# Chapter 5 - Solutions

Matrices : Definitions, special matrices, operations and properties

Collège Sainte-Anne | Fall 2026

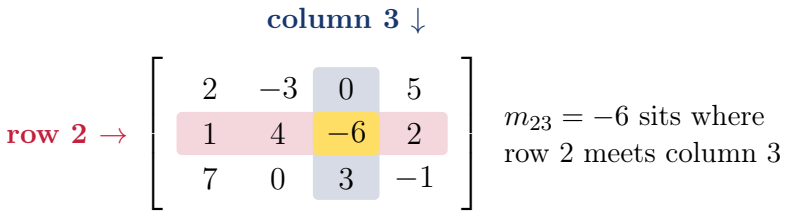
## 1 Reading and building matrices

**What this section trains**

Reading a matrix correctly: size $m \times n$ (rows first), the position $(i, j)$ of an entry, and building a matrix from a rule $a_{ij} = f(i, j)$.

### Solution 1.1 COMPUTE ⋆

**Method.** Count rows first, then columns. To find $m_{ij}$, go to row $i$, then move along it to column $j$.




(a) $M$ has 3 rows and 4 columns, so it is a $3 \times 4$ matrix with $3 \cdot 4 = 12$ entries.

(b) $m_{14} = 5$ (row 1, column 4), $m_{31} = 7$ (row 3, column 1), $m_{23} = -6$ (see the figure). $m_{42}$ would be in row 4, but $M$ has only 3 rows: $m_{42}$ **does not exist**.

(c) Row 2: $1 + 4 + (-6) + 2 = 1$. Column 3: $0 + (-6) + 3 = -3$.

(d) $m_{11} + m_{22} + m_{33} = 2 + 4 + 3 = 9$. These are the entries with equal indices, the "diagonal" of $M$. Since $M$ is not square, this is not a trace; the trace is defined only for square matrices (Section 8).

**Common mistake**

Reading $m_{31}$ as "column 3, row 1" gives 0 instead of 7. The first index is *always* the row.

### Solution 1.2 COMPUTE ⋆

**Method.** Fill the matrix entry by entry: for each position $(i, j)$, substitute $i$ and $j$ into the formula.

(a) $a_{ij} = (-1)^{i+j}$ equals $+1$ when $i + j$ is even and $-1$ when $i + j$ is odd:

$$a_{11} = (-1)^2 = 1, \quad a_{12} = (-1)^3 = -1, \ \ldots \quad \Longrightarrow \quad A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}.$$

| | | |
|---|---|---|
| $+1$ | $-1$ | $+1$ |
| $-1$ | $+1$ | $-1$ |
| $+1$ | $-1$ | $+1$ |

a checkerboard of signs: the sign flips each time you move one step in any direction

(b) $b_{ij} = 2i - j$, with $i \in \{1, 2\}$ and $j \in \{1, 2, 3, 4\}$. Row 1: $2-1$, $2-2$, $2-3$, $2-4$. Row 2: $4-1$, $4-2$, $4-3$, $4-4$.

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

(c) $c_{ij} = \min(i, j)$ is the smaller of the two indices:

$$C = \begin{bmatrix} \min(1,1) & \min(1,2) & \min(1,3) \\ \min(2,1) & \min(2,2) & \min(2,3) \\ \min(3,1) & \min(3,2) & \min(3,3) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

**Check.** Count the entries: $A$ and $C$ have 9, $B$ has $2 \cdot 4 = 8$. Also, $\min(i, j) = \min(j, i)$, so $C$ must be the same on both sides of the diagonal, and it is.

### Solution 1.3 THINK ⋆⋆ *Going backwards*

**Method.** Look for a pattern along a row (how the entries change when $j$ increases) and along a column (when $i$ increases).

(a) Row $i$ is $i$, $2i$, $3i$: row $i$ is $i$ times $(1, 2, 3)$. So $a_{ij} = i \cdot j$. (This is the multiplication table.)

(b) The diagonal is 0, the entries next to it are 1, and the corner entries are 2. Each entry is the *distance to the diagonal*: $a_{ij} = |i - j|$.

(c) Each step to the right adds 1, and each step down adds 1. Since $a_{11} = 2 = 1 + 1$, we get $a_{ij} = i + j$.

**Check.** Always test your formula on one or two entries you did not use to guess it, for example $a_{23}$: (a) $2 \cdot 3 = 6$ ✓, (b) $|2 - 3| = 1$ ✓, (c) $2 + 3 = 5$ ✓.

### Solution 1.4 THINK ⋆⋆

**Method.** A matrix of size $m \times n$ has $mn$ entries, so we are looking for all the ways to write the number of entries as a product $m \times n$, keeping the order.

(a) $12 = 1 \cdot 12 = 2 \cdot 6 = 3 \cdot 4 = 4 \cdot 3 = 6 \cdot 2 = 12 \cdot 1$, so the sizes are $1 \times 12$, $2 \times 6$, $3 \times 4$, $4 \times 3$, $6 \times 2$, $12 \times 1$. A square matrix would need $n \times n = 12$, i.e. $n^2 = 12$, but 12 is not a perfect square: **no**.

(b) 16: $1 \times 16$, $2 \times 8$, $4 \times 4$, $8 \times 2$, $16 \times 1$. Exactly one of them is square: $4 \times 4$, because $16 = 4^2$.

**Takeaway.** $3 \times 4$ and $4 \times 3$ are *different* sizes: they have the same number of entries but not the same shape.

### Solution 1.5 COMPUTE ⋆ *A school supply store*

**Method.** Give each index its meaning: $i$ = store (1 = Laval, 2 = Montréal) and $j$ = item (1 = notebooks, 2 = pens, 3 = calculators).

(a) $p_{23}$ is in row 2 (Montréal), column 3 (calculators): Montréal has $p_{23} = 22$ calculators. Pens in Laval are in row 1, column 2: $p_{12} = 300$.

(b) The pens are the whole column 2: $p_{12} + p_{22} = 300 + 450 = 750$ pens.

**Takeaway.** In a table of data, a **row** gathers everything about one store, and a **column** gathers everything about one item. This interpretation is what makes the product $PC$ meaningful in Exercise 5.4.

# 2 Special matrices and equality

**What this section trains**

Checking definitions *exactly*: which entries must be 0, and the fact that most of these words only apply to square matrices. For equality, *same size* comes before *same entries*.

### Solution 2.1 COMPUTE ⋆

**Method.** First ask "is it square?". If it is, look at where the zeros are: below the diagonal (upper triangular), above it (lower triangular), or both (diagonal). Then check whether the diagonal entries are all equal (scalar) and all equal to 1 (identity).

| | square | zero | identity | diagonal | scalar | upper tri. | lower tri. |
|---|---|---|---|---|---|---|---|
| $A$ | ✓ | | | ✓ | ✓ | ✓ | ✓ |
| $B$ | ✓ | | | ✓ | | ✓ | ✓ |
| $C$ | ✓ | | | | | ✓ | |
| $D$ | ✓ | | | | | | ✓ |
| $E$ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ |
| $F$ | | | | | | | |
| $G$ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ |

- $A = 4I_2$: a scalar matrix. Every scalar matrix is diagonal, and every diagonal matrix is both upper and lower triangular.
- $B$ is diagonal, but *not* scalar: its diagonal entries $1, 0, -2$ are not all equal. A 0 on the diagonal is allowed.
- $C$ has only zeros below the diagonal: upper triangular. Zeros *on* the diagonal are allowed too.
- $D$ has $d_{12} = 0$ above the diagonal: lower triangular. It is not diagonal because $d_{21} = 3 \neq 0$.
- $E = O_{2\times 2} = 0 \cdot I_2$ satisfies every "zero" condition, so it is diagonal, scalar and triangular.
- $F$ is the trap: it *looks* like an identity matrix, but it is $2 \times 3$. It is not the zero matrix, and every other word in the list requires a square matrix, so no word applies.

- $G = [1] = I_1$: the identity of order 1. It is also scalar, diagonal and triangular, since there is nothing off the diagonal.

### Solution 2.2 COMPUTE ⋆

**Method.** The two matrices have the same size $(2 \times 2)$, so they are equal if and only if their four entries are equal one by one. This gives a system of equations.

Comparing entries:

$$\underbrace{x + 2y = 7}_{(1,1)}, \qquad \underbrace{3 = z^2 - 1}_{(1,2)}, \qquad \underbrace{-1 = -1}_{(2,1)\text{: always true}}, \qquad \underbrace{x - y = 1}_{(2,2)}.$$

Subtracting the last equation from the first: $(x + 2y) - (x - y) = 7 - 1$, so $3y = 6$ and $y = 2$; then $x = 1 + y = 3$. From $z^2 - 1 = 3$: $z^2 = 4$, so $z = 2$ or $z = -2$.
**Check.** With $x = 3$, $y = 2$, $z = \pm 2$: $x + 2y = 7$ ✓, $x - y = 1$ ✓, $z^2 - 1 = 3$ ✓. Both values of $z$ work, so there are **two** solutions.

