Linear Algebra

MPhil Preliminary Course 2019–2020

Solution Manual

Daniel Wales
Faculty of Economics
University of Cambridge

Contents

1 Introduction

This solution manual is intended for use alongside the Linear Algebra component of the MPhil course. As such, many of the solutions presented are intentionally brief and more heuristic than those found in complementary textbooks intended for more rigorous study of this subject.

Email me with any typos, corrections and suggestions at ddgw2@cam.ac.uk. I have benefited from suggestions from Alexis De Boeck, and many thanks go to him. Additional explanations and corrected typos appear in red.

2 Preliminary concepts

Exercise 2.1. Given that $\boldsymbol{A}$ is symmetric and $\boldsymbol{B}$ is skew-symmetric ($\boldsymbol{B} = -\boldsymbol{B}'$), find a, b, c, u, v, w, x, y and z.

$$\boldsymbol{A} = \begin{bmatrix} 3 & a & -1 \\ 2 & 5 & c \\ b & 8 & 2 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} u & 3 & v \\ w & x & y \\ -2 & 6 & z \end{bmatrix}.$$

By inspection we have $a = 2$, $b = -1$ and $c = 8$.

As $-\boldsymbol{B}' = \begin{bmatrix} -u & -w & 2 \\ -3 & -x & -6 \\ -v & -y & -z \end{bmatrix}$ we have that $u = x = z = 0$, $v = 2$, $w = -3$ and $y = -6$.

Exercise 2.2. Show that the identity matrix of any dimension may be said to be square, symmetric, diagonal, lower triangular, upper triangular, orthogonal and idempotent.

Let $\boldsymbol{I}_n$ represent the identity matrix of size n.

(i) square, diagonal;

$\boldsymbol{I}_n$ is defined as a square (size $n \times n$), diagonal matrix with ones along the leading diagonal.

(ii) symmetric;

$\boldsymbol{I}_n = \boldsymbol{I}_n'$ is clear, as we have a diagonal matrix.

(iii) lower triangular;

As $\boldsymbol{I}_n$ has zeros on the upper triangular area, $\boldsymbol{I}_n$ may be said to be lower triangular.

(iv) upper triangular;

As $\boldsymbol{I}_n$ has zeros on the lower triangular area, $\boldsymbol{I}_n$ may be said to be upper triangular.

(v) orthogonal;

Orthogonal if $\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{I}_n$, which is trivially satisfied in the case of the identity matrix, with $\boldsymbol{I}_n'\boldsymbol{I}_n = \boldsymbol{I}_n$.

(vi) idempotent;

Idempotent if $\boldsymbol{A} = \boldsymbol{A}^2$, which is again trivially satisfied here, as $\boldsymbol{I}_n^2 = \boldsymbol{I}_n$.

Exercise 2.3. Show, for conformable matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$, the properties of matrix addition.

(i) $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$;

Let $\boldsymbol{D} = \boldsymbol{A} + \boldsymbol{B}$, with a typical element $(\boldsymbol{D})_{ij} = (\boldsymbol{A})_{ij} + (\boldsymbol{B})_{ij}$.

Then $(\boldsymbol{D})_{ij} = (\boldsymbol{A})_{ij} + (\boldsymbol{B})_{ij} = (\boldsymbol{B})_{ij} + (\boldsymbol{A})_{ij}$, and hence the statement is proved.

(ii) $c(\boldsymbol{A} + \boldsymbol{B}) = c\boldsymbol{A} + c\boldsymbol{B}$, for any $c \in \mathbf{R}$;

Again, a typical element would show $c((\boldsymbol{A})_{ij} + (\boldsymbol{B})_{ij}) = c(\boldsymbol{A})_{ij} + c(\boldsymbol{B})_{ij}$

(iii) $\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C}$.

This proof is trivial. Notice that all follow directly from the properties of addition and multiplication in $\mathbf{R}$ because matrix addition is defined element wise.

Exercise 2.4. Show that $(\boldsymbol{A} + \boldsymbol{B})' = \boldsymbol{A}' + \boldsymbol{B}'$ for any two conformable matrices $\boldsymbol{A}$ and $\boldsymbol{B}$.

Let $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$, with a typical element $(\boldsymbol{C})_{ij} = (\boldsymbol{A})_{ij} + (\boldsymbol{B})_{ij}$.

Next, transpose the matrix $\boldsymbol{C}$, such that:

$$\begin{aligned}(\boldsymbol{C}')_{ij} = (\boldsymbol{C})_{ji} &= (\boldsymbol{A})_{ji} + (\boldsymbol{B})_{ji}, \\ &= (\boldsymbol{A}')_{ij} + (\boldsymbol{B}')_{ij},\end{aligned}$$

and thus $(\boldsymbol{A} + \boldsymbol{B})' = \boldsymbol{A}' + \boldsymbol{B}'$.

Exercise 2.5. Show, for conformable vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$, the three properties of an inner product.

(i) $\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{b} \cdot \boldsymbol{a}$;

This follows from commutativity of multiplication in $\mathbf{R}$ as clearly:

$$\boldsymbol{a} \cdot \boldsymbol{b} := \sum_{i=1}^{n} a_i b_i = \sum_{i=1}^{n} b_i a_i = \boldsymbol{b} \cdot \boldsymbol{a}.$$

(ii) $\boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{c}$;

Here we may use the definition and expand brackets:

$$\begin{aligned}\boldsymbol{a}\cdot(\boldsymbol{b}+\boldsymbol{c}) &= \sum_{i=1}^{n} a_i(b_i+c_i),\\ &= \sum_{i=1}^{n} a_i b_i + \sum_{i=1}^{n} a_i c_i,\\ &= \boldsymbol{a}\cdot\boldsymbol{b}+\boldsymbol{a}\cdot\boldsymbol{c}.\end{aligned}$$

(iii) $(c_1\boldsymbol{a})\cdot(c_2\boldsymbol{b}) = c_1c_2(\boldsymbol{a}\cdot\boldsymbol{b})$, for any $c_1, c_2 \in \mathbf{R}$.

$$\begin{aligned}(c_1\boldsymbol{a})\cdot(c_2\boldsymbol{b}) &= \sum_{i=1}^{n}(c_1a_i)(c_2b_i),\\ &= c_1c_2\sum_{i=1}^{n} a_i b_i,\\ &= c_1c_2(\boldsymbol{a}\cdot\boldsymbol{b}).\end{aligned}$$

Exercise 2.6.

(i) Prove $(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}'$. Using this, or otherwise, prove that $(\boldsymbol{ABC})' = \boldsymbol{C}'\boldsymbol{B}'\boldsymbol{A}'$.

The ijth element of $\boldsymbol{AB}$ may be given as $(\boldsymbol{AB})_{ij} = \sum_{s=1}^{k} a_{is}b_{sj}$.

We know that the ijth element of $(\boldsymbol{AB})'$ will be the jith element of $\boldsymbol{AB}$.

$$\begin{aligned}[(\boldsymbol{AB})']_{ij} = (\boldsymbol{AB})_{ji} &= \sum_{s=1}^{k} a_{js}b_{si},\\ &= \sum_{s=1}^{k} b_{si}a_{js},\\ &= \sum_{s=1}^{k} b'_{is}a'_{sj},\\ &= [\boldsymbol{B}'\boldsymbol{A}']_{ij},\end{aligned}$$

where $b'_{is} = (B')_{is}$ and $a'_{is} = (A')_{is}$.

For the further proof, let $\boldsymbol{D} = \boldsymbol{AB}$ such that

$$\begin{aligned}(\boldsymbol{DC})' &= \boldsymbol{C}'\boldsymbol{D}' && \text{by part (i), but}\\ \boldsymbol{D}' &= (\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}' && \text{again, from part (i), and thus}\\ (\boldsymbol{DC})' &= \boldsymbol{C}'\boldsymbol{B}'\boldsymbol{A}'\\ (\boldsymbol{ABC})' &= \boldsymbol{C}'\boldsymbol{B}'\boldsymbol{A}'.\end{aligned}$$

(ii) Under what condition is $(\boldsymbol{AB})' = \boldsymbol{A}'\boldsymbol{B}'$?

We know that both

$$(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}',$$
$$(\boldsymbol{BA})' = \boldsymbol{A}'\boldsymbol{B}',$$

are true. Therefore $(\boldsymbol{AB})' = \boldsymbol{A}'\boldsymbol{B}'$ will require $\boldsymbol{AB} = \boldsymbol{BA}$.

(iii) Find an example of two distinct matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ such that $\boldsymbol{AB} = \boldsymbol{BA}$.

Consider the matrix form of scalar multiplication with:

$$\boldsymbol{A} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \quad \text{and} \quad \boldsymbol{B} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}, \quad \text{such that:}$$

$$\boldsymbol{AB} = \begin{bmatrix} ae & af \\ ag & ah \end{bmatrix} = a\boldsymbol{B}, \quad \text{and} \quad \boldsymbol{BA} = \begin{bmatrix} ea & fa \\ ga & ha \end{bmatrix} = a\boldsymbol{B}.$$

More generally we have already discussed how the identity matrix, $\boldsymbol{I}_n$, will obey this property.

Exercise 2.7. Consider the matrices: $\boldsymbol{A}$ is 3×5, $\boldsymbol{B}$ is 5×3, $\boldsymbol{C}$ is 5×1 and D is 3×1. Which of the following matrix operations are allowed?

(i) $\underset{5\times 3}{\boldsymbol{B}} \; \underset{3\times 5}{\boldsymbol{A}}$;

Allowed (conformable), and results in a square matrix of size 5×5.

(ii) $\underset{3\times 5}{\boldsymbol{A}} \; \underset{5\times 3}{\boldsymbol{B}}$;

Allowed (conformable), and results in a square matrix of size 3×3.

(iii) $\underset{3\times 5}{\boldsymbol{A}} \; \underset{5\times 3}{\boldsymbol{B}} \; \underset{5\times 1}{\boldsymbol{C}}$;

This operation is not allowed (non-conformable), as $\boldsymbol{BC}$ is not permissible.

(iv) $\underset{3\times 1}{\boldsymbol{D}} \; \underset{3\times 5}{\boldsymbol{A}} \; \underset{5\times 3}{\boldsymbol{B}}$;

This operation is not allowed (non-conformable), as $\boldsymbol{DB}$ is not permissible.

(v) $\underset{3\times 5}{\boldsymbol{A}}(\underset{5\times 3}{\boldsymbol{B}} + \underset{5\times 1}{\boldsymbol{C}})$.

This operation is not allowed (non-conformable).

