.124	Other properties of the standard inner product
Exercise 6.4	Sect. 6.1.	pp.125,126	The norm induced by an inner product
Exercise 6.5	Sect. 6.1.	p.125	A property of inner products
Exercise 6.6	Sect. 6.1.	pp.125,126	The parallelogram law etc.
Exercise 6.7	Sect. 6.1.	p.127	The inner products on $\mathbb{R}^2$
Exercise 6.8	Sect. 6.2.	p.129	Orthogonal projections
Exercise 6.9	Sect. 6.2.	p.129	Orthogonal projections on $\mathbb{R}^2$
Exercise 6.10	Sect. 6.2.	p.130	A property of orthogonal projections
Exercise 6.11	Sect. 6.2.	p.134	A projection onto a 2D plane in 3D space*
Exercise 6.12	Sect. 6.2.	p.134	Adjoint of linear mappings
Exercise 6.13	Sect. 6.3.	p.136	Norms and inner products in function spaces
Exercise 6.14	Sect. 6.3.	p.136	Norms and inner products of functions on $[0, 1]$
Exercise 6.15	Sect. 6.4.	p.139	An orthonormal system in function spaces
Exercise 6.16	Sect. 6.4.	p.142	Low-pass filters*
Exercise 6.17	Sect. 6.5.	p.145	Orthogonal function systems by NumPy*
Exercise 6.18	Sect. 6.5.	p.147	Orthogonal function systems by SymPy*
Exercise 6.19	Sect. 6.7.	p.150	The discrete Fourier transform
Exercise 6.20	Sect. 6.7.	p.152	Power spectra of musical scale*
Exercise 6.21	Sect. 6.7.	p.152	A property of power spectrum
Exercise 6.22	Sect. 6.7.	p.152	Power spectra of voices*
Exercise 6.23	Sect. 6.7.	p.154	The inverse Fourier transform*

* Using Python

Exercise 6.1. Prove that $\|\cdot\|_1$ and $\|\cdot\|_\infty$ satisfy Axioms N1 − N3 of norm.

Let $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ be in $\mathbb{K}^n$ and $a \in \mathbb{K}$.

$\|\cdot\|_1$:

$$\|\boldsymbol{x}\|_1 = |x_1| + |x_2| + \cdots + |x_n| \geqq 0$$

where the equality hold if and only if $x_1 = x_2 = \cdots = x_n = 0$.

$$\begin{aligned} \|a\boldsymbol{x}\|_1 &= |ax_1| + |ax_2| + \cdots + |ax_n| \\ &= |a|\,|x_1| + |a|\,|x_2| + \cdots + |a|\,|x_n| \\ &= |a|\big(|x_1| + |x_2| + \cdots + |x_n|\big) \\ &= |a|\,\|\boldsymbol{x}\|_1 . \end{aligned}$$

$$\begin{aligned} \|\boldsymbol{x}+\boldsymbol{y}\|_1 &= |x_1+y_1| + |x_2+y_2| + \cdots + |x_n+y_n| \\ &\leqq \big(|x_1|+|y_1|\big) + \big(|x_2|+|y_2|\big) + \cdots + \big(|x_n|+|y_n|\big) \\ &= \big(|x_1|+|x_2|+\cdots+|x_n|\big) + \big(|y_1|+|y_2|+\cdots+|y_n|\big) \\ &= \|\boldsymbol{x}\|_1 + \|\boldsymbol{y}\|_1 . \end{aligned}$$

$\|\cdot\|_\infty$:

$$\|\boldsymbol{x}\|_\infty \overset{\text{def}}{=} \max\big\{|x_1|, |x_2|, \ldots, |x_n|\big\} \geqq 0,$$

where the equality holds if and only if $x_1 = x_2 = \cdots = x_n = 0$.

$$\begin{aligned} \|a\boldsymbol{x}\|_\infty &= \max\big\{|ax_1|, |ax_2|, \ldots, |ax_n|\big\} \\ &= \max\big\{|a|\,|x_1|, |a|\,|x_2|, \ldots, |a|\,|x_n|\big\} \\ &= |a| \max\big\{|x_1|, |x_2|, \ldots, |x_n|\big\} \\ &= |a|\,\|\boldsymbol{x}\|_\infty . \end{aligned}$$

$$\begin{aligned} \|\boldsymbol{x}+\boldsymbol{y}\|_\infty &= \max\big\{|x_1+y_1|, |x_2+y_2|, \ldots, |x_n+y_n|\big\} \\ &\leqq \max\big\{|x_1|+|y_1|, |x_2|+|y_2|, \ldots, |x_n|+|y_n|\big\} \\ &\leqq \max\big\{|x_1|, |x_2|, \ldots, |x_n|\big\} + \max\big\{|y_1|, |y_2|, \ldots, |y_n|\big\} \\ &= \|\boldsymbol{x}\|_\infty + \|\boldsymbol{y}\|_\infty . \end{aligned}$$

Exercise 6.2. Show that the standard inner product satisfies Axioms of inner product.

I1. **positivity**:

$$\left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle = \sum_{i=1}^{n} \overline{x_i} x_i = \sum_{i=1}^{n} |x_i|^2 \geqq 0,$$

where the equality holds if and only if $x_1 = x_2 = \cdots = x_n = 0$.

I2. **Hermitian property**:

$$\left\langle \boldsymbol{y} \mid \boldsymbol{x} \right\rangle = \sum_{i=1}^{n} \overline{y_i} x_i = \sum_{i=1}^{n} \overline{\overline{x_i} y_i} = \overline{\sum_{i=1}^{n} \overline{x_i} y_i} = \overline{\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle}.$$

I3. **homogeneity**:

$$\left\langle \boldsymbol{x} \mid a\boldsymbol{y} \right\rangle = \sum_{i=1}^{n} \overline{x_i} a y_i = a \sum_{i=1}^{n} \overline{x_i} y_i = a \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle.$$

I4. **additivity** or **distributive low**:

$$\left\langle \boldsymbol{x} \mid \boldsymbol{y} + \boldsymbol{z} \right\rangle = \sum_{i=1}^{n} \overline{x_i} \left(y_i + z_i \right) = \sum_{i=1}^{n} \left(\overline{x_i} y_i + \overline{x_i} z_i \right)$$

$$= \sum_{i=1}^{n} \overline{x_i} y_i + \sum_{i=1}^{n} \overline{x_i} z_i = \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle + \left\langle \boldsymbol{x} \mid \boldsymbol{z} \right\rangle.$$

Exercise 6.3. Prove the above properties.

I5. **linearity**:

$$\begin{aligned}\langle \boldsymbol{x} \mid a\boldsymbol{y} + b\boldsymbol{z} \rangle &= \langle \boldsymbol{x} \mid a\boldsymbol{y} \rangle + \langle \boldsymbol{x} \mid b\boldsymbol{y} \rangle && \text{(by the additivity)} \\ &= a \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle + b \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle && \text{(by the homogenuity).}\end{aligned}$$

I6. **conjugate homogeneity**:

$$\begin{aligned}\langle a\boldsymbol{x} \mid \boldsymbol{y} \rangle &= \overline{\langle \boldsymbol{y} \mid a\boldsymbol{x} \rangle} && \text{(by the Hermitian property)} \\ &= \overline{a \langle \boldsymbol{y} \mid \boldsymbol{x} \rangle} && \text{(by the homogeneity)} \\ &= \overline{a}\overline{\langle \boldsymbol{y} \mid \boldsymbol{x} \rangle} && \text{(by a property of the complex conjugate)} \\ &= \overline{a} \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle && \text{(by the Hermitian property).}\end{aligned}$$

I7. **additivity** or **distributive low**:

$$\begin{aligned}\langle \boldsymbol{x} + \boldsymbol{y} \mid \boldsymbol{z} \rangle &= \overline{\langle \boldsymbol{z} \mid \boldsymbol{x} + \boldsymbol{y} \rangle} && \text{(by the Hermitian property)} \\ &= \overline{\langle \boldsymbol{z} \mid \boldsymbol{x} \rangle + \langle \boldsymbol{z} \mid \boldsymbol{y} \rangle} && \text{(by the additivity)} \\ &= \overline{\langle \boldsymbol{z} \mid \boldsymbol{x} \rangle} + \overline{\langle \boldsymbol{z} \mid \boldsymbol{y} \rangle} && \text{(by a property of the complex conjugate)} \\ &= \langle \boldsymbol{x} \mid \boldsymbol{z} \rangle + \langle \boldsymbol{y} \mid \boldsymbol{z} \rangle && \text{(by the Hermitian property).}\end{aligned}$$

I8. **conjugate linearity**:

$$\begin{aligned}\langle a\boldsymbol{x} + b\boldsymbol{y} \mid \boldsymbol{z} \rangle &= \langle a\boldsymbol{x} \mid \boldsymbol{z} \rangle + \langle b\boldsymbol{y} \mid \boldsymbol{z} \rangle && \text{(by the additivity)} \\ &= \overline{a} \langle \boldsymbol{x} \mid \boldsymbol{z} \rangle + \overline{b} \langle \boldsymbol{y} \mid \boldsymbol{z} \rangle && \text{(by the complex homogeneity).}\end{aligned}$$

Exercise 6.4. Prove the identity

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 \;=\; \|\boldsymbol{x}\|^2 + 2\mathrm{Re}\left(\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle\right) + \|\boldsymbol{y}\|^2 .$$

When $\mathbb{K}=\mathbb{R}$, it becomes

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 \;=\; \|\boldsymbol{x}\|^2 + 2\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle + \|\boldsymbol{y}\|^2 .$$

$$\begin{aligned}
\|\boldsymbol{x}+\boldsymbol{y}\|^2 &= \left\langle \boldsymbol{x}+\boldsymbol{y} \mid \boldsymbol{x}+\boldsymbol{y} \right\rangle \\
&= \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle + \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle + \left\langle \boldsymbol{y} \mid \boldsymbol{x} \right\rangle + \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle \\
&= \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle + \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle + \overline{\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle} + \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle \\
&= \|\boldsymbol{x}\|^2 + 2\mathrm{Re}\left(\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle\right) + \|\boldsymbol{y}\|^2
\end{aligned}$$

When $\mathbb{K}=\mathbb{R}$, note that $\left\langle \boldsymbol{y} \mid \boldsymbol{x} \right\rangle = \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle$.

Exercise 6.5. For $\boldsymbol{x}, \boldsymbol{y} \in V$, prove that $\boldsymbol{x} = \boldsymbol{y}$ if and only if $\langle \boldsymbol{x} \mid \boldsymbol{z} \rangle = \langle \boldsymbol{y} \mid \boldsymbol{z} \rangle$ holds for every $\boldsymbol{z} \in V$.

The if-part is trivial. We show the only-if-part Supose that $\langle \boldsymbol{x} \mid \boldsymbol{z} \rangle = \langle \boldsymbol{y} \mid \boldsymbol{z} \rangle$ holds for every $\boldsymbol{z} \in V$. Then for any $\boldsymbol{z} \in V$,

$$\langle \boldsymbol{x} - \boldsymbol{y} \mid \boldsymbol{z} \rangle \;=\; 0.$$

must hold. Putting $\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{y}$, we have

$$\langle \boldsymbol{x} - \boldsymbol{y} \mid \boldsymbol{x} - \boldsymbol{y} \rangle \;=\; 0.$$

Hence $\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{0}$ and so $\boldsymbol{x} = \boldsymbol{y}$.

Exercise 6.6. By rewriting the square of each norm by inner product, show the formulas:

(1) **The Pythagorean theorem**:

$$\boldsymbol{x} \perp \boldsymbol{y} \;\Rightarrow\; \|\boldsymbol{x}+\boldsymbol{y}\|^2 \;=\; \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 ,$$

(2) **The parallelogram law**:

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2 \;=\; 2\,\|\boldsymbol{x}\|^2 + 2\,\|\boldsymbol{y}\|^2 ,$$

(3) **The polarization identity**:

$$\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle \;=\; \frac{1}{4}\left(\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2 + i\,\|\boldsymbol{x}+i\boldsymbol{y}\|^2 - i\,\|\boldsymbol{x}-i\boldsymbol{y}\|^2\right),$$

and when $\mathbb{K}=\mathbb{R}$,

$$\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle \;=\; \frac{1}{4}\left(\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2\right).$$

(1) Suppose $\boldsymbol{x}\perp\boldsymbol{y}$. Since $\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle = \left\langle \boldsymbol{y} \mid \boldsymbol{x} \right\rangle = 0$,

$$\begin{aligned} &\|\boldsymbol{x}+\boldsymbol{y}\|^2 \;=\; \left\langle \boldsymbol{x}+\boldsymbol{y} \mid \boldsymbol{x}+\boldsymbol{y} \right\rangle \;=\; \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle + \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle + \left\langle \boldsymbol{y} \mid \boldsymbol{x} \right\rangle + \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle \\ =\;& \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle + \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle \;=\; \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 . \end{aligned}$$

(2) Since

$$\left\langle \boldsymbol{x}+\boldsymbol{y} \mid \boldsymbol{x}+\boldsymbol{y} \right\rangle \;=\; \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle + \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle + \left\langle \boldsymbol{y} \mid \boldsymbol{x} \right\rangle + \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle$$

and

$$\left\langle \boldsymbol{x}-\boldsymbol{y} \mid \boldsymbol{x}-\boldsymbol{y} \right\rangle \;=\; \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle - \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle - \left\langle \boldsymbol{y} \mid \boldsymbol{x} \right\rangle + \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle ,$$

adding both sides gives the desired equality.

(3) When $\mathbb{K}=\mathbb{R}$, since $\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle = \left\langle \boldsymbol{y} \mid \boldsymbol{x} \right\rangle$, we get the desired equality by subtractiing both sides of the two equalities in the proof of (2). When $\mathbb{K}=\mathbb{C}$, those equalities are rewritten as

$$\begin{aligned} \left\langle \boldsymbol{x}+\boldsymbol{y} \mid \boldsymbol{x}+\boldsymbol{y} \right\rangle &= \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle + \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle + \overline{\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle} + \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle \\ &= \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle + 2\mathrm{Re}\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle + \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle \\ \left\langle \boldsymbol{x}-\boldsymbol{y} \mid \boldsymbol{x}-\boldsymbol{y} \right\rangle &= \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle - \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle - \overline{\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle} + \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle \\ &= \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle - 2\mathrm{Re}\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle + \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle . \end{aligned}$$

Side-by-side subtraction of these yields

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2 \;=\; 4\mathrm{Re}\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle \qquad \cdots(1).$$

On the other hand, we also get

$$\begin{aligned}
\left\langle \boldsymbol{x} + i\boldsymbol{y} \mid \boldsymbol{x} + i\boldsymbol{y} \right\rangle &= \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle + i \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle + \bar{i} \left\langle \boldsymbol{y} \mid \boldsymbol{x} \right\rangle + \bar{i} i \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle \\
&= \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle + i \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle - i \overline{\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle} - \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle \\
&= \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle - 2\mathrm{Im} \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle - \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle \\
\left\langle \boldsymbol{x} - i\boldsymbol{y} \mid \boldsymbol{x} - i\boldsymbol{y} \right\rangle &= \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle - i \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle - \bar{i} \left\langle \boldsymbol{y} \mid \boldsymbol{x} \right\rangle + \bar{i} i \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle \\
&= \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle - i \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle + i \overline{\left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle} - \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle \\
&= \left\langle \boldsymbol{x} \mid \boldsymbol{x} \right\rangle + 2\mathrm{Im} \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle - \left\langle \boldsymbol{y} \mid \boldsymbol{y} \right\rangle .
\end{aligned}$$

Side-by-side subtraction of these yields

$$\left\| \boldsymbol{x} + i\boldsymbol{y} \right\|^2 - \left\| \boldsymbol{x} - i\boldsymbol{y} \right\|^2 = -4\mathrm{Im} \left\langle \boldsymbol{x} \mid \boldsymbol{y} \right\rangle \qquad \cdots (2).$$

By (1) $-\, i \times$ (2) the desired equality is shown.

Exercise 6.7. On the standard inner product and the Euclidean norm in $\mathbb{R}^2$, prove

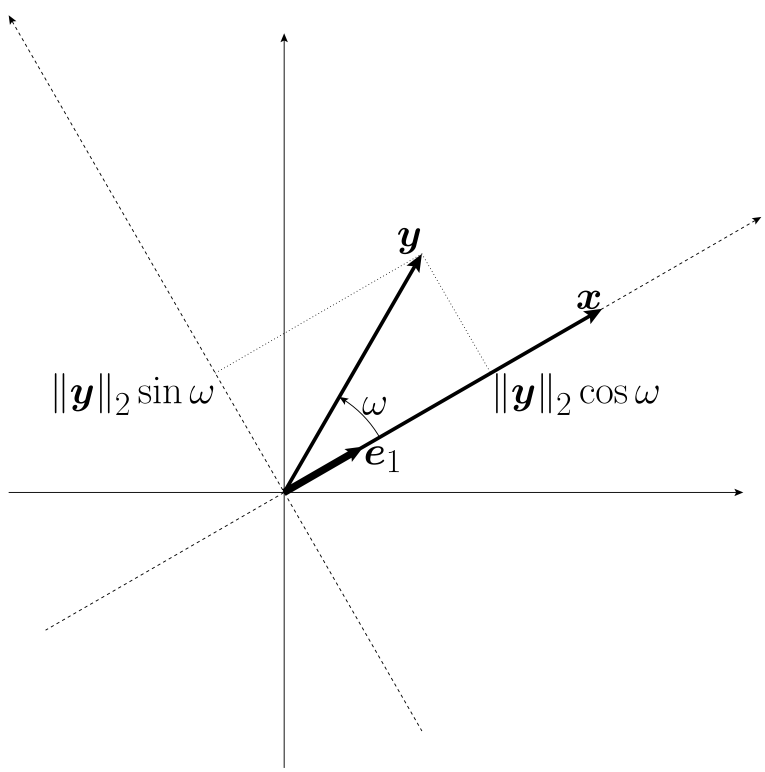
$$\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle = \|\boldsymbol{x}\|_2 \, \|\boldsymbol{y}\|_2 \cos\omega,$$

where ω is the angle between vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^2$.

Let $\boldsymbol{U} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, then

$$\begin{aligned} \langle \boldsymbol{Ux} \mid \boldsymbol{Uy} \rangle &= \left(\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right)^* \left(\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right) \\ &= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^* \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^* \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ &= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^* \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^* \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle \end{aligned}$$

Consider rotation matrix $\boldsymbol{U}$ by which the $\boldsymbol{x}$ and x axes overlap.



Then, $\boldsymbol{x}$ and $\boldsymbol{y}$ are transformed to the following vectors respectively.

$$\boldsymbol{Ux} = \begin{bmatrix} \|\boldsymbol{x}\|_2 \\ 0 \end{bmatrix}, \ \boldsymbol{Uy} = \begin{bmatrix} \|\boldsymbol{y}\|_2 \cos\omega \\ \|\boldsymbol{y}\|_2 \sin\omega \end{bmatrix}$$

Hence

$$\left\langle \boldsymbol{x} \;\middle|\; \boldsymbol{y} \right\rangle \;=\; \left\langle \boldsymbol{U}\boldsymbol{x} \;\middle|\; \boldsymbol{U}\boldsymbol{y} \right\rangle \;=\; \|\boldsymbol{x}\|_2 \, \|\boldsymbol{y}\|_2 \cos\omega + 0 \, \|\boldsymbol{x}\|_2 \, \|\boldsymbol{y}\|_2 \sin\omega \;=\; \|\boldsymbol{x}\|_2 \, \|\boldsymbol{y}\|_2 \cos\omega.$$

Exercise 6.8. Show that the orthogonal projection $\mathbf{proj}_W : V \to V$ is a linear mapping.

Let W be a subspace of V, $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ a orthonormalbasis of W and $\mathbf{proj}_W$ the orthogonal projection onto W. Then $\mathbf{proj}_W$ is expressed by

$$\mathbf{proj}_W(\boldsymbol{x}) = \sum_{i=1}^{n} \left\langle \boldsymbol{e}_i \;\middle|\; \boldsymbol{x} \right\rangle \boldsymbol{e}_i.$$

for any $\boldsymbol{x} \in V$. Since $\boldsymbol{x} \mapsto \left\langle \boldsymbol{e}_i \;\middle|\; \boldsymbol{x} \right\rangle$ is linear, so is $\boldsymbol{x} \mapsto \left\langle \boldsymbol{e}_i \;\middle|\; \boldsymbol{x} \right\rangle \boldsymbol{e}_i$ for every for $i = 1, 2, \ldots, n$. Therefore $\boldsymbol{x} \mapsto \mathbf{proj}_W(\boldsymbol{x})$ is also linear, because the sum of linear mappings is linear.

Exercise 6.9. Let a and b be real numbers with $a^2 + b^2 = 1$ and consider the subspace $W = \left\{ \begin{bmatrix} ax \\ bx \end{bmatrix} \;\middle|\; x \in \mathbb{R} \right\}$ of $\mathbb{R}^2$ generated by $\boldsymbol{e} = \begin{bmatrix} a \\ b \end{bmatrix}$. Prove that the representation matrix of the linear mapping $\mathbf{proj}_W$ on the standard basis of $\mathbb{R}^2$ is given by $\boldsymbol{e}\,\boldsymbol{e}^{\mathrm{T}} = \begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix}$.

Since $\{\boldsymbol{e}\}$ is an orthonomal basis of W, it follows that

$$\mathrm{proj}_W(\boldsymbol{x}) = \big\langle \boldsymbol{e} \mid \boldsymbol{x} \big\rangle \boldsymbol{e} \;=\; (ax+by)\begin{bmatrix} a \\ b \end{bmatrix} \;=\; \begin{bmatrix} a^2x + aby \\ abx + b^2y \end{bmatrix} \;=\; \begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

for all $\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} \in V$.

Exercise 6.10. Show the inequality $\|\mathbf{proj}_W(\boldsymbol{x})\| \leqq \|\boldsymbol{x}\|$ for any vector $\boldsymbol{x} \in V$ and any finite dimensional subspace $W \subseteq V$.

By applying the Pythagorean theorem to the vectors $\boldsymbol{x} - \mathbf{proj}_W(\boldsymbol{x})$ and $\mathbf{proj}_W(\boldsymbol{x})$ which are orthogonal to each other, we have

$$\begin{aligned} \|\boldsymbol{x}\|^2 &= \|\boldsymbol{x} - \mathbf{proj}_W(\boldsymbol{x}) + \mathbf{proj}_W(\boldsymbol{x})\|^2 \\ &= \|\boldsymbol{x} - \mathbf{proj}_W(\boldsymbol{x})\|^2 + \|\mathbf{proj}_W(\boldsymbol{x})\|^2 \\ &\geqq \|\mathbf{proj}_W(\boldsymbol{x})\|^2. \end{aligned}$$

Exercise 6.11. Use VPython to paste the image data created in Section 1 onto a 2D plane in 3D space without distortion (Figure 6.4).

The image data created in Chapter 1 is saved in a file as a list of coordinates of black pixels when a photograph is binarized. Let's project this onto a plane (screen) in the three-dimensional space $\mathbb{R}^3$. To specify the plane of $\mathbb{R}^3$, we give one $\boldsymbol{p} \in \mathbb{R}^3$ and make the plane orthogonal to it. Even if the screen is determined, the way to project the image vertically and horizontally cannot be determined unless the orthonormal basis of the plane is given. We obtain an orthonormal basis $\{\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2\}$ of $\mathbb{R}^3$ from $\{\boldsymbol{p}, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}$ by the Gram-Schmidt orthogonalization method. At this time, we get an orthonormal basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ of the screen, because $\boldsymbol{e}_0$ has the same direction as $\boldsymbol{p}$. We look at vector $\boldsymbol{e}_1$ in the upward direction and vector $\boldsymbol{e}_2$ in the right direction. Plot a black dot at point $x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2$ of $\mathbb{R}^3$ when the pixel at point (x_1, x_2) is black in the image data. The image will be drawn on the screen without distortion.

Program: `mypict6.py`

In [1]:

```
1  from vpython import *
2  from gram_schmidt import gram_schmidt
3  
4  with open('mypict1.txt', 'r') as fd:
5      Data = eval(fd.read())
6  
7  p = [1, 2, 3]
8  x, y, z = [1, 0, 0], [0, 1, 0], [0, 0, 1]
9  E = gram_schmidt([p, x, y, z])
10 
11 vx, vy, vz, vp = vec(*x), vec(*y), vec(*z), vec(*p)
12 e0, e1, e2 = vec(*E[0]), vec(*E[1]), vec(*E[2])
13 canvas(background=color.white, forward = -cross(e1,e2), up = e2)
14 
15 for e in [vx, vy, vz]:
16     curve(pos=[-10*e, 10*e], color=e)
17 arrow(axis=vp, color=color.yellow)
18 
19 curve(pos=[-10*e1, 10*e1], color=color.magenta)
20 curve(pos=[-10*e2, 10*e2], color=color.cyan)
21 
22 for (x1, x2) in Data:
23     points(pos=[10*(x1*e1 + x2*e2)], color=color.black, radius=2)
```

Line 2: In order to use the Gram-Schmidt orthogonalization method, `gram_schmit.py` created in this section is used as a library. Put it in the current folder in whch this program is saved.

Lines 4, 5: Read the binary image data created in Chapter 1. It must be in the current folder too. The data is a list of 2D coordinates, written as a literal of a list of tuples. After reading it as a string, evaluate it with the function `eval` as a list of tuples and refer the list as `Data`.

Lines 7, 8: Represent the vectors $\boldsymbol{p}, \boldsymbol{x}, \boldsymbol{y}$, and $\boldsymbol{z}$ of $\mathbb{R}^3$ as lists `p`, `x`, `y` and `z` in Python respectively.

Line 9: Create an orthonormal system from $\{\boldsymbol{p}, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}$ by the Gram-Schmidt orthogonalization method. The reason why we orthogonalize using four vectors including up to $\boldsymbol{z}$ is because it takes into consideration the case where $\boldsymbol{p}$ is along the x or y axis.

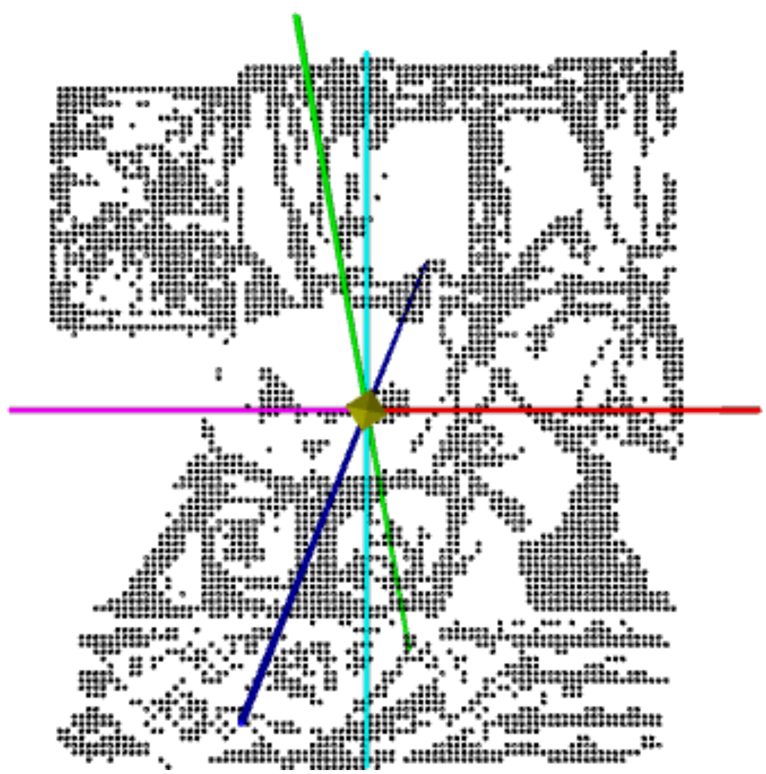
Lines 11,12: Cast $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ and $\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2$ to `vector`'s in VPyton.

Line 13: Set the background color to white when drawing with VPython. We will look at `e1` to the right direction and `e2` to the upwards. VPython's function `cross(e1, e2)` is the *cross product* of `e1` and `e2`.

Lines 15−17: Draw the x, y, and z axes as line segments ranging from -10 to 10 respectively. and the vector $\boldsymbol{p}$ as an arrow.

Lines 19,20: Draw coordinate axes containing $\boldsymbol{e}_1, \boldsymbol{e}_2$ respectively.