### Solution 2.3 THINK ⋆⋆ *When equality is impossible, or too easy*

**Method.** Equal matrices means same size **and** equal entries. Write all the conditions, then check whether they are compatible.

(a) The sizes match. We need $x = 3$ (entry $(1, 1)$) **and** $x^2 = 4$ (entry $(1, 2)$). But $x = 3$ gives $x^2 = 9 \neq 4$. The conditions contradict each other: **no such** $x$.

(b) No, whatever $a, b, c$ are: the left matrix is $1 \times 3$ and the right one is $3 \times 1$. Different sizes, so they are never equal, even though both contain three numbers.

(c) Comparing entries: $a = b$, $b = c$, $c = d$, $d = a$. So $a = b = c = d$, with no other condition. The solutions are

$$\begin{bmatrix} t & t \\ t & t \end{bmatrix} = t \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad t \in \mathbb{R} :$$

infinitely many matrices.

**Takeaway.** A matrix equation is a *system* of equations, one per entry. Like any system, it can have no solution (a), or infinitely many (c). If the sizes differ, it has no solution at all (b).

### Solution 2.4 TRUE / FALSE ⋆⋆

(a) **True.** A $1 \times 1$ matrix $[a]$ has no entry off the diagonal, so the condition "$a_{ij} = 0$ for $i \neq j$" is satisfied (there is nothing to check).

(b) **True.** $O_{n\times n} = 0 \cdot I_n$: it is $kI_n$ with $k = 0$.

(c) **False.** Lower triangular only forces zeros *above* the diagonal. Entries below it can be anything: $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ is lower triangular with a zero diagonal, but it is not $O$.

(d) **False.** $a_{21} = 2 \cdot 1 = 2 \neq 0$, while upper triangular requires $a_{21} = 0$. In fact $[ij] = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}$ (Exercise 1.3(a)) has no zero at all.

(e) **True.** Upper triangular gives $a_{ij} = 0$ for $i > j$, and lower triangular gives $a_{ij} = 0$ for $i < j$. Together, $a_{ij} = 0$ whenever $i \neq j$: that is the definition of diagonal.

(f) **False.** The identity $I_n$ is *square* by definition. There is no "identity of size $2 \times 3$".

### Solution 2.5 THINK ⋆⋆⋆ *Counting*

**Method.** Count the free entries. Each free entry has 2 choices (0 or 1), so $k$ free entries give $2^k$ matrices.

(a) 4 free entries: $2^4 = 16$ matrices.

(b) **Upper triangular:** $a_{21} = 0$ is forced; $a_{11}, a_{12}, a_{22}$ are free, so there are $2^3 = 8$:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

**Diagonal:** only $a_{11}$ and $a_{22}$ are free, so $2^2 = 4$ (the first four above). **Scalar:** we also need $a_{11} = a_{22}$, so only $O$ and $I_2$: **2**.

(c) By Exercise 2.4(e), "upper and lower triangular" means "diagonal": 4, the same number as the diagonal matrices in (b).

**Check.** The chain of inclusions scalar $\subset$ diagonal $\subset$ upper triangular $\subset$ all must give increasing counts, and it does: $2 \leq 4 \leq 8 \leq 16$.

## 3 Addition and multiplication by a scalar

**What this section trains**

Computing entry by entry, checking sizes first, and solving matrix equations *like ordinary equations*: isolate the unknown first and compute only at the end.

Recall: $A = \begin{bmatrix} 1 & -2 & 0 \\ 3 & 1 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 0 & -3 \\ -1 & 4 & 1 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$.

### Solution 3.1 COMPUTE ⋆

**Method.** Check the sizes. $A$ and $B$ are both $2 \times 3$, so any combination of them is defined. $C$ is $2 \times 2$.

(a) $A + B = \begin{bmatrix} 1+2 & -2+0 & 0-3 \\ 3-1 & 1+4 & -1+1 \end{bmatrix} = \begin{bmatrix} 3 & -2 & -3 \\ 2 & 5 & 0 \end{bmatrix}$.

(b) $2A = \begin{bmatrix} 2 & -4 & 0 \\ 6 & 2 & -2 \end{bmatrix}$ and $3B = \begin{bmatrix} 6 & 0 & -9 \\ -3 & 12 & 3 \end{bmatrix}$, so $2A - 3B = \begin{bmatrix} -4 & -4 & 9 \\ 9 & -10 & -5 \end{bmatrix}$.

(c) $-A+\frac{1}{2}B=\begin{bmatrix}-1 & 2 & 0\\ -3 & -1 & 1\end{bmatrix}+\begin{bmatrix}1 & 0 & -\frac{3}{2}\\ -\frac{1}{2} & 2 & \frac{1}{2}\end{bmatrix}=\begin{bmatrix}0 & 2 & -\frac{3}{2}\\ -\frac{7}{2} & 1 & \frac{3}{2}\end{bmatrix}$.

(d) $A+C$ is **not defined**: $A$ is $2\times 3$ and $C$ is $2\times 2$.

(e) $3C-C=(3-1)C=2C=\begin{bmatrix}2 & 2\\ 0 & 4\end{bmatrix}$. Simplifying first, using $(r+s)A=rA+sA$, saves work.

## Solution 3.2 COMPUTE ⋆⋆ *Matrix equations*

**Method.** The rules for + and scalar multiplication are the same as for numbers. Solve for $X$ symbolically, and replace by the matrices only at the end.

(a)

$$3X-2A=B+X \iff 3X-X=B+2A \iff 2X=2A+B \iff X=A+\tfrac{1}{2}B.$$

$$X=\begin{bmatrix}1 & -2 & 0\\ 3 & 1 & -1\end{bmatrix}+\begin{bmatrix}1 & 0 & -\frac{3}{2}\\ -\frac{1}{2} & 2 & \frac{1}{2}\end{bmatrix}=\begin{bmatrix}2 & -2 & -\frac{3}{2}\\ \frac{5}{2} & 3 & -\frac{1}{2}\end{bmatrix}.$$

**Check.** Entry $(1,1)$: $3X-2A$ gives $3(2)-2(1)=4$ and $B+X$ gives $2+2=4$ ✓. Entry $(2,3)$: $3(-\frac{1}{2})-2(-1)=\frac{1}{2}$ and $1+(-\frac{1}{2})=\frac{1}{2}$ ✓.

(b) This is a system of two equations, which we solve exactly as with numbers:

$$\begin{cases}X+Y=A\\ X-Y=B\end{cases} \xrightarrow{\text{add}} 2X=A+B, \qquad \xrightarrow{\text{subtract}} 2Y=A-B.$$

$$X=\tfrac{1}{2}\begin{bmatrix}3 & -2 & -3\\ 2 & 5 & 0\end{bmatrix}=\begin{bmatrix}\frac{3}{2} & -1 & -\frac{3}{2}\\ 1 & \frac{5}{2} & 0\end{bmatrix}, \qquad Y=\tfrac{1}{2}\begin{bmatrix}-1 & -2 & 3\\ 4 & -3 & -2\end{bmatrix}=\begin{bmatrix}-\frac{1}{2} & -1 & \frac{3}{2}\\ 2 & -\frac{3}{2} & -1\end{bmatrix}.$$

**Check.** $X+Y=\begin{bmatrix}1 & -2 & 0\\ 3 & 1 & -1\end{bmatrix}=A$ ✓.

## Solution 3.3 THINK ⋆⋆

**Method.** First compute the general form of $rI_2+sJ$, then compare entries.

$$rI_2+sJ=\begin{bmatrix}r & 0\\ 0 & r\end{bmatrix}+\begin{bmatrix}0 & s\\ s & 0\end{bmatrix}=\begin{bmatrix}r & s\\ s & r\end{bmatrix}.$$

(a) $\begin{bmatrix}r & s\\ s & r\end{bmatrix}=\begin{bmatrix}2 & 3\\ 3 & 2\end{bmatrix}$ gives $r=2$ and $s=3$.

(b) We would need $s=3$ (entry $(1,2)$) and $s=1$ (entry $(2,1)$), which is impossible. No pair $(r,s)$ works.

(c) The matrices $rI_2 + sJ$ are exactly those of the form $\begin{bmatrix} r & s \\ s & r \end{bmatrix}$:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ can be written as } rI_2 + sJ \iff a = d \text{ and } b = c.$$

The two diagonal entries must be equal, and so must the two off-diagonal entries. This explains (b): $3 \neq 1$.

## Solution 3.4 TRUE / FALSE ⋆⋆ *A, B, C of the same size, $k \in \mathbb{R}$*

(a) **True.** Add $-A$ to both sides: $(-A) + A + B = (-A) + A + C$, so $O + B = O + C$, so $B = C$. Addition can always be "cancelled".