Exercise 2.8. Let

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad \text{and} \quad \boldsymbol{B} = \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix}.$$

Compute $\boldsymbol{AB}'$, $\boldsymbol{BA}'$, $\boldsymbol{A}'\boldsymbol{B}$ and $\boldsymbol{B}'\boldsymbol{A}$.

$$AB' = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 5 & 2 \\ 6 & 3 \end{bmatrix} = \begin{bmatrix} 4+10+18 & 1+4+9 \\ 16+25+36 & 4+10+18 \end{bmatrix} = \begin{bmatrix} 32 & 14 \\ 77 & 32 \end{bmatrix}$$

$$BA' = \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 4+10+18 & 16+25+36 \\ 1+4+9 & 4+10+18 \end{bmatrix} = \begin{bmatrix} 32 & 77 \\ 14 & 32 \end{bmatrix}$$

$$A'B = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 4+4 & 5+8 & 6+12 \\ 8+5 & 10+10 & 12+15 \\ 12+6 & 15+12 & 18+18 \end{bmatrix} = \begin{bmatrix} 8 & 13 & 18 \\ 13 & 20 & 27 \\ 18 & 27 & 36 \end{bmatrix}$$

$$B'A = \begin{bmatrix} 4 & 1 \\ 5 & 2 \\ 6 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 4+4 & 8+5 & 12+6 \\ 5+8 & 10+10 & 15+12 \\ 6+12 & 12+15 & 18+18 \end{bmatrix} = \begin{bmatrix} 8 & 13 & 18 \\ 13 & 20 & 27 \\ 18 & 27 & 36 \end{bmatrix}$$

The solution above extensively answers the question here, though using property (v) from the notes we could substantially reduce our work load by noting:

$$(AB')' = BA' \qquad \text{and} \qquad (A'B)' = B'A.$$

and using the transpose of matrices in our answer.

Exercise 2.9. For square matrices A and B, which of the following statements are true, and why?

(i) $(A+B)^2 = (B+A)^2$;

True, as $A+B = B+A$.

(ii) $(A+B)^2 = A^2 + 2AB + B^2$;

False, as generally $AB \neq BA$.

(iii) $(A+B)^2 = A(A+B) + B(A+B)$;

True, direct result of properties.

(iv) $(A+B)^2 = (A+B)(B+A)$;

True, again as $A+B = B+A$.

(v) $(A+B)^2 = A^2 + AB + BA + B^2$.

True, by expanding brackets.

Exercise 2.10. For square matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, which of the following statements are true, and why?

(i) $(\boldsymbol{A}-\boldsymbol{B})^2 = (\boldsymbol{B}-\boldsymbol{A})^2$;

True, expand brackets of both sides and equate.

(ii) $(\boldsymbol{A}-\boldsymbol{B})^2 = \boldsymbol{A}^2 - \boldsymbol{B}^2$;

False, expand brackets

(iii) $(\boldsymbol{A}-\boldsymbol{B})^2 = \boldsymbol{A}^2 - 2\boldsymbol{A}\boldsymbol{B} + \boldsymbol{B}^2$;

Again, false, as generally $\boldsymbol{A}\boldsymbol{B} \neq \boldsymbol{B}\boldsymbol{A}$.

(iv) $(\boldsymbol{A}-\boldsymbol{B})^2 = \boldsymbol{A}(\boldsymbol{A}-\boldsymbol{B}) - \boldsymbol{B}(\boldsymbol{A}-\boldsymbol{B})$;

True, direct result of properties.

(v) $(\boldsymbol{A}-\boldsymbol{B})^2 = \boldsymbol{A}^2 - \boldsymbol{A}\boldsymbol{B} - \boldsymbol{B}\boldsymbol{A} + \boldsymbol{B}^2$.

True, by expanding brackets.

Exercise 2.11. Which rows or columns or matrices do you multiply to find

Firstly consider the following representation of $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{A}\boldsymbol{B}$.

$$\boldsymbol{A} = \begin{bmatrix} \text{———} \boldsymbol{a}^1 \text{———} \\ \text{———} \boldsymbol{a}^2 \text{———} \\ \vdots \\ \text{———} \boldsymbol{a}^n \text{———} \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \begin{pmatrix} | \\ \boldsymbol{b}_1 \\ | \end{pmatrix} & \begin{pmatrix} | \\ \boldsymbol{b}_2 \\ | \end{pmatrix} & \cdots & \begin{pmatrix} | \\ \boldsymbol{b}_m \\ | \end{pmatrix} \end{bmatrix},$$

$$\boldsymbol{A}\boldsymbol{B} = \begin{bmatrix} \boldsymbol{a}^1 \cdot \boldsymbol{b}_1 & \boldsymbol{a}^1 \cdot \boldsymbol{b}_2 & \dots & \boldsymbol{a}^1 \cdot \boldsymbol{b}_m \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}^n \cdot \boldsymbol{b}_1 & \boldsymbol{a}^n \cdot \boldsymbol{b}_2 & \dots & \boldsymbol{a}^n \cdot \boldsymbol{b}_m \end{bmatrix}.$$

(i) the third column of $\boldsymbol{A}\boldsymbol{B}$;

$\boldsymbol{A}\boldsymbol{b}_3$.

(ii) the first row of $\boldsymbol{A}\boldsymbol{B}$;

$\boldsymbol{a}^1\boldsymbol{B}$.

(iii) the entry in row three, column 4 of $\boldsymbol{A}\boldsymbol{B}$;

$a^3 b_4$.

(iv) the entry in row 1, column 1 of CDE.

$(CD)^1 e_1 = c^1 D e_1$. Where we are able to use the result from parts (ii) and (iii).

Exercise 2.12. An econometrician collects data on the number of years of education and the marital status of n different individuals. They organise the data for the ith individual as a 2×1 vector $\boldsymbol{x}_i$ and arranges these n 2-dimensional vectors into an $n \times 2$ matrix $\boldsymbol{X}$. Show that $\boldsymbol{X}'\boldsymbol{X} = \sum_{i=1}^n \boldsymbol{x}_i \boldsymbol{x}_i'$.

This is an important exercise as moving seemly between matrix, vector and scalar notation will be extremely useful during econometrics courses. For clarity we will write the solution to this problem in matrix notation, vector notation and scalar notation. We may first clearly specify the vector $\boldsymbol{x}_i$ as

$$\underset{2\times 1}{\boldsymbol{x}_i} = \begin{bmatrix} e_i \\ m_i \end{bmatrix} \quad \text{with} \quad \underset{1\times 2}{\boldsymbol{x}_i'} = \begin{bmatrix} e_i & m_i \end{bmatrix},$$

where e_i represents the years of education for individual-i and m_i represents the marital status for individual-i and the dimensions are set according to the question. Next we notice the dimension of matrix $\boldsymbol{X}$ implies that we must stack the values of $\boldsymbol{x}_i'$ as

$$\underset{n\times 2}{\boldsymbol{X}} = \begin{bmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \vdots \\ \boldsymbol{x}_n' \end{bmatrix} = \begin{bmatrix} e_1 & m_1 \\ e_2 & m_2 \\ \vdots & \vdots \\ e_n & m_n \end{bmatrix} \quad \text{with} \quad \underset{2\times n}{\boldsymbol{X}'} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_n \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & \dots & e_n \\ m_1 & m_2 & \dots & m_n \end{bmatrix}$$

Finally we observe that

$$\underset{2\times n\, n\times 2}{\boldsymbol{X}'\boldsymbol{X}} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \vdots \\ \boldsymbol{x}_n' \end{bmatrix} = \sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i',$$

$$= \begin{bmatrix} e_1 & e_2 & \dots & e_n \\ m_1 & m_2 & \dots & m_n \end{bmatrix} \begin{bmatrix} e_1 & m_1 \\ e_2 & m_2 \\ \vdots & \vdots \\ e_n & m_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n e_i e_i & \sum_{i=1}^n e_i m_i \\ \sum_{i=1}^n m_i e_i & \sum_{i=1}^n m_i m_i \end{bmatrix}.$$

Exercise 2.13. The econometrician also collects income data on the same n individuals. Let $\boldsymbol{y}$ be the $n \times 1$ vector with ith element y_i. Show that $\boldsymbol{X}'\boldsymbol{y} = \sum_{i=1}^n \boldsymbol{x}_i y_i$.

As above, with the variable y_i as a scalar. Specifically let

$$\underset{n\times 1}{\boldsymbol{y}} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

such that

$$\underset{2\times n}{X'}\underset{n\times 1}{\boldsymbol{y}} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} \boldsymbol{x}_i y_i = \begin{bmatrix} \sum_{i=1}^{n} e_i y_i \\ \sum_{i=1}^{n} m_i y_i \end{bmatrix}.$$

Exercise 2.14. Suppose that for every $i = 1, \dots, n$, the econometrician applies a weight $1/\sigma_i$ to the ith observation and arranges these n 2-dimensional observations into a matrix $\boldsymbol{Z}$ of order $n \times 2$ such that the ijth element of $\boldsymbol{Z}$ is $(\boldsymbol{X})_{ij}/\sigma_i$. Let $\boldsymbol{\Omega}$ be the diagonal matrix with $(\boldsymbol{\Omega})_{ii} = \sigma_i^2$. Show that

$$Z'Z = \sum_{i=1}^{n} (\boldsymbol{x}_i/\sigma_i)(\boldsymbol{x}_i/\sigma_i)' = \sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i'/\sigma_i^2 = \boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}$$

and

$$\sum_{i=1}^{n} (\boldsymbol{x}_i/\sigma_i) y_i/\sigma_i = \sum_{i=1}^{n} \boldsymbol{x}_i y_i/\sigma_i^2 = \boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{y}.$$

We are told that

$$Z = \begin{bmatrix} \boldsymbol{x}_1'/\sigma_1 \\ \vdots \\ \boldsymbol{x}_n'/\sigma_n \end{bmatrix}.$$

Therefore, using the result in previous questions we have that

$$Z'Z = \sum_{i=1}^{n} (\boldsymbol{x}_i/\sigma_i)(\boldsymbol{x}_i/\sigma_i)' = \sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i'/\sigma_i^2 = \boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X},$$

where the final equality uses the fact that $(\Omega^{-1})_{ii} = \sigma_{ii}^{-2}$, which is true from the section on inverses.

Exercise 2.15. Multiply a 3×3 matrix $\boldsymbol{A}$ and $\boldsymbol{I}_3$ using columns of $\boldsymbol{A}$ times rows of $\boldsymbol{I}_3$.

If

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

then

$$AI_3 = \begin{bmatrix} a \\ d \\ g \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} b \\ e \\ h \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} + \begin{bmatrix} c \\ f \\ i \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

$$AI_3 = \begin{bmatrix} a & 0 & 0 \\ d & 0 & 0 \\ g & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & b & 0 \\ 0 & e & 0 \\ 0 & h & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & c \\ 0 & 0 & f \\ 0 & 0 & i \end{bmatrix}$$

$$AI_3 = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = A$$

Exercise 2.16. Multiply AB below using *columns times rows*:

$$AB = \begin{bmatrix} 1 & 0 \\ 2 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 3 & 0 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \begin{bmatrix} 3 & 3 & 0 \end{bmatrix} + \underline{\hspace{4cm}} = \underline{\hspace{4cm}}.$$

$$AB = \begin{bmatrix} 1 & 0 \\ 2 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 3 & 0 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \begin{bmatrix} 3 & 3 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 4 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 0 \\ 6 & 6 & 0 \\ 6 & 6 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 4 & 8 & 4 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 0 \\ 10 & 14 & 4 \\ 7 & 8 & 1 \end{bmatrix}.$$

Exercise 2.17. Construct all the permutation matrices of order $n = 3$. How many permutation matrices of order $n = 4$ are there?

There will be $3! = 6$ permutation matrices of order 3:

$$I_3, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Order 4 will have $4! = 24$ permutation matrices.