Lines 22,23: Plot the black points of a binary image on the screen. However, the screen itself is not drawn.



Exercise 6.12. Prove the following assertions.

(1) $\boldsymbol{f}^*$ is a linear mapping and does not depend on the choice of orthonormal basis of V.

(2) Assume that W is also finite dimensional. If $\boldsymbol{A}$ is the representation matrix of $\boldsymbol{f}$ on orthonormal bases of V and W, then the representation matrix of $\boldsymbol{f}^*$ on the same bases is the adjoint matrix $\boldsymbol{A}^*$ of $\boldsymbol{A}$, and $\left\langle \boldsymbol{A}^*\boldsymbol{y} \mid \boldsymbol{x} \right\rangle_1 = \left\langle \boldsymbol{y} \mid \boldsymbol{A}\boldsymbol{x} \right\rangle_2$ holds for any $\boldsymbol{x} \in V$ and $\boldsymbol{y} \in W$.

(1) Let denote $\boldsymbol{f}^*$ by $\boldsymbol{f}^\circ$ which is the adjoint linear mapping defined by the other orthonormal basis of V. Consider arbitrary but fixed $\boldsymbol{x} \in V$ and $\boldsymbol{y} \in W$. Just like showing $\left\langle \boldsymbol{f}^*(\boldsymbol{y}) \mid \boldsymbol{x} \right\rangle = \left\langle \boldsymbol{y} \mid \boldsymbol{f}(\boldsymbol{x}) \right\rangle$, we can see $\left\langle \boldsymbol{f}^\circ(\boldsymbol{y}) \mid \boldsymbol{x} \right\rangle = \left\langle \boldsymbol{y} \mid \boldsymbol{f}(\boldsymbol{x}) \right\rangle$. Therefore, $\left\langle \boldsymbol{f}^*(\boldsymbol{y}) \mid \boldsymbol{x} \right\rangle = \left\langle \boldsymbol{f}^\circ(\boldsymbol{y}) \mid \boldsymbol{x} \right\rangle$ holds. Since $\boldsymbol{x} \in V$ is arbitrary, it follows that $\boldsymbol{f}^*(\boldsymbol{y}) = \boldsymbol{f}^\circ(\boldsymbol{y})$. Since $\boldsymbol{y} \in W$ is also arbitrary, $\boldsymbol{f}^* : W \to V$ and $\boldsymbol{f}^\circ : W \to V$ are the same.

(2) Let $\boldsymbol{A}$ be an $m \times n$ matrix. Then

$$\left\langle \boldsymbol{y} \mid \boldsymbol{A}\boldsymbol{x} \right\rangle = \boldsymbol{y}^*\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}^*\boldsymbol{A}^{**}\boldsymbol{x} = (\boldsymbol{A}^*\boldsymbol{y})^*\boldsymbol{x} = \left\langle \boldsymbol{A}^*\boldsymbol{y} \mid \boldsymbol{x} \right\rangle$$

for any $\boldsymbol{x} \in \mathbb{K}^n$ and $\boldsymbol{y} \in \mathbb{K}^m$. It follows that the mapping $\boldsymbol{f} : \boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ satisfies

$$\left\langle \boldsymbol{y} \mid \boldsymbol{A}\boldsymbol{x} \right\rangle = \left\langle \boldsymbol{y} \mid \boldsymbol{f}(\boldsymbol{x}) \right\rangle = \left\langle \boldsymbol{f}^*(\boldsymbol{y}) \mid \boldsymbol{x} \right\rangle,$$

and that $\left\langle \boldsymbol{A}^*\boldsymbol{y} \mid \boldsymbol{x} \right\rangle = \left\langle \boldsymbol{f}^*(\boldsymbol{y}) \mid \boldsymbol{x} \right\rangle$. Because $\boldsymbol{x} \in \mathbb{K}^n$ is arbitrary, $\boldsymbol{A}^*\boldsymbol{y} = \boldsymbol{f}^*(\boldsymbol{y})$ for all $\boldsymbol{y} \in W$. Therefore, the matrix representation of $\boldsymbol{f}^* : W \to V$ is no more than $\boldsymbol{A}^*$.

Exercise 6.13. Prove that $\langle \cdot \mid \cdot \rangle$ above satisfies Axioms of inner product. Also prove that each function defined by the following is a norm of $C([a,b],\mathbb{K})$:

$$\|f\|_1 \overset{\text{def}}{=} \int_a^b |f(x)|\,dx,$$
$$\|f\|_2 \overset{\text{def}}{=} \left(\int_a^b |f(x)|^2\,dx\right)^{1/2},$$
$$\|f\|_\infty \overset{\text{def}}{=} \max_{a \leqq x \leqq b} |f(x)|.$$

Assuming that the following facts are known as the basics of calculus: (1) A continuous function defined on a finite close interval is integrable, that is, its definite integration is determined as a finetevalue; (2) If the value of the continuous function is always non-negative in the interval and the definite integration is 0, the function constantly takes value 0 on that interval; (3) A continuous function defined on a finite close interval has the maximum and the minimun on the interval.

$\|\cdot\|_1$: The positivity is clear from the previous comment (1). The absolute homogeneity follows

$$\|cf\|_1 = \int_a^b |cf(x)|\,dx = \int_a^b |c|\,|f(x)|\,dx = |c|\int_a^b |f(x)|\,dx = |c|\,\|f\|_1\,.$$

The triangular inequality (the sub-additivity) follows

$$\|f+g\|_1 = \int_a^b |f(x)+g(x)|\,dx \leqq \int_a^b \big(|f(x)|+|g(x)|\big)dx$$
$$= \int_a^b |f(x)|\,dx + \int_a^b |g(x)|\,dx = \|f\|_1 + \|g\|_1$$

$\|\cdot\|_2$: It suffices to show that

$$\langle f \mid g \rangle = \int_a^b \overline{f(x)}g(x)\,dx$$

satisfies the axioms of inner product. We remark that the definite integration of $f:\mathbb{R}\to\mathbb{C}$ is defined by

$$\int_a^b f(x)\,dx \overset{\text{def}}{=} \int_a^b \mathrm{Re}f(x)\,dx + i\int_a^b \mathrm{Im}f(x)\,dx$$

where i is the imaginary unit. Then, $f \mapsto \displaystyle\int_a^b f(x)\,dx$ is a linear mapping from $C([a,b],\mathbb{C})$ to $\mathbb{C}$. Also we have

$$\overline{\int_a^b f(x)\,dx} = \int_a^b \overline{f(x)}dx.$$

The positivity follows from

$$\langle f \mid f \rangle = \int_a^b \overline{f(x)} f(x)\, dx = \int_a^b |f(x)|^2\, dx$$

and the comment (2). The Hermite property follows as

$$\overline{\langle f \mid g \rangle} = \overline{\int_a^b \overline{f(x)} g(x)\, dx} = \int_a^b \overline{\overline{f(x)} g(x)} dx = \int_a^b f(x)\, \overline{g(x)} dx = \langle g \mid f \rangle .$$

The homogeneity follows as

$$\langle f \mid cg \rangle = \int_a^b \overline{f(x)} cg(x)\, dx = c \int_a^b \overline{f(x)} g(x)\, dx = c \langle f \mid g \rangle .$$

The additivity follows as

$$\langle f \mid g + h \rangle = \int_a^b \overline{f(x)} \big(g(x) + h(x) \big) dx = \int_a^b \big(\overline{f(x)} g(x) + \overline{f(x)} h(x) \big) dx$$
$$= \int_a^b \overline{f(x)} g(x)\, dx + \int_a^b \overline{f(x)} h(x)\, dx = \langle f \mid g \rangle + \langle f \mid h \rangle .$$

$\| \cdot \|_\infty$: The posotivity follows from that there exists x_1 such that $a \leqq x_1 \leqq b$ and $0 \leqq |f(x)| \leqq |f(x_1)| = \|f\|_\infty$ for any $a \leqq x \leqq b$. At this time, the absolute homogeneity follows as

$$\|cf\|_\infty = \max_{a \leqq x \leqq b} |cf(x)| = \max_{a \leqq x \leqq b} |c|\, |f(x)| = |c|\, |f(x_1)| = |c|\, \|f\|_\infty .$$

The triangular inequality (the sub-additivity) follows from that

$$\|f + g\|_\infty = \max_{a \leqq x \leqq b} |f(x) + g(x)| = \max_{a \leqq x \leqq b} \big(|f(x)| + |g(x)| \big)$$
$$= |f(x_3)| + |g(x_3)| \leqq |f(x_1)| + |g(x_2)| = \|f\|_\infty + \|g\|_\infty ,$$

where $x = x_1$ (resp., x_2 and x_3) attains the maximum of $|f(x)|$ (resp., $|g(x)|$ and $|f(x)| + |g(x)|$).

Exercise 6.14. (1) Consider the monomial functions $f_0, f_1, f_2 \in C([0,1], \mathbb{R})$ defined by $f_0(x) \stackrel{\text{def}}{=} 1$, $f_1(x) \stackrel{\text{def}}{=} x$ and $f_2(x) \stackrel{\text{def}}{=} x^2$ for $x \in [0,1]$. Find the inner products $\left\langle f_m \mid f_n \right\rangle$ for $m, n = 0, 1, 2$ and the norms $\|f_n\|_1, \|f_n\|_2, \|f_n\|_\infty$ for $n = 0, 1, 2$.
(2) Consider the exponential functions $e_n(x) \stackrel{\text{def}}{=} e^{2\pi inx} \in C([0,1], \mathbb{C})$ for $x \in [0,1]$ and $n \in \mathbb{Z}$, where i is the imaginary unit. Find the inner products $\left\langle e_m \mid e_n \right\rangle$ for $m, n = 0, \pm 1, \pm 2, \ldots$.

(1)

$$\begin{aligned}
\|f_0\|_1 &= \int_0^1 1dx = [x]_0^1 = 1 - 0 = 1 \\
\|f_1\|_1 &= \int_0^1 |x|\, dx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2} - \frac{0}{2} = \frac{1}{2} \\
\|f_2\|_1 &= \int_0^1 x^2 dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3} - \frac{0}{3} = \frac{1}{3}
\end{aligned}$$

$$\begin{aligned}
\|f_0\|_2 &= \sqrt{\int_0^1 1^2 dx} = \sqrt{[x]_0^1} = \sqrt{1-0} = 1 \\
\|f_1\|_2 &= \sqrt{\int_0^1 x^2 dx} = \sqrt{\left[\frac{x^3}{3}\right]_0^1} = \sqrt{\frac{1}{3} - \frac{0}{3}} = \sqrt{\frac{1}{3}} \\
\|f_2\|_2 &= \sqrt{\int_0^1 (x^2)^2 dx} = \sqrt{\left[\frac{x^5}{5}\right]_0^1} = \sqrt{\frac{1}{5} - \frac{0}{5}} = \sqrt{\frac{1}{5}}
\end{aligned}$$

$$\begin{aligned}
\|f_0\|_\infty &= \max_{0 \leqq x \leqq 1} 1 = 1 \\
\|f_1\|_\infty &= \max_{0 \leqq x \leqq 1} |x| = 1 \\
\|f_2\|_\infty &= \max_{0 \leqq x \leqq 1} x^2 = 1
\end{aligned}$$

(2) Note that

$$\left\langle e_m \mid e_n \right\rangle = \int_0^{2\pi} \overline{e^{imx}} e^{inx} dx = \int_0^{2\pi} e^{-imx} e^{inx} dx = \int_0^{2\pi} e^{i(-m+n)x} dx.$$

When $m = n$,

$$\left\langle e_m \mid e_m \right\rangle = \int_0^{2\pi} 1dx = [x]_0^{2\pi} = 2\pi.$$

When $m \neq n$, putting $k = -m + n \neq 0$,

$$\begin{aligned} \langle e_m \mid e_n \rangle &= \int_0^{2\pi} e^{ikx} dx = \int_0^{2\pi} (\cos kx + i \sin kx)\, dx \\ &= \int_0^{2\pi} \cos kx dx + i \int_0^{2\pi} \sin kx dx = [\cos kx]_0^{2\pi} + i\,[-\cos c]_0^{2\pi} \\ &= 1 - 1 + i(0 - 0) = 0 + 0i = 0. \end{aligned}$$

Note that the following calculation is also valid:

$$\int_0^{2\pi} e^{ikx} dx = \left[\frac{1}{ik} e^{ikx}\right]_0^{2\pi} = \frac{1}{ik}\left(e^{2\pi ik} - e^{0ik}\right) = \frac{1}{ik}(1 - 1) = 0.$$

Exercise 6.15. Prove that E is an orthonormal system of $C([0,1],\mathbb{R})$, that is, $\langle e_i \mid e_j \rangle = \delta_{ij}$ for all $i, j = 0, \pm 1, \pm 2, \ldots$.

We first show each e_i is of norm one. When $k < 0$,

$$\begin{aligned} \|e_k\|_2^2 &= 2\int_0^1 \sin^2(2\pi kt)\,dt = 2\int_0^1 \frac{1-\cos 2(2\pi kt)}{2}dt \\ &= \int_0^1 dt - \int_0^1 \cos(4\pi kt)\,dt = 1 - 0 = 1. \end{aligned}$$

When $k = 1$,

$$\|e_0\|_2^2 = 2\int_0^1 1^2 dt = 1.$$

When $k > 0$,

$$\begin{aligned} \|e_k\|_2^2 &= 2\int_0^1 \cos^2(2\pi kt)\,dt = 2\int_0^1 \frac{1+\cos 2(2\pi kt)}{2}dt \\ &= \int_0^1 dt + \int_0^1 \cos(4\pi kt)\,dt = 1 + 0 = 1. \end{aligned}$$

Next, we show the orthogonality.

$$\langle e_k \mid e_0 \rangle = \int_0^1 \sqrt{2}\sin(2\pi kt)\cdot 1dt = 0,$$

for $k < 0$, and

$$\langle e_k \mid e_0 \rangle = \int_0^1 \sqrt{2}\cos(2\pi kt)\cdot 1dt = 0.$$

for $k > 0$. When $k, l < 0$ and $k \neq l$,

$$\begin{aligned} \langle e_k \mid e_l \rangle &= \int_0^1 \sqrt{2}\sin(2\pi kt)\cdot\sqrt{2}\sin(2\pi lt)\,dt \\ &= \int_0^1 \cos\big(2\pi(k-l)t\big)dt - \int_0^1 \cos\big(2\pi(k+l)t\big)dt = 0 \end{aligned}$$

When $k, l > 0$ and $k \neq l$,

$$\begin{aligned} \langle e_k \mid e_l \rangle &= \int_0^1 \sqrt{2}\cos(2\pi kt)\cdot\sqrt{2}\cos(2\pi lt)\,dt \\ &= \int_0^1 \cos\big(2\pi(k+l)t\big)dt + \int_0^1 \cos\big(2\pi(k-l)t\big)dt = 0. \end{aligned}$$

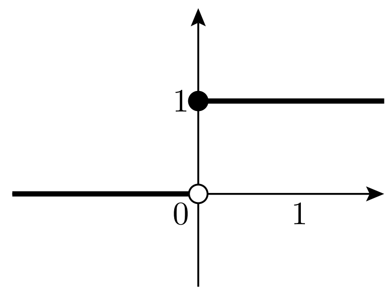
where note that both $k+l$ and $k-l$ can not be 0. When $k<0$ and $l>0$,

$$\begin{aligned}\left\langle e_k \mid e_l \right\rangle &= \int_0^1 \sqrt{2}\sin(2\pi kt)\cdot\sqrt{2}\cos(2\pi lt)\,dt \\ &= \int_0^1 \sin\bigl(2\pi(k+l)t\bigr)dt + \int_0^1 \sin\bigl(2\pi(k-l)t\bigr)dt \;=\; 0\end{aligned}$$

including the case of $k+l=0$.

The following calculation that reduces the number of cases which must be considered. Let

$$\sigma(k) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } k \geqq 0 \\ 0 & \text{if } k<0 \end{cases}$$



and

$$s_k(t) = \sin\left(2\pi kt + \frac{\sigma(k)}{2}\pi\right) \qquad k=0,\pm1,\pm2,\ldots.$$

Note that $e_0 = s_0$ and that if $k\neq 0$, then $\sqrt{2}s_k = e_k$.

$$\begin{aligned}\left\langle s_k \mid s_l \right\rangle &= \int_0^1 \sin\left(2\pi kt + \frac{\sigma(k)}{2}\pi\right)\sin\left(2\pi kt + \frac{\sigma(l)}{2}\pi\right)dt \\ &= -\frac{1}{2}\int_0^1 \cos\left(2\pi(k+l)t + \frac{\sigma(k)+\sigma(l)}{2}\pi\right)dt \\ &\quad -\frac{1}{2}\int_0^1 \cos\left(2\pi(k-l)t + \frac{\sigma(k)-\sigma(l)}{2}\pi\right)dt \\ &= -\frac{1}{2}I(k,l) + \frac{1}{2}J(k,l) \quad \text{(we denote)},\end{aligned}$$

where $\cos x$ has period2π and the definite interation form 0 to 2π is 0.

1. Case $k\neq l$: Note that $J(k,l)=0$. Furthermore,

   (a) Case $k+l\neq 0$: Since $I(k,l)=0$, it follows that $\left\langle s_k \mid s_l \right\rangle = 0$.

   (b) Case $k+l=0$: Since $\sigma(k)+\sigma(l)=1$, it follows that

   $$I(k,l) = \int_0^1 \cos\frac{\pi}{2} = 0,$$

   $\left\langle s_k \mid s_l \right\rangle = 0$ holds.

2. Case $k = l$: Note that

$$J(k,k) \;=\; \cos 0 = 1.$$

   Furtermore,

   (a) Case $k \neq 0$: Since $I(k,k) = 0$, it follows that $\|s_k\|^2 = \left\langle s_k \mid s_k \right\rangle = \frac{1}{2}$.

   (b) Case $k = 0$: Since

$$I(0,0) \;=\; \cos\pi \;=\; -1,$$

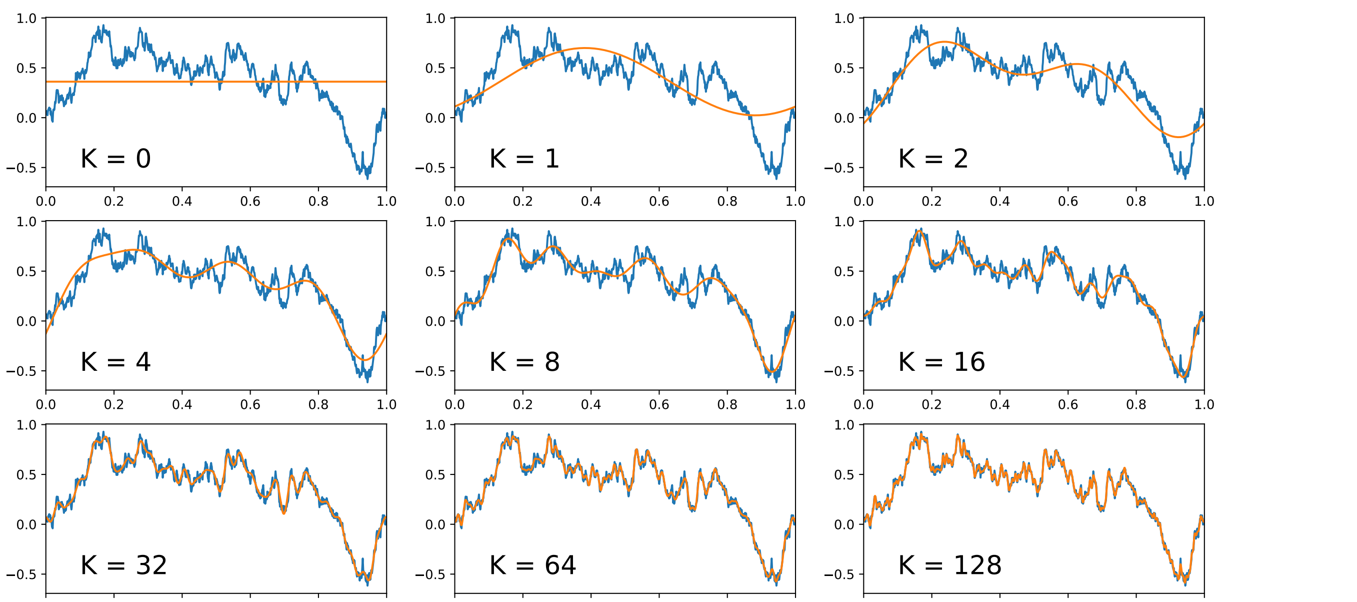
   it follows that $\|s_0\|^2 = \left\langle s_0 \mid s_0 \right\rangle = 1$.

Therefore, $\{s_0, s_{\pm1}, s_{\pm2}, \dots\}$ is an orthogonal system and by normalizing it we get E.

Exercise 6.16. In program `brown.py`, in order to use this program `fourier.py` as a library and import this `lowpass`, comment out Line 3 and uncomment Line 4, and execute the program. The modified `brown.py` will display a warning at Line 17 because `f_K` takes a complex value. If we use `f_K.real` instead of `f_K` at Line 17, the warning will disappear. Display also `f_K.imag` at the same time.

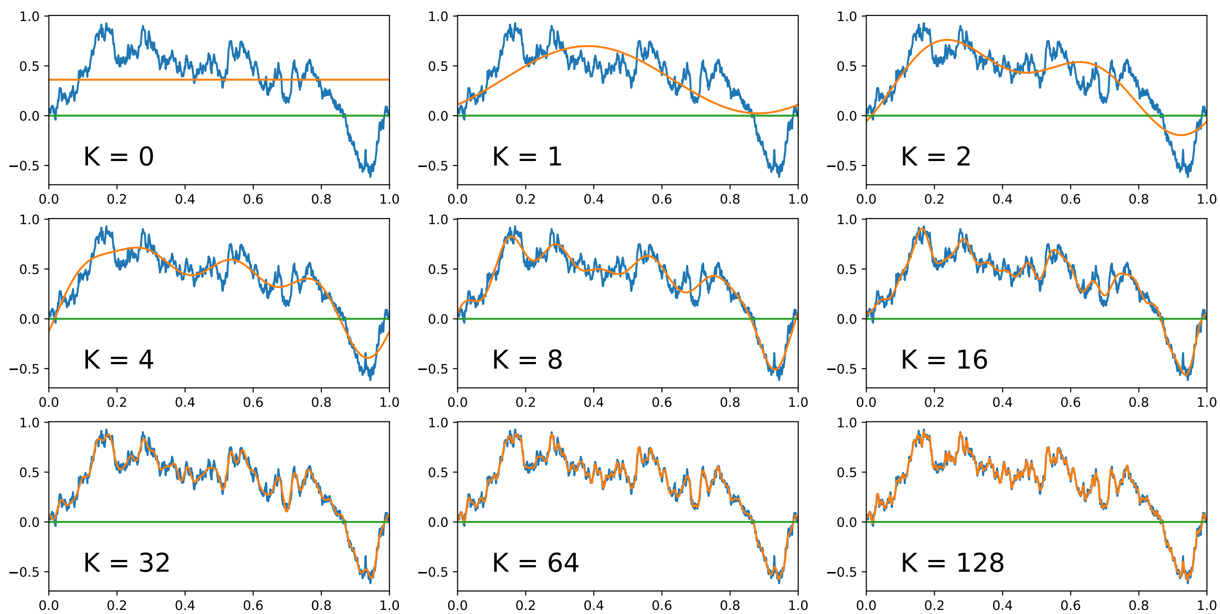
Program: `brown.py`

In [1]:
```
1  from numpy import arange, cumsum, sqrt
2  from numpy.random import seed, normal
3  from trigonometric import lowpass
4  import matplotlib.pyplot as plt
5
6  seed(2021)
7
8  n = 1000
9  dt = 1 / n
10 t = arange(0, 1, dt)
11 f = cumsum(normal(0, sqrt(dt), n))
12
13 fig, ax = plt.subplots(3, 3, figsize=(16, 8))
14 for k, K in enumerate([0, 1, 2, 4, 8, 16, 32, 64, 128]):
15     i, j = divmod(k, 3)
16     f_K = lowpass(K, t, f)
17     ax[i][j].plot(t, f), ax[i][j].plot(t, f_K)
18     ax[i][j].text(0.1, min(f), f'K = {K}', fontsize = 20)
19     ax[i][j].set_xlim(0, 1)
20 plt.show()
```



Program: `brown2.py`

In [1]:
```
from numpy.random import seed, normal
from fourier import lowpass
import matplotlib.pyplot as plt

seed(2021)

n = 1000
dt = 1 / n
t = arange(0, 1, dt)
f = cumsum(normal(0, sqrt(dt), n))

fig, ax = plt.subplots(3, 3, figsize=(16, 8))
for k, K in enumerate([0, 1, 2, 4, 8, 16, 32, 64, 128]):
    i, j = divmod(k, 3)
    f_K = lowpass(K, t, f)
    ax[i][j].plot(t, f), ax[i][j].plot(t, f_K.real)
    ax[i][j].plot(t, f_K.imag)
    ax[i][j].text(0.1, min(f), f'K = {K}', fontsize = 20)
    ax[i][j].set_xlim(0, 1)
plt.show()
```



The imaginary part of `f_K` is the constant function 0, i.e., `f_K` is a real valued function. When the values of `K` in `trigonometric.py` and `fourier.py` are equal , the graphs of each `f_K` draw the same shape. Indeed, since

$$
\begin{aligned}
& \left\langle e_k \mid f \right\rangle_{\mathbb{C}} e_k(t) \\
= & \int_0^1 e^{-2\pi i k s} f(s)\, ds\, e^{2\pi i k t} \\
= & \left(\int_0^1 \cos(2\pi k s) f(s)\, ds - i \int_0^1 \sin(2\pi k s) f(s)\, ds \right) \left(\cos(2\pi k t) + i \sin(2\pi k t) \right) \\
= & \int_0^1 \cos(2\pi k s) f(s)\, ds \cos(2\pi k t) + \int_0^1 \sin(2\pi k s) f(s)\, ds \sin(2\pi k t) \\
& + i \left(\int_0^1 \cos(2\pi k s) f(s)\, ds \sin(2\pi k t) - \int_0^1 \sin(2\pi k s) f(s)\, ds \cos(2\pi k t) \right),
\end{aligned}
$$

it follows that

$$
\begin{aligned}
& \left\langle e_k \mid f \right\rangle_{\mathbb{C}} e_k(t) + \left\langle e_{-k} \mid f \right\rangle e_{-k}(t) \\
= & \ 2 \int_0^1 \cos(2\pi k s) f(s)\, ds \cos(2\pi k t) + 2 \int_0^1 \sin(2\pi k s) f(s)\, ds \sin(2\pi k t) \\
= & \left\langle e'_k \mid f \right\rangle_{\mathbb{R}} e'_k(t) + \left\langle e'_{-k} \mid f \right\rangle_{\mathbb{R}} e'_{-k}(t)
\end{aligned}
$$

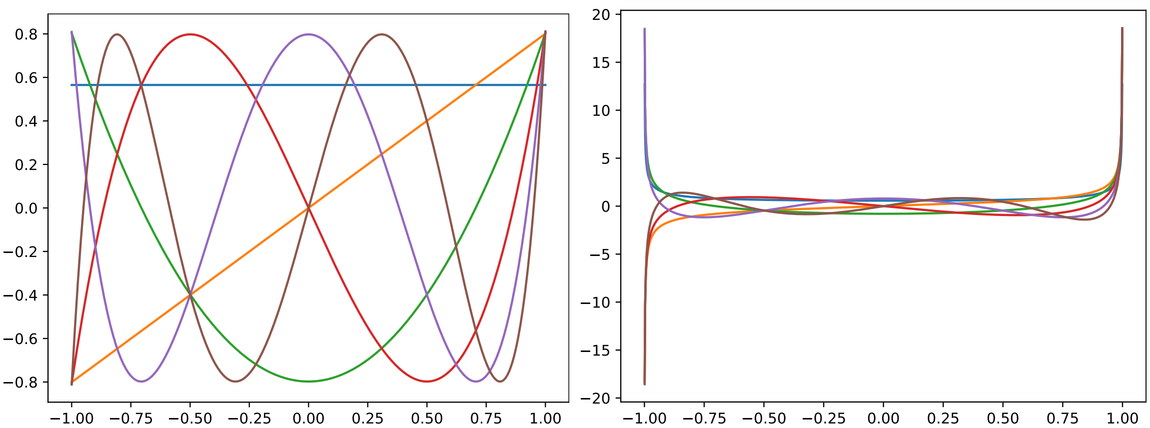
for $k > 0$ and also

$$
\left\langle e_0 \mid f \right\rangle_{\mathbb{C}} e_0(t) \ = \ \left\langle e'_0 \mid f \right\rangle_{\mathbb{R}} e'_0(t).
$$

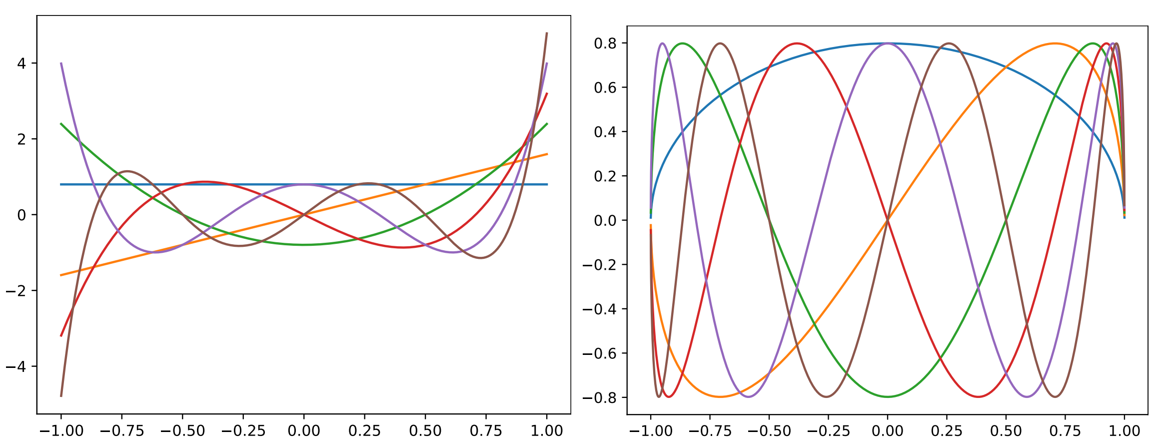
Here, $\left\langle \cdot \mid \cdot \right\rangle_{\mathbb{C}}$ is the inner product on the complex linear space $C([0,1], \mathbb{C})$, and $\left\langle \cdot \mid \cdot \right\rangle_{\mathbb{R}}$ is the one on the real linear space $C([0,1], \mathbb{R})$. On the other hand, $\{e_k\}_{k=-\infty}^{\infty}$ is the Fourier series on $C([0,1], \mathbb{C})$ and $\{e'_k\}_{k=-\infty}^{\infty}$ is the trigonometric series on $C([0,1], \mathbb{R})$.