(b) **True.** Suppose $A \neq O$. Then some entry $a_{ij} \neq 0$. The $(i, j)$ entry of $kA = O$ gives $k\, a_{ij} = 0$, and since $a_{ij} \neq 0$, $k = 0$.

(c) **False.** Take $k = 0$: then $kA = O = kB$ for *any* $A, B$, for instance $A = I_2$, $B = O$. The statement is true only when $k \neq 0$, because then we can multiply by $\frac{1}{k}$.

(d) **False.** $A - B = B - A \iff 2A = 2B \iff A = B$. So it fails as soon as $A \neq B$: with $A = I_2$ and $B = O$, $A - B = I_2$ but $B - A = -I_2$.

(e) **True.** Add $-A$ to both sides of $A + B = O$.

**Takeaway.** (a) and (b) are "cancellation" and "no zero divisors" for addition and scalars. Section 6 shows that **both fail for the matrix product**.

## Solution 3.5 PROVE ⋆⋆

**Method.** Two matrices of the same size are equal when their entries in every position $(i, j)$ are equal. Compare the $(i, j)$ entries of both sides, and use the rules for real numbers.

**$(r + s)A = rA + sA$.** Both sides have the size of $A$. For every $(i, j)$:

$$\big((r + s)A\big)_{ij} = (r + s)\, a_{ij} \overset{\text{numbers}}{=} r\, a_{ij} + s\, a_{ij} = (rA)_{ij} + (sA)_{ij} = (rA + sA)_{ij}.$$

**$A + (-A) = O$.** For every $(i, j)$: $\big(A + (-A)\big)_{ij} = a_{ij} + (-a_{ij}) = 0 = O_{ij}$. ■

**Takeaway.** Every property of addition and scalar multiplication is proved this way: work in *one* position $(i, j)$, and use the corresponding property of real numbers there.

# 4 The transpose

**What this section trains**

Transposing (rows ↔ columns), recognising symmetric and antisymmetric matrices, and the decomposition $A = S + K$ from the course.

### Solution 4.1 COMPUTE ⋆

**Method.** Row $i$ of $A$ becomes column $i$ of $A^T$.

$$A = \begin{bmatrix} 1 & -2 & 0 \\ 3 & 1 & -1 \end{bmatrix} \longrightarrow A^T = \begin{bmatrix} 1 & 3 \\ -2 & 1 \\ 0 & -1 \end{bmatrix}$$

$$B^T = \begin{bmatrix} 2 & -1 \\ 0 & 4 \\ -3 & 1 \end{bmatrix}, \qquad A + 2B = \begin{bmatrix} 5 & -2 & -6 \\ 1 & 9 & 1 \end{bmatrix} \implies (A+2B)^T = \begin{bmatrix} 5 & 1 \\ -2 & 9 \\ -6 & 1 \end{bmatrix},$$

$$A^T + 2B^T = \begin{bmatrix} 1 & 3 \\ -2 & 1 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 4 & -2 \\ 0 & 8 \\ -6 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ -2 & 9 \\ -6 & 1 \end{bmatrix}.$$

**Observation:** $(A+2B)^T = A^T + 2B^T$. This illustrates two properties from the course: $(A+B)^T = A^T + B^T$ and $(kA)^T = kA^T$. Transposing "passes through" sums and scalars.

### Solution 4.2 COMPUTE ⋆

**Method.** Symmetric: each entry equals its mirror image across the diagonal ($a_{ji} = a_{ij}$). Antisymmetric: each entry is the *opposite* of its mirror image ($a_{ji} = -a_{ij}$), which forces a zero diagonal.

(a) Pair up mirror positions:

$$(1,2) \leftrightarrow (2,1): \; x+1 = 4, \qquad (2,3) \leftrightarrow (3,2): \; y = 5, \qquad (1,3) \leftrightarrow (3,1): \; 3 = z - 1.$$

So $x = 3$, $y = 5$, $z = 4$. The diagonal entries $2, 0, 1$ are not constrained.

(b) Diagonal: $a = -a$, $d = -d$, $f = -f$, so $a = d = f = 0$. Off the diagonal:

$$c = -2, \qquad b = -1, \qquad e = -(-4) = 4.$$

**Check.** The answer to (b) is $\begin{bmatrix} 0 & 2 & -1 \\ -2 & 0 & -4 \\ 1 & 4 & 0 \end{bmatrix}$: each pair of mirror entries adds up to 0 ✓.

### Solution 4.3 TRUE / FALSE ⋆⋆

(a) **False.** $A$ is $2 \times 3$ and $A^T$ is $3 \times 2$: different sizes, so $A + A^T$ is not defined. (It is defined only when $A$ is square.)

(b) **True.** $A^T = A$ requires $A^T$ and $A$ to have the same size, i.e. $n \times m = m \times n$, so $m = n$.

(c) **True.** For $i = j$, $a_{ii} = -a_{ii}$, so $2a_{ii} = 0$ and $a_{ii} = 0$.

(d) **False.** Transposing reflects the matrix across the diagonal, so zeros below become zeros above: $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ is lower triangular, not upper.

(e) **True.** If $A^T = A$ and $B^T = B$, then $(A+B)^T = A^T + B^T = A + B$.

(f) **True.** The only nonzero entries are on the diagonal, which does not move when we transpose: $D^T = D$.

(g) **True.** $A = A^T = -A$ gives $2A = O$, so $A = O$.

### Solution 4.4 COMPUTE ⋆⋆

**Method.** Use the formulas from the course: $S = \frac{1}{2}(A + A^T)$ and $K = \frac{1}{2}(A - A^T)$.

$$A = \begin{bmatrix} 1 & 4 & -2 \\ 0 & 3 & 5 \\ 6 & 1 & 0 \end{bmatrix}, \qquad A^T = \begin{bmatrix} 1 & 0 & 6 \\ 4 & 3 & 1 \\ -2 & 5 & 0 \end{bmatrix}.$$

$$A + A^T = \begin{bmatrix} 2 & 4 & 4 \\ 4 & 6 & 6 \\ 4 & 6 & 0 \end{bmatrix} \Rightarrow S = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 3 & 3 \\ 2 & 3 & 0 \end{bmatrix},$$

$$A - A^T = \begin{bmatrix} 0 & 4 & -8 \\ -4 & 0 & 4 \\ 8 & -4 & 0 \end{bmatrix} \Rightarrow K = \begin{bmatrix} 0 & 2 & -4 \\ -2 & 0 & 2 \\ 4 & -2 & 0 \end{bmatrix}.$$

**Check.** (1) $S + K = \begin{bmatrix} 1 & 4 & -2 \\ 0 & 3 & 5 \\ 6 & 1 & 0 \end{bmatrix} = A$ ✓; (2) $S$ is a mirror image across the diagonal ✓; (3) $K$ has a zero diagonal and opposite mirror entries ✓.

> **Takeaway.** Each entry of $S$ is the *average* of $a_{ij}$ and its mirror $a_{ji}$: for instance $s_{13} = \frac{-2+6}{2} = 2$. $K$ takes *half the difference*: $k_{13} = \frac{-2-6}{2} = -4$.

### Solution 4.5 PROVE ⋆⋆⋆ *The decomposition is unique*

**Method.** Show that the difference of the two decompositions is a matrix that is both symmetric and antisymmetric, then use 4.3(g).

From $S + K = S' + K'$ we get

$$N := S - S' = K' - K.$$

- $N = S - S'$ is symmetric: $(S - S')^T = S^T - S'^T = S - S'$.
- $N = K' - K$ is antisymmetric: $(K' - K)^T = K'^T - K^T = -K' + K = -(K' - K)$.

So $N$ is both symmetric and antisymmetric; by Exercise 4.3(g), $N = O$. Hence $S = S'$ and $K = K'$. ■

**Takeaway.** "Show uniqueness" usually means: take two solutions, look at their difference, and show that it is zero.

### Solution 4.6 THINK ⋆⋆⋆ *How much freedom?*

**Method.** Once the entries on and above the diagonal are chosen, the symmetric (or antisymmetric) condition determines the entries below it.

| | | |
|---|---|---|
| free | free | free |
| $= a_{12}$ | free | free |
| $= a_{13}$ | $= a_{23}$ | free |

symmetric: $3 + 3 = 6$ free

| | | |
|---|---|---|
| 0 | free | free |
| $= -a_{12}$ | 0 | free |
| $= -a_{13}$ | $= -a_{23}$ | 0 |

antisymmetric: 3 free

- **Symmetric $3 \times 3$:** 3 diagonal entries + 3 entries above it, so 6 free entries.
- **Antisymmetric $3 \times 3$:** the diagonal is forced to 0 (4.3(c)), and only the 3 entries above it are free: 3.
- **$n \times n$:** above the diagonal there are $\frac{n^2-n}{2} = \frac{n(n-1)}{2}$ entries (all entries minus the diagonal, halved). So

$$\text{symmetric: } n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}, \qquad \text{antisymmetric: } \frac{n(n-1)}{2},$$

$$\frac{n(n+1)}{2} + \frac{n(n-1)}{2} = \frac{2n^2}{2} = n^2.$$

**Link with 4.4–4.5.** A general $n \times n$ matrix has $n^2$ free entries. Splitting it as $S + K$ (in exactly one way) distributes these $n^2$ degrees of freedom: $\frac{n(n+1)}{2}$ go to $S$ and $\frac{n(n-1)}{2}$ go to $K$. Nothing is lost and nothing is counted twice.