Exercise 2.18. Suppose you are given 101 points on the surface of a sphere. What is the dimension of the matrix you would require in order to map those points to 101 points on the surface of a 3-dimensional ellipsoid?

This will require a 3×3 matrix, since we transform points on a sphere to points on an ellipsoid. Again by changing the (3-dimensional) basis.

Exercise 2.19. Consider the matrix

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \\ x_{51} & x_{52} & x_{53} \end{bmatrix}.$$

Construct a matrix A such that

$$AX = \begin{bmatrix} x_{21} - x_{11} & x_{22} - x_{12} & x_{23} - x_{13} \\ x_{31} - x_{21} & x_{32} - x_{22} & x_{33} - x_{23} \\ x_{41} - x_{31} & x_{42} - x_{32} & x_{43} - x_{33} \\ x_{51} - x_{41} & x_{52} - x_{42} & x_{53} - x_{43} \end{bmatrix}.$$

A must be of dimension 4×5, and gives the operation of subtracting the previous row from the current row.

$$A = \begin{bmatrix} r_2 - r_1 \\ r_3 - r_2 \\ r_4 - r_3 \\ r_5 - r_4 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}.$$

Exercise 2.20. Decompose the following matrices into the product of series of elementary matrices and state the operations each matrix represents.

(i) The matrix A represents four times the second row, and is already an elementary matrix;

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} = E_2(4).$$

(ii) The matrix B is not expressible in terms of elementary matrices.;

(iii) The matrix C firstly[1] swaps rows 2 and 3, then multiplies row 4 by 3, and finally adds rows 2 to row 1;

$$C = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = E_1(1|2)E_4(3)E_{23}.$$

[1] Of course, the concept of 'firstly' depends whether we consider the operation Cx of $x'C$, we assume the former here.

3 Systems of linear equations

Exercise 3.1. Let $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{z}$ be vectors of the same order. Show that:

(i) $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$;

$$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix} = \begin{bmatrix} v_1 + u_1 \\ \vdots \\ v_n + v_n \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$$

This follows from the addition property in $\mathbf{R}$ as this operation is defined element wise.

(ii) $(\boldsymbol{u} + \boldsymbol{v}) + z = \boldsymbol{v} + (\boldsymbol{u} + z)$.

Again, this follows from the addition property in $\mathbf{R}$ as this operation is defined element wise.

Exercise 3.2. For vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ of the same order and scalars c and d. Show that

(i) $(c + d)(\boldsymbol{u} + \boldsymbol{v}) = c\boldsymbol{v} + c\boldsymbol{u} + d\boldsymbol{v} + d\boldsymbol{u}$;

By expanding brackets for scaler quantities. Follows from properties of additional and multiplication in $\mathbf{R}$.

(ii) (Harder) the zero vector is uniquely determined by the condition that $c\mathbf{0} = \mathbf{0}$ for all scalars $c \in \mathbf{R}$.

Case 1. For $c = 1$. Obvious statement. $\mathbf{0} = \mathbf{0}$.

Case 2. For $c \neq 1$ The statement is equivalent to (we may rewrite the statement for each element as):

$$\begin{aligned} &\forall i \in \{1, \cdots, n\}, \forall c \in \mathbf{R}\backslash\{1\} : \\ &\qquad (c-1)\boldsymbol{v}_i = 0 \rightarrow \boldsymbol{v}_i = 0. \end{aligned}$$

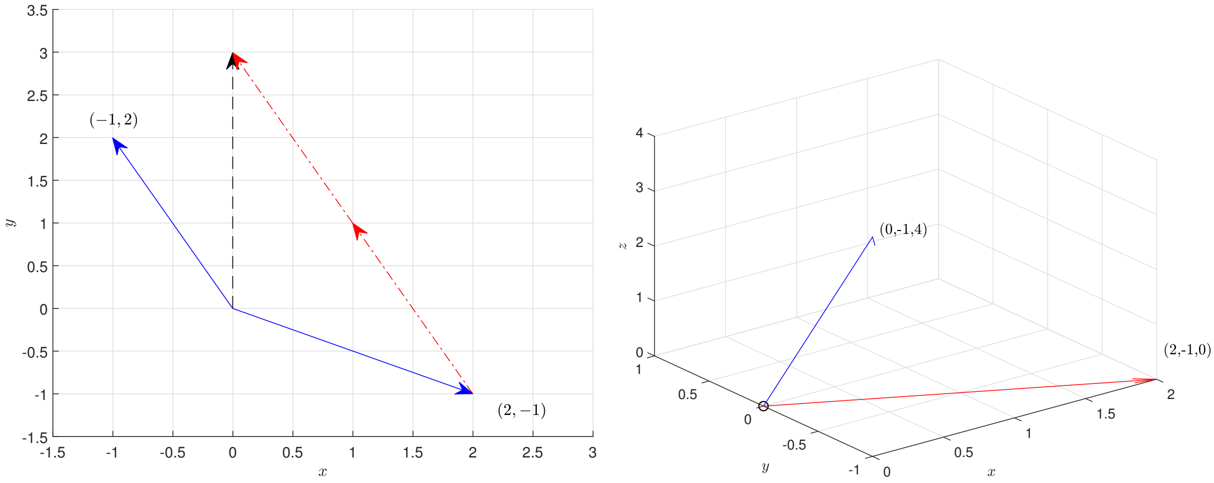
Assume not. This means

$$\begin{aligned} &\exists i, \exists c \in \mathbf{R}\backslash\{1\} \\ &(c-1)\boldsymbol{v}_i = 0 \rightarrow \boldsymbol{v}_i \neq 0. \end{aligned}$$

However, we then have $c - 1 = 0 \rightarrow c = 1$, and arrive at a contradiction.

Exercise 3.3. Assume that a unique solution exits to the system $\boldsymbol{Ax} = \boldsymbol{b}$, where:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 6 & -3 & 1 \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{and} \quad \boldsymbol{b} = \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}.$$



(a) Exercise 3.4

(b) Exercise 3.5

Notes: Second vector removed from solutoion in (b) to prevent negative axis in x and z-axes. Becomes difficult to visualise.

Figure 1: Solution to linear system (unique)

It is immediately revealed (possibly via picture in notes) that $\boldsymbol{Ax} = \boldsymbol{b} \leftrightarrow \boldsymbol{a}_1 x + \boldsymbol{a}_2 y + \boldsymbol{a}_3 z$, and that $2\boldsymbol{a}_3 = \boldsymbol{b}$. The unique solution must therefore be $\boldsymbol{x} = \{0, 0, 2\}$.

Exercise 3.4. Write the following system in matrix form and draw the corresponding column picture. By simple inspection, identify the $\boldsymbol{x}$ vector that solves the system

$$\begin{aligned} 2x - y &= 0 \\ -x + 2y &= 3. \end{aligned}$$

The system in form $\boldsymbol{Ax} = \boldsymbol{b}$ yields:

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix},$$

in which we notice that $\boldsymbol{A}$ is a symmetric matrix. The unique solution is clearly at $\boldsymbol{x} = \{(1, 2)\}$, as shown in Figure 1a.

Exercise 3.5. Write the following system in matrix form and draw the corresponding column picture. By simple inspection, identify the $\boldsymbol{x}$ vector that solves the system

$$\begin{aligned} 2x - y &= 0 \\ -x + 2y - z &= -1 \\ -3y + 4z &= 4. \end{aligned}$$

The system in form $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ yields:

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix},$$

in which we notice that only the final column of $\boldsymbol{A}$ matters for the (unique) solution of $\boldsymbol{x} = \{(0, 0, 1)\}$. This shown in Figure 1b.

Exercise 3.6. Consider the following sets of vectors, $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ and state whether they are linear independent. If they are linearly dependent write down a set of associated scalars c_1, c_2, c_3 to show this.

(i) $\boldsymbol{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$;

Clearly linearly dependant as

$$\boldsymbol{u} + \frac{1}{2}\boldsymbol{v} - \boldsymbol{w} = 0,$$

such that $c_1 = 1$, $c_2 = \frac{1}{2}$ and $c_3 = -1$.

(ii) $\boldsymbol{u} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$;

These vectors are linearly independent (indeed they may be used to form the identity matrix $\boldsymbol{I}_3$). Observe the only possible combination $c_1\boldsymbol{u} + c_2\boldsymbol{v} + c_3\boldsymbol{w} = \boldsymbol{0}$ is with $c_1 = c_2 = c_3 = 0$.

(iii) $\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$;

Again, clearly linearly dependant as $2\boldsymbol{u} = \boldsymbol{v}$ and hence

$$2\boldsymbol{u} - \boldsymbol{v} + 0\boldsymbol{w} = 0,$$

such that $c_1 = 2$, $c_2 = -1$ and $c_3 = 0$.

(iv) $\boldsymbol{u} = \begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} 3 \\ 6 \\ 8 \end{bmatrix}$;

These vectors are linearly independent. Assume not, then we show need to show that for some set of c_1, c_2, c_3

$$c_1\boldsymbol{u} + c_2\boldsymbol{v} + c_3\boldsymbol{w} = \boldsymbol{0},$$

which gives the system

$$\begin{aligned} c_1 + c_2 + 3c_3 &= 0, \\ 4c_1 + c_2 + 6c_3 &= 0, \\ 5c_1 + 8c_3 &= 0, \end{aligned}$$

Subtract the first from the second equation to give two linear equations in two unknowns

$$\begin{aligned} 3c_1 + 3c_3 &= 0, \\ 5c_1 + 8c_3 &= 0, \end{aligned}$$

rewrite to ready eliminate c_1

$$\begin{aligned} c_1 &= c_3, \\ c_1 &= -\frac{8}{5}c_3, \end{aligned}$$

which implies $c_1 = c_3 = 0$. Then, from either of the first two (original) equations observe that $c_2 = 0$ is be the only solution. Hence $c_1 = c_2 = c_3 = 0$ and these vectors are linearly independent.

(v) $\boldsymbol{u} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 3 \\ 7 \\ -7 \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} -1 \\ -1 \\ 5 \end{bmatrix}$;

These vectors are linearly dependent. Observe $2\boldsymbol{u} - \boldsymbol{v} = \boldsymbol{w}$, such that

$$2\boldsymbol{u} - \boldsymbol{v} + \boldsymbol{w} = \boldsymbol{0},$$

and we have $c_1 = 2$, $c_2 = -1$ and $c_3 = 1$.

(vi) $\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 2 \\ 2 \\ -4 \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$;

These vectors are linearly independent. Assume not, then we show need to show that for some set of c_1, c_2, c_3

$$c_1\boldsymbol{u} + c_2\boldsymbol{v} + c_3\boldsymbol{w} = \boldsymbol{0},$$

which gives the system

$$\begin{aligned} c_1 + 2c_2 + c_3 &= 0, \\ 2c_1 + 2c_2 + 2c_3 &= 0, \\ -c_1 - 4c_2 + 3c_3 &= 0, \end{aligned}$$

Subtract two times the first equation from the second to give

$$-2c_2 = 0 \longrightarrow c_2 = 0,$$

use the result $c_2 = 0$ in the other equations to make the system

$$\begin{aligned} c_1 + c_3 &= 0, \\ 2c_1 + 2c_3 &= 0, \\ -c_1 + 3c_3 &= 0, \end{aligned}$$

the first and second equations are clearly multiples of one another which imply $c_1 = -c_3$. The final equation, however implies $c_1 = 3c_3$, such that the only solution must be with $c_1 = c_2 = c_3 = 0$, and hence these vectors are linearly independent.