Exercise 6.17. Draw the graphs of other polynomials and compare them with each other.

(1) `poly_np1.py` (using the Gram-Schmidt orthogonalization method)

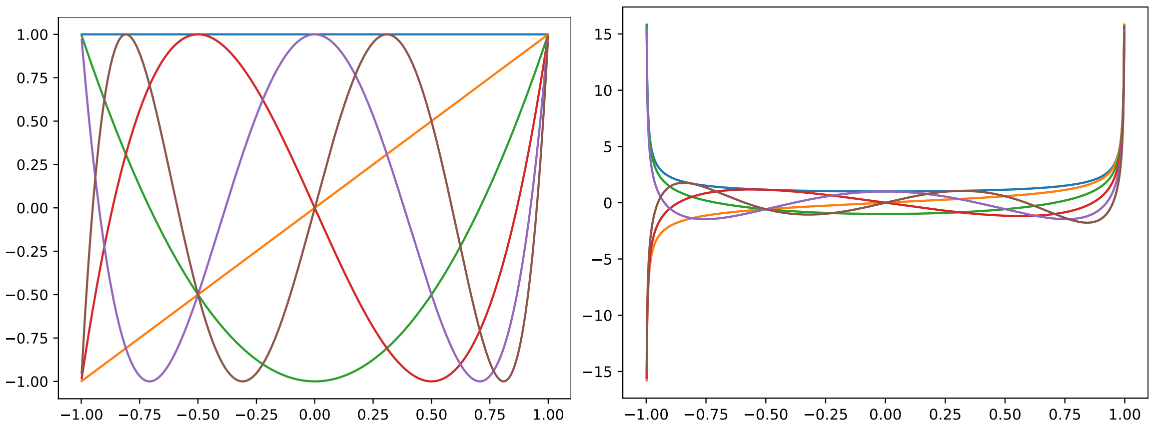


Chebyshev polynomial of the first kind (left), Chebyshev polynomial of the first kind with weights (right)

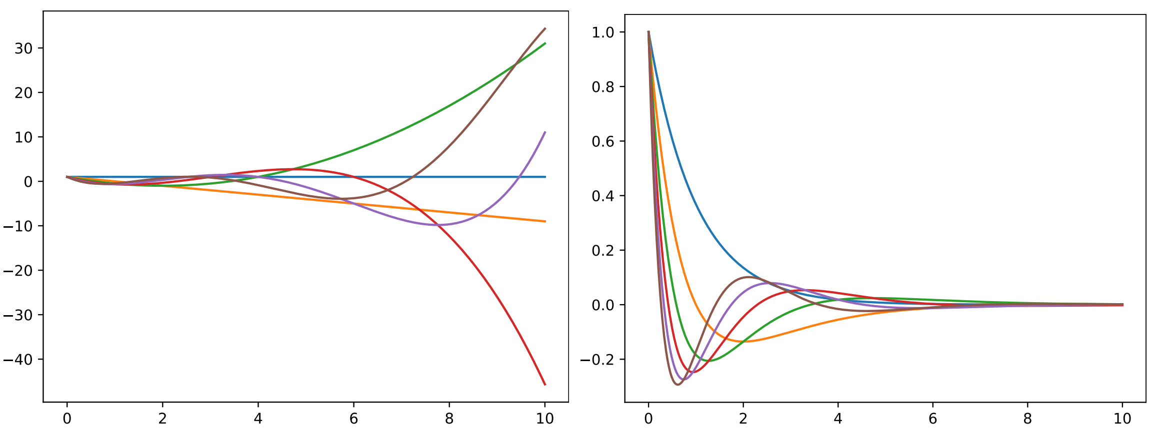


Chebyshev polynomial of the 2nd kind (left), Chebyshev polynomial of the 2nd kind with weights (right)

`poly_np2.py` (using orthogonal polynomials defined in `numpy.polynomial`)

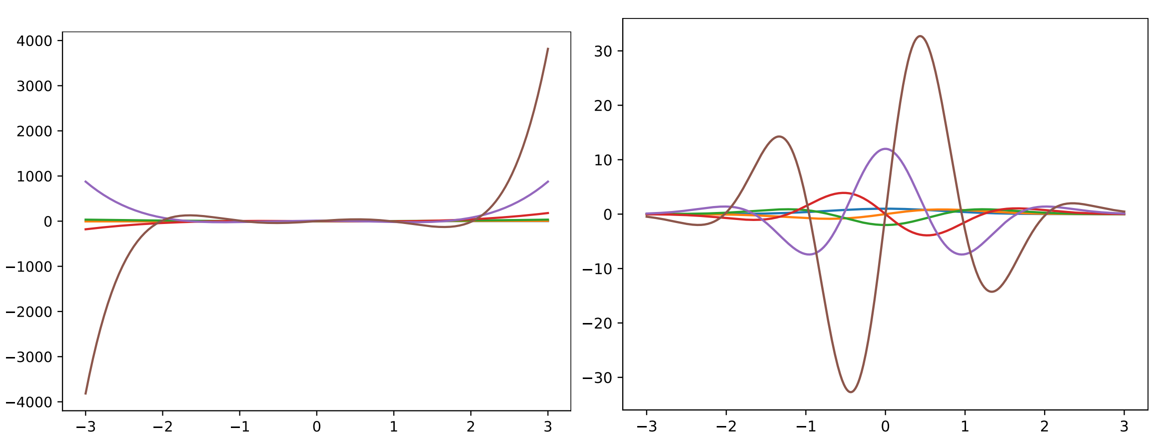


Chebyshev polynomial of the first kind (left), Chebyshev polynomial of the first kind with weights (right)



Laguerre polynomial (left), Laguerre polynomial weighted (right)

(2) `poly_np2.py` (using orthogonal polynomials defined in `numpy.polynomial`)



Hermite polynomial (left), Hermite polynomial weighted (right)

Orthogonal polynomials are also defined in SciPy's Special functions module (`scipy.special`).

Exercise 6.18. Experiment with the other polynomials than Legendre's. Compare the results by NumPy and by SymPy.

Program: `poly_sp1.py` (using the Gram-Schmidt orthogonalization method)

In [1]:
```
1  from sympy import *
2  from sympy.abc import x
3  
4  D = {
5      'Ledendre': ((x, -1, 1), 1),
6      'Chebyshev1': ((x, -1, 1), 1 / sqrt(1 - x**2)),
7      'Chebyshev2': ((x, -1, 1), sqrt(1 - x**2)),
8      'Laguerre': ((x, 0, oo), exp(-x)),
9      'Hermite': ((x, -oo, oo), exp(-x**2)),
10 }
11 dom, weight = D['Chebyshev2']
12 
13 def inner(f1, f2):
14     f = f1 * f2 * weight
15     return integrate(f.expand(), dom)
16 
17 def norm(f):
18     return sqrt(inner(f, f))
19 
20 def gram_schmidt(A):
21     E = []
22     while A != []:
23         a = A.pop(0)
24         b = a - sum([inner(e, a) * e for e in E])
25         E.append(b / norm(b))
26     return E
27 
28 E = gram_schmidt([1, x, x**2, x**3])
29 for n, e in enumerate(E):
30     print(f'e{n}(x) = {e}')
```

Lines 13−15: If you directly write an expression that contains a fraction in the integrand or the integral area, the fraction will be converted to a real number in the calculation.

```
>>> integrate(x**2 -1/4, (x, -1/2, 1/2))
-0.166666666666667
```

In this case, if you pass it as a character string, it will calculate it as a fraction.

```
>>> integrate('x**2 - 1/4', (x, '-1/2', '1/2'))
-1/6
```

This is the same as the following.

```
>>> integrate(x**2 - Rational(1, 4), (x, -Rational(1, 2), Rational(1, 2)))
-1/6
```

Program: By using `poly_sp2.py` (with `sympy.polys.orthopolys`)

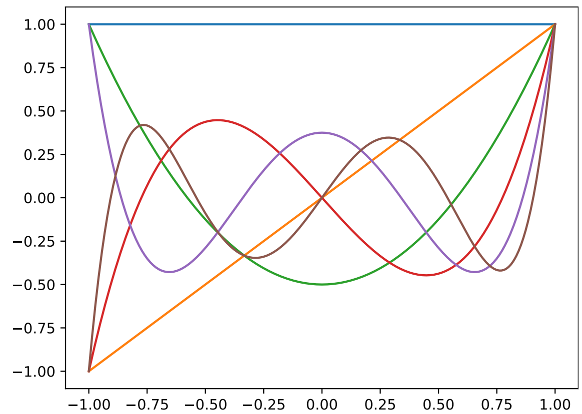
In [1]:
```
from sympy.polys.orthopolys import (
    legendre_poly,
    chebyshevt_poly,
    chebyshevu_poly,
    laguerre_poly,
    hermite_poly,
)
from sympy.abc import x
from sympy import Lambda
import matplotlib.pyplot as plt
import numpy as np

x1 = np.linspace(-1, 1, 1001)
x2 = x1[1:-1]
x3 = np.linspace(0, 10, 1001)
x4 = np.linspace(-3, 3, 1001)

Poly = {
    'Legendre': (legendre_poly, x1, 1),
    'Chebyshev1': (chebyshevt_poly, x2, 1),
    'Chebyshev2': (chebyshevu_poly, x2, np.sqrt(1 - x2**2)),
    'Laguerre': (laguerre_poly, x3, np.exp(-x3)),
    'Hermite': (hermite_poly, x4, np.exp(-x4**2))
}

poly = Poly['Hermite']
e = poly[0]
dom = poly[1]
weight = poly[2]
for n in range(6):
    print(f'e{n}(x) = {e(n, x)}')
    f = np.vectorize(Lambda(x, e(n, x)))
    plt.plot(dom, f(dom) * weight)

plt.show()
```

- Lagrange polynomials
  - Results of `poly_sp1.py`:

    ```
    e0(x) = 1/sqrt(pi)
    e1(x) = sqrt(2)*x/sqrt(pi)
    e2(x) = 2*sqrt(2)*(x**2 - 1/2)/sqrt(pi)
    e3(x) = 4*sqrt(2)*(x**3 - 3*x/4)/sqrt(pi)
    ```

  - Results of `poly_sp2.py`:

```
e0(x) = 1
e1(x) = x
e2(x) = 3*x**2/2 - 1/2
e3(x) = 5*x**3/2 - 3*x/2
e4(x) = 35*x**4/8 - 15*x**2/4 + 3/8
e5(x) = 63*x**5/8 - 35*x**3/4 + 15*x/8
```

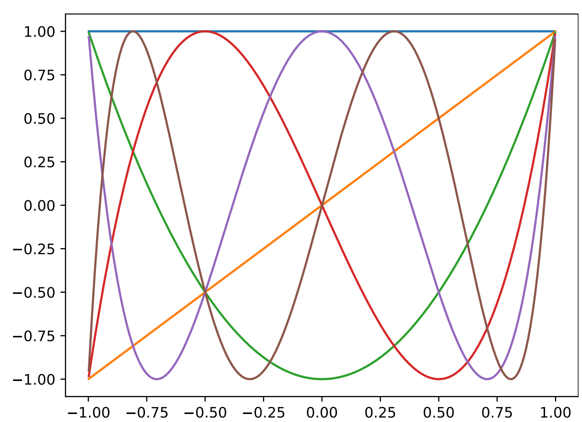


- Chebyshev polynomials of the first kind
  - Results of `poly_sp1.py`:

```
e0(x) = 1/sqrt(pi)
e1(x) = sqrt(2)*x/sqrt(pi)
e2(x) = 2*sqrt(2)*(x**2 - 1/2)/sqrt(pi)
e3(x) = 4*sqrt(2)*(x**3 - 3*x/4)/sqrt(pi)
```

  - Results of `poly_sp2.py`:

```
e0(x) = 1
e1(x) = x
e2(x) = 2*x**2 - 1
e3(x) = 4*x**3 - 3*x
e4(x) = 8*x**4 - 8*x**2 + 1
e5(x) = 16*x**5 - 20*x**3 + 5*x
```



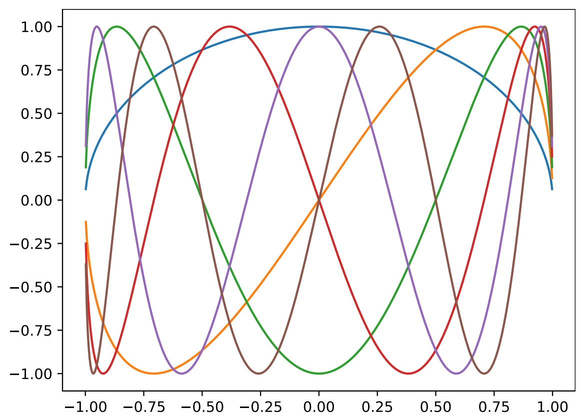
- Chebyshev polynomials of the second kind

  - Results of `poly_sp1.py`:

```
e0(x) = sqrt(2)/sqrt(pi)
e1(x) = 2*sqrt(2)*x/sqrt(pi)
e2(x) = 4*sqrt(2)*(x**2 - 1/4)/sqrt(pi)
e3(x) = 8*sqrt(2)*(x**3 - x/2)/sqrt(pi)
```

  - Results of `poly_sp2.py` (The graph is multiplied by a weight function.):

```
e0(x) = 1
e1(x) = 2*x
e2(x) = 4*x**2 - 1
e3(x) = 8*x**3 - 4*x
e4(x) = 16*x**4 - 12*x**2 + 1
e5(x) = 32*x**5 - 32*x**3 + 6*x
```



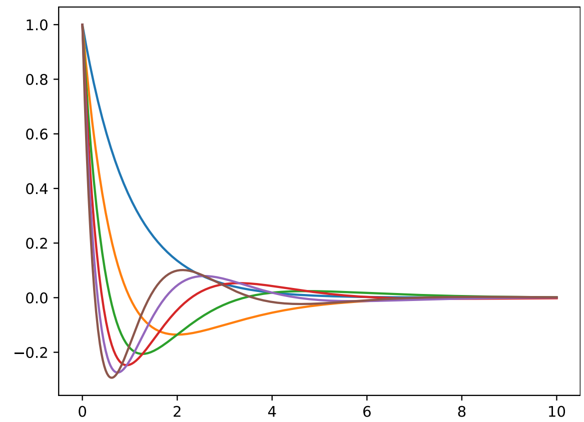
- Laguerre polynomials

  - Results of `poly_sp1.py`:

```
e0(x) = 1
e1(x) = x - 1
e2(x) = x**2/2 - 2*x + 1
e3(x) = x**3/6 - 3*x**2/2 + 3*x - 1
```

  - Results of `poly_sp2.py` (The graph is multiplied by a weight function.):

```
e0(x) = 1
e1(x) = 1 - x
e2(x) = x**2/2 - 2*x + 1
e3(x) = -x**3/6 + 3*x**2/2 - 3*x + 1
e4(x) = x**4/24 - 2*x**3/3 + 3*x**2 - 4*x + 1
e5(x) = -x**5/120 + 5*x**4/24 - 5*x**3/3 + 5*x**2 - 5*x + 1
```

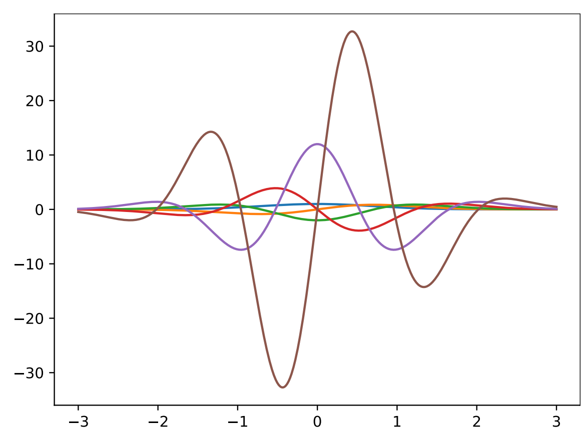


- Hermite polynomials
  - Results of `poly_sp1.py`:

```
e0(x) = pi**(-1/4)
e1(x) = sqrt(2)*x/pi**(1/4)
e2(x) = sqrt(2)*(x**2 - 1/2)/pi**(1/4)
e3(x) = 2*sqrt(3)*(x**3 - 3*x/2)/(3*pi**(1/4))
```

  - Results of `poly_sp2.py` (The graph is multiplied by a weight function.):

```
e0(x) = 1
e1(x) = 2*x
e2(x) = 4*x**2 - 2
e3(x) = 8*x**3 - 12*x
e4(x) = 16*x**4 - 48*x**2 + 12
e5(x) = 32*x**5 - 160*x**3 + 120*x
```



Exercise 6.19. Find the matrix representation of the discrete Fourier transform, that is, the basis change matrix from the standard basis to $\{\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_{n-1}\}$.

Let denote

$$\boldsymbol{e}_k = \frac{1}{\sqrt{n}} \begin{bmatrix} e^{2\pi i k t_0} \\ e^{2\pi i k t_1} \\ \vdots \\ e^{2\pi i k t_{n-1}} \end{bmatrix} \qquad (k = 0, 1, 2, \ldots, n-1).$$

Since

$$\hat{\boldsymbol{x}} = \begin{bmatrix} \langle \boldsymbol{e}_0 \mid \boldsymbol{x} \rangle \\ \langle \boldsymbol{e}_1 \mid \boldsymbol{x} \rangle \\ \vdots \\ \langle \boldsymbol{e}_{n-1} \mid \boldsymbol{x} \rangle \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_0 & \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_{n-1} \end{bmatrix}^* \boldsymbol{x},$$

it follows that $\begin{bmatrix} \boldsymbol{e}_0 & \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_{n-1} \end{bmatrix}^* = \begin{bmatrix} \boldsymbol{e}_0 & \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_{n-1} \end{bmatrix}^{-1}$ is the representation matrix of the Fourier tarnform, which is

$$\begin{bmatrix} \boldsymbol{e}_0 & \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_{n-1} \end{bmatrix}^* = \begin{bmatrix} \boldsymbol{e}_0^* \\ \boldsymbol{e}_1^* \\ \boldsymbol{e}_2^* \\ \vdots \\ \boldsymbol{e}_{n-1}^* \end{bmatrix}$$

$$= \frac{1}{\sqrt{n}} \begin{bmatrix} e^{-2\pi i\cdot 0\cdot t_0} & e^{-2\pi i\cdot 1\cdot t_0} & e^{-2\pi i\cdot 2\cdot t_0} & \cdots & e^{-2\pi i(n-1)t_0} \\ e^{-2\pi i\cdot 0\cdot t_1} & e^{-2\pi i\cdot 1\cdot t_1} & e^{-2\pi i\cdot 2\cdot t_1} & \cdots & e^{-2\pi i(n-1)t_1} \\ e^{-2\pi i\cdot 0\cdot t_2} & e^{-2\pi i\cdot 1\cdot t_2} & e^{-2\pi i\cdot 2\cdot t_2} & \cdots & e^{-2\pi i(n-1)t_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{-2\pi i\cdot 0\cdot t_{n-1}} & e^{-2\pi i\cdot 1\cdot t_{n-1}} & e^{-2\pi i\cdot 2\cdot t_{n-1}} & \cdots & e^{-2\pi i(n-1)t_{n-1}} \end{bmatrix}.$$

Especially, if $t_0, t_1, \ldots, t_{n-1}$ are $0, \dfrac{1}{n}, \ldots, \dfrac{n-1}{n}$ respectively, then it is

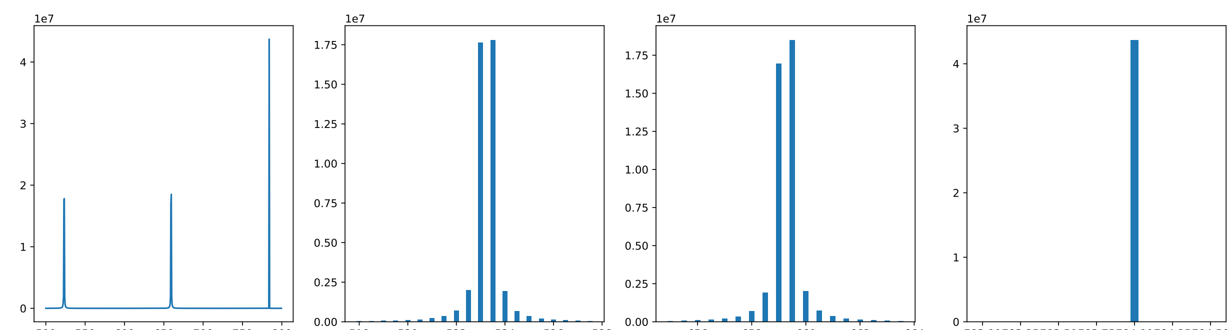
$$= \frac{1}{\sqrt{n}} \begin{bmatrix} e^{-2\pi i 0\cdot 0/n} & e^{-2\pi i\cdot 1\cdot 0/n} & e^{-2\pi i\cdot 2\cdot 0/n} & \cdots & e^{-2\pi i(n-1)\cdot 0/n} \\ e^{-2\pi i\cdot 0\cdot 1/n} & e^{-2\pi i\cdot 1\cdot 1/n} & e^{-2\pi i\cdot 2\cdot 1/n} & \cdots & e^{-2\pi i(n-1)1/n} \\ e^{-2\pi i\cdot 0\cdot 2/n} & e^{-2\pi i\cdot 1\cdot 2/n} & e^{-2\pi i\cdot 2\cdot 2/n} & \cdots & e^{-2\pi i(n-1)2/n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{-2\pi i\cdot 0(n-1)/n} & e^{-2\pi i\cdot 1(n-1)/n} & e^{-2\pi i\cdot 2(n-1)/n} & \cdots & e^{-2\pi i(n-1)(n-1)/n} \end{bmatrix}.$$

Exercise 6.20. Though the sounds of the notes C, E, and G were added at the same ratio, the peak pitches are different. This is because their frequencies are not integers. Change Line 8 of the above program to zoom in near the peaks of frequency and observe.

Program: `power_spectrum.py`

In [1]:
```
1  import matplotlib.pyplot as plt
2  from sound import Sound
3
4  sound = Sound('CEG')
5  fig, ax = plt.subplots(1, 4, figsize=(20, 5))
6  ax[0].plot(*sound.power_spectrum((500, 800)))
7  ax[1].bar(*sound.power_spectrum((518, 528)), width=0.2)
8  ax[2].bar(*sound.power_spectrum((655, 664)), width=0.2)
9  ax[3].bar(*sound.power_spectrum((783, 785)), width=0.05)
10 plt.show()
```

Line 2: We use the program `sound.py` given later in this section as a library.
Line 4: Load the wav file `CEG.wav` of the cord CEG created by the program `chord.py` in Section 2.4.
Line 5: Display four graphs side by side.
Line 6: The frequencies of the C(do), E(mi) and G(sol) scales are 523.251131 Hertz, 659.255114 Hertz and 783.990872 respectively. Check the graphs to see that the spectrum appears around these frequencies (the first graph from the left in the figure below).
Line 7: Check the spectrum around 523.251131 Hertz on the graph (the second one).
Line 8: Check the spectrum around 659.255114 Hertz on the graph (the third one).
Line 9: Check the spectrum around 783.990872 Hertz on the graph (the fourth one).



Since the frequency of the G sound is close to an integer value, only one spectrum of the G sound appears. The frequencies of the other two sounds are apart from integer values, so the spectrum appears distributed around their respective frequencies. Adding up the lengths of the scattered sticks for both the C and E sounds should equal the height of the stick for the G sound (since the strengths of the sounds are the same, according to the Riesz-Fisher equality).

Exercise 6.21. Give a reason why the power spectrum appears symmetrically around 0. Hint: This is generally true for the discrete Fourier transform $\widehat{\boldsymbol{x}} \in \mathbb{C}^n$ of real vector $\boldsymbol{x} \in \mathbb{R}^n$ (but not for complex vector $\boldsymbol{x} \in \mathbb{C}^n$).

Since f is real-valued,

$$\left\langle e_k \mid f \right\rangle = \int_0^1 e^{-2\pi iks} f(s)\, ds = \int_0^1 \cos(2\pi ks) f(s)\, ds - i \int_0^1 \sin(2\pi ks) f(s)\, ds,$$

it follows that

$$\begin{aligned}
\left|\left\langle e_k \mid f \right\rangle\right|^2 &= \left(\int_0^1 \cos(2\pi ks) f(s)\, ds\right)^2 + \left(\int_0^1 \sin(2\pi ks) f(s)\, ds\right)^2 \\
&= \left(\int_0^1 \cos(2\pi ks) f(s)\, ds\right)^2 + \left(-\int_0^1 \sin(2\pi ks) f(s)\, ds\right)^2 \\
&= \left(\int_0^1 \cos(-2\pi ks) f(s)\, ds\right)^2 + \left(\int_0^1 \sin(-2\pi ks) f(s)\, ds\right)^2 \\
&= \left|\left\langle e_{-k} \mid f \right\rangle\right|^2 .
\end{aligned}$$

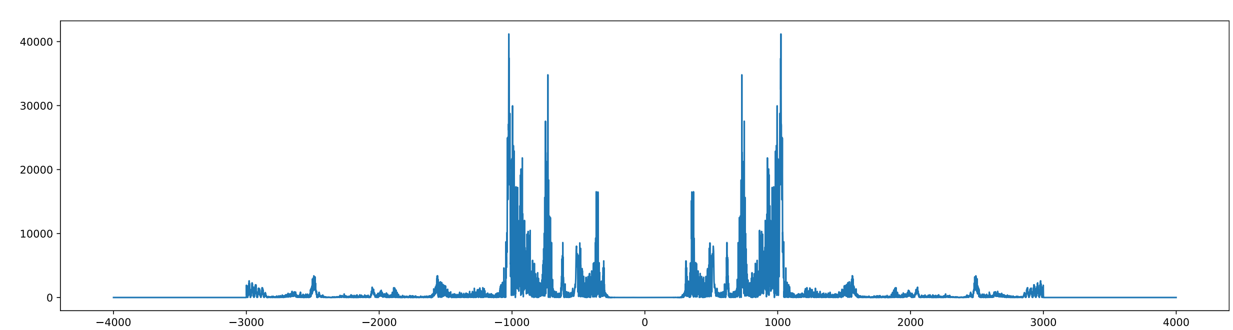
Exercise 6.22. Listen to the sound of `mono3000.wav` and compare it with the original sound `mono.wav`. Also, display the power spectrum of the sound of `mono3000.wav`. Moreover, experiment with different cutoff frequencies to observe what happens.