# 5 The matrix product: computing

**What this section trains**

The size rule $(m \times n)(n \times p) = m \times p$; the rule "row $i$ · column $j$"; the *column view* $A\vec{\mathbf{x}} = x_1\vec{\mathbf{a}}_1 + \cdots + x_n\vec{\mathbf{a}}_n$; and the meaning of a product in a concrete situation.

### Solution 5.1 COMPUTE ⋆ *Size first*

**Method.** Write the sizes next to each other. The product is defined when the *inner* sizes match; the result has the *outer* sizes. A transpose swaps the two sizes. For three factors, multiply two at a time.

| Product | Sizes | Result |
|---|---|---|
| $AB$ | $(3\times\mathbf{2})(\mathbf{2}\times 4)$ | $3\times 4$ |
| $BA$ | $(2\times\mathbf{4})(\mathbf{3}\times 2)$ | not defined |
| $BC$ | $(2\times\mathbf{4})(\mathbf{4}\times 4)$ | $2\times 4$ |
| $CD$ | $(4\times\mathbf{4})(\mathbf{4}\times 1)$ | $4\times 1$ |
| $DE$ | $(4\times\mathbf{1})(\mathbf{1}\times 3)$ | $4\times 3$ |
| $ED$ | $(1\times\mathbf{3})(\mathbf{4}\times 1)$ | not defined |
| $EA$ | $(1\times\mathbf{3})(\mathbf{3}\times 2)$ | $1\times 2$ |
| $AE$ | $(3\times\mathbf{2})(\mathbf{1}\times 3)$ | not defined |

| Product | Sizes | Result |
|---|---|---|
| $ABD$ | $(AB)D=(3\times\mathbf{4})(\mathbf{4}\times 1)$ | $3\times 1$ |
| $C^2$ | $(4\times\mathbf{4})(\mathbf{4}\times 4)$ | $4\times 4$ |
| $D^TD$ | $(1\times\mathbf{4})(\mathbf{4}\times 1)$ | $1\times 1$ |
| $DD^T$ | $(4\times\mathbf{1})(\mathbf{1}\times 4)$ | $4\times 4$ |
| $EAB$ | $(EA)B=(1\times\mathbf{2})(\mathbf{2}\times 4)$ | $1\times 4$ |

**Takeaway.** $DE$ is defined but $ED$ is not; $D^TD$ is a single number while $DD^T$ is $4\times 4$. The **order** of the factors changes everything, even before computing any entry.

## Solution 5.2 COMPUTE ⋆

**Method.** Determine the size of the result, then fill it with "row · column" dot products.

(a) $PQ$ is $(2\times 3)(3\times 2)=2\times 2$:

$$PQ=\begin{bmatrix}1\cdot 2+2(-1)+(-1)4 & 1\cdot 1+2\cdot 0+(-1)(-2)\\ 0\cdot 2+3(-1)+1\cdot 4 & 0\cdot 1+3\cdot 0+1(-2)\end{bmatrix}=\begin{bmatrix}-4 & 3\\ 1 & -2\end{bmatrix}.$$

$QP$ is $(3\times 2)(2\times 3)=3\times 3$:

$$QP=\begin{bmatrix}2\cdot 1+1\cdot 0 & 2\cdot 2+1\cdot 3 & 2(-1)+1\cdot 1\\ -1\cdot 1+0\cdot 0 & -1\cdot 2+0\cdot 3 & -1(-1)+0\cdot 1\\ 4\cdot 1-2\cdot 0 & 4\cdot 2-2\cdot 3 & 4(-1)-2\cdot 1\end{bmatrix}=\begin{bmatrix}2 & 7 & -1\\ -1 & -2 & 1\\ 4 & 2 & -6\end{bmatrix}.$$

The two products do not even have the same size: $2\times 2$ versus $3\times 3$.

(b) $\vec{\mathbf{u}}^T\vec{\mathbf{v}}$ is $(1\times 3)(3\times 1)=1\times 1$: $\vec{\mathbf{u}}^T\vec{\mathbf{v}}=\begin{bmatrix}1 & -2 & 3\end{bmatrix}\begin{bmatrix}4\\0\\-1\end{bmatrix}=[\,4+0-3\,]=[\,1\,]$. This is the **dot product** $\vec{\mathbf{u}}\cdot\vec{\mathbf{v}}=1$.

$\vec{\mathbf{u}}\,\vec{\mathbf{v}}^T$ is $(3\times 1)(1\times 3)=3\times 3$; each entry is $u_iv_j$:

$$\vec{\mathbf{u}}\,\vec{\mathbf{v}}^T=\begin{bmatrix}1\\-2\\3\end{bmatrix}\begin{bmatrix}4 & 0 & -1\end{bmatrix}=\begin{bmatrix}4 & 0 & -1\\ -8 & 0 & 2\\ 12 & 0 & -3\end{bmatrix}.$$

Each row is a multiple of $\vec{\mathbf{v}}^T$, and each column is a multiple of $\vec{\mathbf{u}}$.

**Check.** The diagonal entries of $\vec{\mathbf{u}}\vec{\mathbf{v}}^T$ are $u_1v_1,\ u_2v_2,\ u_3v_3$, so their sum is $4+0-3=1=\vec{\mathbf{u}}\cdot\vec{\mathbf{v}}$. This is not a coincidence: the diagonal of $\vec{\mathbf{u}}\vec{\mathbf{v}}^T$ contains exactly the terms of the dot product.

## Solution 5.3 COMPUTE ⋆⋆ *One entry at a time*

**Method.** $(AB)_{ij}$ only needs **row** $i$ **of** $A$ and **column** $j$ **of** $B$. There is no need to compute the other 8 entries.

$$\underset{A}{\begin{bmatrix} 2 & -1 & 3 \\ 0 & 4 & 1 \\ 5 & 2 & -2 \end{bmatrix}} \underset{B}{\begin{bmatrix} 1 & 0 & 2 \\ -3 & 1 & 4 \\ 2 & 5 & -1 \end{bmatrix}} \qquad \begin{aligned}(AB)_{23} &= (0,4,1)\cdot(2,4,-1) \\ &= 0+16-1 = \mathbf{15}\end{aligned}$$

- $(AB)_{23} = (\text{row 2 of } A)\cdot(\text{col. 3 of } B) = (0,4,1)\cdot(2,4,-1) = 0+16-1 = 15.$
- $(AB)_{31} = (\text{row 3 of } A)\cdot(\text{col. 1 of } B) = (5,2,-2)\cdot(1,-3,2) = 5-6-4 = -5.$
- $(BA)_{23} = (\text{row 2 of } \mathbf{B})\cdot(\text{col. 3 of } \mathbf{A}) = (-3,1,4)\cdot(3,1,-2) = -9+1-8 = -16.$

**Common mistake**

In $(BA)_{23}$ the *left* factor is $B$, so the row comes from $B$. Using row 2 of $A$ and column 3 of $B$ would give $(AB)_{23} = 15$ again.

**Second row of $AB$.** Each of its entries is (row 2 of $A$) · (a column of $B$). So we need **only row 2 of** $A$, together with all of $B$:

$$\text{row}_2(AB) = \text{row}_2(A)\,B.$$

### Solution 5.4 COMPUTE ⋆⋆ *The store, continued*

**Method.** Follow the meaning of the indices: $P$ is (store × item) and $C$ is (item × {cost, price}). The common index "item" is summed over, and what remains is (store × {cost, price}).

(a) $P$ is $2\times 3$ and $C$ is $3\times 2$: the inner sizes are both 3 (the three items), so $PC$ is defined and is $2\times 2$.

(b) Each entry is "quantities of a store" · "unit amounts of a column":

$$\begin{aligned}(PC)_{11} &= 120(2.00)+300(0.60)+15(11.00) = 240+180+165 = 585,\\ (PC)_{12} &= 120(3.50)+300(1.25)+15(18.00) = 420+375+270 = 1065,\\ (PC)_{21} &= 80(2.00)+450(0.60)+22(11.00) = 160+270+242 = 672,\\ (PC)_{22} &= 80(3.50)+450(1.25)+22(18.00) = 280+562.50+396 = 1238.50.\end{aligned}$$

| $PC$ | **total cost** | **total selling value** |
|---|---|---|
| **Laval** | \$585 | \$1065 |
| **Montréal** | \$672 | \$1238.50 |

$(PC)_{11}$ is what the Laval stock cost the company, and $(PC)_{12}$ is what it brings in if everything is sold. The second row gives the same two numbers for Montréal.