Exercise 3.7. Let $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ be linearly independent vectors of order n, show that $(\boldsymbol{x} + \boldsymbol{y})$, $(\boldsymbol{x} + \boldsymbol{z})$, and $(\boldsymbol{y} + \boldsymbol{z})$ are also linearly independent.

Observe the systems, applying the definitions of linear independence:

$$c_1\boldsymbol{x} + c_2\boldsymbol{y} + c_3\boldsymbol{z} = \boldsymbol{0}, \quad \text{and} \quad a_1(\boldsymbol{x} + \boldsymbol{y}) + a_2(\boldsymbol{x} + \boldsymbol{z}) + a_3(\boldsymbol{y} + \boldsymbol{z}) = \boldsymbol{0},$$

and notice the unique correspondence between these systems as through the relationship:

$$\begin{aligned} a_1 + a_2 &= c_1, \\ a_1 + a_3 &= c_2, \\ a_2 + a_3 &= c_3, \end{aligned}$$

This system may be inverted (in terms of a_1, a_2 and a_3) to show

$$\begin{aligned} a_1 &= \frac{c_1 + c_2 - c_3}{2}, \\ a_2 &= \frac{c_1 + c_3 - c_2}{2}, \\ a_3 &= \frac{c_2 + c_3 - c_1}{2}. \end{aligned}$$

such that given coefficient triples c_1, c_2 and c_3, then a_1, a_2 and a_3 may be found. With $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ a linearly independent set of vectors, we know that the only solution to the first system must be $c_1 = 0$, $c_2 = 0$ and $c_3 = 0$. Hence, using the correspondence, $a_1 = 0$, $a_2 = 0$ and $a_3 = 0$ may be inferred as the unique solution to the second system, and linear independence is shown.

Exercise 3.8. Let $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ be a set of linearly independent vectors. Consider another vector $\boldsymbol{v}$ such that the vectors in $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p, \boldsymbol{v}\}$ are linearly dependent. Show that $\boldsymbol{v}$ can be expressed as a unique linear combination of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$.

Consider the set-up. Linear *independence* between the vectors in the set $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ infers that the only set of scalars $c_1, \cdots, c_p$ such that

$$c_1\boldsymbol{v}_1 + \cdots + c_p\boldsymbol{v}_p = \mathbf{0}.$$

will be the zero vector, $\boldsymbol{c} = \mathbf{0}$. In addition, as the set $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p, \boldsymbol{v}\}$ are linearly *dependent* we have that there exists scalars $a_1, \cdots, a_p, a$ such that

$$a_1\boldsymbol{v}_1 + \cdots + a_p\boldsymbol{v}_p + a\boldsymbol{v} = \mathbf{0}.$$

where the scalars are not all equal to zero, such that $\boldsymbol{a} \neq \mathbf{0}$. In particular, notice that if $a = 0$, this would contract the first statement, and thus $a \neq 0$. Hence we may write

$$\boldsymbol{v} = -\frac{a_1}{a}\boldsymbol{v}_1 - \cdots - \frac{a_p}{a}\boldsymbol{v}_p.$$

which shows that $\boldsymbol{v}$ may be written as a linear combination of the vectors in the set $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$.

In a second step we now prove uniqueness. Suppose there were another set of scalars $b_1, \cdots, b_p$ such that

$$\boldsymbol{v} = b_1\boldsymbol{v}_1 + \cdots + b_p\boldsymbol{v}_p.$$

We may then subtract the first linear combination from the second and show that

$$\mathbf{0} = \left(b_1 + \frac{a_1}{a}\right)\boldsymbol{v}_1 + \cdots + \left(b_p + \frac{a_p}{a}\right)\boldsymbol{v}_p.$$

We know from the linear independence of the set $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ that the only scalars for which this is true are zero. Such that $b_i + \frac{a_i}{a} = 0\ \forall i$ or $b_i = -\frac{a_i}{a}$, and hence the linear combination must be unique.

Exercise 3.9. Find the rank of the following matrices

$$\text{(i)}\ \ \boldsymbol{A} = \begin{bmatrix} 0 & 3 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 7 \end{bmatrix}, \quad \text{(ii)}\ \ \boldsymbol{B} = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 2 & 4 & 2 \\ 0 & 2 & 2 & 1 \end{bmatrix}, \quad \text{(iii)}\ \ \boldsymbol{C} = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 0 & 1 \\ 0 & 2 & -1 \\ 1 & 1 & 4 \end{bmatrix}, \quad \text{(iv)}\ \ \boldsymbol{D} = \begin{bmatrix} 1 & 4 & 0 \\ 0 & 3 & 12 \\ 0 & 2 & 8 \end{bmatrix}.$$

(i) As the matrix is square, $\text{rk}(\boldsymbol{A}) \leq 3$, and in this case $\text{rk}(\boldsymbol{A}) = 3$ as each column is clearly independent of each other;

(ii) First, as $\boldsymbol{B}$ is of size 3×4 the $\text{rk}(\boldsymbol{B}) \leq 3$. Then notice $\text{rk}(\boldsymbol{B}) = 3$ as columns $\boldsymbol{b}_4 = 2\boldsymbol{b}_3$ while columns $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ and $\boldsymbol{b}_4$ are clearly independent;

(iii) As $\boldsymbol{C}$ is of size 4×3 we know $\text{rk}(\boldsymbol{C}) \leq 3$. Here $\text{rk}(\boldsymbol{C}) = 3$ as all columns are linearly independent;

(iv) We know $\text{rk}(\boldsymbol{D}) \leq 3$, but since $\boldsymbol{d}_3 = 4(\boldsymbol{d}_2 - 4\boldsymbol{d}_1)$, these columns are not linearly independent, and hence we have $\text{rk}(\boldsymbol{D}) = 2$.

Of course since the row rank and column rank are the same we may have examined row vectors instead. This may have been easier to immediately verify the correct solution in the case of matrix $\boldsymbol{B}$

Exercise 3.10. Let $\boldsymbol{A}$ be an $n \times p$ matrix. Suppose you know that there exists an $n \times 1$ vector $\boldsymbol{a}$ such that if $\boldsymbol{a}$ is added as an additional column to $\boldsymbol{A}$, the rank of increases by 1, i.e. $\text{rk}(\boldsymbol{A}|\boldsymbol{a}) = \text{rk}(\boldsymbol{A}) + 1$. Show that this implies the rows of $\boldsymbol{A}$ are linearly dependent. *Hint: prove by contradiction.*

Suppose not, such that the rows are linearly *independent* but adding a column to A yields $\text{rk}(\boldsymbol{A}|\boldsymbol{a}) = \text{rk}(\boldsymbol{A}) + 1$. If the rows are linearly independent then we have than $\text{rk}(\boldsymbol{A}) = n$ (full row rank). Adding an additional column will not alter the number of rows, and therefore the matrix $\boldsymbol{A}|\boldsymbol{a}$ must still have rank at most n. Or $\text{rk}(\boldsymbol{A}|\boldsymbol{a}) \leq \min(n, p+1) = n$. This contradicts the statement that $\text{rk}(\boldsymbol{A}|\boldsymbol{a}) = n + 1 > n$, and therefore $\text{rk}(\boldsymbol{A}) < n$ and the rows are linearly dependent.

Exercise 3.11. Show that it is not true in general that $\text{rk}(\boldsymbol{AB}) = \text{rk}(\boldsymbol{BA})$ for two square matrices $\boldsymbol{A}$ and $\boldsymbol{B}$.

We know that the correct relationship is $\text{rk}(\boldsymbol{AB}) \leq \min(\text{rk}(\boldsymbol{A}), \text{rk}(\boldsymbol{B}))$. One counterexample is sufficient to disprove the statement. For simplicity, consider the case of matrices with size 2×2 matrices, both with rank deficiency.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{B} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix},$$

such that $\text{rk}(\boldsymbol{A}) = \text{rk}(\boldsymbol{B}) = 1$ is clear with

$$\boldsymbol{AB} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{BA} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

and therefore $\text{rk}(\boldsymbol{AB}) = 1$ and $\text{rk}(\boldsymbol{BA}) = 0$ and we have found a contradiction which disproves the statement.

Exercise 3.12. Suppose two states of nature exist $S \in \{\text{Rain}, \text{Sun}\}$, such that $n = 2$. Two financial assets exist, one which pays out 2 units in rain and another which pays out 1 unit in any state of the world. By considering the returns matrix, $\boldsymbol{R}$, argue why financial markets are complete. If a household desires 4 units of consumption in the sunny state, but none when it rains, what are the quantities of each asset they much purchase for their portfolio?

The returns matrix will be given by $\boldsymbol{R} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$, which clearly has full column rank, $\text{rk}\,\boldsymbol{R} = 2$, and hence financial markets are complete.

Intuitively a consumer can achieve any level of consumption across states of nature in two steps:

(i) Use the second asset to purchase as much consumption in the sunny states as desired;

(ii) Then, buy or sell units of the first asset to achieve the correct level of consumption in the rainy state of the world;

This strategy answers the second part of the question. The household's portfolio will contain 4 units of the second asset and short two of the first.

4 Square matrices

Exercise 4.1. Evaluate the determinants of the following matrices using the cofactor expansion method, along an appropriate row or column

(i) For matrix $\boldsymbol{A}$, choose to expand along either row 2 or column 2;

$$\det \boldsymbol{A} = \begin{vmatrix} 2 & 5 & 1 \\ 1 & 0 & 2 \\ 7 & 1 & 1 \end{vmatrix} = -1 \begin{vmatrix} 5 & 1 \\ 1 & 1 \end{vmatrix} - 2 \begin{vmatrix} 2 & 5 \\ 7 & 1 \end{vmatrix} = -(5-1) - 2(2-35) = -4 + 66 = 62.$$

(ii) For matrix $\boldsymbol{B}$, choose to expand along row 2, then row 2 again;

$$\det \boldsymbol{B} = \begin{vmatrix} 7 & 5 & 2 & 3 \\ 2 & 0 & 0 & 0 \\ 11 & 2 & 0 & 0 \\ 23 & 57 & 1 & -1 \end{vmatrix} = -2 \begin{vmatrix} 5 & 2 & 3 \\ 2 & 0 & 0 \\ 57 & 1 & -1 \end{vmatrix} = 4 \begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} = 4(-2-3) = -20.$$

(iii) For matrix $\boldsymbol{C}$, choose to expand along either column 1 and then column 3's, or column 4 and then column 1's;

$$\det \boldsymbol{C} = \begin{vmatrix} 1 & 2 & 1 & 0 \\ 3 & 2 & 1 & 0 \\ 0 & 1 & 6 & 5 \\ 0 & 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 6 & 5 \\ 1 & 1 & 1 \end{vmatrix} - 3 \begin{vmatrix} 2 & 1 & 0 \\ 1 & 6 & 5 \\ 1 & 1 & 1 \end{vmatrix} \equiv -2 \begin{vmatrix} 2 & 1 & 0 \\ 1 & 6 & 5 \\ 1 & 1 & 1 \end{vmatrix}$$

$$\det \boldsymbol{C} = -2\Big(-5 \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} + \begin{vmatrix} 2 & 1 \\ 1 & 6 \end{vmatrix} \Big) = -2\Big(-5(2-1) + (12-1) \Big) = -2(-5+11) = -12.$$