Program: `lowpass.py`

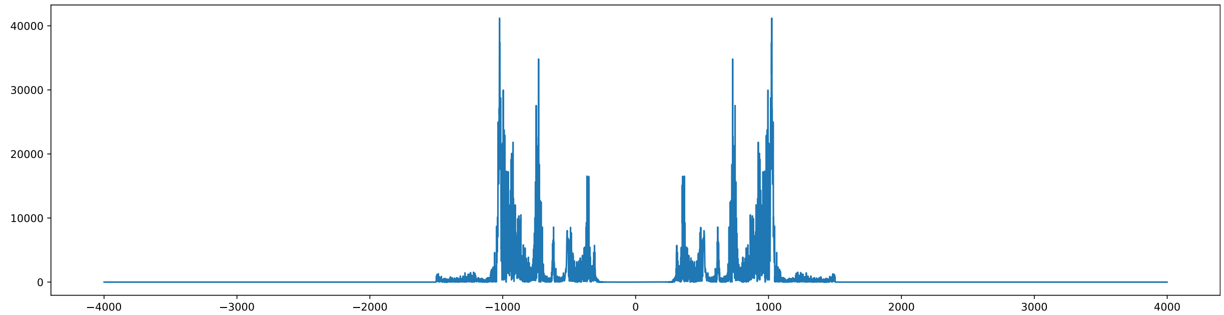
In [1]:
```
1  import matplotlib.pyplot as plt
2  from sound import Sound
3
4  K = 3000
5  sound = Sound('mono')
6  X, Y = sound.time, sound.data
7  Y_K = sound.lowpass(K)
8
9  fig, ax = plt.subplots(1, 2, figsize=(20, 5), dpi=100)
10 ax[0].plot(X, Y), ax[0].plot(X, Y_K)
11 ax[0].set_ylim(-1, 1)
12 ax[1].plot(X, Y), ax[1].plot(X, Y_K)
13 ax[1].set_xlim(0.2, 0.21), ax[1].set_ylim(-1, 1)
14 plt.show()
```

By changing the value of `K` in Line 4 of the program (`lowpass.py`), we can observe the waveform of sound that has passed through a low-pass filter with various cutoff frequencies, and can also create a wav file of that sound. The program `sound.py` of Section 6.7 is used as a library, so copy and put it in the current folder.

In [1]:
```
import matplotlib.pyplot as plt
from sound import Sound

sound = Sound('mono3000')
plt.figure(figsize=(20, 5))
plt.plot(*sound.power_spectrum((-4000, 4000)))
plt.show()
```



Cutoff frequency 3000



Cutoff frequency 1500

Exercise 6.23. Check that the inverse Fourier transforms of `U` in Line 28 is an array of complex numbers with imaginary part 0 by drawing the graph of `fft.ifft(U).imag`. Consider the mathematical reason why the low-pass filter of a real vector by the Fourier transform and its inverse are also real vectors.

Program: `sound.py`

In [1]:

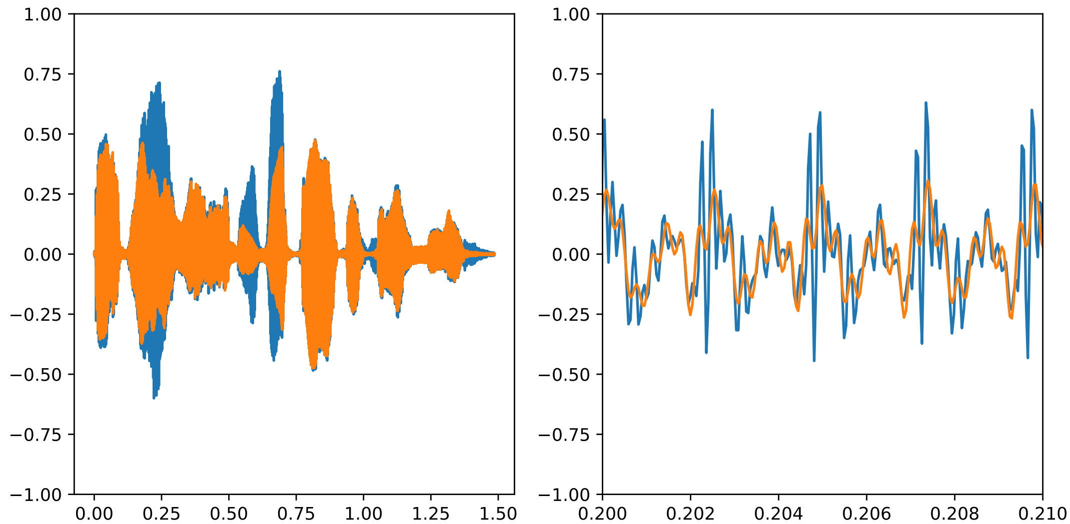
```
1  from numpy import arange, fft
2  import scipy.io.wavfile as wav
3  
4  
5  class Sound:
6      def __init__(self, wavfile):
7          self.file = wavfile
8          self.rate, Data = wav.read(f'{wavfile}.wav')
9          dt = 1 / self.rate
10         self.len = len(Data)
11         self.tmax = self.len / self.rate
12         self.time = arange(0, self.tmax, dt)
13         self.data = Data.astype('float') / 32768
14         self.fft = fft.fft(self.data)
15 
16     def power_spectrum(self, rng=None):
17         spectrum = abs(self.fft) ** 2
18         if rng is None:
19             r1, r2 = -self.len / 2, self.len / 2
20         else:
21             r1, r2 = rng[0] * self.tmax, rng[1] * self.tmax
22         R = arange(int(r1), int(r2))
23         return R / self.tmax, spectrum[R]
24 
25     def lowpass(self, K):
26         k = int(K * self.tmax)
27         U = self.fft.copy()
28         U[range(k + 1, self.len - k)] = 0
29         V = fft.ifft(U)
30         Data = (V.real * 32768).astype('int16')
31         wav.write(f'{self.file}{K}.wav', self.rate, Data)
32         return V.real, V.imag
```

Program: `lowpass.py`

In [1]:

```
import matplotlib.pyplot as plt
from sound import Sound

sound = Sound('mono')
X, Y = sound.time, sound.data
```

In [1]:
```
Y3000, Z3000 = sound.lowpass(3000)

fig, ax = plt.subplots(1, 2, figsize=(10, 5))
ax[0].set_ylim(-1, 1)
ax[0].plot(X, Y), ax[0].plot(X, Y3000), ax[0].plot(X, Z3000)
ax[1].set_xlim(0.2, 0.21), ax[1].set_ylim(-1, 1)
ax[1].plot(X, Y), ax[1].plot(X, Y3000), ax[1].plot(X, Z3000)
plt.show()
```



In the graphs of the original data, the real part of a low-pass filter with a cutoff frequency of 3000, and its imaginary part are overlaid. The one on the right is an enlarged part. It can be seen that the imaginary parts are all identically the zero function.

Since f is real-valued,

$$\langle e_k \mid f \rangle = \int_0^1 e^{-2\pi iks} f(s)\, ds = \int_0^1 \cos(2\pi ks) f(s)\, ds - i \int_0^1 \sin(2\pi ks) f(s)\, ds$$

it follows that

$$\begin{aligned} & \langle e_k \mid f \rangle e_k(t) \\ = \; & \cos(2\pi kt) \int_0^1 \cos(2\pi ks) f(s)\, ds + \sin(2\pi kt) \int_0^1 \sin(2\pi ks) f(s)\, ds \\ & -i \cos(2\pi kt) \int_0^1 \sin(2\pi ks) f(s)\, ds + i \sin(2\pi kt) \int_0^1 \cos(2\pi ks) f(s)\, ds. \end{aligned}$$

Hence, we have

$$\begin{aligned} & \langle e_k \mid f \rangle e_k(t) + \langle e_{-k} \mid f \rangle e_{-k}(t) \\ = \; & 2\cos(2\pi kt) \int_0^1 \cos(2\pi ks) f(s)\, ds + 2\sin(2\pi kt) \int_0^1 \sin(2\pi ks) f(s)\, ds \end{aligned}$$

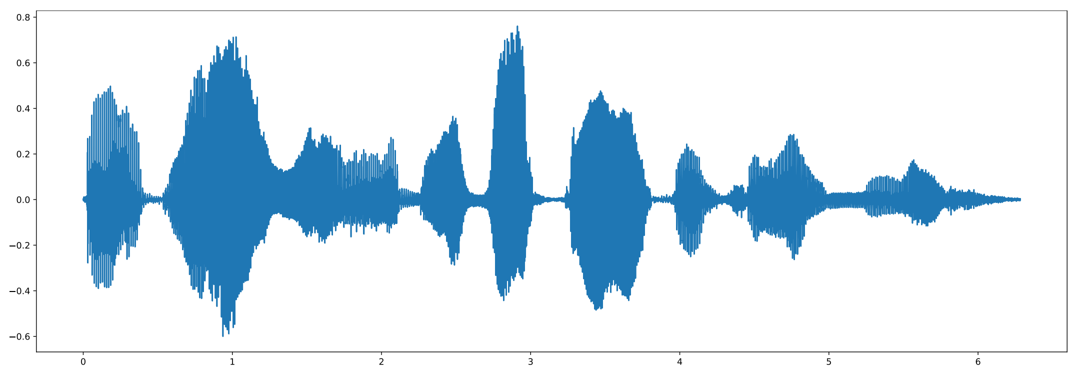
for $k > 0$. On the other hand, becouse of

$$\left\langle e_0 \mid f \right\rangle e_0(t) = \int_0^1 f(s)\, ds$$

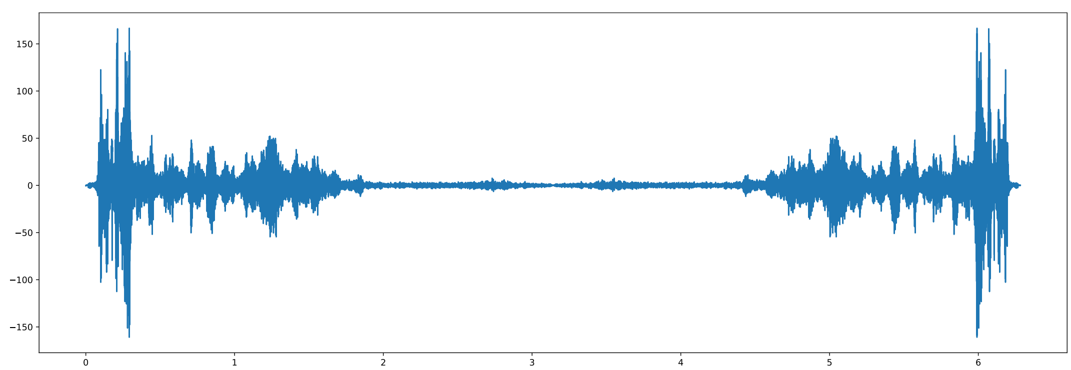
we can conclude that

$$\sum_{k=-k}^{k} \left\langle e_k \mid f \right\rangle e_k(t) = \text{real.}$$

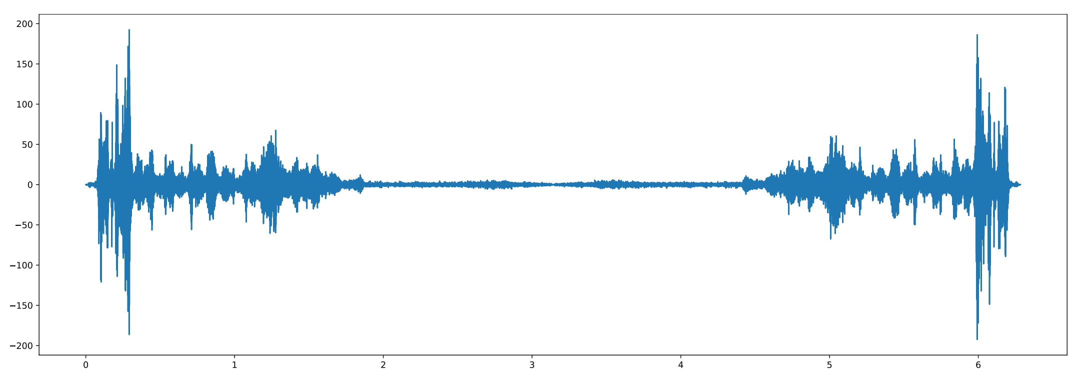
Observe the graph after the Fourier transform.



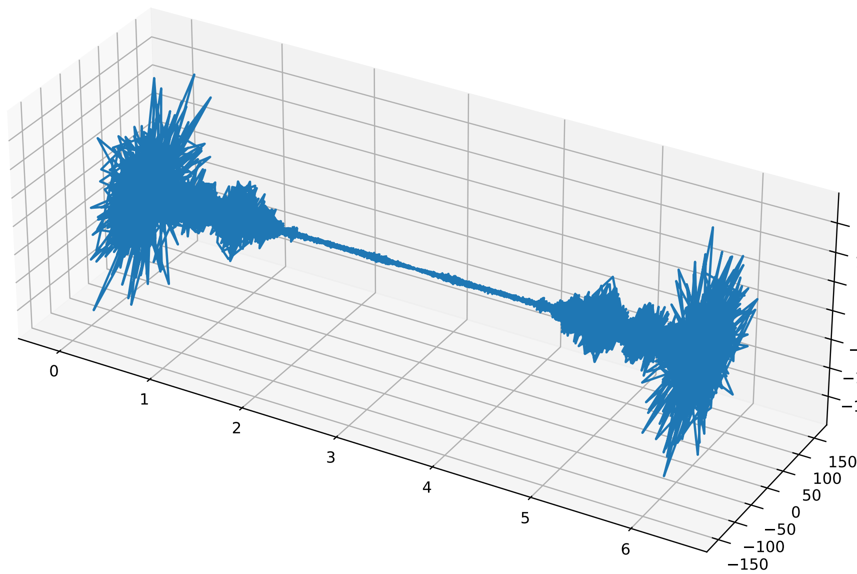
Graph of original data



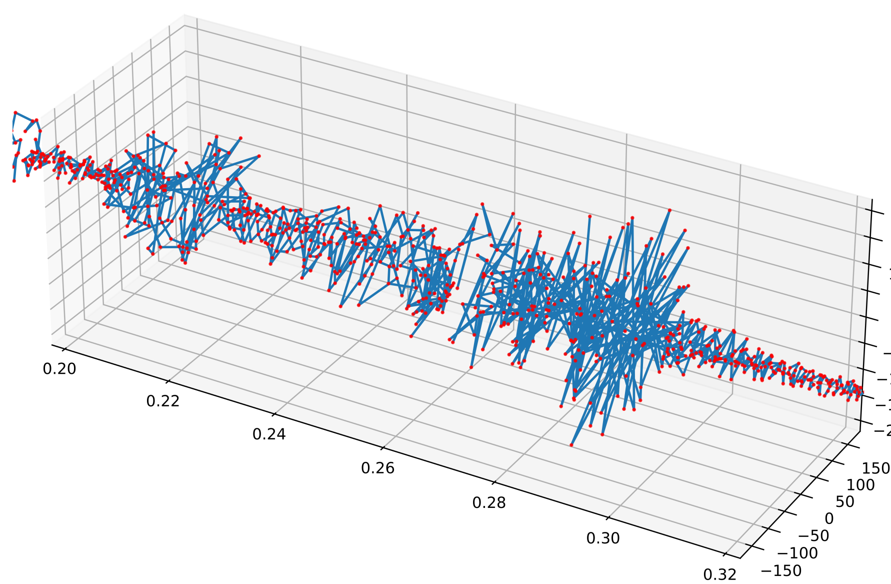
Graph of the real part of the Fourier transform



Graph of the imaginary part of the Fourier transform



Graph of the Fourier transform

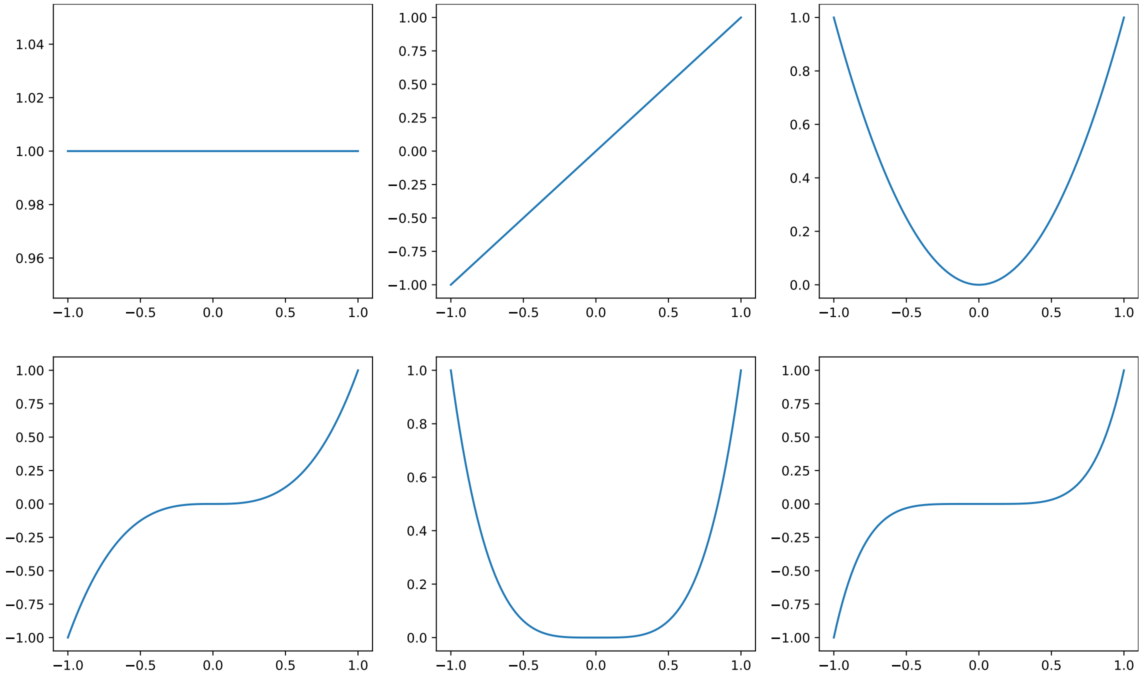


Enlarged part of the graph of the Fourier transform

Actually, the red dots represents the values of the function. The blue line connects them to form a line graph.

Supplement: How to divide the figure and draw multiple graphs with Matplotlib

Suppose we want to draw a graph like this:



It's a little confusing, but there are two ways: using the `subplots` method and using the `subplot` method.

1. using the `subplots`:

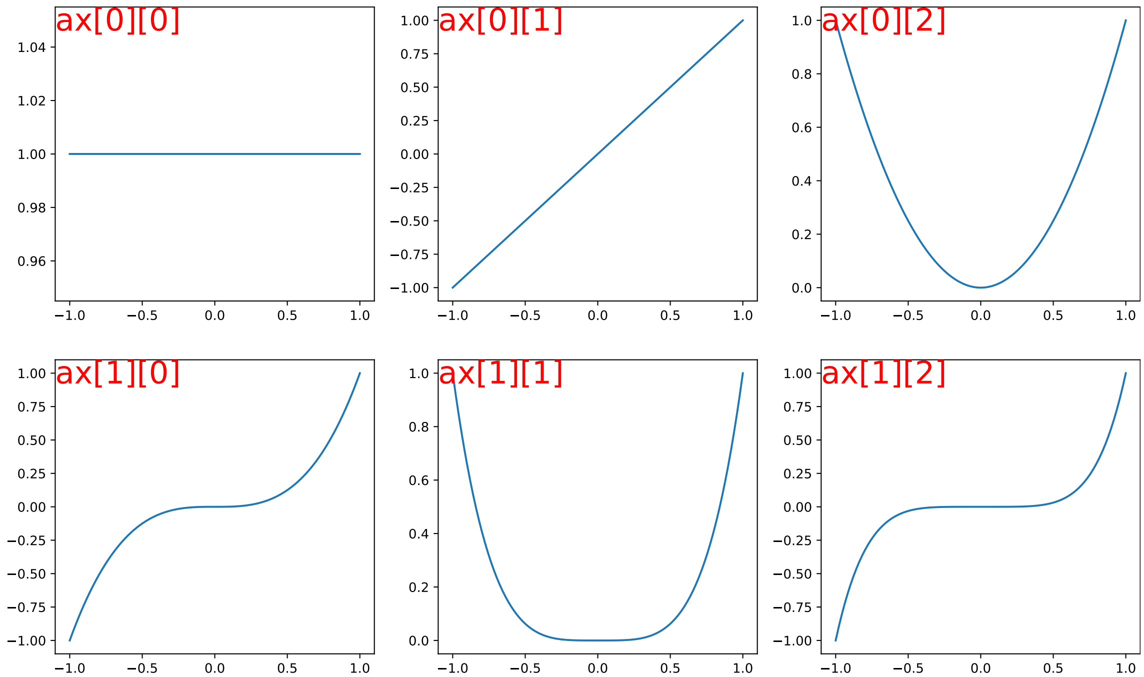
   Program: `graph1.py`

In [1]:

```
1  from numpy import *
2  import matplotlib.pyplot as plt
3
4  fig, ax = plt.subplots(2, 3, figsize=(9, 6))
5  x = linspace(-1, 1, 100)
6  for i in range(2):
7      for j in range(3):
8          ax[i][j].plot(x, x**(i*3+j))
9          xmin, xmax = ax[i][j].get_xlim()
10         ymin, ymax = ax[i][j].get_ylim()
11         ax[i][j].text(xmin, ymax, f'ax[{i}][{j}]', color='r',
12                       fontsize=24, ha='left', va='top')
13 plt.show()
```

Line 4: Prepare to draw the figure by dividing it into 2 rows and 3 columns. The figure size specification `(9, 6)` means that the width is 9, the height is 6, and the unit is inches. Note that the vertical and horizontal sides are reversed.

Line 8: Plot the graphs on the area of the i-th row and the j-th column for $i = 0, 1$ and $j = 0, 1, 2$.



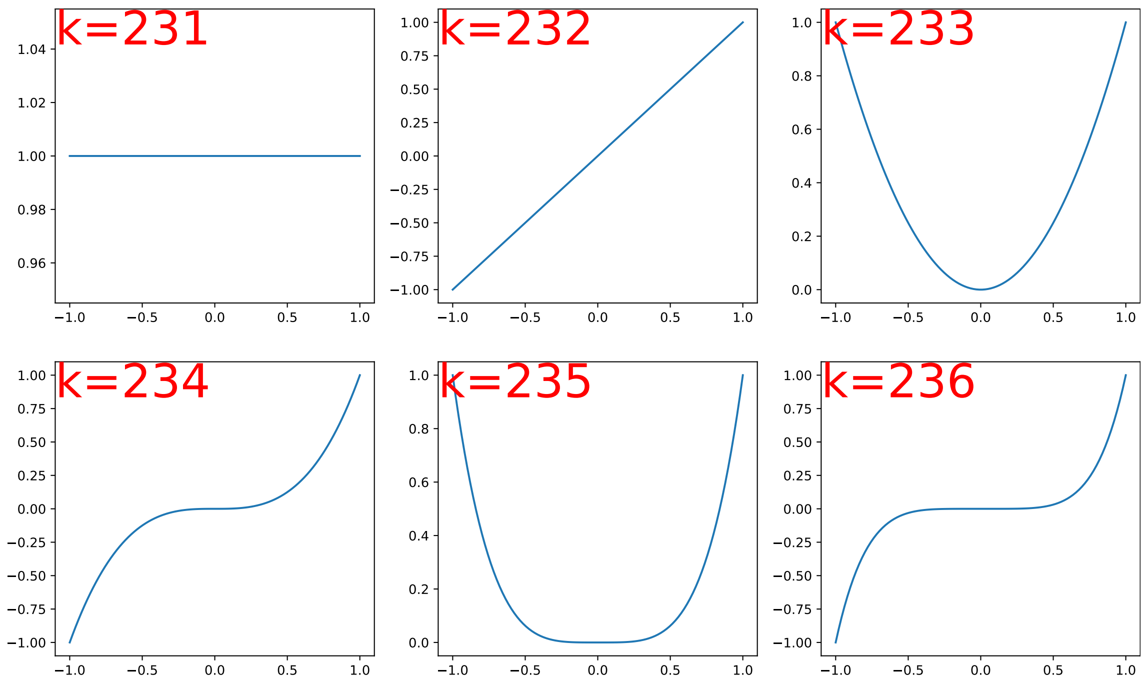
2. using the `subplot` method:

   Program: `graph2.py`

In [1]:
```
1  from numpy import *
2  import matplotlib.pyplot as plt
3
4  fig = plt.figure(figsize=(9, 6))
5  x = linspace(-1, 1, 100)
6  k = 230
7  for i in range(2):
8      for j in range(3):
9          k += 1
10         plt.subplot(k)
11         plt.plot(x, x**(i*3+j))
12         xmin, xmax = plt.xlim()
13         ymin, ymax = plt.ylim()
14         plt.text(xmin, ymax, f'k={k}', color='r',
15                  fontsize=36, ha='left', va='top')
16 plt.show()
```

Line 4: Specifies the size of the figure.

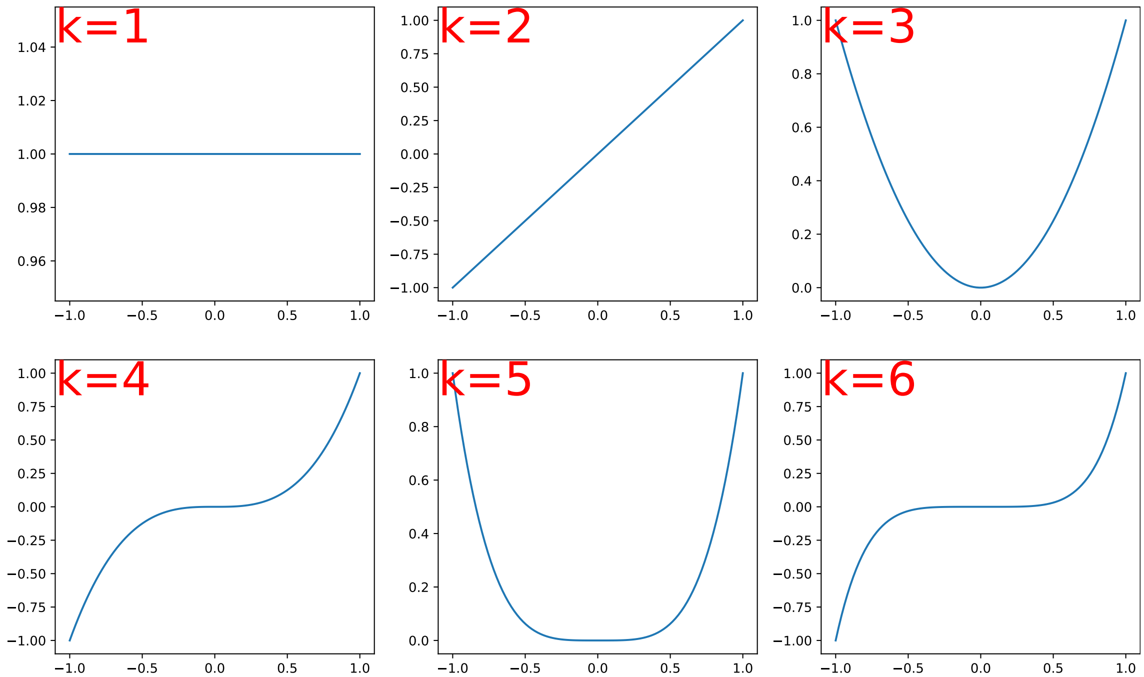
Line 10, 11: `plt.subplot(k)` switches the drawing region. In this example, $k = 231, 232, 233, 234, 235, 236$. The drawing area is specified by this 3-digit integer $d_1d_2d_3$, which means the d_3th region from the upper left corner to the lower right corner with horizontal priority in the region divided into d_1 rows and d_2 columns. d_1, d_2, d_3 must be integers between 1 and 9. After switching, draw the graphs with the `plot` method.