(c) Profit = selling value − cost: Laval: $1065-585 = \$480$; Montréal: $1238.50-672 = \$566.50$.

**Check.** Distributivity gives another way: profit $= P\big(\text{price column} - \text{cost column}\big)$, where the margins per item are $(1.50,\ 0.65,\ 7.00)$:

$$P\begin{bmatrix}1.50\\0.65\\7.00\end{bmatrix} = \begin{bmatrix}180+195+105\\120+292.50+154\end{bmatrix} = \begin{bmatrix}480\\566.50\end{bmatrix}\ \checkmark$$

(d) $CP$ is $(3\times 2)(2\times 3)$: defined, $3\times 3$. But its entries would be, for example, $(CP)_{11} = c_{11}p_{11}+c_{12}p_{21}$, which is (cost of a notebook) $\times$ (notebooks in Laval) + (price of a notebook) $\times$ (notebooks in Montréal). This mixes costs and prices: it has **no meaning**.

**Takeaway.** A product is *meaningful* when the summed index (the inner one) means the same thing in both factors: here, "item". Being *defined* (matching sizes) is necessary, but it is not enough to have a meaning.

## Solution 5.5 COMPUTE ⋆ *Systems and matrices*

**Method.** Row $i$ of the coefficient matrix holds the coefficients of equation $i$, in the order $x, y, z$, with a 0 for each missing variable.

(a)

$$\begin{cases} x-2y+3z=4\\ 2x+\mathbf{0}y+z=-1\\ \mathbf{0}x+y-z=0\end{cases} \iff \underbrace{\begin{bmatrix}1&-2&3\\2&0&1\\0&1&-1\end{bmatrix}}_{A}\underbrace{\begin{bmatrix}x\\y\\z\end{bmatrix}}_{\vec{\mathbf{x}}} = \underbrace{\begin{bmatrix}4\\-1\\0\end{bmatrix}}_{\vec{\mathbf{b}}}.$$

(b) Multiplying out (row · column):

$$\begin{cases} 3x-y=5\\ 2y=4\\ x+y=3\end{cases}$$

3 equations (rows), 2 unknowns (columns).

$(2,1)$: $3(2)-1=5$ ✓, $2(1)=2\neq 4$ ×. So $(2,1)$ is **not** a solution: a solution must satisfy *all* the equations.

**No solution.** The second equation forces $y=2$. Then the first gives $3x=7$, so $x=\frac{7}{3}$. But then $x+y=\frac{7}{3}+2=\frac{13}{3}\neq 3$, so the third equation fails. The values forced by the first two equations do not satisfy the third: the system has no solution.

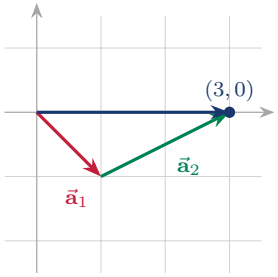
## Solution 5.6 THINK ⋆⋆ *The column view*

**Method.** $A\vec{\mathbf{x}} = x_1\vec{\mathbf{a}}_1 + x_2\vec{\mathbf{a}}_2 + x_3\vec{\mathbf{a}}_3$ with $\vec{\mathbf{a}}_1=(1,-1)$, $\vec{\mathbf{a}}_2=(2,1)$, $\vec{\mathbf{a}}_3=(0,3)$.

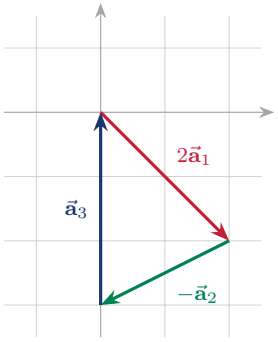
(a) $A\begin{bmatrix}1\\0\\0\end{bmatrix} = 1\vec{\mathbf{a}}_1+0\vec{\mathbf{a}}_2+0\vec{\mathbf{a}}_3 = \begin{bmatrix}1\\-1\end{bmatrix}$, and similarly we get $\begin{bmatrix}2\\1\end{bmatrix}$ and $\begin{bmatrix}0\\3\end{bmatrix}$: **the three columns of** $A$.

(b) We want a combination of the columns equal to $(3,0)$. Since $\vec{\mathbf{a}}_1 + \vec{\mathbf{a}}_2 = (1,-1) + (2,1) = (3,0)$, the vector $\vec{\mathbf{x}} = (1,1,0)$ works (left figure below).

(c) $A \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} = 2\vec{\mathbf{a}}_1 - \vec{\mathbf{a}}_2 + \vec{\mathbf{a}}_3 = (2,-2) - (2,1) + (0,3) = (0,0)$. Placed tip to tail, these three vectors form a **closed path** (right figure).




(b) $\vec{\mathbf{a}}_1 + \vec{\mathbf{a}}_2 = (3,0)$




(c) back to the start: $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$

**Consequence.** Let $\vec{\mathbf{x}} = (2,-1,1)$ and $\vec{\mathbf{y}} = \vec{\mathbf{0}}$: then $A\vec{\mathbf{x}} = \vec{\mathbf{0}} = A\vec{\mathbf{y}}$ but $\vec{\mathbf{x}} \neq \vec{\mathbf{y}}$. Adding this vector to the answer of (b) gives another solution: $A(3,0,1) = A(1,1,0) + A(2,-1,1) = (3,0)$. So $A\vec{\mathbf{x}} = A\vec{\mathbf{y}}$ **does not imply** $\vec{\mathbf{x}} = \vec{\mathbf{y}}$: here the three columns of $A$ can be combined, with coefficients that are not all zero, to cancel out.

### Solution 5.7 PROVE ⋆⋆⋆

**Method.** Use the column view with the vector $\vec{\mathbf{e}}_j$, whose only nonzero entry is a 1 in position $j$.

(a) By the column view,

$$A\vec{\mathbf{e}}_j = 0 \cdot \vec{\mathbf{a}}_1 + \cdots + 0 \cdot \vec{\mathbf{a}}_{j-1} + 1 \cdot \vec{\mathbf{a}}_j + 0 \cdot \vec{\mathbf{a}}_{j+1} + \cdots + 0 \cdot \vec{\mathbf{a}}_n = \vec{\mathbf{a}}_j.$$

Exercise 5.6(a) is this fact for $n = 3$.

(b) If $A\vec{\mathbf{x}} = B\vec{\mathbf{x}}$ for *every* $\vec{\mathbf{x}}$, it holds in particular for $\vec{\mathbf{x}} = \vec{\mathbf{e}}_1, \ldots, \vec{\mathbf{e}}_n$:

$$\vec{\mathbf{a}}_j = A\vec{\mathbf{e}}_j = B\vec{\mathbf{e}}_j = \vec{\mathbf{b}}_j \qquad \text{for } j = 1, \ldots, n.$$

All the columns are equal, so $A = B$. ■

**Takeaway.** A matrix is completely determined by what it does to $\vec{\mathbf{e}}_1, \ldots, \vec{\mathbf{e}}_n$. Section 7 uses exactly this idea: the columns are the images of $\vec{\mathbf{i}}$ and $\vec{\mathbf{j}}$.

# 6 The matrix product: properties and traps

**What this section trains**

Knowing which rules of numbers survive for matrices (associativity, distributivity in order, $IA = AI = A$), which fail (commutativity, cancellation, no zero divisors), and how to *find* a counterexample.

## Solution 6.1 TRUE / FALSE ⋆⋆

(a) **False.** With $A$ of size $2\times 3$ and $B$ of size $3\times 2$, both $AB$ ($2\times 2$) and $BA$ ($3\times 3$) are defined, but neither factor is square.

(b) **True.** Let $A$ be $m\times n$. For $AB$ to be defined, $B$ is $n\times p$; for $BA$ to be defined, $p = m$. So $AB$ is $m\times m$ and $BA$ is $n\times n$. Same size means $m = n$: $A$ and $B$ are both $n\times n$.

(c) **False.** $\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$, with neither factor equal to $O$.

(d) **False.** $N = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$ gives $N^2 = \begin{bmatrix}0\cdot 0+1\cdot 0 & 0\cdot 1+1\cdot 0\\ 0 & 0\end{bmatrix} = O$. (This is a nilpotent matrix, as in the course.)

(e) **False.** The rule is $(AB)^T = B^TA^T$, with the order reversed. Sizes are enough to see it: with $A$ of size $2\times 3$ and $B$ of size $3\times 2$, $(AB)^T$ is $2\times 2$ but $A^TB^T$ is $(3\times 2)(2\times 3) = 3\times 3$.

(f) **True.** Row $i$ of $AB$ is (row $i$ of $A$) $B$. Each of its entries is a dot product with the zero row, so it is 0.

(g) **False.** $\begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} = \begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$. A zero column 2 in $A$ only kills the terms $a_{i2}b_{2j}$ of each dot product; the terms $a_{i1}b_{1j}$ survive.