(iv) A multitude of expansion options exist for matrix $\boldsymbol{D}$. Here column 1 is repeatedly used;

$$\det \boldsymbol{D} = \begin{vmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} = -\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = -\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = -1.$$

(v) For matrix $\boldsymbol{E}$, choose to expand always along row 1's;

$$\det \boldsymbol{E} = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 2 & 9 & 3 \\ 0 & 0 & 1 & 0 & 7 & 4 \\ 0 & 6 & 9 & 8 & 7 & 5 \\ 1 & 3 & 4 & 2 & 9 & 6 \end{vmatrix} = -\begin{vmatrix} 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 2 & 9 \\ 0 & 0 & 1 & 0 & 7 \\ 0 & 6 & 9 & 8 & 7 \\ 1 & 3 & 4 & 2 & 9 \end{vmatrix} = -3\begin{vmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 6 & 9 & 8 \\ 1 & 3 & 4 & 2 \end{vmatrix} = 6\begin{vmatrix} 0 & 0 & 1 \\ 0 & 6 & 9 \\ 1 & 3 & 4 \end{vmatrix} = 6\begin{vmatrix} 0 & 6 \\ 1 & 3 \end{vmatrix} = -36.$$

(vi) Matrix $\boldsymbol{F}$ has no preferred column or row for cofactor expansion. Here column 1 is chosen;

$$\det \boldsymbol{F} = \begin{vmatrix} 3 & t & -2 \\ -1 & 5 & 3 \\ 2 & 1 & 1 \end{vmatrix} = 3\begin{vmatrix} 5 & 3 \\ 1 & 1 \end{vmatrix} + \begin{vmatrix} t & -2 \\ 1 & 1 \end{vmatrix} + 2\begin{vmatrix} t & -2 \\ 5 & 3 \end{vmatrix} = 3(5-3) + (t+2) + 2(3t+10) = 28 + 7t.$$

Exercise 4.2. Using property (vii) of determinants [that the determinant of a matrix $\boldsymbol{A}$ changes sign when two rows are interchanged], prove that if two rows of a matrix $\boldsymbol{A}$ are equal then $\det \boldsymbol{A} = 0$. *Hint: proof by contradiction by supposing that* $\boldsymbol{A}$ *has two equal rows, but* $\det(A) \neq 0$.

Use the Hint. Suppose $\boldsymbol{A}$ has two equal rows, but that $\det \boldsymbol{A} = a \neq 0$. Let $\boldsymbol{B}$ be the matrix with the two equal rows exchanged. Trivially we have that $\boldsymbol{A} = \boldsymbol{B}$, but that while $\det \boldsymbol{A} = a$, property (vii) will mean $\det \boldsymbol{B} = -a$. We therefore arrive at a contradiction, which may be eradicated by changing the original claim to read $a = 0$.

Exercise 4.3. Prove property (viii) of determinants [that subtracting a multiple of one row of $\boldsymbol{A}$ from another leaves $\det \boldsymbol{A}$ unchanged] by using the result from exercise 4.2 and property (x) of determinants [that the determinant is a function of each row separately]. *Hint: Use both parts of property (x) to separate the relevant determinant into two components and apply exercise 4.2.*

Let $\boldsymbol{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$ and for some $c \in \mathbf{R}$, we have $\tilde{\boldsymbol{A}} = \begin{bmatrix} a_{11} + ca_{21} & \dots & a_{1n} + ca_{2n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$

Then, $\det \tilde{\boldsymbol{A}} = \det \boldsymbol{A} + c \begin{vmatrix} a_{21} & \dots & a_{2n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{vmatrix}$.

The using exercise exercise 4.2, as two rows of this component are the same, $\det \tilde{\boldsymbol{A}} = \det \boldsymbol{A}$, which proves property (viii)

Exercise 4.4. Prove property (ix) of determinants [that if $\boldsymbol{A}$ has a row of zeros then $\det \boldsymbol{A} = 0$] by using property (x) of determinants [that the determinant is a function of each row separately].

This property follows directly by multiplying any matrix by zero, as we have:

$$0 = 0 \times \det \boldsymbol{A} = 0 \times \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{vmatrix} = \begin{vmatrix} 0 \times a_{11} & \dots & 0 \times a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{vmatrix} = \begin{vmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{vmatrix}.$$

Exercise 4.5. Prove or disprove the statement: $\det(\boldsymbol{A}+\boldsymbol{B}) = \det \boldsymbol{A} + \det \boldsymbol{B}$.

This statement may be disproved using a simple counterexample.

Let $\boldsymbol{A} = \boldsymbol{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\boldsymbol{B} = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}$, such that $\boldsymbol{A}+\boldsymbol{B} = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}$.

Then we have that $\det \boldsymbol{A} = 1$, while $\det \boldsymbol{B} = -2$ and $\det(\boldsymbol{A}+\boldsymbol{B}) = 2 \neq -1 = \det \boldsymbol{A} + \det \boldsymbol{B}$.

Exercise 4.6. By considering the matrices

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{B} = \begin{bmatrix} 0 & 0 \\ 3 & 4 \end{bmatrix}$$

and the properties of determinants, show that $\boldsymbol{AB} = \boldsymbol{O}$ does not imply that either $\boldsymbol{A}$ or $\boldsymbol{B}$ is the zero matrix, but that it does imply that at least one of them is singular.

Notice here that $\boldsymbol{AB} = \boldsymbol{O}$ applies, with neither matrix itself the zero matrix. The example therefore demonstrates immediately that $\boldsymbol{AB} = \boldsymbol{O} \nrightarrow \boldsymbol{A} = \boldsymbol{O}$ or $\boldsymbol{B} = \boldsymbol{O}$. However, by property (iv) and (ix) we have that $0 = \det(\boldsymbol{AB}) = (\det \boldsymbol{A})(\det \boldsymbol{B}) \longrightarrow \det(\boldsymbol{A}) = 0$ or $\det(\boldsymbol{B}) = 0$.

Exercise 4.7. Prove the properties above.

Properties (i) and (ii) follow directly from the properties of addition and multiplication in **R**.

To prove $\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA})$ we first consider a typical (diagonal) element in this matrix. $(\boldsymbol{AB})_{ii} = \sum_{k=1}^{n} a_{ik} b_{ki}$, such that:

$$\begin{aligned} \operatorname{tr}(\boldsymbol{AB}) &= \sum_{i=1}^{n} \Big(\sum_{k=1}^{n} a_{ik} b_{ki} \Big), \\ &= \sum_{i=1}^{n} \Big(\sum_{k=1}^{n} b_{ki} a_{ik} \Big), \\ &= \sum_{k=1}^{n} \Big(\sum_{i=1}^{n} b_{ki} a_{ik} \Big), \\ &= \sum_{k=1}^{n} (\boldsymbol{BA})_{kk} \Big), \\ &= \operatorname{tr}(\boldsymbol{BA}). \end{aligned}$$

Property (iv) is proved by noting that the diagonal elements will be unchanged when taking the transpose of a matrix.

Finally (v) simply states that the scalar, c, is a 1×1 matrix which is its own trace.

Exercise 4.8. Using property (iii) of the trace [that $\text{tr}(\boldsymbol{AB}) = \text{tr}(\boldsymbol{BA})$], show that for three conformable matrices we have[2]

$$\text{tr}(\boldsymbol{ABC}) = \text{tr}(\boldsymbol{CAB}) = \text{tr}(\boldsymbol{BCA}).$$

Let $\boldsymbol{D} = \boldsymbol{AB}$ such that property (iii) may be used to show the first equality:

$$\begin{aligned} \text{tr}(\boldsymbol{ABC}) &= \text{tr}(\boldsymbol{DC}), \\ &= \text{tr}(\boldsymbol{CD}), \\ &= \text{tr}(\boldsymbol{CAB}) \end{aligned}$$

The same procedure is then used again with $\boldsymbol{E} = \boldsymbol{CA}$ to show the final equality:

$$\begin{aligned} \text{tr}(\boldsymbol{CAB}) &= \text{tr}(\boldsymbol{EB}), \\ &= \text{tr}(\boldsymbol{BE}), \\ &= \text{tr}(\boldsymbol{BCA}). \end{aligned}$$

Exercise 4.9. Referring to your computations in exercise 2.8 show that

$$\text{tr}(\boldsymbol{A'B}) = \text{tr}(\boldsymbol{BA'}) = \text{tr}(\boldsymbol{AB'}) = \text{tr}(\boldsymbol{B'A}).$$

First, recall previous answers:

$$\begin{aligned} \boldsymbol{A'B} &= \begin{bmatrix} 8 & 13 & 18 \\ 13 & 20 & 27 \\ 18 & 27 & 36 \end{bmatrix}, \\ \boldsymbol{BA'} &= \begin{bmatrix} 32 & 77 \\ 14 & 32 \end{bmatrix}, \\ \boldsymbol{AB'} &= \begin{bmatrix} 32 & 14 \\ 77 & 32 \end{bmatrix}, \\ \boldsymbol{B'A} &= \begin{bmatrix} 8 & 13 & 18 \\ 13 & 20 & 27 \\ 18 & 27 & 36 \end{bmatrix}. \end{aligned}$$

[2]Note that this property will be used in econometrics to estimate σ^2, after first noting that $\text{tr}(\varepsilon' M \varepsilon)$ is a scalar.

Such that we have:

$$\begin{aligned}
\text{tr}(\boldsymbol{A}'\boldsymbol{B}) &= 8 + 20 + 36 = 64 \\
\text{tr}(\boldsymbol{B}\boldsymbol{A}') &= 32 + 32 = 64, \\
\text{tr}(\boldsymbol{A}\boldsymbol{B}') &= 32 + 32 = 64, \\
\text{tr}(\boldsymbol{B}'\boldsymbol{A}) &= 8 + 20 + 36 = 64.
\end{aligned}$$

Exercise 4.10. Show (by multiplying $\boldsymbol{A}\boldsymbol{A}^{-1}$) that in general if $\boldsymbol{A}$ is a 2×2 matrix given by $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, and $(ad - cb) \neq 0$ then $\boldsymbol{A}^{-1} = \frac{1}{ad-cb}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, noting that $(ad - cb)$ is the determinant of $\boldsymbol{A}$.

In general we then have

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{1}{ad - cb} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad - cb} \begin{bmatrix} ad - bc & -ab + ab \\ cd - cd & ad - bc \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \boldsymbol{I}_2.$$

and, in the same way, $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}_2$. Note that this formula then tells you how to find the inverse of a 2×2 matrix.

$$\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{and} \quad ad - bc \neq 0 \rightarrow \boldsymbol{A}^{-1} = \frac{1}{ad - cb} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

which may be stated as:

(i) Switch the entries on the main diagonal;

(ii) Change the signs of the entries off the main diagonal;

(iii) Multiply the matrix by $\frac{1}{\det \boldsymbol{A}}$.