Program: `graph3.py`

In [1]:
```
1  import matplotlib.pyplot as plt
2
3  fig = plt.figure(figsize=(9, 6))
4  x = linspace(-1, 1, 100)
5  k = 0
6  for i in range(2):
7      for j in range(3):
8          k += 1
9          plt.subplot(2, 3, k)
10         plt.plot(x, x**(i*3+j))
11         xmin, xmax = plt.xlim()
12         ymin, ymax = plt.ylim()
13         plt.text(xmin, ymax, f'k={k}', color='r',
14                  fontsize=36, ha='left', va='top')
15 plt.show()
```

Line 10: If the drawing location cannot be represented by a 3-digit integer, use this line. In this example, $k = 1, 2, 3, 4, 5, 6$. The drawing location is the kth area from the upper left corner to the lower right corner with horizontal priority. Note that k starts at 1.



Therefore, using the second method described above, `lowpass.py` after line 8 can also be rewritten as follows.

Program: `lowpass.py`

In [1]:
```
1 fig = plt.figure(figsize=(10, 5))
2 plt.subplot(121)
3 plt.ylim(-1, 1)
4 plt.plot(X, Y), plt.plot(X, Y3000), plt.plot(X, Z3000)
5 plt.subplot(122)
6 plt.xlim(0.2, 0.21), plt.ylim(-1, 1)
7 plt.plot(X, Y), plt.plot(X, Y3000), plt.plot(X, Z3000)
8 plt.show()
```

Note that lines 10 and 13 specify the axis limits differently.

Chapter 7

Eigenvalues and Eigenvectors

Exercise 7.1	Sect. 7.1. p.156	Properties of unitary matrices
Exercise 7.2	Sect. 7.1. p.156	Diagonal matrices
Exercise 7.3	Sect. 7.1. p.157	Positive definiteness of real matrices of order 2
Exercise 7.4	Sect. 7.1. p.158	Eigenvalues of rthogonal projevtions
Exercise 7.5	Sect. 7.2. p.160	Eigenvalues and regularity of matrices
Exercise 7.6	Sect. 7.2. p.162	Computing eigenvalues and eigenvectors in 2D
Exercise 7.7	Sect. 7.2. p.165	Computing eigenvalues and eigenvectors in 3D
Exercise 7.8	Sect. 7.3. p.167	Computation of diagonalization
Exercise 7.9	Sect. 7.3. pp.167,168	Eigenvalues and invariants of matrices
Exercise 7.10	Sect. 7.3. p.168	When is a real matrix of order 2 normal?
Exercise 7.11	Sect. 7.3. p.170	Unitary invariants
Exercise 7.12	Sect. 7.3. p.172	Diagonalization of randomly generated matrices
Exercise 7.13	Sect. 7.4. p.179	The exponential of matrices
Exercise 7.14	Sect. 7.4. p.180	Computing powers and the exponential of a matrix
Exercise 7.15	Sect. 7.4. pp.180,181	Power series of a matrix

Exercise 7.1. Prove that $\boldsymbol{A}$ is a unitary matrix if it preserves the inner product.

Let $\boldsymbol{A}$ be a square matrix of order n. If $\boldsymbol{A}$ preserves the inner product, then

$$\left\langle \boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x} \;\middle|\; \boldsymbol{y} \right\rangle = \left\langle \boldsymbol{A}\boldsymbol{x} \;\middle|\; \boldsymbol{A}\boldsymbol{y} \right\rangle = \left\langle \boldsymbol{x} \;\middle|\; \boldsymbol{y} \right\rangle$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{K}^n$. Therefore $\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{I}$. From this equality, we have $\boldsymbol{A}^* = \boldsymbol{A}^{-1}$.

$\boldsymbol{A}$ is a unitary matrix if $\boldsymbol{A}$ preserves the norm, using the polarization identity given in Exercise 6.6.

Exercise 7.2. For a diagonal matrix $\boldsymbol{A} = \mathbf{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n)$, show the following equivalences:

(1) $\boldsymbol{A}$ is a Hermitian matrix $\Leftrightarrow$ $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$,

(2) $\boldsymbol{A}$ is a unitary matrix $\Leftrightarrow$ $|\lambda_1| = \cdots = |\lambda_n| = 1$,

(3) $\boldsymbol{A}$ is a positive definite matrix $\Leftrightarrow$ $\lambda_1, \ldots, \lambda_n > 0$,

(4) $\boldsymbol{A}$ is a positive semi-definite matrix $\Leftrightarrow$ $\lambda_1, \ldots, \lambda_n \geqq 0$.

(1)

$$\begin{aligned}\boldsymbol{A} \text{ is a Hermitian matrix } &\Leftrightarrow \mathbf{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n) = \mathbf{diag}\left(\overline{\lambda_1}, \overline{\lambda_2}, \ldots, \overline{\lambda_n}\right) \\ &\Leftrightarrow \lambda_1 = \overline{\lambda_1},\ \lambda_2 = \overline{\lambda_2},\ \ldots,\ \lambda_n = \overline{\lambda_n} \\ &\Leftrightarrow \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}.\end{aligned}$$

(2)

$$\begin{aligned}\boldsymbol{A} \text{ is a unitary matrix } &\Leftrightarrow \mathbf{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n)\,\mathbf{diag}\left(\overline{\lambda_1}, \overline{\lambda_2}, \ldots, \overline{\lambda_n}\right) = \boldsymbol{I} \\ &\Leftrightarrow \mathbf{diag}\left(\lambda_1\overline{\lambda_1},\ \lambda_2\overline{\lambda_2},\ \ldots,\ \lambda_n\overline{\lambda_n}\right) = \boldsymbol{I} \\ &\Leftrightarrow |\lambda_1|^2 = |\lambda_2|^2 = \cdots = |\lambda_n|^2 = 1 \\ &\Leftrightarrow |\lambda_1| = |\lambda_2| = \cdots = |\lambda_n| = 1.\end{aligned}$$

(3)

$$\begin{aligned}&\boldsymbol{A} \text{ is a positive definite matrix} \\ \Leftrightarrow\ &\left\langle \boldsymbol{x} \,\middle|\, \mathbf{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n)\,\boldsymbol{x} \right\rangle > 0 \text{ for all } \boldsymbol{x} \in \mathbb{K}^n \text{ if } \boldsymbol{x} \neq \boldsymbol{0} \\ \Leftrightarrow\ &\lambda_1 |x_1|^2 + \lambda_2 |x_2|^2 + \cdots + \lambda_n |x_n|^2 > 0 \text{ for all } x_1, x_2, \ldots, x_n \in \mathbb{K} \text{ if } x_1 x_2 \cdots x_n \neq 0 \\ \Leftrightarrow\ &\lambda_1, \lambda_2, \ldots, \lambda_n > 0.\end{aligned}$$

(4)

$$\begin{aligned}&\boldsymbol{A} \text{ is a positive semi-definite matrix} \\ \Leftrightarrow\ &\left\langle \boldsymbol{x} \,\middle|\, \mathbf{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n)\,\boldsymbol{x} \right\rangle \geqq 0 \text{ for all } \boldsymbol{x} \in \mathbb{K}^n \\ \Leftrightarrow\ &\lambda_1 |x_1|^2 + \lambda_2 |x_2|^2 + \cdots + \lambda_n |x_n|^2 \geqq 0 \text{ for all } x_1, x_2, \ldots, x_n \in \mathbb{K} \\ \Leftrightarrow\ &\lambda_1, \lambda_2, \ldots, \lambda_n \geqq 0.\end{aligned}$$

Exercise 7.3. Let $a, b, c \in \mathbb{R}$ and consider a symmetric matrix $\boldsymbol{A} = \begin{bmatrix} a & c \\ c & b \end{bmatrix}$. Prove that $\boldsymbol{A}$ is positive definite over $\mathbb{R}$ if and only if $a > 0, b > 0$ and $ab > c^2$.

$$\begin{aligned} & \boldsymbol{A} \text{ is positive definite over } \mathbb{R} \\ \Leftrightarrow \quad & \langle \boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} \rangle > 0 \text{ for all } \boldsymbol{x} \in \mathbb{R}^2 \text{ with } \boldsymbol{x} \neq \boldsymbol{0} \\ \Leftrightarrow \quad & \left\langle \begin{bmatrix} x \\ y \end{bmatrix} \middle| \begin{bmatrix} a & c \\ c & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right\rangle > 0 \text{ for all } x, y \in \mathbb{R} \text{ with } x^2 + y^2 \neq 0 \\ \Leftrightarrow \quad & ax^2 + 2cxy + by^2 > 0 \text{ for all } x, y \in \mathbb{R} \text{ with } x^2 + y^2 \neq 0 \end{aligned}$$

We obtain the desired condition by applying the quadratic discriminant $D = 4\left(c^2 - ab\right)$ to the last condition.

Exercise 7.4. Let $\boldsymbol{P}$ be an orthogonal projection in $\mathbb{K}^n$ and let $\boldsymbol{x}\,(\neq \mathbf{0}) \in \mathbb{K}^n$. Show that a scalar $a \in \mathbb{K}$ satisfying $\boldsymbol{Px} = a\boldsymbol{x}$ is only 0 or 1.

Put $\boldsymbol{P} = \mathbf{proj}_W$ and consider the orthogonal decomposition $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$ by $\boldsymbol{x}_1 \in W$ and $\boldsymbol{x}_2 \in W^{\perp}$. Because of $\boldsymbol{x} \neq \mathbf{0}$, at least one of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is not a zero vector. Suppose $\boldsymbol{Px} = a\boldsymbol{x}$. Since $\boldsymbol{Px} = \boldsymbol{x}_1$, it follows that $\boldsymbol{x}_1 = a\boldsymbol{x}_1 + a\boldsymbol{x}_2$. Therefore, we have

$$(a-1)\,\boldsymbol{x}_1 + a\boldsymbol{x}_2 \;=\; \mathbf{0}.$$

Assume that both $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are nonzero vectors. Since these must be linearly independent, we have $a - 1 = a = 0$, which is a contradiction. Therefore, one of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is the zero vector and the other is not. If $\boldsymbol{x}_1 = \mathbf{0}$, then $a\boldsymbol{x}_2 = \mathbf{0}$ and $\boldsymbol{x}_2 \neq \mathbf{0}$, so we get $a = 0$. If $\boldsymbol{x}_2 = \mathbf{0}$, then$(a-1)\,\boldsymbol{x}_2 = \mathbf{0}$ and $\boldsymbol{x}_1 \neq \mathbf{0}$, so we get $a - 1 = 0$, that is, $a = 1$.

Exercise 7.5. Prove that a matrix $\boldsymbol{A}$ is regular if and only if $\boldsymbol{A}$ does not have zero eigenvalue.

Let $\boldsymbol{A}$ be a square matrix of order n. Then, $\boldsymbol{f} : \boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ is a linear mapping of $\boldsymbol{f} : \mathbb{K}^n \to \mathbb{K}^n$ and satisfies

$$\begin{aligned}
\boldsymbol{A} \text{ does not have zero eigenvalue} &\Leftrightarrow \boldsymbol{A}\boldsymbol{x} = 0\boldsymbol{x} = \mathbf{0} \text{ implies } \mathbf{0} \\
&\Leftrightarrow \operatorname{kernel}(\boldsymbol{f}) = \{\mathbf{0}\} \\
&\Leftrightarrow \boldsymbol{f} \text{ is injective} \\
&\Leftrightarrow \boldsymbol{f} \text{ is bijective (Remark after the dimension theorem)} \\
&\Leftrightarrow \boldsymbol{A} \text{ is regular.}
\end{aligned}$$

Exercise 7.6. Compute the eigenvalues and eigenvectors of a matrix generated by the program above without the aid of a computer. Next, solve the same problem with the aid of a computer.

Solve for $\begin{bmatrix} 3 & 2 \\ 1 & 3 \end{bmatrix}$. From

$$\begin{vmatrix} 3-\lambda & 2 \\ 1 & 3-\lambda \end{vmatrix} = 0,$$

the eigenpolynomial is $\lambda^2 - 6\lambda + 7 = 0$. Solving this gives eigenvalues $\lambda = 3 - \sqrt{2}, 3 + \sqrt{2}$.

Find the eigenvector for the eigenvalue $3 - \sqrt{2}$. For this we solve

$$\begin{bmatrix} 3 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \left(3 - \sqrt{2}\right) \begin{bmatrix} x \\ y \end{bmatrix}$$

and get $x = -\sqrt{2}y$, where y is an arbtrary constant. For instance, putting $y = 1$, we have an eigenvector $\begin{bmatrix} -\sqrt{2} \\ 1 \end{bmatrix}$ for the eigenvalue $3 - \sqrt{2}$.

Find the eigenvector for the eigenvalue $3 + \sqrt{2}$. For this we solve

$$\begin{bmatrix} 3 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \left(3 + \sqrt{2}\right) \begin{bmatrix} x \\ y \end{bmatrix}$$

and get $x = \sqrt{2}y$, where y is an arbtrary constant. For instance, putting $y = 1$, we have an eigenvector $\begin{bmatrix} \sqrt{2} \\ 1 \end{bmatrix}$ for the eigenvalue $3 + \sqrt{2}$.

Next, solve for $\begin{bmatrix} 3 & 3 \\ -3 & 2 \end{bmatrix}$. From

$$\begin{vmatrix} 3-\lambda & 3 \\ -3 & 2-\lambda \end{vmatrix} = 0,$$

the eigenpolynomial is $\lambda^2 - 5\lambda + 15 = 0$. Solving this gives eigenvalues $\lambda = \dfrac{5 - \sqrt{35}i}{2}, \dfrac{5 + \sqrt{35}i}{2}$.

To find the eigenvector for the eigenvalue $\dfrac{5 - \sqrt{35}i}{2}$, we solve

$$\begin{bmatrix} 3 & 3 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \frac{5 - \sqrt{35}i}{2} \begin{bmatrix} x \\ y \end{bmatrix}$$

and get $x = \dfrac{-y + \sqrt{35}iy}{6}$, where y is an arbitrary constant. Therefore, putting $y = 1$, we have an eigenvector $\begin{bmatrix} \frac{-1+\sqrt{35}i}{6} \\ 1 \end{bmatrix}$ for the eigenvalue $\dfrac{5 - \sqrt{35}i}{2}$.

To find the eigenvector for the eigenvalue $\dfrac{5+\sqrt{35}i}{2}$, we solve

$$\begin{bmatrix} 3 & 3 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \frac{5+\sqrt{35}i}{2} \begin{bmatrix} x \\ y \end{bmatrix}$$

and get $x = \dfrac{-y-\sqrt{35}iy}{6}$, where y is an arbitrary constant. Therefore, putting $y = 1$, we have an eigenvector $\begin{bmatrix} \frac{-1-\sqrt{35}i}{6} \\ 1 \end{bmatrix}$ for the eigenvalue $\dfrac{5+\sqrt{35}i}{2}$.

Exercise 7.7. The following program randomly generates a square matrix of order 3 whose characteristic polynomial can be factorized into linear factors over $\mathbb{Z}$, and so the eigenvalues are all integers. Compute the eigenvalues and eigenvectors of a matrix generated by the program without aid of a computer.

Solve for $\begin{bmatrix} -2 & -1 & -1 \\ -2 & 3 & -3 \\ -2 & -2 & -3 \end{bmatrix}$. From

$$\begin{vmatrix} -2-\lambda & -1 & -1 \\ -2 & 3-\lambda & -3 \\ -2 & -2 & -3-\lambda \end{vmatrix} = 0,$$

the eigenpolynomial is $-\lambda^3-2\lambda^2+19\lambda+20=0$. Foctorizing this, we have $-(\lambda-4)(\lambda+1)(\lambda+5)=0$ and then we get eigenvalues $\lambda=-5,-1,4$.

Find an eigenvector for eigenvalue -5. Considering

$$\begin{bmatrix} -2 & -1 & -1 \\ -2 & 3 & -3 \\ -2 & -2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = -5 \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

we get $x=\dfrac{z}{2}$, $y=\dfrac{z}{2}$ where z is an arbitrary constant. Therefore, if $y=2$, then $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ is an eigenvector for eigenvalue -5.

Find an eigenvector for eigenvalue -1. Considering

$$\begin{bmatrix} -2 & -1 & -1 \\ -2 & 3 & -3 \\ -2 & -2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = - \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

we get $x=-\dfrac{7z}{6}$, $y=\dfrac{z}{6}$ where z is an arbitrary constant. Therefore, if $z=6$, then we get an eigenvector $\begin{bmatrix} -7 \\ 1 \\ 6 \end{bmatrix}$ for eigenvalue -1.

Find an eigenvector for eigenvalue 4. Considering

$$\begin{bmatrix} -2 & -1 & -1 \\ -2 & 3 & -3 \\ -2 & -2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 4 \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

we get $x=\dfrac{z}{2}$, $y=-4z$ where z is an arbitrary constant. Therefore, if $z=2$, then we get a eigenvector $\begin{bmatrix} 1 \\ -8 \\ 2 \end{bmatrix}$ for eigenvalue 4.

Exercise 7.8. Make another matrix with `prob2.py`, and diagonalize it in the same manner as above.

Here, we will diagonalize the matrix $\begin{bmatrix} -2 & -1 & -1 \\ -2 & 3 & -3 \\ -2 & -2 & -3 \end{bmatrix}$ used in Question Exercise 7.7. Make $\boldsymbol{V}$ with the three eigenvectors as column vectors obtained above:

$$\boldsymbol{V} = \begin{bmatrix} 1 & -7 & 1 \\ 1 & 1 & -8 \\ 2 & 6 & 2 \end{bmatrix}.$$

Calculate its inverse matrix:

$$\boldsymbol{V}^{-1} = \frac{1}{180} \begin{bmatrix} 50 & 20 & 55 \\ -18 & 0 & 9 \\ 4 & -20 & 8 \end{bmatrix}.$$

Hence, we get the diagonalization

$$\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V} = \begin{bmatrix} -5 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

of $\boldsymbol{A}$.

The arrangement of diagonal components is optional. Making $\boldsymbol{V}$ by arranging the eigenvector corresponding to 4, -1, -5,

$$\boldsymbol{V} = \begin{bmatrix} 1 & -7 & 1 \\ -8 & 1 & 1 \\ 2 & 6 & 2 \end{bmatrix}.$$

Its inverse matrix is

$$\boldsymbol{V}^{-1} = \frac{1}{180} \begin{bmatrix} 4 & -20 & 8 \\ -18 & 0 & 9 \\ 50 & 20 & 55 \end{bmatrix}.$$

Hence, we get diagonalization

$$\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -5 \end{bmatrix}.$$

Exercise 7.9. Suppose that a matrix $\boldsymbol{A}$ is diagonalized as $\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V} = \mathbf{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ by a regular matrix $\boldsymbol{V}$. Prove the following results from (1) to (4):

(1) $\mathrm{rank}(\boldsymbol{A})$ is equal to the number of nonzero elements of $\lambda_1, \lambda_2, \ldots, \lambda_n$,

(2) $\det(\boldsymbol{A}) = \lambda_1\lambda_2\cdots\lambda_n$,

(3) $\mathrm{Tr}(\boldsymbol{A}) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$,

(4) $\boldsymbol{A}^{-1} = \boldsymbol{V}\,\mathbf{diag}\left(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1}\right)\boldsymbol{V}^{-1}$, if $\boldsymbol{A}$ is regular.

(1) Because rank is an invariant, $\mathrm{rank}\,\boldsymbol{A} = \mathrm{rank}\,\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Since

$$\mathrm{range}\left(\mathbf{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)\right) = \left\{ \begin{bmatrix} \lambda_1 x_1 \\ \lambda_2 x_2 \\ \vdots \\ \lambda_n x_n \end{bmatrix} \middle| \; x_1, x_2, \ldots, x_n \in \mathbb{K} \right\},$$

the dimension of this subspace is the number of non-zero values among $\lambda_1, \lambda_2, \ldots$ and λ_n.

(2) Because det is an invariant,

$$\det(\boldsymbol{A}) = \det\left(\mathbf{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)\right) = \lambda_1\lambda_2\cdots\lambda_n.$$

(3) Because Tr is an invariant,

$$\mathrm{Tr}(\boldsymbol{A}) = \mathrm{Tr}\left(\mathbf{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)\right) = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

(4) If $\boldsymbol{A}$ be a regular matrix, none of $\lambda_1, \lambda_2, \ldots, \lambda_n$ is 0, and we get

$$\mathbf{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)^{-1} = \mathbf{diag}\left(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1}\right).$$

Hence

$$\begin{aligned} \boldsymbol{A}^{-1} &= \left(\boldsymbol{A}\,\mathbf{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)\,\boldsymbol{A}^{-1}\right)^{-1} \\ &= \left(\boldsymbol{A}^{-1}\right)^{-1}\mathbf{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)^{-1}\,\boldsymbol{A}^{-1} \\ &= \boldsymbol{A}\,\mathbf{diag}\left(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1}\right)\boldsymbol{A}^{-1}. \end{aligned}$$

Exercise 7.10. Confirm the result above without using SymPy.

Since

$$\begin{aligned} \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} &= \begin{bmatrix} a & c \\ b & d \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{bmatrix}, \\ \boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} &= \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{bmatrix}, \end{aligned}$$

if $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$, then we get

$$c^2 = b^2, \quad ab+cd = ac+bd.$$

From $c^2 = b^2$, $c = b$ or $c = -b$. If $c = b$, $ab+cd = ac+bd$ always holds. Then, $\boldsymbol{A} = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ is a normal matrix. If $c = -b$, from $ab - bd = -ab + bd$ we get $ab = bd$, and so either $b = 0$ or $a = d$ holds. When $c = -b$ and $a = d$, $\boldsymbol{A} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ is a normal matrix. When $c = -b$ and $b = 0$, $\boldsymbol{A} = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$ is a diagonal matrix. Therefore, a normal matrix $\boldsymbol{A}$ of order 2 takes one of the following shapes:

$$\begin{bmatrix} a & b \\ b & d \end{bmatrix} \text{(a symmetric matrix) and } \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \text{(a scalar multiple of an orthogonal matrix).}$$

Exercise 7.11. Prove that the properties of being (1) normal, (2) Hermitian, (3) unitary, (4) positive (semi-)definite and (5) an orthogonal projection are all unitarily invariant over $\mathbb{C}$.

Let $\boldsymbol{A}$ be a square matrix and $\boldsymbol{U}$ a unitary matrix.

(1) If $\boldsymbol{A}$ is a normal matrix, that is $\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^*$, then, since

$$(\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U})^*(\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}) = \boldsymbol{U}^*\boldsymbol{A}^*\boldsymbol{U}\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U} = \boldsymbol{U}^*\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{U} = \boldsymbol{U}^*\boldsymbol{A}\boldsymbol{A}^*\boldsymbol{U}$$
$$= \boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}\boldsymbol{U}^*\boldsymbol{A}^*\boldsymbol{U} = (\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U})(\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U})^*,$$

$\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}$ is a normal matrix.

(2) If $\boldsymbol{A}$ is a Hermitian matrix, that is $\boldsymbol{A}^* = \boldsymbol{A}$, then, since

$$(\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U})^* = \boldsymbol{U}^*\boldsymbol{A}^*\boldsymbol{U} = \boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U},$$

$\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}$ is a Hermitian matrix.

(3) If $\boldsymbol{A}$ is a unitary matrix, that is $\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{I}$, then, since

$$(\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U})^*(\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}) = \boldsymbol{U}^*\boldsymbol{A}^*\boldsymbol{U}\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U} = \boldsymbol{U}^*\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{U} = \boldsymbol{U}^*\boldsymbol{U} = \boldsymbol{I},$$

$\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}$ is a unitary matrix.

(4) Let $\boldsymbol{A}$ be a positive definite matrix and $\boldsymbol{x}$ be a nonzero vector, then

$$\langle \boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} \rangle > 0.$$

Because $\boldsymbol{U}\boldsymbol{x}$ is a nonzero vector, we have

$$\langle \boldsymbol{x} \mid \boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}\boldsymbol{x} \rangle = \langle \boldsymbol{U}\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{U}\boldsymbol{x} \rangle > 0.$$

Threrfore, $\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}$ is a positive definite matrix. Similarly, if $\boldsymbol{A}$ is a positive semi-definite matrix, so is $\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}$.

(5) Note that $\boldsymbol{A}$ be an orthogonal projection if and only if $\boldsymbol{A}^* = \boldsymbol{A}^2 = \boldsymbol{A}$. Let $\boldsymbol{A}$ be an orthogonal projection. Since $\boldsymbol{A}$ is a Hermitian matrix, so is $\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}$. On the other hand, we have

$$(\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U})^2 = \boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U} = \boldsymbol{U}^*\boldsymbol{A}^2\boldsymbol{U} = \boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}.$$

We can conclude that $\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}$ is also an orthogonal projection.

Exercise 7.12. The following program randomly generates a real normal matrix of order 2. Diagonalize a matrix generated by this program by a unitary matrix.

Matrix $\begin{bmatrix} 2 & 3 \\ 3 & -2 \end{bmatrix}$ has eigenvalues $\pm\sqrt{13}$ and these associated eigenvectors are $\begin{bmatrix} 2 \pm \sqrt{13} \\ 3 \end{bmatrix}$, whose norms are $\sqrt{26 \pm 4\sqrt{13}}$ respectively. By normalizing these vectors we make a unitary (orthogonal) matrix

$$\boldsymbol{U} = \begin{bmatrix} \frac{2-\sqrt{13}}{\sqrt{26-4\sqrt{13}}} & \frac{2+\sqrt{13}}{\sqrt{26+4\sqrt{13}}} \\ \frac{3}{\sqrt{26-4\sqrt{13}}} & \frac{3}{\sqrt{26+4\sqrt{13}}} \end{bmatrix}.$$

By this uniatry matrix we get diagonalization

$$\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U} = \begin{bmatrix} -\sqrt{13} & 0 \\ 0 & \sqrt{13} \end{bmatrix}.$$

Matrix $\begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix}$ has eigenvalues $2 \pm 3i$ and these associated eigenvectors are $\begin{bmatrix} \mp i \\ 1 \end{bmatrix}$ whose norms as same $\sqrt{2}$. By normalizing these vectors we make a unitary (orthogonal) matrix

$$\boldsymbol{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}.$$

By this uniatry matrix we get diagonalization

$$\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U} = \begin{bmatrix} 2-3i & 0 \\ 0 & 2+3i \end{bmatrix}.$$

Exercise 7.13. Let $\boldsymbol{A}$ and $\boldsymbol{V}$ be square matrices of order n, where $\boldsymbol{V}$ is regular. Prove the equality
$$e^{\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V}} = \boldsymbol{V}^{-1}e^{\boldsymbol{A}}\boldsymbol{V}.$$

We need some technique of analysis.

$$\begin{aligned}
&\left\| e^{\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V}} - \boldsymbol{V}^{-1}e^{\boldsymbol{A}}\boldsymbol{V} \right\| \\
= &\left\| e^{\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V}} - \sum_{k=0}^{n} \frac{\left(\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V}\right)^k}{k!} + \sum_{k=0}^{n} \frac{\left(\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V}\right)^k}{k!} - \boldsymbol{V}^{-1}e^{\boldsymbol{A}}\boldsymbol{V} \right\| \\
\leqq &\left\| e^{\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V}} - \sum_{k=0}^{n} \frac{\left(\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V}\right)^k}{k!} \right\| + \left\| \sum_{k=0}^{n} \frac{\left(\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V}\right)^k}{k!} - \boldsymbol{V}^{-1}e^{\boldsymbol{A}}\boldsymbol{V} \right\| \qquad \cdots (*)
\end{aligned}$$

The first term of $(*)$ tends to 0 as $n \to \infty$. Since

$$\begin{aligned}
\left\| \sum_{k=0}^{n} \frac{\boldsymbol{V}^{-1}\boldsymbol{A}^k\boldsymbol{V}}{k!} - \boldsymbol{V}^{-1}e^{\boldsymbol{A}}\boldsymbol{V} \right\| &= \left\| \boldsymbol{V}^{-1} \sum_{k=0}^{n} \frac{\boldsymbol{A}^k}{k!}\boldsymbol{V} - \boldsymbol{V}^{-1}e^{\boldsymbol{A}}\boldsymbol{V} \right\| \\
&= \left\| \boldsymbol{V}^{-1} \left(\sum_{k=0}^{n} \frac{\boldsymbol{A}^k}{k!} - e^{\boldsymbol{A}} \right) \boldsymbol{V} \right\| \\
&\leqq \left\| \boldsymbol{V}^{-1} \right\| \left\| \sum_{k=0}^{n} \frac{\boldsymbol{A}^k}{k!} - e^{\boldsymbol{A}} \right\| \left\| \boldsymbol{V} \right\|,
\end{aligned}$$

the second term of $(*)$ also tends to 0 as $n \to \infty$. Hence, we have the desired equality.