(h) **True.** $(AB)_{ij}$ = (row $i$ of $A$) $\cdot$ (column $j$ of $B$) $= (\ldots)\cdot\vec{\mathbf{0}} = 0$ for every $i$. So column $j$ of $AB$ is zero.

(i) **True.** Distribute, keeping the order: $(A-I)(A+I) = A^2+AI-IA-I^2 = A^2+A-A-I = A^2-I$. It works because $A$ commutes with $I$.

(j) **True.** If $A$ is $m\times n$, then $A^T$ is $n\times m$, and $(n\times \mathbf{m})(\mathbf{m}\times n)$ always matches. $A^TA$ is $n\times n$.

**Takeaway.** (f) and (h) are true, but (g) is false: zero **rows** pass from the *left* factor to the product, and zero **columns** pass from the *right* factor. This is one more way the product treats its two factors differently.

## Solution 6.2 COMPUTE ⋆⋆ *Who commutes?*

**Method.** Compute both products with a general $B$ and compare entries.

(a)
$$DB = \begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix} = \begin{bmatrix}a & b\\ 2c & 2d\end{bmatrix}, \qquad BD = \begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix} = \begin{bmatrix}a & 2b\\ c & 2d\end{bmatrix}.$$

Multiplying by $D$ on the *left* scales the *rows* (row 2 by 2); on the *right*, it scales the *columns* (column 2 by 2). Equality requires $b = 2b$ and $2c = c$, i.e. $b = 0$ and $c = 0$, while $a$ and $d$ are free:

$$DB = BD \iff B = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} \quad \text{(the diagonal matrices)}.$$

(b) For any $2 \times 2$ matrix $B$, using the rule $k(AB) = (kA)B = A(kB)$ and $IB = BI = B$:

$$(3I_2)B = 3(I_2B) = 3B = 3(BI_2) = B(3I_2).$$

## Solution 6.3 THINK ⋆⋆ *Equations that behave strangely*

(a) $A^2 = \begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 4-4 & -8+8 \\ 2-2 & -4+4 \end{bmatrix} = O$, although $A \neq O$.

(b) $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ✓and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0+1 & 0 \\ 0 & 1+0 \end{bmatrix} = I_2$ ✓. For every $t$:

$$X_t^2 = \begin{bmatrix} 1 & t \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & t \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & t-t \\ 0 & 1 \end{bmatrix} = I_2.$$

Each real value of $t$ gives a different solution, so $X^2 = I_2$ **has infinitely many solutions**, not just $\pm I_2$. Geometrically, each of them is a transformation that comes back to where it started when applied twice, like a reflection. For example, $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is the reflection across $y = x$ (Exercise 7.1(d)).

(c) $AB = AC \iff A(B - C) = O$. We need a nonzero matrix $N = B - C$ with $AN = O$. Every row of $A$ is $(1, 1)$, so $AN = O$ exactly when each column of $N$ has entries adding up to 0:

$$N = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}: \qquad AN = \begin{bmatrix} 1-1 & 1-1 \\ 1-1 & 1-1 \end{bmatrix} = O.$$

Take $B = N$ and $C = O$: then $AB = O = AC$, but $B \neq C$. (Any $C$ works, with $B = C + N$; for example, $C = I_2$ and $B = \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$ give $AB = AC = A$.)

**The root of (a), (b) and (c)**

With numbers, $xy = 0$ forces $x = 0$ or $y = 0$; that is how we solve $x^2 = 1$ and simplify $xy = xz$. With matrices, $AB = O$ is possible with $A, B \neq O$ (zero divisors), so these familiar arguments break down.

## Solution 6.4 COMPUTE ⋆⋆ *Expanding correctly*

**Method.** $(A + B)^2$ means $(A + B)(A + B)$. Distribute, keeping each product in its order.

(a) $A + B = \begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix}$, so $(A + B)^2 = \begin{bmatrix} 4+2 & 4+4 \\ 2+2 & 2+4 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 4 & 6 \end{bmatrix}$.

(b) The four pieces:

$$A^2 = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}, \quad B^2 = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \quad AB = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}, \quad BA = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \ (\neq AB).$$

$$A^2 + 2AB + B^2 = \begin{bmatrix} 8 & 8 \\ 4 & 4 \end{bmatrix} \neq (A+B)^2, \qquad A^2 + AB + BA + B^2 = \begin{bmatrix} 6 & 8 \\ 4 & 6 \end{bmatrix} = (A+B)^2.$$

**Why:** $(A+B)(A+B) = AA+AB+BA+BB$ (four products, each in the order it appears). Grouping $AB + BA$ into $2AB$ requires $BA = AB$, which is false here.

(c) First, $(A+I)^2 = A^2 + AI + IA + I^2 = A^2 + 2A + I$. Then

$$(A+I)^3 = (A^2 + 2A + I)(A + I) = A^3 + A^2 + 2A^2 + 2A + A + I = A^3 + 3A^2 + 3A + I.$$

It is correct because $A$ *does* commute with $I$ ($AI = IA = A$), so $AI$ and $IA$ can be grouped. The binomial formulas hold for any two matrices that commute.

**Takeaway.** Commutativity is a property of a *pair* of matrices. Some pairs commute ($A$ and $I$, two rotations, two diagonal matrices), and most do not. The identities from high-school algebra hold exactly for the pairs that commute.

## Solution 6.5 COMPUTE ⋆⋆ *$AA^T$ and $A^TA$*

**Method.** $(AA^T)_{ij}$ = (row $i$ of $A$)·(row $j$ of $A$) and $(A^TA)_{ij}$ = (col. $i$ of $A$)·(col. $j$ of $A$). Every entry is a dot product between rows (or between columns) of $A$.

(a) The rows of $A$ are $\vec{\mathbf{r}}_1 = (1, -2, 0)$ and $\vec{\mathbf{r}}_2 = (3, 1, -1)$; its columns are $\vec{\mathbf{c}}_1 = (1, 3)$, $\vec{\mathbf{c}}_2 = (-2, 1)$, $\vec{\mathbf{c}}_3 = (0, -1)$.

$$AA^T = \begin{bmatrix} \vec{\mathbf{r}}_1 \cdot \vec{\mathbf{r}}_1 & \vec{\mathbf{r}}_1 \cdot \vec{\mathbf{r}}_2 \\ \vec{\mathbf{r}}_2 \cdot \vec{\mathbf{r}}_1 & \vec{\mathbf{r}}_2 \cdot \vec{\mathbf{r}}_2 \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ 1 & 11 \end{bmatrix} \ (2 \times 2), \qquad A^TA = \begin{bmatrix} 10 & 1 & -3 \\ 1 & 5 & -1 \\ -3 & -1 & 1 \end{bmatrix} \ (3 \times 3).$$

Both are **symmetric**. This is natural: $\vec{\mathbf{r}}_1 \cdot \vec{\mathbf{r}}_2 = \vec{\mathbf{r}}_2 \cdot \vec{\mathbf{r}}_1$, since the dot product is commutative.

(b) $\operatorname{tr}(AA^T) = 5 + 11 = 16$ and $\operatorname{tr}(A^TA) = 10 + 5 + 1 = 16$. The sum of the squares of the entries is $1 + 4 + 0 + 9 + 1 + 1 = 16$. All three are equal: the diagonal entries are $\|\vec{\mathbf{r}}_i\|^2$ (or $\|\vec{\mathbf{c}}_j\|^2$), and adding them up adds the square of every entry once. Exercise 8.5 proves this in general.

## Solution 6.6 PROVE ⋆⋆

(a) Using $(XY)^T = Y^TX^T$ and $(X^T)^T = X$:

$$(AA^T)^T = (A^T)^TA^T = AA^T, \qquad (A^TA)^T = A^T(A^T)^T = A^TA.$$

(b) Group the factors (associativity) and apply the rule twice:

$$(ABC)^T = \big((AB)C\big)^T = C^T(AB)^T = C^T\big(B^TA^T\big) = C^TB^TA^T.$$

As with socks and shoes, the last matrix applied is the first one to come off.

(c) If row $i$ of $A$ is zero, then $a_{ik} = 0$ for every $k$, so for every $j$:

$$(AB)_{ij} = \sum_k a_{ik} b_{kj} = \sum_k 0 \cdot b_{kj} = 0.$$

The whole row $i$ of $AB$ is zero. ■

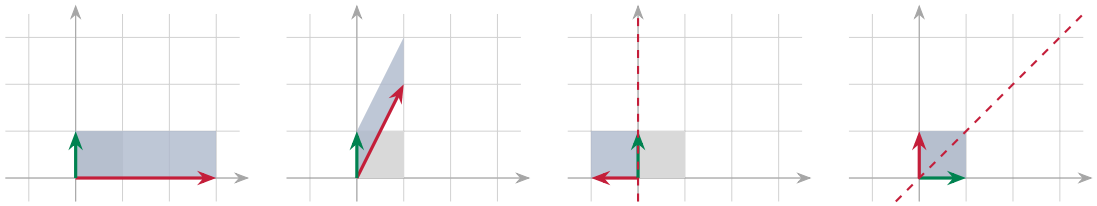
# 7 Matrices as transformations of the plane

**What this section trains**

Reading a $2 \times 2$ matrix as a transformation: the columns are the images of $\vec{\mathbf{i}}$ and $\vec{\mathbf{j}}$. The product is composition: $AB$ means "first $B$, then $A$".