This formula should be committed to memory.

Exercise 4.11. Write down the inverse of the matrices

$$\text{(i)} \quad \boldsymbol{A} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}, \quad \text{(ii)} \quad \boldsymbol{B} = \begin{bmatrix} 5 & -3 \\ 2 & 1 \end{bmatrix}, \quad \text{(iii)} \quad \boldsymbol{C} = \begin{bmatrix} 6 & 1 \\ 1 & 2 \end{bmatrix}.$$

By inspection, applying the formulae.

$$\text{(i)} \quad \boldsymbol{A}^{-1} = -\frac{1}{2}\begin{bmatrix} 4 & -3 \\ -2 & 1 \end{bmatrix}, \quad \text{(ii)} \quad \boldsymbol{B}^{-1} = \frac{1}{11}\begin{bmatrix} 1 & 3 \\ -2 & 5 \end{bmatrix}, \quad \text{(iii)} \quad \boldsymbol{C}^{-1} = \frac{1}{11}\begin{bmatrix} 2 & -1 \\ -1 & 6 \end{bmatrix}.$$

But should also check these results (quickly).

$$\text{(i)} \quad \boldsymbol{A}\boldsymbol{A}^{-1} = -\frac{1}{2}\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 4 & -3 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} = \boldsymbol{I}_2$$

$$\text{(ii)} \quad \boldsymbol{B}\boldsymbol{B}^{-1} = \frac{1}{11}\begin{bmatrix} 5 & -3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ -2 & 5 \end{bmatrix} = \frac{1}{11}\begin{bmatrix} 11 & 0 \\ 0 & 11 \end{bmatrix} = \boldsymbol{I}_2,$$

$$\text{(iii)} \quad \boldsymbol{C}C^{-1} = \frac{1}{11}\begin{bmatrix} 6 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ -1 & 6 \end{bmatrix} = \frac{1}{11}\begin{bmatrix} 11 & 0 \\ 0 & 11 \end{bmatrix} = \boldsymbol{I}_2.$$

Exercise 4.12. Use Theorem 2 [uniqueness of an inverse] to prove Theorem 4 [unique solution for each choice of b_i].

Rewrite the system in Theorem 4 by

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

Suppose that $\boldsymbol{A}$ is invertible. Then, by Theorem 2, $\boldsymbol{A}^{-1}$ will be unique and hence the unique solution to the system will be:

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}.$$

Exercise 4.13. Construct a matrix $\boldsymbol{A}$ that multiplies the vector $(3, -1)'$ to produce the zero vector $(0, 0)'$. What do you notice about the matrix $\boldsymbol{A}$? Compute its determinant.

Let $\boldsymbol{A}\begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ such that $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ will generate the system:

$$3a - b = 0,$$
$$3c - d = 0,$$

and therefore one such matrix to solve the system will be $\boldsymbol{A} = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$ will solve the system. By Theorem 4 [unique solution for each choice of b_i], the matrix $\boldsymbol{A}$ needs to be non-singular with a zero determinant.

Exercise 4.14. Consider the matrix

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 7 \\ 3 & 6 & 2 \end{bmatrix}.$$

Show, using Theorem 6 [invertible if full rank] that the matrix $\boldsymbol{A}$ is not invertible.

Observe that $2\boldsymbol{a}_1 = \boldsymbol{a}_3$. Therefore $\boldsymbol{A}$ is rank deficient, with $\operatorname{rk} \boldsymbol{A} = 2 < 3$. The matrix $\boldsymbol{A}$ is therefore singular and not invertible.

Exercise 4.15. Suppose that $\boldsymbol{A}$ and $\boldsymbol{B}$ are non-singular matrices of the same order. Show that

$$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1},$$

and thereby proving Theorem 5 [product $\boldsymbol{AB}$ is invertible if both $\boldsymbol{A}$ and $\boldsymbol{B}$ are non-singular and of the same dimension].

Let $\boldsymbol{C} = (\boldsymbol{AB})^{-1}$ be the inverse of $\boldsymbol{AB}$ such that

$$\begin{aligned} (\boldsymbol{AB})\boldsymbol{C} &= \boldsymbol{I}, \\ \boldsymbol{A}^{-1}\boldsymbol{ABC} &= \boldsymbol{A}^{-1}, \\ \boldsymbol{BC} &= \boldsymbol{A}^{-1}, \\ \boldsymbol{B}^{-1}\boldsymbol{BC} &= \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}, \\ \boldsymbol{C} &= \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}, \\ (\boldsymbol{AB})^{-1} &= \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}. \end{aligned}$$

Theorem 6: If $\boldsymbol{A}, \boldsymbol{B}$ are non-singular $\rightarrow \boldsymbol{AB}$ is invertible follows from this result. If $(\boldsymbol{AB})^{-1}$ exists, this implies $\boldsymbol{A}$ and $\boldsymbol{B}$ are non-singular. $\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$ exists as the unique inverse for $\boldsymbol{AB}$, and therefore $\boldsymbol{A}^{-1}$ and $\boldsymbol{B}^{-1}$ need to exist.

Exercise 4.16. Prove the properties above.

For a non-singular matrix $\boldsymbol{A}$:

(i) $(c\boldsymbol{A})^{-1} = (1/c)\boldsymbol{A}^{-1}$, for any $c \in \mathbb{R}_0$;

Let $\boldsymbol{B} = (c\boldsymbol{A})^{-1}$, such that:

$$c\boldsymbol{AB} = \boldsymbol{I} \rightarrow \boldsymbol{B} = \frac{1}{c}\boldsymbol{A}^{-1},$$

(ii) $(\boldsymbol{A}')^{-1} = (\boldsymbol{A}^{-1})'$;

If $\boldsymbol{A}$ is invertible, then by definition $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$. As $\boldsymbol{I}$ is symmetric, we may take the transpose of this statement $(\boldsymbol{A}^{-1}\boldsymbol{A})' = \boldsymbol{I}$ which, by properties of the transpose, gives $\boldsymbol{A}'(\boldsymbol{A}^{-1})' = \boldsymbol{I}$. Finally we may premultiply by $(\boldsymbol{A}')^{-1}$ to give the result $(\boldsymbol{A}^{-1})' = (\boldsymbol{A}')^{-1}$.

(iii) $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$.

Initially we observe that $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$, now set $\boldsymbol{B} = (\boldsymbol{A}^{-1})^{-1}$. Notice that $\boldsymbol{B}\boldsymbol{A}^{-1} = \boldsymbol{I}$, and therefore $\boldsymbol{B} = \boldsymbol{A}$, to give the desired result.

Exercise 4.17. Find the inverse of

$$\text{(i)}\quad \boldsymbol{A} = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{(ii)}\quad \boldsymbol{B} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{(iii) (Harder)}\quad \boldsymbol{C} = \begin{bmatrix} -2 & 1.6 & -0.4 \\ 1.6 & -0.9 & 0.4 \\ 1.3 & -1.2 & 1.6 \end{bmatrix}.$$

(i) First, compute the matrix of minors, $\boldsymbol{M}$, for each element of $\boldsymbol{A}$ as

$$\boldsymbol{M} = \begin{bmatrix} -1 & 2 & 2 \\ -1 & 1 & 1 \\ 0 & -2 & -1 \end{bmatrix}$$

Next, use cofactors, and transpose, to convert this matrix to become the adjoint matrix, $\boldsymbol{A}^{\#}$.

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

Then, calculate the determinant of the matrix (cofactor expansion using minor along first column).

$$\det \boldsymbol{A} = 1 \times -1 + 2 \times 1 = 1$$

Finally, combine the adjoint matrix and the determinant to give the inverse.

$$\boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}} \boldsymbol{A}^{\#} = \begin{bmatrix} -1 & 1 & 0 \\ -2 & 1 & 2 \\ 2 & -1 & -1 \end{bmatrix},$$

We should also check this matrix.

$$\boldsymbol{A}^{-1}\boldsymbol{A} = \begin{bmatrix} -1 & 1 & 0 \\ -2 & 1 & 2 \\ 2 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix} = I_3.$$

(ii) First, compute the matrix of minors, $\boldsymbol{M}$, for each element of $\boldsymbol{B}$ as

$$\boldsymbol{M} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 4 \end{bmatrix}$$

Next, use cofactors, and transpose, to convert this matrix to become the adjoint matrix, $\boldsymbol{B}^{\#}$.

$$\boldsymbol{B}^{\#} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & -2 & 4 \end{bmatrix}$$

Then, calculate the determinant of the matrix (cofactor expansion using minor along first column).

$$\det \boldsymbol{B} = 2 \times 1 = 2$$

Finally, combine the adjoint matrix and the determinant to give the inverse.

$$\boldsymbol{B}^{-1} = \frac{1}{\det \boldsymbol{B}} \boldsymbol{B}^{\#} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & -2 & 4 \end{bmatrix},$$

We should also check this matrix.

$$\boldsymbol{B}^{-1}\boldsymbol{B} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & -2 & 4 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = I_3.$$

(iii) First, compute the matrix of minors, $\boldsymbol{M}$, for each element of $\boldsymbol{C}$ as

$$\boldsymbol{M} = \begin{bmatrix} -0.96 & 2.04 & -0.75 \\ 2.08 & -2.68 & 0.32 \\ 0.28 & -0.16 & -0.76 \end{bmatrix}$$

Next, use cofactors, and transpose, to convert this matrix to become the adjoint matrix, $\boldsymbol{C}^{\#}$.

$$\boldsymbol{A}^{\#} = \begin{bmatrix} -0.96 & -2.08 & 0.28 \\ -2.04 & -2.68 & 0.16 \\ -0.75 & -0.32 & -0.76 \end{bmatrix}$$

Then, calculate the determinant of the matrix.

$$\det \boldsymbol{C} = -2 \times -0.96 - 1.6 \times 2.04 - 0.4 \times -0.75 = -1.044 = -\frac{261}{250}$$

Finally, combine the adjoint matrix and the determinant to give the inverse.

$$C^{-1} = \frac{1}{\det C} C^{\#} = -\frac{250}{261} \begin{bmatrix} -0.96 & -2.08 & 0.28 \\ -2.04 & -2.68 & 0.16 \\ -0.75 & -0.32 & -0.76 \end{bmatrix} = \begin{bmatrix} 0.9195 & 1.9923 & -0.2682 \\ 1.954 & 2.567 & -0.1533 \\ 0.7184 & 0.3065 & 0.728 \end{bmatrix} = \begin{bmatrix} \frac{80}{87} & \frac{520}{261} & -\frac{70}{261} \\ \frac{170}{87} & \frac{670}{261} & -\frac{40}{261} \\ \frac{125}{174} & \frac{80}{261} & \frac{190}{261} \end{bmatrix}.$$

As always, this matrix should be checked.

Exercise 4.18. Show that a matrix with a column (or row) of zeros is not invertible.

This follows from the determinant being zero.

Exercise 4.19. Suppose $\boldsymbol{A}$ is invertible and you exchange its first two rows to obtain $\boldsymbol{B}$. Explain why the new matrix is invertible. How would you find $\boldsymbol{B}^{-1}$ from $\boldsymbol{A}^{-1}$.