Exercise 7.14. Diagonalize $\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ first, and then compute a general term of $\boldsymbol{A}^n$ and $\exp(\boldsymbol{A})$.

The eigenpolynomial of $\boldsymbol{A}$ is

$$\begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = 0.$$

Since $\lambda^2 - 2\lambda - 3 = (\lambda - 3)(\lambda + 1) = 0$, the eigenvalues of $\boldsymbol{A}$ are 3 and -1. Let $\begin{bmatrix} x \\ y \end{bmatrix}$ be a eigenvector associated with 3. It follows that

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 3 \begin{bmatrix} x \\ y \end{bmatrix},$$

and that $2x - 2y = 0$. Hence, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector associated with 3. On the other hand, let $\begin{bmatrix} x \\ y \end{bmatrix}$ be a eigenvector associated with-1. It follows that

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = - \begin{bmatrix} x \\ y \end{bmatrix}$$

and that $2x + 2y = 0$. Hence, $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is an eigenvector associated with -1. Sincec $\boldsymbol{A}$ is a Hermitian matrix, it can be diagonalized by a unitary matrix. Normalizing both vectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, we have vectors $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ of norm 1 respectively. Using these vectors, we make a unitary (orthogonal) matrix $\boldsymbol{U}$ such as

$$\boldsymbol{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \ \boldsymbol{U}^* = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Then,

$$\boldsymbol{U}^* \boldsymbol{A} \boldsymbol{U} = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}.$$

is the diagonalization of $\boldsymbol{A}$ by unitary matrix $\boldsymbol{U}$. Since

$$\boldsymbol{U}^* \boldsymbol{A}^n \boldsymbol{U} = (\boldsymbol{U}^* \boldsymbol{A} \boldsymbol{U})^n = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}^n = \begin{bmatrix} 3^n & 0 \\ 0 & (-1)^n \end{bmatrix},$$

we get

$$\begin{aligned} \boldsymbol{A}^n &= \boldsymbol{U} \begin{bmatrix} 3^n & 0 \\ 0 & (-1)^n \end{bmatrix} \boldsymbol{U}^* = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3^n & 0 \\ 0 & (-1)^n \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 3^n + (-1)^n & 3^n - (-1)^n \\ 3^n - (-1)^n & 3^n + (-1)^n \end{bmatrix}. \end{aligned}$$

Moreover, since

$$\boldsymbol{U}^* \exp(\boldsymbol{A})\, \boldsymbol{U} \;=\; \exp(\boldsymbol{U}^* \boldsymbol{A} \boldsymbol{U}) \;=\; \exp\left(\begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}\right) \;=\; \begin{bmatrix} e^3 & 0 \\ 0 & e^{-1} \end{bmatrix},$$

we get

$$\begin{aligned} \exp(\boldsymbol{A}) &= \boldsymbol{U} \begin{bmatrix} e^3 & 0 \\ 0 & e^{-1} \end{bmatrix} \boldsymbol{U}^* = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^3 & 0 \\ 0 & e^{-1} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} e^3 + e^{-1} & e^3 - e^{-1} \\ e^3 - e^{-1} & e^3 + e^{-1} \end{bmatrix}. \end{aligned}$$

Exercise 7.15. For a square matrix $\boldsymbol{A}$ with $\|\boldsymbol{A}\| < 1$, show that $\boldsymbol{I} - \boldsymbol{A}$ is regular and

$$(\boldsymbol{I} - \boldsymbol{A})^{-1} = \sum_{n=0}^{\infty} \boldsymbol{A}^n$$

holds (Hint: first prove that $\sum_{n=0}^{\infty} \boldsymbol{A}^n$ converges and then show the equality $(\boldsymbol{I} - \boldsymbol{A}) \sum_{n=0}^{\infty} \boldsymbol{A}^n = \boldsymbol{I}$).

Because $1 > \|\boldsymbol{A}\|$,

$$\sum_{n=0}^{\infty} \|\boldsymbol{A}^n\| \leqq \sum_{n=0}^{\infty} \|\boldsymbol{A}\|^n = \frac{1}{1 - \|\boldsymbol{A}\|}$$

is satisfied and $\sum_{n=0}^{\infty} \boldsymbol{A}^n$ exists. Then,

$$\begin{aligned} 0 &\leqq \left\| (\boldsymbol{I} - \boldsymbol{A}) \sum_{n=0}^{\infty} \boldsymbol{A}^n - \boldsymbol{I} \right\| \\ &= \left\| (\boldsymbol{I} - \boldsymbol{A}) \left(\sum_{n=0}^{\infty} \boldsymbol{A}^n - \sum_{n=0}^{k} \boldsymbol{A}^n \right) - \boldsymbol{A}^{k+1} \right\| \\ &\leqq \left\| (\boldsymbol{I} - \boldsymbol{A}) \left(\sum_{n=0}^{\infty} \boldsymbol{A}^n - \sum_{n=0}^{k} \boldsymbol{A}^n \right) \right\| + \left\| \boldsymbol{A}^{k+1} \right\| \\ &\leqq \|\boldsymbol{I} - \boldsymbol{A}\| \left\| \sum_{n=0}^{\infty} \boldsymbol{A}^n - \sum_{n=0}^{k} \boldsymbol{A}^n \right\| + \|\boldsymbol{A}\|^{k+1}. \end{aligned}$$

The last side tends to 0 as $k \to \infty$, so we have

$$\left\| (\boldsymbol{I} - \boldsymbol{A}) \sum_{n=0}^{\infty} \boldsymbol{A}^n - \boldsymbol{I} \right\| = 0.$$

This says

$$(\boldsymbol{I} - \boldsymbol{A}) \sum_{n=0}^{\infty} \boldsymbol{A}^n = \boldsymbol{I}$$

and shows $\sum_{n=0}^{\infty} \boldsymbol{A}^n$ is the inverse matrix of $\boldsymbol{I} - \boldsymbol{A}$.

Chapter 8

Jordan Normal Form and Spectrum

Exercise 8.1	Sect. 8.1.	p.184	Invariant subspaces
Exercise 8.2	Sect. 8.2.	p.193	Computation of the Jordan normal form*
Exercise 8.3	Sect. 8.4.	p.201	Gelfand's formula*
Exercise 8.4	Sect. 8.5.	p.204	The Peron–Frobenius eigenvalue*

* Using Python

Exercise 8.1. Prove that K and R are invariant subspaces of $\boldsymbol{f}$.

If $\boldsymbol{x} \in K$, there exists k such that $\boldsymbol{x} \in K^{(k)}$. Then, since $\boldsymbol{f}^{k-1}\big(\boldsymbol{f}(\boldsymbol{x})\big) = \boldsymbol{f}^k(\boldsymbol{x}) = \mathbf{0}$, we have $\boldsymbol{f}(\boldsymbol{x}) \in K^{(k-1)} \subseteq K$. Hence, K is invariant subspace of $\boldsymbol{f}$.

If $\boldsymbol{y} \in R$, then $\boldsymbol{y} \in R^{(k)}$ for every $k = 0, 1, 2, \ldots$. Since $\boldsymbol{f}(\boldsymbol{y}) \in R^{(k+1)} \subseteq R^{(k)}$ for arbitrary k, it follows that $\boldsymbol{f}(\boldsymbol{y}) \in R$. Therefore, R is an invariant subspace of $\boldsymbol{f}$.

Exercise 8.2. The following program generates an exercise for Jordan normal form and Jordan decomposition (see the next section). Solve the problem generated by this program by hand.

The following program is `jordan2.py` modified to output in LaTeX format.

Program: `jordan2.py`

In [1]:

```
from sympy import *
from numpy.random import seed, permutation

seed(2021)
def latex_output(a, x):
    A = f'{latex(x)}'
    B =A.replace(r'\begin{matrix}', r'\begin{array}{rrr}\\')
    C = B.replace(r'\end{matrix}', r'\\\end{array}')
    L = C.split(r'\\')
    print(a + L[0])
    print(' '*4 + L[1] + r'\\')
    print(' '*4 + L[2] + r'\\')
    print(' '*4 + L[3] + r'\\')
    print(L[4])

X = Matrix([[1, 1, 0], [0, 1, 0], [0, 0, 2]])
Y = Matrix([[2, 1, 0], [0, 2, 1], [0, 0, 2]])
Z = Matrix([[2, 1, 0], [0, 2, 0], [0, 0, 2]])

while True:
    A = X.copy()
    while 0 in A:
        i, j, _  = permutation(3)
        A[:, j] += A[:, i]
        A[i, :] -= A[j, :]
        if max(abs(A)) >= 10:
            break
    if max(abs(A)) < 10:
        break
U, J = A.jordan_form()

latex_output('A = ', A)
latex_output('U = ', U)
latex_output('U**(-1)*A*U = ', J)
C = U * diag(J[0, 0], J[1, 1], J[2, 2]) * U**(-1)
B = A - C
latex_output('B = ', B)
latex_output('C = ', C)
```

```
A = \left[\begin{array}{rrr}
    2 & 4 & 4\\
    -4 & 3 & -1\\
    2 & -4 & -1\\
\end{array}\right]
U = \left[\begin{array}{rrr}
    \frac{24}{7} & - \frac{4}{7} & - \frac{1}{2}\\
    \frac{30}{7} & 1 & -1\\
    - \frac{36}{7} & 0 & 1\\
\end{array}\right]
U**(-1)*A*U = \left[\begin{array}{rrr}
    1 & 1 & 0\\
    0 & 1 & 0\\
    0 & 0 & 2\\
\end{array}\right]
B = \left[\begin{array}{rrr}
    8 & 8 & 12\\
    10 & 10 & 15\\
    -12 & -12 & -18\\
\end{array}\right]
C = \left[\begin{array}{rrr}
    -6 & -4 & -8\\
    -14 & -7 & -16\\
    14 & 8 & 17\\
\end{array}\right]
```

Find a Jotdan normal form of $\boldsymbol{A} = \begin{bmatrix} 2 & 4 & 4 \\ -4 & 3 & -1 \\ 2 & -4 & -1 \end{bmatrix}$. The eigenpolynomial of this matrix is $-\lambda^3 + 4\lambda^2 - 5\lambda + 2 = 0$.

In [2]:
```
lmd = Symbol('lambda')
f = det(A - lmd * eye(3)); f
```

Out[2]:
```
-lambda**3 + 4*lambda**2 - 5*lambda + 2
```

This can be factorized as $(\lambda - 2)(\lambda - 1)^2 = 0$, and we get 1 and 2 as eigenvalues of $\boldsymbol{A}$ Their multiplicity are 2 and 1 respectively.

In [3]:
```
factor(f)
```

Out[3]:
```
-(lambda - 2)*(lambda - 1)**2
```

In order to find a eigenvector associated with the egenvalue 1, we solve the following equation.

$$\left(\begin{bmatrix} 2 & 4 & 4 \\ -4 & 3 & -1 \\ 2 & -4 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

We get $x = -2z/3$, $y = -5z/6$, where z is an arbitrary constant. Hence, for example putting

$z = 6$, we have a eigenvector

$$\boldsymbol{v}_{11} = \begin{bmatrix} -4 \\ -5 \\ 6 \end{bmatrix}$$

associated with the eigenvalue 1.

In [4]:
```
x, y, z = symbols('x y z')
v = Matrix([x, y, z]); v
```

Out[4]:
```
Matrix([
[x],
[y],
[z]])
```

In [5]:
```
ans = solve((A - val[0]*eye(3))*v, [x, y]); ans
```

Out[5]:
```
{x: -2*z/3, y: -5*z/6}
```

In [6]:
```
v11 = v.subs([(x, ans[x]), (y, ans[y]), (z, 6)]); v11
```

Out[6]:
```
Matrix([
[-4],
[-5],
[ 6]])
```

The multiplicity of eigenvalue 1 was 2, so there is another vector linearly independent of this eigenvector in the general eigenspace. To find it, solve

$$\left(\begin{bmatrix} 2 & 4 & 4 \\ -4 & 3 & -1 \\ 2 & -4 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -4 \\ -5 \\ 6 \end{bmatrix}.$$

Then, we get $x = 2/3 - 2z/3, y = -5z/6 - 7/6$, whrer z is an arbitrary constant. Therefore, for example putting $z = 1$, we get a vector

$$\boldsymbol{v}_{12} = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}$$

belonging to the generarized eigenspace associated with the eigenvector 1 which not a eigenvector.

In [7]:
```
ans = solve((A - val[0]*eye(3))*v - v11, [x, y]); ans
```

Out[7]:
```
{x: 2/3 - 2*z/3, y: -5*z/6 - 7/6}
```

In [8]:
```
v12 = v.subs([(x, ans[x]), (y, ans[y]), (z, 1)]); v12
```

Out[8]:
```
Matrix([
[ 0],
[-2],
[ 1]])
```

Next, find an eigenvector associated with the egenvalue 2. By solving

$$\left(\begin{bmatrix} 2 & 4 & 4 \\ -4 & 3 & -1 \\ 2 & -4 & -1 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}\right)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

we get $= -z/2$, $y = -z$, where z is an arbitrary constant. Therefore, putting $z = 2$ we get a eigenvector

$$\boldsymbol{v}_2 = \begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix}$$

assiciated with the eigenvalue 2. Since the multiplicity of the eigenvalue 2 is 1, the eigenspace associated with the egenvalue 2 is the one dimentional subspace generated by this eigenvector.

In [9]:
```
ans = solve((A - val[0]*eye(3))*v - v11, [x, y]); ans
```

Out[9]:
```
{x: 2/3 - 2*z/3, y: -5*z/6 - 7/6}
```

In [10]:
```
v12 = v.subs([(x, ans[x]), (y, ans[y]), (z, 1)]); v12
```

Out[10]:
```
Matrix([
[ 0],
[-2],
[ 1]])
```

In [11]:
```
ans = solve((A - val[1]*eye(3))*v, [x, y]); ans
```

Out[11]:
```
{x: -z/2, y: -z}
```

In [12]:
```
v2 = v.subs([(x, ans[x]), (y, ans[y]), (z, 2)]); v2
```

Out[12]:
```
Matrix([
[-1],
[-2],
[ 2]])
```

Put

$$\boldsymbol{V} \stackrel{\text{def}}{=} \begin{bmatrix} \boldsymbol{v}_{11} & \boldsymbol{v}_{12} & \boldsymbol{v}_2 \end{bmatrix} = \begin{bmatrix} -4 & 0 & -1 \\ -5 & -2 & -2 \\ 6 & 1 & 2 \end{bmatrix}.$$

In [13]:
```
V = Matrix([v11.T, v12.T, v2.T]).T; V
```

Out[13]:
```
Matrix([
[-4,  0, -1],
[-5, -2, -2],
[ 6,  1,  2]])
```

Then

$$V^{-1} = \begin{bmatrix} -2 & -1 & -2 \\ -2 & -2 & -3 \\ 7 & 4 & 8 \end{bmatrix}.$$

In [14]:
```
W = V**(-1); W
```

Out[14]:
```
Matrix([
[-2, -1, -2],
[-2, -2, -3],
[ 7,  4,  8]])
```

Therefore, we get a Jordan normal form

$$V^{-1}AV = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

In [15]:
```
W*A*V
```

Out[15]:
```
Matrix([
[1, 1, 0],
[0, 1, 0],
[0, 0, 2]])
```

Let

$$B = V\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}V^{-1} = \begin{bmatrix} -6 & -4 & -8 \\ -14 & -7 & -16 \\ 14 & 8 & 17 \end{bmatrix},$$

$$C = V\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}V^{-1} = \begin{bmatrix} 8 & 8 & 12 \\ 10 & 10 & 15 \\ -12 & -12 & -18 \end{bmatrix}.$$

Then, B is a diagonalizable matrix, and C is a nilpotent matrix($C^2 = O$). Thus we get the Jordan decomposition

$$A = B + C.$$

In [16]:
```
D = diag(val[0], val[0], val[1]); D
```

Out[16]:
```
Matrix([
[1, 0, 0],
[0, 1, 0],
[0, 0, 2]])
```

In [17]:
```
E = W*A*V - D; E
```

Out[17]:
```
Matrix([
[0, 1, 0],
[0, 0, 0],
[0, 0, 0]])
```

In [18]:
```
B = V*D*W; B
```

Out[18]:
```
Matrix([
[ -6, -4,  -8],
[-14, -7, -16],
[ 14,  8,  17]])
```

In [19]:
```
C = V*E*W; C
```

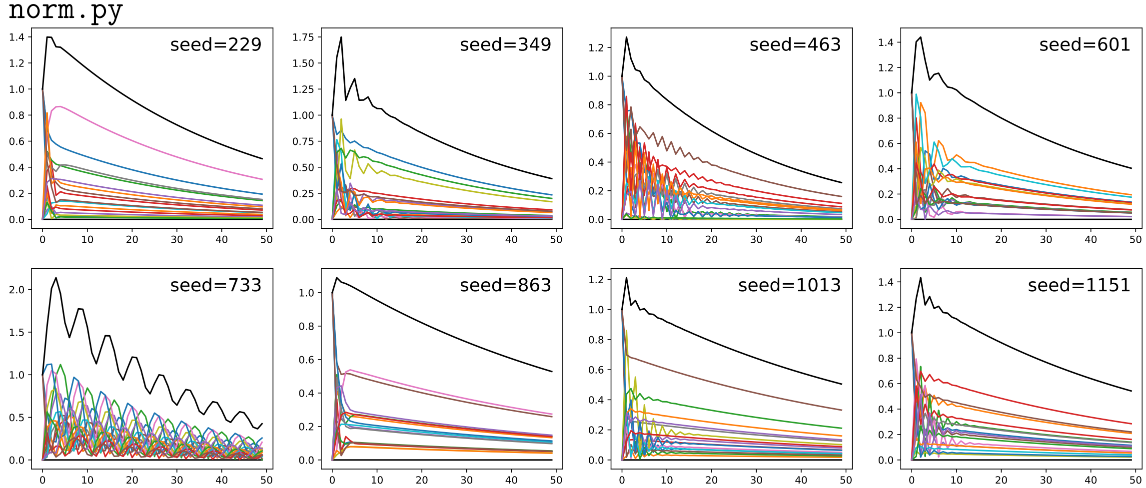
In [20]:
```
Matrix([
[  8,   8,  12],
[ 10,  10,  15],
[-12, -12, -18]])
\end{pyout}
\begin{pyin}
B + C
```

Out[20]:
```
Matrix([
[ 2,  4,  4],
[-4,  3, -1],
[ 2, -4, -1]])
```

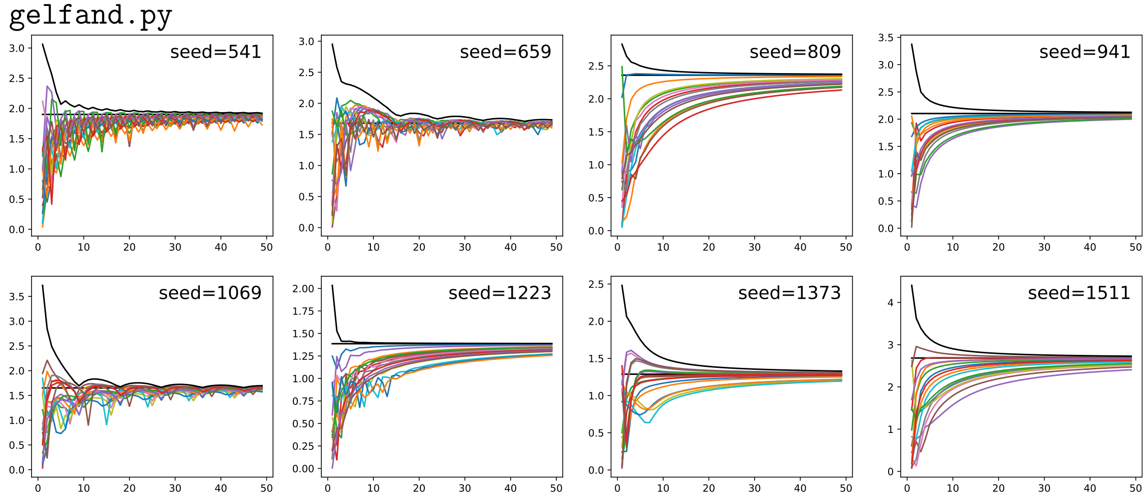
Exercise 8.3. Experiment with changing seeds in Programs norm.py and gelfand.py and observe the convergences.

The following are examples of executing for square matrices of order 4 getting by appropriate seeds of random numbers.

`norm.py`



`gelfand.py`



Exercise 8.4. Let $\boldsymbol{A}$ be the square matrix of order n whose elements are all 1. Find the Peron-Frobenius eigenvalue of $\boldsymbol{A}$ and a positive eigenvector associated with it.

Let λ be the Perron-Frobenius eigenvalue and $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ an eigenvector associated it .

Then,

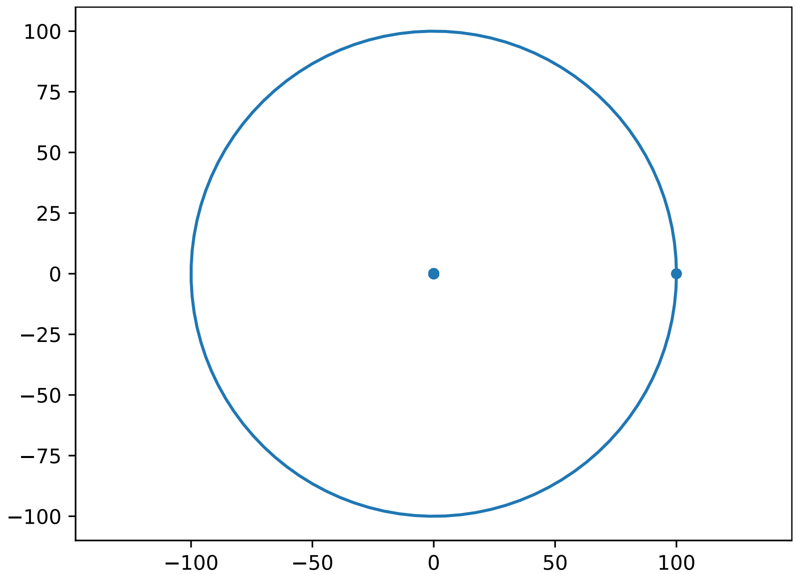
$$\boldsymbol{A}\boldsymbol{x} \;=\; (x_1 + x_2 + \cdots + x_n) \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \;=\; \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

It follows that $\lambda = n$ and $x_1 = x_2 = \cdots = x_n$. Hence, n is the Perron-Frobenius eigenvalue and $\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ is an eigenvector associated with it.

The following program verifies this when $n = 100$.

Program: `exercise8_4.py`

In [1]:
```
from numpy import pi, sin, cos, linspace, ones
from numpy.linalg import eig, eigh
import matplotlib.pylab as plt

N = 100
X = ones((N, N))
Lmd = eig(X)[0]
r = max(abs(Lmd))
T = linspace(0, 2 * pi, 100)
plt.axis('equal')
plt.plot(r * cos(T), r * sin(T))
plt.scatter(Lmd.real, Lmd.imag, s=20)
plt.show()
```


100
75
50
25
0
−25
−50
−75
−100
−100
−50
0
50
100

Chapter 9

Dynamical Systems

Exercise 9.1	Sect. 9.3.	p.215	A linear differential equation*
Exercise 9.2	Sect. 9.3.	p.217	Simulation of a dynamical system*
Exercise 9.3	Sect. 9.4.	p.219	Numerical semigroups*
Exercise 9.4	Sect. 9.4.	p.219	Irredusible and appeoriodic transition matrices
Exercise 9.5	Sect. 9.6.	p.229	One parameter semigroups

* Using Python

Exercise 9.1. Let $\boldsymbol{A} = \begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix}$. Solve the differential equation $x'(t) = \boldsymbol{A}\, x(t)$ and draw the trajectories of the solutions for several initial values.

For the eigenpolynomial $\begin{vmatrix} 1-\lambda & -2 \\ 3 & 1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda + 7 = 0$, we have eigenvalues $1 \pm \sqrt{6}i$. In order to find eigenvectors, we solve

$$\begin{bmatrix} 1 & -2 \\ 3 & 1a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \left(1 \pm \sqrt{6}i\right) \begin{bmatrix} x \\ y \end{bmatrix},$$

that is,

$$\begin{bmatrix} \pm\sqrt{6}ix + 2y \\ -3x \mp \sqrt{6}iy \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Puting $x = 1$, we get eigenvectors $\begin{bmatrix} 1 \\ \mp\frac{\sqrt{6}i}{2} \end{bmatrix}$ associated with eigenvalues $1 \pm \sqrt{6}i$ respectively. Let

$$\boldsymbol{V} \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 1 \\ \frac{\sqrt{6}i}{2} & -\frac{\sqrt{6}i}{2} \end{bmatrix}.$$

Then

$$\boldsymbol{V}^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{6}i}{6} \\ \frac{1}{2} & \frac{\sqrt{6}i}{6} \end{bmatrix}$$

and

$$\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V} = \begin{bmatrix} 1-\sqrt{6}i & 0 \\ 0 & 1+\sqrt{6}i \end{bmatrix}.$$

Hence

$$\boldsymbol{V}^{-1}e^{t\boldsymbol{V}}\boldsymbol{V} = \begin{bmatrix} e^{t\left(1-\sqrt{6}i\right)} & 0 \\ 0 & e^{t\left(1+\sqrt{6}i\right)} \end{bmatrix} = e^t \begin{bmatrix} e^{-\sqrt{6}it} & 0 \\ 0 & e^{\sqrt{6}it} \end{bmatrix}.$$

Finally, we have

$$\begin{aligned} e^{t\boldsymbol{V}} &= e^t \boldsymbol{V} \begin{bmatrix} e^{-\sqrt{6}it} & 0 \\ 0 & e^{\sqrt{6}it} \end{bmatrix} \boldsymbol{V}^{-1} \\ &= e^t \begin{bmatrix} \frac{e^{-\sqrt{6}it}}{2} + \frac{e^{\sqrt{6}it}}{2} & -\frac{\sqrt{6}ie^{-\sqrt{6}it}}{6} + \frac{\sqrt{6}ie^{\sqrt{6}it}}{6} \\ \frac{\sqrt{6}ie^{-\sqrt{6}it}}{4} - \frac{\sqrt{6}ie^{\sqrt{6}it}}{4} & \frac{e^{-\sqrt{6}it}}{2} + \frac{e^{\sqrt{6}it}}{2} \end{bmatrix} \\ &= e^t \begin{bmatrix} \cos\left(\sqrt{6}t\right) & -\frac{\sqrt{6}\sin\left(\sqrt{6}t\right)}{3} \\ \frac{\sqrt{6}\sin\left(\sqrt{6}t\right)}{2} & \cos\left(\sqrt{6}t\right) \end{bmatrix}. \end{aligned}$$

Try using sympy to find the exponential function of the above matrix and compare the results.