### Solution 7.1 DRAW ⋆

**Method.** Draw column 1 (the image of $\vec{\mathbf{i}}$, red) and column 2 (the image of $\vec{\mathbf{j}}$, green). The image of the unit square is the parallelogram they span.



**(a)** **(b)** **(c)** **(d)**

(a) $\vec{\mathbf{i}} \mapsto (3,0)$, $\vec{\mathbf{j}} \mapsto (0,1)$: the square becomes a $3 \times 1$ rectangle. **Horizontal stretch by a factor** 3: $(x, y) \mapsto (3x, y)$.

(b) $\vec{\mathbf{i}} \mapsto (1,2)$, $\vec{\mathbf{j}} \mapsto (0,1)$: $(x, y) \mapsto (x,\ 2x + y)$. Each vertical line $x = c$ is mapped onto itself, but its points slide vertically by $2c$ (up if $c > 0$, down if $c < 0$): a **vertical shear**.

(c) $\vec{\mathbf{i}} \mapsto (-1,0)$, $\vec{\mathbf{j}} \mapsto (0,1)$: $(x, y) \mapsto (-x, y)$. **Reflection across the $y$-axis** (dashed line).

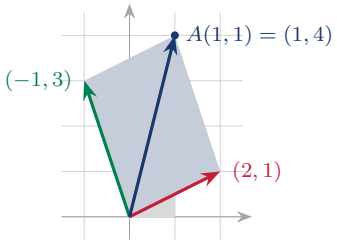
(d) $\vec{\mathbf{i}} \mapsto (0,1)$, $\vec{\mathbf{j}} \mapsto (1,0)$: $(x, y) \mapsto (y, x)$. **Reflection across the line $y = x$** (dashed). The square lands on itself, but $\vec{\mathbf{i}}$ and $\vec{\mathbf{j}}$ are swapped: it is flipped over.

### Solution 7.2 COMPUTE ⋆ *From images to matrix*

**Method.** The columns of $A$ are the images of $\vec{\mathbf{i}}$ and $\vec{\mathbf{j}}$. Then use the column view: $A(x, y) = x(\text{col. } 1) + y(\text{col. } 2)$.

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

- $A\begin{bmatrix}1\\1\end{bmatrix} = 1\begin{bmatrix}2\\1\end{bmatrix} + 1\begin{bmatrix}-1\\3\end{bmatrix} = \begin{bmatrix}1\\4\end{bmatrix}$.
  $(1,1)$ is the far corner of the unit square, so its image is the far corner of the parallelogram (figure).




- $A\begin{bmatrix}3\\-2\end{bmatrix} = 3\begin{bmatrix}2\\1\end{bmatrix} - 2\begin{bmatrix}-1\\3\end{bmatrix} = \begin{bmatrix}6+2\\3-6\end{bmatrix} = \begin{bmatrix}8\\-3\end{bmatrix}$.

**Check.** Row · column: $2(3) + (-1)(-2) = 8$ ✓, $1(3) + 3(-2) = -3$ ✓.

### Solution 7.3 THINK ⋆⋆ *Order matters*

**Method.** Find where $F$ and $R$ send $\vec{\mathbf{i}}$ and $\vec{\mathbf{j}}$; those images are the columns. For a composition, remember that the matrix on the **right** acts first.

(a) $F$ (reflection across the $x$-axis): $\vec{\mathbf{i}} \mapsto (1,0)$, $\vec{\mathbf{j}} \mapsto (0,-1)$. $R$ (rotation by 90°): $\vec{\mathbf{i}} \mapsto (0,1)$, $\vec{\mathbf{j}} \mapsto (-1,0)$.

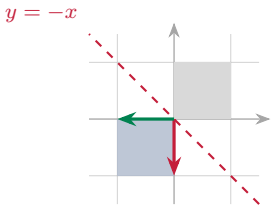
$$F = \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}, \qquad R = \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}.$$

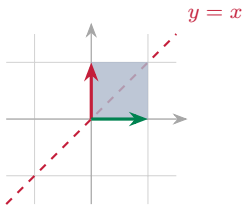
(b)

$$FR = \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix} = \begin{bmatrix}0 & -1\\-1 & 0\end{bmatrix}, \qquad RF = \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix} = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}.$$

$FR \neq RF$: **the order matters**.

(c) $FR =$ "$R$ first, then $F$": $(1,0) \xrightarrow{R} (0,1) \xrightarrow{F} (0,-1)$, the first column of $FR$ ✓.
$RF =$ "$F$ first, then $R$": $(1,0) \xrightarrow{F} (1,0) \xrightarrow{R} (0,1)$, the first column of $RF$ ✓.



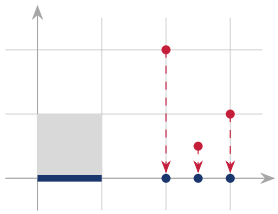

$FR$: reflection across $y = -x$ $RF$: reflection across $y = x$

red: image of $\vec{\mathbf{i}}$ green: image of $\vec{\mathbf{j}}$ grey: unit square

**Identification.** $FR(x,y) = (-y,-x)$: the reflection across the line $y = -x$. $RF(x,y) = (y,x)$: the reflection across the line $y = x$, the matrix of 7.1(d). Composing a reflection and a rotation gives a reflection, but *which* reflection depends on the order.

### Solution 7.4 THINK ⋆⋆ *A projection*

(a) $P\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x\\0\end{bmatrix}$: each point drops vertically onto the $x$-axis. This is the **orthogonal projection** onto the $x$-axis. The unit square is flattened onto the segment from $(0,0)$ to $(1,0)$.



every point of a vertical line lands on the same point: the height $y$ is lost

(b) $P^2 = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} = P$. **Geometrically:** after the first projection the point is already on the $x$-axis, so projecting it a second time does not move it. Projecting twice is the same as projecting once.

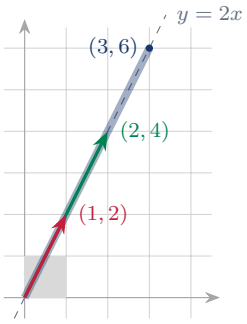
(c) Many points (for instance $(2,0)$, $(2,1)$ and $(2,5)$) have the same image $(2,0)$. Knowing only the image, we cannot tell which point we started from, so the transformation cannot be undone, and $P$ has no inverse.

**Check.** $ad - bc = 1 \cdot 0 - 0 \cdot 0 = 0$ ✓.

**Takeaway.** $P^2 = P$ with $P \neq O$ and $P \neq I$. This shows again that $A^2 = A$ does not force $A = O$ or $A = I$ (see 10.1(a)).

## Solution 7.5 THINK ⋆⋆⋆ *When the square collapses*

(a) $\vec{\mathbf{i}} \mapsto (1,2)$ and $\vec{\mathbf{j}} \mapsto (2,4) = 2(1,2)$. The two images are **parallel**, so the "parallelogram" they span is flat: the unit square becomes the **segment** from $(0,0)$ to $(1,2)+(2,4) = (3,6)$.




(b) By the column view, for every $\vec{\mathbf{x}} = (x_1, x_2)$:

$$A\vec{\mathbf{x}} = x_1\begin{bmatrix}1\\2\end{bmatrix} + x_2\begin{bmatrix}2\\4\end{bmatrix} = (x_1 + 2x_2)\begin{bmatrix}1\\2\end{bmatrix}.$$

Every image is a multiple of $(1,2)$, so it lies on the line $y = 2x$ through the origin. The whole plane is squashed onto this line.

(c) $ad - bc = 1 \cdot 4 - 2 \cdot 2 = 0$. $ad - bc = 0$ exactly when the two columns are parallel (or one of them is zero). Then the square collapses, as in (a), and $A$ is not invertible. Many points go to the same image, as in 7.4.

## Solution 7.6 PROVE ⋆⋆⋆ *Rotations and trigonometry*

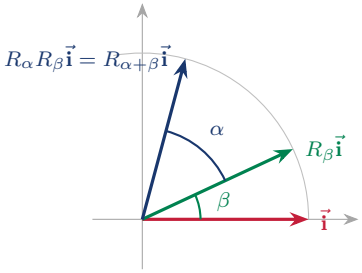
(a) $\cos 90° = 0$ and $\sin 90° = 1$, so $R_{90°} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = R$ ✓.