Follows from exchange of rows.

Exercise 4.20. Find the two eigenvalues of the matrix

$$\boldsymbol{B} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

Deduce two eigenvectors of $\boldsymbol{B}$ by inspection.

$$|\boldsymbol{B} - \lambda \boldsymbol{I}| = \begin{vmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 - 1 = \lambda^2 - 6\lambda + \lambda^2 + 9 - 1 = (\lambda - 2)(\lambda - 4)$$

The two eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = 4$. The eigenvectors we will deduce by inspection. We are looking for the condition $\boldsymbol{B}\boldsymbol{x} = \lambda \boldsymbol{x}$ to hold in each case. The easiest way will be to normalise one element of the vector $\boldsymbol{x}$ for each case.

In case 1, we have $\lambda_1 = 2$ and therefore guess $\boldsymbol{x} = \begin{bmatrix} x_1 \\ 1 \end{bmatrix}$ as a candidate eigenvector. From the relationship we know that this must therefore satisfy

$$\begin{aligned} 3x_1 + 1 &= 2x_1, \\ x_1 + 3 &= 2, \end{aligned}$$

hence $x_1 = -1$ and we have that

$$\boldsymbol{x} = \alpha \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

where $\alpha \in \mathbf{R}$.

In case 2, we have $\lambda_2 = 4$ and proceed similarly to generate

$$\boldsymbol{x} = \beta \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

where $\beta \in \mathbf{R}$.

Exercise 4.21. Let $\boldsymbol{A}$ be the permutation matrix in example 4.4, and write the matrix $\boldsymbol{B}$ of exercise 4.20 as $\boldsymbol{B} = \boldsymbol{A} + 3\boldsymbol{I}$. Algebraically, deduce that the eigenvalues of $\boldsymbol{A}$ are three less than the eigenvalues of $\boldsymbol{B}$ and that the eigenvectors are unchanged.

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

Rewrite as $\boldsymbol{A} = \boldsymbol{B} - 3\boldsymbol{I}$, and let μ be an eigenvalue of $\boldsymbol{A}$ and λ be an eigenvalue of $\boldsymbol{B}$. We have that

$$\begin{aligned} |\boldsymbol{A} - \mu\boldsymbol{I}| &= |\boldsymbol{B} - (\mu + 3)\boldsymbol{I}|, \\ &= |\boldsymbol{B} - \lambda\boldsymbol{I}| \end{aligned}$$

thus the eigenvalues of $\boldsymbol{A}$ and $\boldsymbol{B}$ are related by the relationship $\mu + 3 = \lambda$ and hence $\lambda = \mu - 3$, as required.

We therefore have that the eigenvalues are $\mu_1 = 2 - 3 = -1$ and $\mu_2 = 4 - 3 = 1$.

We now consider the eigenvectors to maintain the relationship $\boldsymbol{Ax} = \mu\boldsymbol{x}$.

In case 1, with $\mu_1 = -1$, we observe the relationship $x_1 = -x_2$, such that the eigenvector must be

$$\boldsymbol{x} = \alpha \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

where $\alpha \in \mathbf{R}$.

In case 2, with $\mu_2 = 1$, we observe the relationship $x_1 = x_2$, such that the eigenvector must be

$$\boldsymbol{x} = \beta \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

where $\beta \in \mathbf{R}$.
The eigenvectors therefore remain unchanged from the previous example.

Exercise 4.22. Show that:

(i) λ^2 is an eigenvalue of $\boldsymbol{A}^2$;

$$\begin{aligned} \boldsymbol{Ax} &= \lambda\boldsymbol{x}, \\ \boldsymbol{AAx} &= \boldsymbol{A}\lambda\boldsymbol{x}, \\ \boldsymbol{A}^2\boldsymbol{x} &= \lambda\boldsymbol{Ax} = \lambda\lambda\boldsymbol{x} = \lambda^2\boldsymbol{x}. \end{aligned}$$

(ii) λ^{-1} is an eigenvalue of $\boldsymbol{A}^{-1}$;

Follows analogously

$$\begin{aligned}
\boldsymbol{Ax} &= \lambda\boldsymbol{x}, \\
\boldsymbol{A}^{-1}\boldsymbol{Ax} &= \boldsymbol{A}^{-1}\lambda\boldsymbol{x}, \\
\boldsymbol{Ix} &= \lambda\boldsymbol{A}^{-1}\boldsymbol{x}, \\
\lambda^{-1}\boldsymbol{x} &= \boldsymbol{A}^{-1}\boldsymbol{x}.
\end{aligned}$$

(iii) $\lambda + 1$ is an eigenvalue of $\boldsymbol{A} + \boldsymbol{I}$.

$$\begin{aligned}
(\boldsymbol{A} + \boldsymbol{I})\boldsymbol{x} &= \mu\boldsymbol{x}, \\
\boldsymbol{Ax} + \boldsymbol{x} &= \mu\boldsymbol{x}, \\
\lambda\boldsymbol{x} + \boldsymbol{x} &= \mu\boldsymbol{x}, \\
(\lambda + 1)\boldsymbol{x} &= \mu\boldsymbol{x}, \\
\mu &= \lambda + 1.
\end{aligned}$$

Exercise 4.23. Show that the eigenvalues of idempotent matrices are either 0 or 1.

The eigenvalues and eigenvectors are defined obey the equation $\boldsymbol{Ax} = \lambda\boldsymbol{x}$. However as $\boldsymbol{A}$ is idempotent, we have that $\boldsymbol{A}^2 = \boldsymbol{A}$, such that $\boldsymbol{A}^2\boldsymbol{x} = \lambda\boldsymbol{x}$. From exercise 4.19 we know that λ^2 is an eigenvalue of $\boldsymbol{A}^2$ such that $\boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{Ax} = \lambda\boldsymbol{x} = \lambda^2\boldsymbol{x}$. Thus $\lambda(\lambda - 1)\boldsymbol{x} = 0$ and therefore $\lambda = 0$ or $\lambda = 1$.

Exercise 4.24. Let

$$\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

(i) Show that $\det(\boldsymbol{A} - \lambda\boldsymbol{I}) = \lambda^2 - \text{tr}(\boldsymbol{A})\lambda + \det \boldsymbol{A}$;

$$\begin{aligned}
|\boldsymbol{A} - \lambda\boldsymbol{I}| = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} &= (a - \lambda)(d - \lambda) - bc, \\
&= \lambda^2 - (a + d)\lambda + ad - bc, \\
&= \lambda^2 - (a + d)\lambda + \det(\boldsymbol{A}), \\
&= \lambda^2 - \text{tr}(\boldsymbol{A})\lambda + \det(\boldsymbol{A}).
\end{aligned}$$

(ii) Find an expression for the two eigenvalues of $\boldsymbol{A}$ in terms of a, b, c and d. These characteristic roots satisfy

$$(\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - \operatorname{tr}(\boldsymbol{A})\lambda + \det \boldsymbol{A} = 0,$$

and $\lambda_1\lambda_2 = \det \boldsymbol{A}$ and $\lambda_1 + \lambda_2 = \operatorname{tr} \boldsymbol{A}$.

Clearly here the discriminant of the relationship is given as $\Delta = \operatorname{tr}^2(\boldsymbol{A}) - 4\det(\boldsymbol{A})$, such that

$$\lambda_{1,2} = \frac{\operatorname{tr}(A) \pm \sqrt{\Delta}}{2} = \frac{a + d \pm \sqrt{(a+d)^2 - 4(ad - bc)}}{2}.$$

5 Diagonalisation and powers of A

Exercise 5.1. Prove Lemma 1 [Diagonalisable determinants and trace].

Let $\boldsymbol{A}$ be diagonalisable such that $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}$. We therefore have

$$\begin{aligned}
\det \boldsymbol{A} &= \det(\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}),\\
&= \det(\boldsymbol{Q})\det(\boldsymbol{\Lambda})\det(\boldsymbol{Q}^{-1}),\\
&= \det(\boldsymbol{\Lambda})\det(\boldsymbol{Q}\boldsymbol{Q}^{-1}),\\
&= \det(\boldsymbol{\Lambda}),\\
&= \prod_{i=1}^{n} \lambda_i,
\end{aligned}$$

where the final step uses the fact that $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues.

Secondly,

$$\begin{aligned}
\operatorname{tr} \boldsymbol{A} &= \operatorname{tr}(\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}),\\
&= \operatorname{tr}(\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}\boldsymbol{Q}),\\
&= \operatorname{tr}(\boldsymbol{\Lambda}),\\
&= \sum_{i=1}^{n} \lambda_i
\end{aligned}$$

where, again, the final step uses the fact that $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues.

Exercise 5.2. Assuming you can write the initialisation vector $\boldsymbol{u}_0$ as a linear combination of the eigenvectors, find an expression for the kth element of an arbitrary sequence of the form $\boldsymbol{u}_{k+1} = \boldsymbol{A}\boldsymbol{u}_k$.

Assume $\mu_0 = \sum_{i=1}^{n} c_i q_i = \boldsymbol{Q}\boldsymbol{c}$, with $\boldsymbol{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_2 \end{bmatrix}$ and then we have that

$$\begin{aligned}
\mu_{k+1} &= \boldsymbol{A}^{k+1}\mu_0,\\
&= \boldsymbol{Q}\boldsymbol{\Lambda}^{k+1}\boldsymbol{Q}^{-1}\mu_0,\\
&= \boldsymbol{Q}\boldsymbol{\Lambda}^{k+1}\boldsymbol{Q}^{-1}\boldsymbol{Q}\boldsymbol{c},\\
&= \boldsymbol{Q}\boldsymbol{\Lambda}^{k+1}\boldsymbol{c},\\
&= \begin{bmatrix} q_1 & \dots & q_n \end{bmatrix} \begin{bmatrix} \lambda^{k+1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n^{k+1} \end{bmatrix} \boldsymbol{c},\\
&= \begin{bmatrix} q_1\lambda_1^{k+1} & \dots & q_n\lambda_n^{k+1} \end{bmatrix} \boldsymbol{c},\\
&= \sum_{i=1}^{n} \lambda_i^{k+1} c_i q_i
\end{aligned}$$

Exercise 5.3. Using the expression you found in exercise 5.2, find the 100th element of the Fibonacci sequence in example 5.1.

This exercise is provided for the benefit of the most able students only.

Exercise 5.4. Show the properties above.

For a symmetric positive definite matrix, $\boldsymbol{A}$, we have that:

(i) $\det \boldsymbol{A} > 0$ such that it is also non-singular;

This follows from $\det \boldsymbol{A} = \prod_{i=1}^{n} \lambda_i$, and $\lambda_i > 0$ if $\boldsymbol{A}$ is positive definite. The converse is not true.

(ii) $\operatorname{tr} \boldsymbol{A} > 0$;

Similar to above, this follows from $\operatorname{tr} \boldsymbol{A} = \sum_{i=1}^{n} \lambda_i$, and $\lambda_i > 0$ if $\boldsymbol{A}$ is positive definite. The converse is not true.