In [1]:
```
from sympy import *

A = Matrix([[1, -2], [3, 1]])
t = var('t', real=True)
E = exp(t*A)
```

In [2]:
```
print(E)
```

```
Matrix([[exp(t)*cos(sqrt(6)*t), -sqrt(6)*exp(t)*sin(sqrt(6)*t)/3],
 [sqrt(6)*exp(t)*sin(sqrt(6)*t)/2, exp(t)*cos(sqrt(6)*t)]])
```

In [3]:
```
print(latex(E))
```

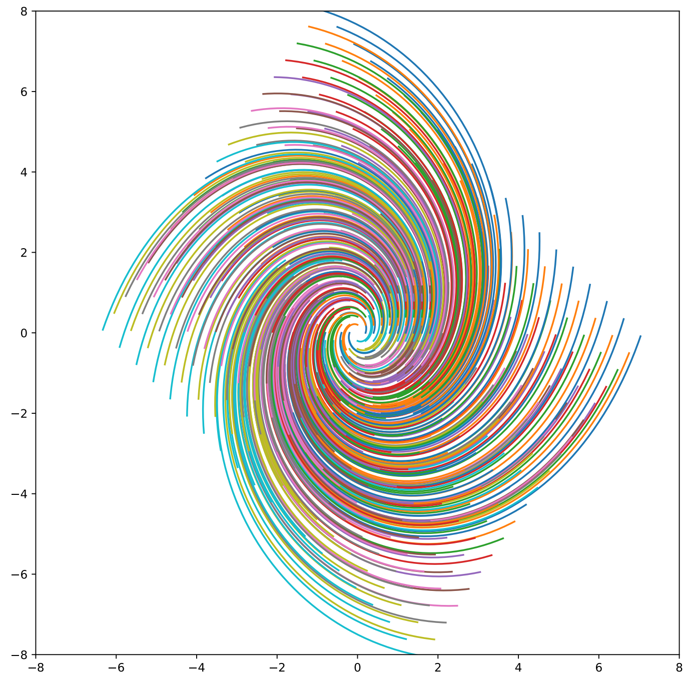
```
\left[\begin{matrix}e^{t} \cos{\left(\sqrt{6} t \right)} & - \frac{\sqrt{6}
    e^{t} \sin{\left(\sqrt{6} t \right)}}{3}\\\frac{\sqrt{6} e^{t} \sin{\
   left(\sqrt{6} t \right)}}{2} & e^{t} \cos{\left(\sqrt{6} t \right)}\end{
   matrix}\right]
```

$$\left[\begin{matrix}e^{t} \cos{\left(\sqrt{6} t \right)} & - \frac{\sqrt{6} e^{t} \sin{\left(\sqrt{6} t \right)}}{3}\\\frac{\sqrt{6} e^{t} \sin{\left(\sqrt{6} t \right)}}{2} & e^{t} \cos{\left(\sqrt{6} t \right)}\end{matrix}\right]$$

Solve the differential equation with the Euler method and look at the phase space.

Program: `exercise9_1_1`

In [1]:
```
from numpy import array, arange
import matplotlib.pyplot as plt

A = array([[1, -2],
           [3,  1]])

def update(x, v, dt):
    dx = A[0, 0] * x * dt + A[0, 1] * y * dt
    dy = A[1, 0] * x * dt + A[1, 1] * y * dt
    return x + dx, y + dy

dt = 0.001
fig = plt.figure(figsize=(10, 10))
plt.axis('scaled'), plt.xlim(-8.0, 8.0), plt.ylim(-8.0, 8.0)
for x0 in arange(-2, 2, 0.2):
    for y0 in arange(-2, 2, 0.2):
        path = [(x0, y0)]
        for t in arange(0, 1, dt):
            x, y = path[-1]
            path.append(update(x, y, dt))
        plt.plot(*zip(*path))
plt.show()
```

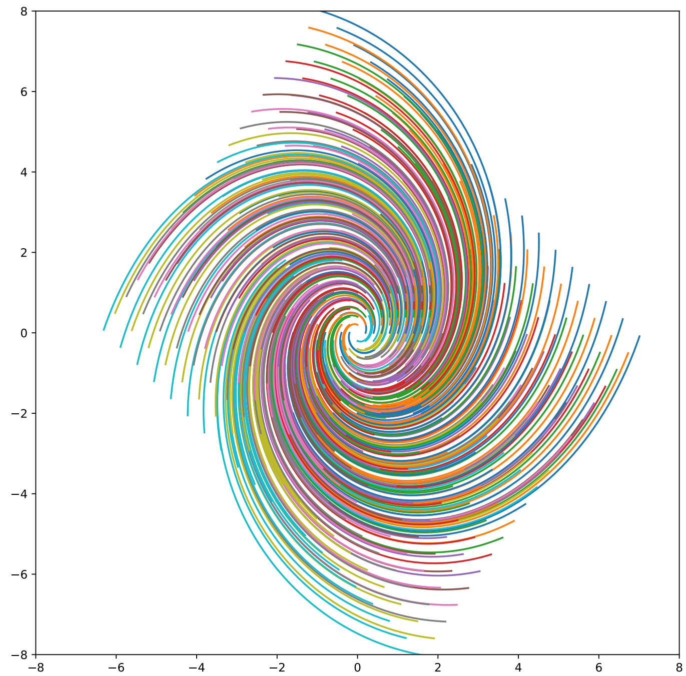


The following program uses the exact solution.

Program: `exercise9_1_2`

In [1]:

```
from numpy import *
import matplotlib.pyplot as plt

A = array([[1, -2],
           [3,  1]])

def f(t):
    return exp(t)*array([[cos(sqrt(6)*t), -sqrt(6)*sin(sqrt(6)*t)/3],
                         [sqrt(6)*sin(sqrt(6)*t)/2, cos(sqrt(6)*t)]])

dt = 0.001
fig = plt.figure(figsize=(10, 10))
plt.axis('scaled'), plt.xlim(-8.0, 8.0), plt.ylim(-8.0, 8.0)
for x0 in arange(-2, 2, 0.2):
    for y0 in arange(-2, 2, 0.2):
        path = [(x0, y0)]
        for t in arange(0, 1, dt):
            x, y = dot(f(t), path[0])
            path.append((x, y))
        plt.plot(*zip(*path))
plt.show()
```


8
6
4
2
0
−2
−4
−6
−8
−8
−6
−4
−2
0
2
4
6
8

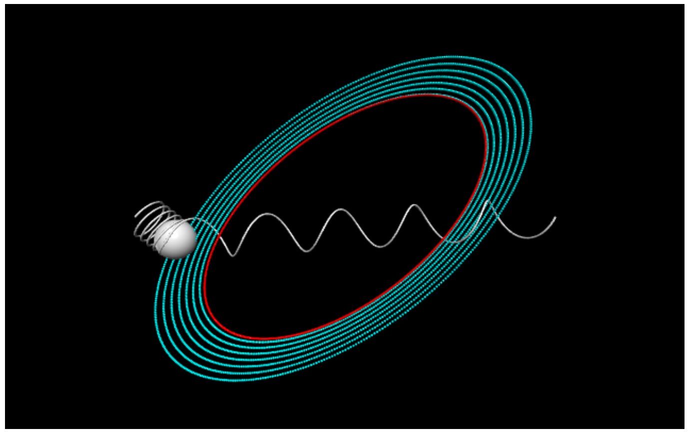
Exercise 9.2. (1) Solve the same problem above by Euler's method, and compare the approximate solution with the exact solution above.
(2) Investigate the movement of an object in $\mathbb{R}^3$ attached to three or more springs.

Program: `exercise9_2_1`

In [1]:
```
1  from vpython import *
2  
3  p = vec(0, 0, 0)
4  q = vec(2, 0, 0)
5  x = vec(1, 0.5, 0)
6  v = vec(1, 0.5, 0)
7  
8  P = sphere(pos=x, radius=0.1)
9  h1 = helix(pos=p, radius=0.1, axis=x-p)
10 h2 = helix(pos=q, radius=0.1, axis=x-q)
11 
12 scene.center = vec(1, 0, 0)
13 
14 def update(x, v, dt):
15     x1 = x + v * dt
16     v1 = v + (p + q - 2 * x) * dt
17     return x1, v1
18 
19 def pos(t):
20     return vec(1+sin(sqrt(2)*t)/sqrt(2),
21                (sqrt(2)*cos(sqrt(2)*t)+sin(sqrt(2)*t))/2/sqrt(2),0)
22 
23 t = 0
24 dt = 0.01
25 while True:
26     rate(100)
27     x, v = update(x, v, dt)
28     P.pos = x
29     points(pos=[x], radius=2, color=color.cyan)
30     points(pos=[pos(t)], radius=2, color=color.red)
31     h1.axis = x-p
32     h2.axis = x-q - v * dt**2
33     t += dt
```

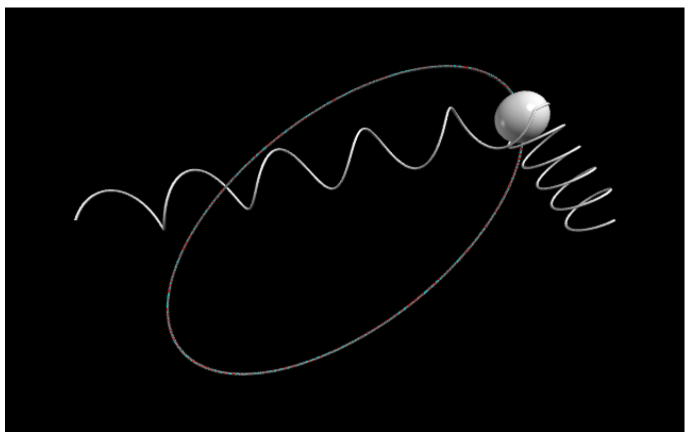
Lines 14 – 17: Update formula by first-order approximation of Taylor expansion.

Lines 19 – 21: Exact solution.



Errors accumulate in the first-order approximation of Taylor's expansion. Use the second-order approximation of the Taylor expansion.

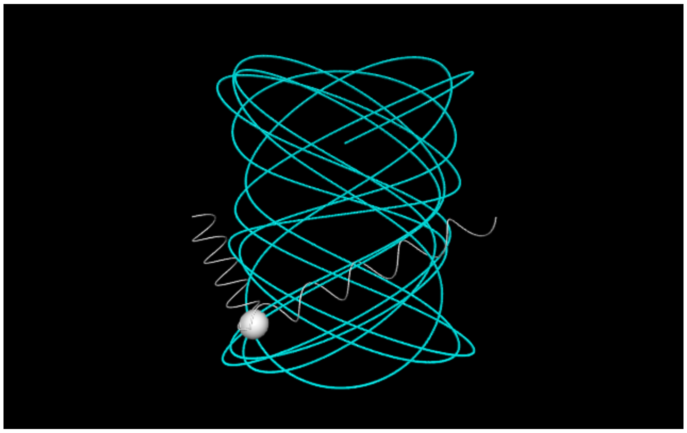
In [ ]:
```
def update(x, v, dt):
    x1 = x + v * dt + (p + q - 2 * x) / 2 * dt**2
    v1 = v + (p + q - 2 * x) * dt - v * dt**2
    return x1, v1
```



Program: `exercise9_2_2`

In [1]:
```
from vpython import *

p = vec(0, 0, 0)
q = vec(2, 0, 0)
x = vec(1, 0.5, 0)
v = vec(1, 0.5, 0)

P = sphere(pos=x, radius=0.1)
h1 = helix(pos=p, radius=0.1, axis=x-p)
h2 = helix(pos=q, radius=0.1, axis=x-q)

scene.center = vec(1, 0, 0)

def update(x, v, dt):
```

In [1]:
```
    e1 = hat(x - p) * (1 - mag(p - x))
    e2 = hat(x - q) * (1 - mag(q - x))
    x1 = x + v * dt + (e1 + e2) / 2 * dt**2
    v1 = v + (e1 + e2) * dt - v * dt**2
    return x1, v1

t = 0
dt = 0.01
while True:
    rate(100)
    x, v = update(x, v, dt)
    P.pos = x
    points(pos=x, radius=2, color=color.cyan)
    h1.axis=x-p
    h2.axis=x-q
    t += dt
```

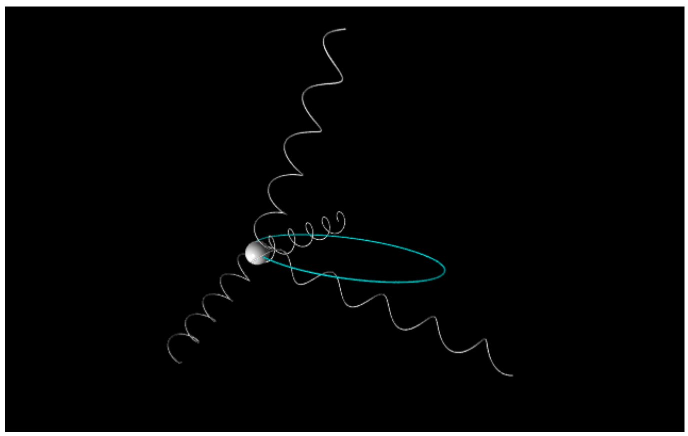


Program: `exercise9_2_3`

In [1]:
```
from vpython import *

p = vec(0, 0, 0)
q = vec(2, 0, 0)
r = vec(0, 2, 0)
s = vec(0, 0, 2)
x = vec(1, 0.5, 0)
v = vec(1, 0.5, 0)

P = sphere(pos=x, radius=0.1)
h1 = helix(pos=p, radius=0.1, axis=x-p)
h2 = helix(pos=q, radius=0.1, axis=x-q)
h3 = helix(pos=r, radius=0.1, axis=x-r)
h4 = helix(pos=s, radius=0.1, axis=x-s)

def update(x, v, dt):
    x1 = x + v * dt + (p + q + r + s - 4 * x) / 2 * dt**2
    v1 = v + (p + q + r + s - 4 * x) * dt - 2 * v * dt**2
    return x1, v1

```

In [1]:

```

t = 0
dt = 0.01
while True:
    rate(100)
    x, v = update(x, v, dt)
    P.pos = x
    points(pos=[x], radius=1, color=color.cyan)
    h1.axis = x-p
    h2.axis = x-q
    h3.axis = x-r
    h4.axis = x-s
    t += dt
```



Exercise 9.3. Let X be a numerical semigroup containing p and q whose greatest common divisor is 1. Prove that all numbers greater than or equal to $(p-1)(q-1)$ belong to X. (Hint: Prove that the remainders of σq for $\sigma = 0, 1, \ldots, p-1$ divided by p are all different, and for any $m \geqq (p-1)(q-1)$, $m - \sigma q$ is a multiple of p for some σ.)

Remark that X is closed under summation and also multiplication by natural numbers. We first show the former half of the Hint. Conside natural numbers σ_1 and σ_2 such that $0 \leqq \sigma_1 < \sigma_2 < p$. Suppose the $\sigma_1 q$ and $\sigma_2 q$ have a commom remainder r devided by p. Then there exist m and n such that $\sigma_1 q = mp + r$ and $\sigma_2 q = np + r$. Hence

$$(\sigma_2 - \sigma_1)\, q \;=\; \sigma_2 q - \sigma_1 q \;=\; (n-m)\, p.$$

Since $0 < \sigma_2 - \sigma_1 < p$ and $\mathrm{GCD}\,(p, q) = 1$, the left hand side above can not be a multiple of p. This is contradiction.

Next, we first show the latter half of the Hint. Consider an integer m with $m \geqq (p-1)\,(q-1)$. Let r be the remainder of m by p and choose σ such that $0 \leqq \sigma < p$ and the remainder of σq by p is r. Then, $m - \sigma q$ is a multiple of p. Thus, $m > (p-1)\,q - p \geqq \sigma q - p$, and we have $m - \sigma q > -p$. Hence, $m - \sigma q \geqq 0$ and it is a multiple of p. Since

$$m \;=\; (m - \sigma q) + \sigma q \;=\; ap + bq$$

for some nonnegative integers a and b where a and b are not zero at the same time, it follows that m belongs to X.

The following program verifies the above proof with some concrete numbers.

Program: `exercise9_3`

In [1]:

```
p, q = 5, 9

for k in range(p):
    m = (p - 1)*(q - 1) + k

    r = m % p
    s = [0] * p
    for i in range(p):
        j = i*q % p
        s[j] = i

    sigma = s[r]
    a = (m - sigma*q) // p
    print(f'{m} = {a}*{p} + {sigma}*{q} = {a*p} + {sigma*q}')
```

```
32 = 1*5 + 3*9 = 5 + 27
33 = 3*5 + 2*9 = 15 + 18
34 = 5*5 + 1*9 = 25 + 9
```

```
35 = 7*5 + 0*9 = 35 + 0
36 = 0*5 + 4*9 = 0 + 36
```

Supplimentary Excercise: Give a proof of Lemma 9.1 using this exercise. (Hint: use mathematical induction.)

Exercise 9.4. Prove the above results (Hint: Use Lemma 9.1).

It is easy to see that $\boldsymbol{P}$ is positive, then it is irrecucible and aperiodic. Conversely, suppose that $\boldsymbol{P}$ is irreducible and aperiodic. We shall prove that $\boldsymbol{P}^m$ is pisitive for all sufficiently large m.

Let $i, j \in I = \{1, 2, \dots, n\}$. First suppose that $i \neq j$. Let $(s_1, s_2, \dots, s_n)$ be a permutaion of $(1, 2, \dots, n)$ such that $s_1 = i$ and $s_n = j$. Because $\boldsymbol{P}$ is irreducible, for each$k \in I \setminus \{n\}$, there is$m_k \in \mathbb{N}$ such that the (s_k, s_{k+1})-component of $\boldsymbol{P}^{m_k}$ is positive. Because $\boldsymbol{P}$ is aperiodic, for each $k \in I$ there is $\xi_k \in \mathbb{N}$ such that the greatest common divisor of $\xi_1, \xi_2, \dots, \xi_n$ is 1 and the (s_k, s_{k+1})-component od $\boldsymbol{P}^{\xi_k}$ is positive. Let

$$\alpha \;=\; m_1 + m_2 + \cdots + m_{n-1} + a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n$$

for $a_1, a_2, \dots, a_n \in \mathbb{N}$. Then the (i, j)-component of

$$\boldsymbol{P}^{\alpha} \;=\; \boldsymbol{P}^{a_1\xi_1}\boldsymbol{P}^{m_1}\boldsymbol{P}^{a_2\xi_2}\boldsymbol{P}^{m_2}\cdots\boldsymbol{P}^{a_{n-1}\xi_{n-1}}\boldsymbol{P}^{m_{n-1}}\boldsymbol{P}^{a_n\xi_n}$$

is positive. The set of all integers expressed as $a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n$ is a numerical semigroup, and by Lemma 9.1, it contains all sufficient large numbers. Hence, there is $n_{ij} \in \mathbb{N}$ such that (i, j)-component of $\boldsymbol{P}^m$ is positive for any $m \geqq n_{ij}$.

Next, suppose $i = j$. Let $(s_1, s_2, \dots, s_n)$ be a permutaion of $(1, 2, \dots, n)$ such that $s_1 = i$. Let $m_1, m_2, \dots, m_{n-1}$ and $\xi_1, \xi_2, \dots, \xi_n$ be the same as above. There is $m_n \in \mathbb{N}$ such that the (s_n, i)-component od $\boldsymbol{P}^{m_n}$ is positive. Let

$$\beta \;=\; m_1 + m_2 + \cdots + m_n + a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n$$

for $a_1, a_2, \dots, a_n \in \mathbb{N}$ Then, the (i, i)-component of $\boldsymbol{P}^{\beta}$ is positive. For the same reason as above, there is $n_i \in \mathbb{N}$ such that the(i, i)-component of $\boldsymbol{P}^m$ is positive for all $m \geqq n_i$.

Consequently, $\boldsymbol{P}^m$ is positive for all $m \geqq \max\limits_{i,j}\{n_{ij}, n_i\}$.

Exercise 9.5. Prove the above properties (2)–(4).

(2)

$$\boldsymbol{P}_{s+t} = e^{(s+t)\boldsymbol{G}} = e^{s\boldsymbol{G}+t\boldsymbol{G}} = e^{s\boldsymbol{G}} + e^{t\boldsymbol{G}} = \boldsymbol{P}_s + \boldsymbol{P}_t$$

because $s\boldsymbol{G}$ and $t\boldsymbol{G}$ commute (see Set. 7.4.).

(3)

$$\begin{aligned} \|\boldsymbol{P}_t - \boldsymbol{I}\| &= \left\| \frac{\boldsymbol{G}t}{1!} + \frac{(\boldsymbol{G}t)^2}{2!} + \cdots \right\| \\ &\leqq \left\| \frac{\boldsymbol{G}t}{1!} \right\| + \left\| \frac{(\boldsymbol{G}t)^2}{2!} \right\| + \cdots \\ &\leqq \frac{\|\boldsymbol{G}\| t}{1!} + \frac{\|\boldsymbol{G}\|^2 t^2}{2!} + \cdots \\ &\to 0 \quad (t \downarrow 0) \end{aligned}$$

(4)

$$\begin{aligned} \left\| \frac{\boldsymbol{P}_t - \boldsymbol{I}}{t} - \boldsymbol{G} \right\| &= \left\| \frac{\boldsymbol{G}^2 t}{2!} + \frac{\boldsymbol{G}^3 t^2}{3!} \cdots \right\| \\ &\leqq \frac{\|\boldsymbol{G}\|^2 t}{2!} + \frac{\|\boldsymbol{G}\|^3 t^2}{3!} + \cdots \\ &\to 0 \quad (t \downarrow 0) \end{aligned}$$

Here, we use the fact that, if $\sum_{n=1}^{\infty} \|\boldsymbol{x}_n\| < \infty$, then $\sum_{n=1}^{\infty} \boldsymbol{x}_n$ exists and

$$\left\| \sum_{n=1}^{\infty} \boldsymbol{x}_n \right\| \leqq \sum_{n=1}^{\infty} \|\boldsymbol{x}_n\|$$

holds for a vector sequence $\{\boldsymbol{x}_n\}_{n=1}^{\infty}$ in a normed linear space.

Chapter 10

Applications and Development of Linear Algebra

Exercise 10.1	Sect. 10.2.	p.242	Properties of the polar expression of matrices
Exercise 10.2	Sect. 10.2.	p.243	Properties of the trace norm
Exercise 10.3	Sect. 10.2.	p.243	Properties of the Hilbert–Schmidt norm
Exercise 10.4	Sect. 10.2.	p.243	The maximum singular value
Exercise 10.5	Sect. 10.3.	p.245	The equivalence relation induced by a subspace
Exercise 10.6	Sect. 10.3.	p.245	Equivalence classes induced by a subspace
Exercise 10.7	Sect. 10.3.	p.246	The inner product on a tensor product space
Exercise 10.8	Sect. 10.3.	p.250	Computation of singular value decompositions*
Exercise 10.9	Sect. 10.3.	p.251	A quantum system*
Exercise 10.10	Sect. 10.4.	p.254	The space of vector valued random variables
Exercise 10.11	Sect. 10.4.	p.255	Expectation of matrix valued random variables
Exercise 10.12	Sect. 10.5.	p.262	An inner product on a space of matrices

* Using Python

Exercise 10.1. Prove the above equalities 1–4. (Hint: use the singular value decomposition $\boldsymbol{A}^* = \boldsymbol{U^*}_2 \boldsymbol{\Sigma} \boldsymbol{U}_1^*$.)

1.
$$\text{RHS} \;=\; \boldsymbol{U}_1 \boldsymbol{U}_2 {\boldsymbol{U}_2}^* \boldsymbol{\Sigma} \boldsymbol{U}_2 \;=\; \boldsymbol{U}_1 \boldsymbol{\Sigma} \boldsymbol{U}_2 \;=\; \text{LHS.}$$

2.
$$\text{RHS} \;=\; {\boldsymbol{U}_2}^* {\boldsymbol{U}_1}^* \boldsymbol{U}_1 \boldsymbol{\Sigma} \boldsymbol{U}_2 \;=\; {\boldsymbol{U}_2}^* \boldsymbol{\Sigma} \boldsymbol{U}_2 \;=\; \text{LHS.}$$

3. Since $\boldsymbol{A}^* = {\boldsymbol{U}_2}^* \boldsymbol{\Sigma} {\boldsymbol{U}_1}^*$, it follows that $[\![\boldsymbol{A}^*]\!] = \boldsymbol{U}_1 \boldsymbol{\Sigma} {\boldsymbol{U}_1}^*$. Hence,

$$\text{RHS} \;=\; \boldsymbol{U}_1 \boldsymbol{U}_2 {\boldsymbol{U}_2}^* \boldsymbol{\Sigma} \boldsymbol{U}_2 {\boldsymbol{U}_2}^* {\boldsymbol{U}_1}^* \;=\; \boldsymbol{U}_1 \boldsymbol{\Sigma} {\boldsymbol{U}_1}^* \;=\; \text{LHS.}$$

4.
$$\text{RHS} \;=\; {\boldsymbol{U}_2}^* {\boldsymbol{U}_1}^* \boldsymbol{U}_1 \boldsymbol{\Sigma} {\boldsymbol{U}_1}^* \;=\; {\boldsymbol{U}_2}^* \boldsymbol{\Sigma} {\boldsymbol{U}_1}^* \;=\; \text{LHS.}$$

Exercise 10.2. Prove the following properties:

1. $\|\boldsymbol{A}\|_{\mathrm{Tr}} = \sum_{i=1}^{k} \sigma_i$, where $\sigma_1, \sigma_2, \ldots, \sigma_k$ are all of the singular values of $\boldsymbol{A}$.
2. $\|\boldsymbol{A}\|_{\mathrm{Tr}} \geqq 0$ and the equality holds if and only if $\boldsymbol{A} = \boldsymbol{O}$
3. $\|c\boldsymbol{A}\|_{\mathrm{Tr}} = |c|\, \|\boldsymbol{A}\|_{\mathrm{Tr}}$.
4. $\|\boldsymbol{A}+\boldsymbol{B}\|_{\mathrm{Tr}} \leqq \|\boldsymbol{A}\|_{\mathrm{Tr}} + \|\boldsymbol{B}\|_{\mathrm{Tr}}$.

We present some remarks about the trace of a matrix. Let $\boldsymbol{A}$ be a square matrix of order n. Then

$$\mathrm{Tr}\,(\boldsymbol{A}) \;=\; \sum_{i=1}^{n} \left\langle \boldsymbol{e}_i \;\middle|\; \boldsymbol{A}\boldsymbol{e}_i \right\rangle$$

holds for an arbitrary orthonomal basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ of $\boldsymbol{K}^n$. This means that the right-hand side does not depend on the orthonormal basis chosen. In fact, if $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ is the standard basis of $\boldsymbol{K}^n$, the right hand side is the sum of all diagonal components of $\boldsymbol{A}$, so the equality holds. For any orthonomal basis $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ of $\boldsymbol{K}^n$, consider the unitary matrix $\boldsymbol{U} = \begin{bmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n\end{bmatrix}$. Then

$$\mathrm{Tr}\,(\boldsymbol{A}) \;=\; \mathrm{Tr}\,(\boldsymbol{U}\boldsymbol{U}^*\boldsymbol{A}) \;=\; \mathrm{Tr}\,(\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}) \;=\; \sum_{i=1}^{n} \left\langle \boldsymbol{U}\boldsymbol{e}_i \;\middle|\; \boldsymbol{A}\boldsymbol{U}\boldsymbol{e}_i \right\rangle \;=\; \sum_{i=1}^{n} \left\langle \boldsymbol{u}_i \;\middle|\; \boldsymbol{A}\boldsymbol{u}_i \right\rangle$$

holds and the desired equality is true for any orthonomal basis.