(iii) if $\boldsymbol{A} \geq \boldsymbol{O}$ and $\boldsymbol{B} \geq \boldsymbol{O}$, then $\boldsymbol{A} + \boldsymbol{B}$ is at least positive semi-definite.

For any $\boldsymbol{x} \in \mathbf{R}^n$ we have that both

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \geq 0,$$

and

$$\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x} \geq 0,$$

which, by simple addition, implies

$$\boldsymbol{x}'(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{x} \geq 0.$$

(iv) $\boldsymbol{A} > \boldsymbol{O}$ if and only if $\boldsymbol{A}^{-1} > \boldsymbol{O}$.

If $\boldsymbol{A}$ is a symmetric positive semi-definite matrix, then by Lemma 2 we have that $\lambda_i > 0$. We have shown earlier than λ_i^{-1} is an eigenvalue of $\boldsymbol{A}^{-1}$, thus $\lambda_i^{-1} > 0$ hence again by Lemma 2 we have that $\boldsymbol{A}^{-1}$ is positive definite. Full proof requires a repetition for the opposite direction.

Exercise 5.5. Show that there exists a matrix which is both positive semi-definite and negative semi-definite, but that a matrix cannot be positive and negative definite at the same time.

Find a matrix such that for all $\boldsymbol{x} \in \mathbf{R}^n$ both $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \leq 0$ and $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \geq 0$, this simply requires $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = 0 \leftrightarrow \boldsymbol{A} = \mathbf{O}$ will work.

It is not possible to reconcile both $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} < 0$ and $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} > 0$ simultaneously.

Exercise 5.6. Let $\boldsymbol{\Omega}$ be a symmetric positive definite matrix. Use the spectral decomposition to show that there exists a non-singular matrix $\boldsymbol{L}$ such that $\boldsymbol{\Omega} = \boldsymbol{L}'\boldsymbol{L}$ and such that $\boldsymbol{L}^{-1} = (\boldsymbol{L}')^{-1}$. What is $\boldsymbol{L}^{-1}$?

As $\boldsymbol{\Omega}$ is symmetric it will have spectral decomposition with $\boldsymbol{\Omega} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}'$.

It is also positive definite, such that $\boldsymbol{\Lambda}^{\frac{1}{2}} = \text{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$ and $\boldsymbol{\Lambda}^{-\frac{1}{2}} = \text{diag}(\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_n}})$ exist

Notice that $\boldsymbol{L} = \boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{Q}'$ would work to generate

$$\begin{aligned}\boldsymbol{\Omega} = \boldsymbol{L}'\boldsymbol{L} &= (\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{Q}')'\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{Q}'.\\ &= \boldsymbol{Q}'\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{Q}'.\end{aligned}$$

However, this will not obey the requirement $\boldsymbol{L}^{-1} = (\boldsymbol{L}')^{-1}$ as $\boldsymbol{Q}\boldsymbol{\Lambda}^{-\frac{1}{2}} \neq \boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{Q}'$.

Instead we require $\boldsymbol{L} = \boldsymbol{Q}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{Q}' = \boldsymbol{L}'$. In this case

$$\begin{aligned}\boldsymbol{\Omega} = \boldsymbol{L}'\boldsymbol{L} &= (\boldsymbol{Q}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{Q}')'\boldsymbol{Q}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{Q}'.\\ &= \boldsymbol{Q}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{Q}'\boldsymbol{Q}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{Q}'.\\ &= \boldsymbol{Q}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{Q}'.\\ &= \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}'.\end{aligned}$$

Therefore $\boldsymbol{L}^{-1} = \boldsymbol{Q}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{Q}' = (\boldsymbol{L}')^{-1}$, as required.

Exercise 5.7. Find the $\boldsymbol{X}$ that solves

$$2\boldsymbol{X}^2 - 3\boldsymbol{X} + \boldsymbol{I}_n = \boldsymbol{O}_n,$$

where $\boldsymbol{X}$ is a real symmetric matrix.

As $\boldsymbol{X}$ is symmetric it has the spectral decomposition with $\boldsymbol{X} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}'$, which may be plugged into the equation with $\boldsymbol{Q}\boldsymbol{Q}' = \boldsymbol{I}_n$ such that

$$\begin{aligned}2\boldsymbol{Q}\boldsymbol{\Lambda}^2\boldsymbol{Q}' - 3\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}' + \boldsymbol{I}_n &= \boldsymbol{O}_n,\\ 2\boldsymbol{Q}\boldsymbol{\Lambda}^2\boldsymbol{Q}' - 3\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}' + \boldsymbol{Q}\boldsymbol{Q}' &= \boldsymbol{O}_n,\\ \boldsymbol{Q}(2\boldsymbol{\Lambda}^2 - 3\boldsymbol{\Lambda} + \boldsymbol{I}_n)\boldsymbol{Q}' &= \boldsymbol{O}_n\end{aligned}$$

and therefore the problem reduces to solving $2\boldsymbol{\Lambda}^2 - 3\boldsymbol{\Lambda} + \boldsymbol{I}_n = \boldsymbol{O}_n$, which further reduces to become $2\lambda_i^2 - 3\lambda_i + 1 = 0 \; \forall_{i=1\ldots n}$. This may be factorised to become $(2\lambda_i - 1)(\lambda_i - 1) = 0$ such that $\lambda_i = \frac{1}{2}$ or $\lambda_i = 1$.

The set of solutions is therefore $\boldsymbol{V} = \{\boldsymbol{X} \in S^n : \lambda_i \in \{\frac{1}{2}, 1\}\}$ where S^n are the set of square symmetric $n \times n$ with elements $\in \mathbf{R}^n$.

6 Vector and matrix calculus

Exercise 6.1. Assume $\varphi(\boldsymbol{x}) = \varphi(x_1, x_2) = 2x_1^2 x_2^3$, find the Jacobian and Hessian matrix of this (scalar) function (which has the vector input $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$).

$$\frac{d\varphi(\boldsymbol{x})}{d\boldsymbol{x}'} = \begin{bmatrix} \frac{\partial \varphi(\boldsymbol{x})}{\partial x_1}, & \frac{\partial \varphi(\boldsymbol{x})}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 4x_1 x_2^3, & 6x_1^2 x_2^2 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{H}(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^2 \varphi(\boldsymbol{x})}{\partial x_1^2} & \frac{\partial^2 \varphi(\boldsymbol{x})}{\partial x_1 \partial x_2} \\ \frac{\partial^2 \varphi(\boldsymbol{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 \varphi(\boldsymbol{x})}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 4x_2^3 & 12x_1 x_2^2 \\ 12x_1 x_2^2 & 12x_1^2 x_2 \end{bmatrix}$$

Exercise 6.2. Assuming an input an n-vector $\boldsymbol{x}$, find both the Jacobian and Hessian matrix of the scalar function $\varphi(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{x}$.

Firstly we may rewrite the function, accounting for the definition of the inner product as

$$\varphi(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{x} = x_1^2 + \cdots + x_n^2.$$

Next note the derivative, second derivative and cross-derivative of each element as

$$\frac{\partial \varphi(\boldsymbol{x})}{\partial x_j} = 2x_j, \quad \frac{\partial^2 \varphi(\boldsymbol{x})}{\partial x_j^2} = 2, \quad \text{and} \quad \frac{\partial^2 \varphi(\boldsymbol{x})}{\partial x_j x_k} = 0 \quad \text{for} \quad j \neq k.$$

Finally combine into a Jacobian vector and Hessian matrix.

$$\frac{d\varphi(\boldsymbol{x})}{d\boldsymbol{x}'} = \begin{bmatrix} \frac{\partial \varphi(\boldsymbol{x})}{\partial x_1} & \dots & \frac{\partial \varphi(\boldsymbol{x})}{\partial x_n} \end{bmatrix} = \begin{bmatrix} 2x_1 & \dots & 2x_n \end{bmatrix} = 2\boldsymbol{x}', \quad \text{and} \quad \boldsymbol{H}(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^2 \varphi(\boldsymbol{x})}{\partial x_1^2} & \dots & \frac{\partial^2 \varphi(\boldsymbol{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \varphi(\boldsymbol{x})}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 \varphi(\boldsymbol{x})}{\partial x_n^2} \end{bmatrix} = 2\boldsymbol{I}_n.$$

Exercise 6.3. Assume $\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} x + \cos(y) \\ y + \cos(x) \end{bmatrix}$ and find the Jacobian of this function.

Firstly, note that the input vector may be written as, $\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix}$.

$$\frac{d\boldsymbol{f}(\boldsymbol{x})}{d\boldsymbol{x}'} = \begin{bmatrix} \frac{\partial f_1(\boldsymbol{x})}{\partial x} & \frac{\partial f_1(\boldsymbol{x})}{\partial y} \\ \frac{\partial f_2(\boldsymbol{x})}{\partial x} & \frac{\partial f_2(\boldsymbol{x})}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & -\sin(y) \\ -\sin(x) & 1 \end{bmatrix}.$$

Exercise 6.4. Consider the scalar function $\varphi(x, y) = \sin(x)\sin(y)$. Find all its stationary points and determine whether these are local minima, maxima or saddle points by computing the Hessian matrix.

Hint: what does the definiteness of the Hessian matrix tell you about the concavity or convexity of φ?

This may be done without vector or matrix calculus. Try after the optimisation module.

Exercise 6.5. In statistics, the problem of ridge regression is formulated as follows

$$\hat{\boldsymbol{\beta}}_R(\lambda) := \arg\min_{\beta \in \mathbf{R}^k} (\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta}) + \lambda \|\boldsymbol{\beta}\|^2,$$

where $\lambda \in \mathbf{R}$ is a parameter that needs to be chosen beforehand. Notice, that this is a least-squares problem but where $\boldsymbol{\beta}$'s which are "too large" are penalised. By computing the first-order conditions, find $\hat{\boldsymbol{\beta}}_R(\lambda)$.

Generate the loss function

$$L(\boldsymbol{\beta}) = (\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta}) + \lambda \boldsymbol{\beta}'\boldsymbol{\beta},$$

which has a first order condition of

$$dL(\boldsymbol{\beta}) = [-2(\boldsymbol{y} - \boldsymbol{X\beta})'X + 2\lambda\boldsymbol{\beta}']d(\boldsymbol{\beta})$$

Hence we set this expression equal to zero and solve for $\boldsymbol{\beta}$.

$$\begin{aligned}
-2(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'X + 2\lambda\hat{\boldsymbol{\beta}}' &= \boldsymbol{O}', \\
-(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'X + \lambda\hat{\boldsymbol{\beta}}' &= \boldsymbol{O}', \\
-\boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) + \lambda\hat{\boldsymbol{\beta}} &= \boldsymbol{O}, \\
-\boldsymbol{X}'\boldsymbol{y} + \boldsymbol{X}'X\hat{\boldsymbol{\beta}} + \lambda\hat{\boldsymbol{\beta}} &= \boldsymbol{O}, \\
-\boldsymbol{X}'\boldsymbol{y} + (\lambda + \boldsymbol{X}'X)\hat{\boldsymbol{\beta}} &= \boldsymbol{O}, \\
\hat{\boldsymbol{\beta}} &= (\lambda + \boldsymbol{X}'X)^{-1}\boldsymbol{X}'\boldsymbol{y},
\end{aligned}$$