For $n \times n$-matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, we denote by $\boldsymbol{A} \leqq \boldsymbol{B}$ the property that $\boldsymbol{B} - \boldsymbol{A}$ is a positive semi-definite matrix. Then, for an arbitrary unitary matrix $\boldsymbol{U}$, we have $\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U} \leqq \boldsymbol{U}^*\boldsymbol{B}\boldsymbol{U}$. This is bacause for any $\boldsymbol{x} \in \mathbb{K}^n$

$$\left\langle \boldsymbol{x} \;\middle|\; (\boldsymbol{U}^*\boldsymbol{B}\boldsymbol{U} - \boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U})\,\boldsymbol{x} \right\rangle \;=\; \left\langle \boldsymbol{x} \;\middle|\; \boldsymbol{U}^*(\boldsymbol{B}-\boldsymbol{A})\,\boldsymbol{U}\boldsymbol{x} \right\rangle \;=\; \left\langle \boldsymbol{U}\boldsymbol{x} \;\middle|\; (\boldsymbol{B}-\boldsymbol{A})\,\boldsymbol{U}\boldsymbol{x} \right\rangle \;\geqq\; 0.$$

We also have $\mathrm{Tr}\,(\boldsymbol{A}) \leqq \mathrm{Tr}\,(\boldsymbol{B})$, bacause

$$\mathrm{Tr}\,(\boldsymbol{B}) - \mathrm{Tr}\,(\boldsymbol{A}) \;=\; \mathrm{Tr}\,(\boldsymbol{B}-\boldsymbol{A}) \;=\; \sum_{i=1}^{n} \left\langle \boldsymbol{e}_i \;\middle|\; (\boldsymbol{B}-\boldsymbol{A})\,\boldsymbol{e}_i \right\rangle \;\geqq\; 0.$$

Next, we show that

$$\left|\mathrm{Tr}\,(\boldsymbol{X}\boldsymbol{A})\right| \;\leqq\; \|\boldsymbol{X}\|\,\|\boldsymbol{A}\|_{\mathrm{Tr}}\,.$$

for $n \times n$-matrices $\boldsymbol{A}$ and $\boldsymbol{X}$. Note that

$$\boldsymbol{O} \;\leqq\; \boldsymbol{X}^*\boldsymbol{X} \;\leqq\; \|\boldsymbol{X}^*\boldsymbol{X}\|\,\boldsymbol{I} \;=\; \|\boldsymbol{X}\|^2\,\boldsymbol{I},$$

where we use the facts shown in Sect. 7.4. Since $\boldsymbol{A} = \boldsymbol{U} [\![\boldsymbol{A}]\!]$ for some unitary matrix $\boldsymbol{U}$, it follows that

$$\begin{aligned}
\left|\mathrm{Tr}\,(\boldsymbol{X}\boldsymbol{A})\right| &= \left|\mathrm{Tr}\,(\boldsymbol{X}\boldsymbol{U}\,[\![\boldsymbol{A}]\!])\right| \\
&= \left|\mathrm{Tr}\left(\boldsymbol{X}\boldsymbol{U}\sqrt{[\![\boldsymbol{A}]\!]}\sqrt{[\![\boldsymbol{A}]\!]}\right)\right| \\
&\leqq \sqrt{\mathrm{Tr}\left(\left(\boldsymbol{X}\boldsymbol{U}\sqrt{[\![\boldsymbol{A}]\!]}\right)^{*}\boldsymbol{X}\boldsymbol{U}\sqrt{[\![\boldsymbol{A}]\!]}\right)}\sqrt{\mathrm{Tr}\left(\left(\sqrt{[\![\boldsymbol{A}]\!]}\right)^{*}\sqrt{[\![\boldsymbol{A}]\!]}\right)} \\
&= \sqrt{\mathrm{Tr}\left(\sqrt{[\![\boldsymbol{A}]\!]}\boldsymbol{U}^{*}\boldsymbol{X}^{*}\boldsymbol{X}\boldsymbol{U}\sqrt{[\![\boldsymbol{A}]\!]}\right)}\sqrt{\mathrm{Tr}\,([\![\boldsymbol{A}]\!])} \\
&\leqq \sqrt{\mathrm{Tr}\left(\sqrt{[\![\boldsymbol{A}]\!]}\boldsymbol{U}^{*}\left\|\boldsymbol{X}^{*}\boldsymbol{X}\right\|\boldsymbol{I}\boldsymbol{U}\sqrt{[\![\boldsymbol{A}]\!]}\right)}\sqrt{\mathrm{Tr}\,([\![\boldsymbol{A}]\!])} \\
&= \left\|\boldsymbol{X}\right\|\sqrt{\mathrm{Tr}\,([\![\boldsymbol{A}]\!])}\sqrt{\mathrm{Tr}\,([\![\boldsymbol{A}]\!])} \\
&= \left\|\boldsymbol{X}\right\|\left\|\boldsymbol{A}\right\|_{\mathrm{Tr}},
\end{aligned}$$

where the first inequality is the Schwarz inequality for the inner produnt $(\boldsymbol{X}, \boldsymbol{Y}) \mapsto \mathrm{Tr}\,(\boldsymbol{X}^{*}\boldsymbol{Y})$ (we will see this in Exercise 10.3.1) and the second inequality is due to the previous remarks.

1.
$$\left\|\boldsymbol{A}\right\|_{\mathrm{Tr}} = \mathrm{Tr}\,(\boldsymbol{U}_2{}^{*}\boldsymbol{\Sigma}\boldsymbol{U}_2) = \mathrm{Tr}\,(\boldsymbol{U}_2\boldsymbol{U}_2{}^{*}\boldsymbol{\Sigma}) = \mathrm{Tr}\,(\boldsymbol{\Sigma}) = \sum_{i=1}^{k}\sigma_i.$$

2. The first half of 2 is shown by 1 and the fact that all singular values of $\boldsymbol{A}$ are nonnegative. The second half is due to the fact that $\boldsymbol{A} = \boldsymbol{O}$ if and only if all of the singular values of $\boldsymbol{A}$ are zero.

3. Note that the singular values of $c\boldsymbol{A}$ are $|c|\,\sigma_1$, $|c|\,\sigma_2$, $\ldots$, $|c|\,\sigma_k$.

4. Since $[\![\boldsymbol{A}+\boldsymbol{B}]\!] = \boldsymbol{V}(\boldsymbol{A}+\boldsymbol{B})$ for some unitary matrix $\boldsymbol{V}$, it follows that

$$\begin{aligned}
\left\|\boldsymbol{A}+\boldsymbol{B}\right\|_{\mathrm{Tr}} &= \mathrm{Tr}\left(\boldsymbol{V}(\boldsymbol{A}+\boldsymbol{B})\right) \\
&\leqq \left|\mathrm{Tr}\,(\boldsymbol{V}\boldsymbol{A})\right| + \left|\mathrm{Tr}\,(\boldsymbol{V}\boldsymbol{B})\right| \\
&\leqq \left\|\boldsymbol{V}\right\|\left\|\boldsymbol{A}\right\|_{\mathrm{Tr}} + \left\|\boldsymbol{V}\right\|\left\|\boldsymbol{B}\right\|_{\mathrm{Tr}} \\
&= \left\|\boldsymbol{A}\right\|_{\mathrm{Tr}} + \left\|\boldsymbol{B}\right\|_{\mathrm{Tr}}.
\end{aligned}$$

Exercise 10.3. Prove the following properties:

1. $\langle \cdot \mid \cdot \rangle_{\mathrm{HS}}$ is an inner product on the space of all square matrices of order n.
2. $\|\boldsymbol{A}\|_{\mathrm{HS}} = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$, where a_{ij} is the (i,j)-element of $\boldsymbol{A}$.
3. $\|\boldsymbol{A}\|_{\mathrm{HS}} = \left(\sum_{i=1}^{k} |\sigma_i|^2 \right)^{1/2}$, where $\sigma_1, \sigma_2, \ldots, \sigma_k$ are all of the singular values of $\boldsymbol{A}$.

$\|\boldsymbol{A}\|_{\mathrm{HS}}$ is called the *Hilbert–Schmidt norm* of $\boldsymbol{A}$.

1. Positivity: Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ be an orthogonal basis (for example, thestandard basis) of $\mathbb{K}^n$. Then

$$\langle \boldsymbol{A} \mid \boldsymbol{A} \rangle_{\mathrm{HS}} = \mathrm{Tr}\,(\boldsymbol{A}^*\boldsymbol{A}) = \sum_{i=1}^{n} \langle \boldsymbol{e}_i \mid \boldsymbol{A}^*\boldsymbol{A}\boldsymbol{e}_i \rangle = \sum_{i=1}^{n} \langle \boldsymbol{A}\boldsymbol{e}_i \mid \boldsymbol{A}\boldsymbol{e}_i \rangle \geqq 0.$$

If $\langle \boldsymbol{A} \mid \boldsymbol{A} \rangle_{\mathrm{HS}} = 0$, then $\boldsymbol{A}\boldsymbol{e}_i = \boldsymbol{0}$ for every $i = 1, 2, \ldots, n$. Since for any $\boldsymbol{x} \in \mathbb{K}^n$, which is expressed as

$$\boldsymbol{x} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{e}_n \quad (a_1, a_2, \ldots, a_n \in \mathbb{K}),$$

it follows that

$$\boldsymbol{A}\boldsymbol{x} = a_1\boldsymbol{A}\boldsymbol{e}_1 + a_2\boldsymbol{A}\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{A}\boldsymbol{e}_n = \boldsymbol{0}.$$

Hence $\boldsymbol{A} = \boldsymbol{O}$.

The Hermitian property:

$$\overline{\langle \boldsymbol{A} \mid \boldsymbol{B} \rangle_{\mathrm{HS}}} = \overline{\mathrm{Tr}\,(\boldsymbol{A}^*\boldsymbol{B})} = \mathrm{Tr}\left((\boldsymbol{A}^*\boldsymbol{B})^*\right) = \mathrm{Tr}\,(\boldsymbol{B}^*\boldsymbol{A}) = \langle \boldsymbol{B} \mid \boldsymbol{A} \rangle_{\mathrm{HS}}.$$

Here, the properties shown in Sect. 5.4 are used. Other axioms of the inner product follow from the linearity of Tr.

2. Note that the i-th diagonal component of $\boldsymbol{A}^*\boldsymbol{A}$ is expressed as $\sum_{j=1}^{n} |a_{ij}|^2$ for $i = 1, 2, \ldots, n$.

3.

$$\langle \boldsymbol{A} \mid \boldsymbol{A} \rangle_{\mathrm{HS}} = \mathrm{Tr}\,(\boldsymbol{A}^*\boldsymbol{A}) = \mathrm{Tr}\,(\boldsymbol{U}_2{}^*\boldsymbol{\Sigma}^*\boldsymbol{\Sigma}\boldsymbol{U}_2) = \mathrm{Tr}\,(\boldsymbol{\Sigma}^*\boldsymbol{\Sigma}) = \sum_{i=1}^{k} |\sigma_i|^2.$$

Exercise 10.4. Prove that the maximum singular value of square matrix $\boldsymbol{A}$ equals the spectral radius (the matrix norm) of $\boldsymbol{A}$. (Hint: use $\|\boldsymbol{A}^*\boldsymbol{A}\| = \|\boldsymbol{A}\|^2$ shown in Section 7.4.)

Since $\boldsymbol{A}^*\boldsymbol{A}$ is a Hermitian matrix, its norm equals the maximum of the absolute values of the eigenvalues, but since all of its eigenvalues are nonnegative, the maximum (it is the square of the largest singular value) is the norm of $\|\boldsymbol{A}^*\boldsymbol{A}\| = \|\boldsymbol{A}\|^2$.

Exercise 10.5. Prove the above properties 1−3 of $\simeq_W$.

1. Since $\boldsymbol{x} - \boldsymbol{x} = \mathbf{0} \in W$, we have $\boldsymbol{x} \simeq_W \boldsymbol{x}$.

2. Suppose $\boldsymbol{x} \simeq_W \boldsymbol{y}$. Then, $\boldsymbol{x} - \boldsymbol{y} \in W$ and $\boldsymbol{y} - \boldsymbol{x} = -(\boldsymbol{x} - \boldsymbol{y}) \in W$. Therefore, $\boldsymbol{y} \simeq_W \boldsymbol{x}$.

3. Suppose $\boldsymbol{x} \simeq_W \boldsymbol{y}$ and $\boldsymbol{y} \simeq_W \boldsymbol{z}$. Then, $\boldsymbol{x} - \boldsymbol{y} \in W$ and $\boldsymbol{y} - \boldsymbol{z} \in W$. Since

$$\boldsymbol{x} - \boldsymbol{x} \ = \ (\boldsymbol{x} - \boldsymbol{y}) + (\boldsymbol{y} - \boldsymbol{z}) \in W,$$

   we have $\boldsymbol{x} \simeq_W \boldsymbol{z}$.

Exercise 10.6. Show that the addition and scalar multiplication above are well-defined, that is, they do not depend on the choice of representatives $\boldsymbol{x}$ and $\boldsymbol{y}$ of the classes.

Suppose $[\boldsymbol{x}]_W = [\boldsymbol{u}]_W$ and $[\boldsymbol{y}]_W = [\boldsymbol{v}]_W$. Then, $\boldsymbol{x} \simeq_W \boldsymbol{u}$ and $\boldsymbol{y} \simeq_W \boldsymbol{v}$. That is, $\boldsymbol{x} - \boldsymbol{u} \in W$ and $\boldsymbol{y} - \boldsymbol{v} \in W$. Since

$$(\boldsymbol{x} + \boldsymbol{y}) - (\boldsymbol{u} + \boldsymbol{v}) \;=\; (\boldsymbol{x} - \boldsymbol{u}) - (\boldsymbol{y} - \boldsymbol{v}) \in W,$$

we have $\boldsymbol{x} + \boldsymbol{y} \simeq_W \boldsymbol{u} + \boldsymbol{v}$. Therefore, $[\boldsymbol{x} + \boldsymbol{y}]_W = [\boldsymbol{u} + \boldsymbol{v}]_W$. This says that addition does not depends on on the choice of representatives $\boldsymbol{x}$ and $\boldsymbol{y}$ of the classes. The well-definedness of scalar multiplication is left to the reader.

Exercise 10.7. Prove the results discussed above in detail.

(1) φ is additive, homogenous and Hermitian: Let $\boldsymbol{\xi} = \sum_{i=1}^{k} a_i (\boldsymbol{x}_i, \boldsymbol{y}_i)$, $\boldsymbol{\eta} = \sum_{j=1}^{l} b_j \left(\boldsymbol{x}'_j, \boldsymbol{y}'_j\right)$, $\boldsymbol{\zeta} = \sum_{n=1}^{m} c_n (\boldsymbol{x}''_n, \boldsymbol{y}''_n)$ $(a_i, b_j, c_n \in \mathbb{K}; \boldsymbol{x}_i, \boldsymbol{x}'_j, \boldsymbol{x}''_n \in V; \boldsymbol{y}_i, \boldsymbol{y}'_j, \boldsymbol{y}''_n \in W)$ and $a, b \in \mathbb{K}$. By the definition of φ, we have

$$\begin{aligned} & \varphi(\boldsymbol{\xi}, a\boldsymbol{\eta} + b\boldsymbol{\zeta}) \\ = \;& \sum_{i=1}^{k} \left(\sum_{j=l}^{k} \overline{a_i} a b_j \left\langle \boldsymbol{x}_i \mid \boldsymbol{x}'_j \right\rangle_V \left\langle \boldsymbol{y}_i \mid \boldsymbol{y}'_j \right\rangle_W + \sum_{n=1}^{m} \overline{a_i} b c_n \left\langle \boldsymbol{x}_i \mid \boldsymbol{x}''_n \right\rangle_V \left\langle \boldsymbol{y}_i \mid \boldsymbol{y}''_n \right\rangle_W \right) \\ = \;& a \sum_{i=1}^{k} \sum_{j=l}^{k} \overline{a_i} b_j \left\langle \boldsymbol{x}_i \mid \boldsymbol{x}'_j \right\rangle_V \left\langle \boldsymbol{y}_i \mid \boldsymbol{y}'_j \right\rangle_W + b \sum_{i=1}^{k} \sum_{n=l}^{m} \overline{a_i} c_n \left\langle \boldsymbol{x}_i \mid \boldsymbol{x}''_n \right\rangle_V \left\langle \boldsymbol{y}_i \mid \boldsymbol{y}''_m \right\rangle_W \\ = \;& a\varphi(\boldsymbol{\xi}, \boldsymbol{\eta}) + b\varphi(\boldsymbol{\xi}, \boldsymbol{\zeta}) \end{aligned}$$

Also we have

$$\begin{aligned} \overline{\varphi(\boldsymbol{\xi}, \boldsymbol{\eta})} &= \sum_{i=1}^{k} \sum_{j=1}^{l} \overline{\overline{a_i} b_j \left\langle \boldsymbol{x}_i \mid \boldsymbol{x}'_j \right\rangle_V \left\langle \boldsymbol{y}_i \mid \boldsymbol{y}'_j \right\rangle_W} \\ &= \sum_{i=1}^{k} \sum_{j=1}^{l} \overline{b_j} a_i \left\langle \boldsymbol{x}'_i \mid \boldsymbol{x}_j \right\rangle_V \left\langle \boldsymbol{y}'_i \mid \boldsymbol{y}_j \right\rangle_W \\ &= \varphi(\boldsymbol{\eta}, \boldsymbol{\xi}). \end{aligned}$$

Note that by the above results we have

$$\begin{aligned} \varphi(a\boldsymbol{\xi} + b\boldsymbol{\eta}, \boldsymbol{\zeta}) &= \overline{\varphi(\boldsymbol{\zeta}, a\boldsymbol{\xi} + b\boldsymbol{\eta})} = \overline{a\varphi(\boldsymbol{\zeta}, \boldsymbol{\xi}) + b\varphi(\boldsymbol{\zeta}, \boldsymbol{\eta})} = \overline{a}\overline{\varphi(\boldsymbol{\zeta}, \boldsymbol{\xi})} + \overline{b}\overline{\varphi(\boldsymbol{\zeta}, \boldsymbol{\eta})} \\ &= \overline{a}\varphi(\boldsymbol{\xi}, \boldsymbol{\zeta}) + \overline{b}\varphi(\boldsymbol{\eta}, \boldsymbol{\zeta}) \end{aligned}$$

(2) φ is zero on $\mathcal{K} \times \mathcal{L}$ and $\mathcal{L} \times \mathcal{K}$: For $\boldsymbol{\xi} = (\boldsymbol{x}_1, \boldsymbol{y}) + (\boldsymbol{x}_2, \boldsymbol{y}) - (\boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y})$ we have

$$\varphi(\boldsymbol{\eta}, \boldsymbol{\xi}) = \overline{\varphi(\boldsymbol{\xi}, \boldsymbol{\eta})} = 0.$$

Similarly, for $\boldsymbol{\xi} = (\boldsymbol{x}, \boldsymbol{y}_1) + (\boldsymbol{x}, \boldsymbol{y}_2) - (\boldsymbol{x}, \boldsymbol{y}_1 + \boldsymbol{y}_2)$ we have

$$\varphi(\boldsymbol{\xi}, \boldsymbol{\eta}) = \varphi(\boldsymbol{\eta}, \boldsymbol{\xi}) = 0.$$

For $\boldsymbol{\xi} = a(\boldsymbol{x}, \boldsymbol{y}) - (a\boldsymbol{x}, \boldsymbol{y})$ we have

$$\varphi(\boldsymbol{\eta}, \boldsymbol{\xi}) = a\varphi\big(\boldsymbol{\eta}, (\boldsymbol{x}, \boldsymbol{y})\big) - \varphi\big(\boldsymbol{\eta}, a(\boldsymbol{x}, \boldsymbol{y})\big) = 0,$$

and

$$\varphi(\boldsymbol{\xi},\boldsymbol{\eta}) = \overline{\varphi(\boldsymbol{\eta},\boldsymbol{\xi})} \;=\; 0.$$

Similarly, for $\boldsymbol{\xi} = a(\boldsymbol{x},\boldsymbol{y}) - (\boldsymbol{x}, a\boldsymbol{y})$ we have

$$\varphi(\boldsymbol{\xi},\boldsymbol{\eta}) \;=\; \varphi(\boldsymbol{\eta},\boldsymbol{\xi}) \;=\; 0.$$

(3) $\langle \cdot \mid \cdot \rangle$ is well-defined: Consider $\boldsymbol{\xi},\boldsymbol{\xi}',\boldsymbol{\eta},\boldsymbol{\eta}' \in \mathcal{L}$ with $\boldsymbol{\xi}-\boldsymbol{\xi}', \boldsymbol{\eta}-\boldsymbol{\eta}' \in \mathcal{K}$. Then

$$\begin{aligned}
\varphi(\boldsymbol{\xi},\boldsymbol{\eta}) &= \varphi\big(\boldsymbol{\xi}' + (\boldsymbol{\xi}-\boldsymbol{\xi}'), \boldsymbol{\eta}' + (\boldsymbol{\eta}-\boldsymbol{\eta}')\big) \\
&= \varphi(\boldsymbol{\xi}',\boldsymbol{\eta}') + \varphi\big(\boldsymbol{\xi}', (\boldsymbol{\eta}-\boldsymbol{\eta}')\big) + \varphi\big(\boldsymbol{\xi}-\boldsymbol{\xi}', (\boldsymbol{\eta})\big) \\
&= \varphi(\boldsymbol{\xi}',\boldsymbol{\eta}').
\end{aligned}$$

Thus, $\varphi(\boldsymbol{\xi},\boldsymbol{\eta})$ does notdepend on the choice of the representatives $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$.

Exercise 10.8. Find the spectral decomposition of $\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ and the singular value decomposition of $\boldsymbol{B} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix}$. Use NumPy or SymPy if necessary.

Program: `exercise10_8_1.py`

In [1]:
```
from sympy import *

A = Matrix([[1, 2],
            [2, 1]])

ans = A.eigenvects()

lmd1 = ans[0][0]
lmd2 = ans[1][0]

v1 = Matrix(ans[0][2])
v2 = Matrix(ans[1][2])

e1 = v1 / v1.norm()
e2 = v2 / v2.norm()

print(lmd1, lmd2)

E1 = e1 * e1.T
E2 = e2 * e2.T
print(E1, E2)
print(lmd1*E1 + lmd2*E2)
```

```
-1 3
Matrix([[1/2, -1/2], [-1/2, 1/2]]) Matrix([[1/2, 1/2], [1/2, 1/2]])
Matrix([[1, 2], [2, 1]])
```

The spectral decomposition is $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = (-1)\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} + 3\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

Program: `exercise10_8_2.py`

In [1]:
```
from numpy import *

A = array([[1, 2, 3],
           [2, 3, 4]])

W, sigma, U = linalg.svd(A)
Sigma = diag([sigma[0], sigma[1]])
VT = U[:2,:]
```

In [1]:
```

print(f'W = \n{W}')
print(f'Sigma = \n{Sigma}')
print(f'VT = \n{VT}')

print(f'W*Sigma*VT = \n{W.dot(Sigma.dot(VT))}')
```

```
W =
[[-0.56959484 -0.82192562]
 [-0.82192562  0.56959484]]
Sigma =
[[6.54675564 0.        ]
 [0.         0.37415323]]
VT =
[[-0.33809817 -0.55064932 -0.76320047]
 [ 0.84795222  0.17354729 -0.50085764]]
W*Sigma*VT =
[[1. 2. 3.]
 [2. 3. 4.]]
```

Program: exercise10_8_3.py

In [1]:
```
from sympy import *
from functools import reduce

A = Matrix([[1, 2, 3],
            [2, 3, 4]])

B = A.T * A
ans = B.eigenvects()

s0 = sqrt(ans[0][0])
s1 = sqrt(ans[1][0])
s2 = sqrt(ans[2][0])

v0 = ans[0][2][0]
v1 = ans[1][2][0]
v2 = ans[2][2][0]

v1 /= v1.norm()
v2 /= v2.norm()

w1 = A*v1 / s1
w2 = A*v2 / s2

V = reduce(lambda x, y: x.row_join(y), [v1, v2])
W = reduce(lambda x, y: x.row_join(y), [w1, w2])
S = diag(s1, s2)

C = W * S * V.T
```

In [2]:
```
simplify(C)
```

Out[2]:
```
Matrix([
[1, 2, 3],
[2, 3, 4]])
```

In [3]:
```
W.evalf()
```

```
Matrix([
[ 0.821925617555625, 0.569594837762601],
[-0.569594837762601, 0.821925617555625]])
```

In [4]:
```
S.evalf()
```

```
Matrix([
[0.374153226240496,                 0],
[                0, 6.54675563644267]])
```

In [5]:
```
V.T.evalf()
```

```
Matrix([
[-0.847952217751687, -0.173547289245653, 0.500857639260381],
[ 0.338098165838458,  0.550649318285984,  0.76320047073351]])
```

$$
\begin{aligned}
\boldsymbol{W} &= \begin{bmatrix} 0.569594837762601 & 0.821925617555625 \\ 0.821925617555625 & -0.569594837762601 \end{bmatrix} \\
\boldsymbol{\Sigma} &= \begin{bmatrix} 6.54675563644267 & 0 \\ 0 & 0.374153226240496 \end{bmatrix} \\
\boldsymbol{V}^* &= \begin{bmatrix} 0.338098165838458 & 0.550649318285984 & 0.76320047073351 \\ -0.847952217751687 & -0.173547289245653 & 0.500857639260381 \end{bmatrix}
\end{aligned}
$$

Exercise 10.9. Calculate $\boldsymbol{PQ} - \boldsymbol{QP}$ in the case of $n = 2$, $\lambda_1 = \mu_1$, $\lambda_2 = \mu_2$.

Program: `exercise10_9.py`

In [1]:
```
from sympy import *

lmd1, lmd2 = symbols('lmd1, lmd2', real=True)
mu1, mu2 = lmd1, lmd2

e1 = Matrix([[1], [0]])
e2 = Matrix([[0], [1]])

f1 = Matrix([[1/sqrt(2)], [1/sqrt(2)]])
f2 = Matrix([[1/sqrt(2)], [-1/sqrt(2)]])

P = lmd1*e1*e1.T + lmd2*e2*e2.T
Q = mu1*f1*f1.T + mu2*f2*f2.T

print(simplify(P*Q - Q*P))
```

```
Matrix([[0, (lmd1 - lmd2)**2/2], [-(lmd1 - lmd2)**2/2, 0]])
```

Supplementary Problem. Prove that no pair of square matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ can satisfy the equality $\boldsymbol{PQ} - \boldsymbol{QP} = \boldsymbol{I}$. (Hint: Consider the trace of both sides.)

Exercise 10.10. Prove that the above mapping $\overline{f}$ is bijective.

(1) $\overline{f}$ is surjective: Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ be a basis of V, and let $\boldsymbol{Y} \in L(\Omega, p; V)$ be arbitrary. Then, for $\omega \in \Omega$, $\boldsymbol{Y}(\omega)$ is iniquely written as

$$\boldsymbol{Y}(\omega) = \sum_{i=1}^{n} a_i(\omega)\, \boldsymbol{e}_i$$

with $a_i(\omega) \in \mathbb{K}$. Then $a_i \in L(\Omega, p)$ for each $i = 1, 2, \ldots, n$, and we have (see page 247)

$$\overline{f}\left(\sum_{i=1}^{n} a_i \otimes \boldsymbol{e}_i\right) = \sum_{i=1}^{n} a_i \boldsymbol{e}_i = \boldsymbol{Y}.$$

(2) $\overline{f}$ is injective: Any element $\boldsymbol{X}$ of $L(\Omega, p) \otimes V$ is (by the construction of the tensor product) expressed as

$$\boldsymbol{X} = \sum_{i=1}^{n} a_i \otimes \boldsymbol{e}_i$$

with $a_i \in L(\Omega, p)$. Suppose

$$\overline{f}(\boldsymbol{X}) = \sum_{i=1}^{n} a_i \boldsymbol{e}_i = \mathbf{0}.$$

Then, for $\omega \in \Omega$ we have

$$\sum_{i=1}^{n} a_i(\omega)\, \boldsymbol{e}_i = \mathbf{0}.$$

Ti follows that $a_i(\omega) = 0$ for all $i = 1, 2, \ldots, n>$ Because ω is arbitrary, we see $a_i = 0$ for all $i = 1, 2, \ldots, n$. Therefore, $\boldsymbol{x} = \mathbf{0}$, and $\ker\left(\overline{f}\right) = \{\mathbf{0}\}$.

Exercise 10.11. Prove the linearity of expectation above (Hint: first prove it for the case where both $\boldsymbol{A}$ and $\boldsymbol{B}$ are matrix units).

We consider an arbitrary but fixed $n \times k$-matrix $\boldsymbol{B}$. If $\boldsymbol{A}$ be a matrix unit, then we can easily verify that

$$E(\boldsymbol{AXB}) = \boldsymbol{A}E(\boldsymbol{XB})$$

holds. Hence, this equality holds for any $l \times m$-matrix $\boldsymbol{A}$. Next, by the similar discussion,

$$E(\boldsymbol{XB}) = E(\boldsymbol{X})\boldsymbol{B}$$

holds for any $n \times k$-matrix $\boldsymbol{B}$. Thus, the desired equality holds.

Exercise 10.12. Prove that the above $\langle \cdot \mid \cdot \rangle$ satisfies axioms of the inner product. Moreover, the mapping sending a matrix $\boldsymbol{A} \in M_{\mathbb{K}}(m,n)$ to the vector $\boldsymbol{v} \in \mathbb{K}^{mn}$ consisting of the elements of $\boldsymbol{A}$ is a linear isomorphism preserving the inner product.

This problem is essentially the same as 1 in Exercuse 10.3. There we considered the case of square matrices, but here we are considering arbitrary rectangular matrices.