(b)

$$R_\alpha R_\beta = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix}$$
$$= \begin{bmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -(\cos\alpha\sin\beta + \sin\alpha\cos\beta) \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta & \cos\alpha\cos\beta - \sin\alpha\sin\beta \end{bmatrix}.$$

This must equal $R_{\alpha+\beta} = \begin{bmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{bmatrix}$. Comparing the first column:

$$\boxed{\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta}, \qquad \boxed{\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta}.$$




(c) Yes: $R_\alpha R_\beta = R_{\alpha+\beta} = R_{\beta+\alpha} = R_\beta R_\alpha$, because $\alpha + \beta = \beta + \alpha$. Rotating by $\alpha$ then $\beta$ or by $\beta$ then $\alpha$ gives the same total angle. Unlike $F$ and $R$ in 7.3, two rotations *commute*.

(d)

$$R_\theta^T R_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos^2\theta + \sin^2\theta & 0 \\ 0 & \sin^2\theta + \cos^2\theta \end{bmatrix} = I_2.$$

(The off-diagonal entries are $-\cos\theta\sin\theta + \sin\theta\cos\theta = 0$.) The product $R_\theta R_\theta^T$ also equals $I_2$, by the same computation. Hence

$$R_\theta^{-1} = R_\theta^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = R_{-\theta}$$

(using $\cos(-\theta) = \cos\theta$ and $\sin(-\theta) = -\sin\theta$). To undo a rotation by $\theta$, rotate by $-\theta$.

# 8 Powers and trace

**What this section trains**

Computing a few powers, spotting a pattern and proving it (by induction or with an algebraic shortcut), and using the properties of the trace, in particular $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

**Solution 8.1** COMPUTE ⋆⋆ *Guess, then prove*

(a)
$$A^2 = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 0 & 4 \end{bmatrix}, \qquad A^3 = A^2A = \begin{bmatrix} 4 & 4 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 8 & 12 \\ 0 & 8 \end{bmatrix}.$$

(b) **Guess.** The diagonal entries are $2, 4, 8$, i.e. $2^n$. The upper-right entries are $1, 4, 12$; dividing by $2^{n-1}$ gives $1, 2, 3$, so they are $n\,2^{n-1}$.

$$\text{Conjecture:} \quad A^n = \begin{bmatrix} 2^n & n\,2^{n-1} \\ 0 & 2^n \end{bmatrix} \quad (n \geq 1).$$

**Proof by induction.**
*Base case* $n = 1$: $\begin{bmatrix} 2 & 1\cdot 2^0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} = A$ ✓.
*Induction step:* assume the formula holds for some $n$. Then

$$A^{n+1} = A^nA = \begin{bmatrix} 2^n & n\,2^{n-1} \\ 0 & 2^n \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 2^{n+1} & 2^n + n\,2^n \\ 0 & 2^{n+1} \end{bmatrix} = \begin{bmatrix} 2^{n+1} & (n+1)\,2^n \\ 0 & 2^{n+1} \end{bmatrix},$$

which is the formula for $n+1$. By induction, it holds for every $n \geq 1$. ■

(c) $n = 10$: $A^{10} = \begin{bmatrix} 2^{10} & 10\cdot 2^9 \\ 0 & 2^{10} \end{bmatrix} = \begin{bmatrix} 1024 & 5120 \\ 0 & 1024 \end{bmatrix}$.

**Check.** The formula gives $A^3$ correctly: $\begin{bmatrix} 8 & 3\cdot 4 \\ 0 & 8 \end{bmatrix} = \begin{bmatrix} 8 & 12 \\ 0 & 8 \end{bmatrix}$ ✓.

**Common mistake**

Raising each entry to the power: $\begin{bmatrix} 2^{10} & 1^{10} \\ 0 & 2^{10} \end{bmatrix}$ is **wrong**. Only *diagonal* matrices can be powered entry by entry (as in the course).

## Solution 8.2 COMPUTE ⋆⋆ *A shortcut*

(a) $J^2 = \begin{bmatrix} 1+1 & 1+1 \\ 1+1 & 1+1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} = 2J.$

(b) Use *only* $J^2 = 2J$ and the rules for scalars:

$$J^3 = J^2J = (2J)J = 2J^2 = 2(2J) = 4J, \qquad J^4 = J^3J = (4J)J = 4J^2 = 8J.$$

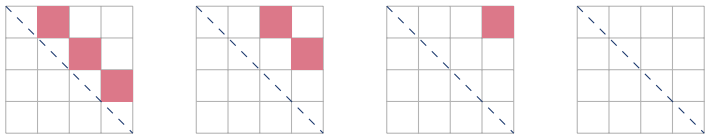
Each multiplication by $J$ doubles the coefficient: $J^n = 2^{n-1}J = \begin{bmatrix} 2^{n-1} & 2^{n-1} \\ 2^{n-1} & 2^{n-1} \end{bmatrix}$. (Induction: if $J^n = 2^{n-1}J$, then $J^{n+1} = 2^{n-1}J^2 = 2^{n-1}\cdot 2J = 2^nJ$.)

**Takeaway.** A *relation* satisfied by a matrix ($J^2 = 2J$ here) lets us compute any power without multiplying any more matrices. The same idea gives an inverse without a formula in Exercise 9.5.

## Solution 8.3 COMPUTE ⋆⋆

**Method.** Multiplying by $N$ on the right shifts the columns: column $j$ of $XN$ is column $j-1$ of $X$, and the first column becomes zero. So every power moves the line of 1's one column to the right, i.e. one diagonal further from the main diagonal.

$$N = \begin{bmatrix} 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow N^2 = \begin{bmatrix} 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow N^3 = \begin{bmatrix} 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow N^4 = O.$$



$N$ $N^2$ $N^3$ $N^4$

The 1's go from the first diagonal above the main one (3 entries) to the second (2 entries), then the third (1 entry), and then they leave the matrix. The smallest $k$ with $N^k = O$ is $\mathbf{k = 4}$: $N$ is **nilpotent**, like the strictly upper triangular matrices in the course.

## Solution 8.4 TRUE / FALSE ⋆⋆ *A, B, C square of the same order*

(a) **True.** The diagonal of $A+B$ is $a_{ii}+b_{ii}$; summing over $i$ gives $\operatorname{tr} A + \operatorname{tr} B$.

(b) **False.** $A = B = I_2$: $\operatorname{tr}(I_2 I_2) = \operatorname{tr}(I_2) = 2$, but $\operatorname{tr}(I_2)\operatorname{tr}(I_2) = 2 \cdot 2 = 4$.

(c) **True.** Transposing does not move the diagonal entries.

(d) **True.** Apply $\operatorname{tr}(XY) = \operatorname{tr}(YX)$ with $X = AB$ and $Y = C$: $\operatorname{tr}\big((AB)C\big) = \operatorname{tr}\big(C(AB)\big) = \operatorname{tr}(CAB)$.

(e) **False.** We may move the *last* factor to the front (cyclic shift), but we may not swap two factors. Counterexample:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ C = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}:$$

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ ABC = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ \operatorname{tr}(ABC) = 1,$$

$$BA = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},\ BAC = O,\ \operatorname{tr}(BAC) = 0 \neq 1.$$

(f) **False.** $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ has trace 0, but it is not $O$.

**Takeaway.** The trace does not change under a *cyclic* shift of the factors: $\operatorname{tr}(ABC) = \operatorname{tr}(BCA) = \operatorname{tr}(CAB)$. Any other reordering may change it.

## Solution 8.5 PROVE ⋆⋆⋆

(a) The $j$-th diagonal entry of $A^TA$ is (row $j$ of $A^T$) $\cdot$ (column $j$ of $A$), i.e. (column $j$ of $A$) $\cdot$ (column $j$ of $A$):

$$(A^TA)_{jj} = \sum_i a_{ij}\, a_{ij} = \sum_i a_{ij}^2 = \big\|\text{column } j \text{ of } A\big\|^2.$$

Summing over $j$: $\operatorname{tr}(A^TA) = \sum_j \sum_i a_{ij}^2$, the sum of the squares of *all* the entries of $A$. With the matrix of Exercise 6.5: $\operatorname{tr}(A^TA) = 10 + 5 + 1 = 16$ and $1^2 + (-2)^2 + 0^2 + 3^2 + 1^2 + (-1)^2 = 16$ ✓.

(b) If $A^TA = O$, then $\operatorname{tr}(A^TA) = 0$, so $\sum_{i,j} a_{ij}^2 = 0$. A sum of squares of real numbers is zero only if every term is zero, so every $a_{ij} = 0$ and $A = O$. ■

**Takeaway.** Compare with $A^2 = O$, which does *not* force $A = O$ (6.3(a)). Putting $A^T$ in front changes everything: each diagonal entry of $A^TA$ is a sum of squares, which cannot cancel out.

# 9 The inverse of a matrix

**What this section trains**

Computing $2 \times 2$ inverses, recognising non-invertible matrices ($ad - bc = 0$), solving $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$, $AX = B$ and $XA = B$ by multiplying on the correct side, and reasoning with the definition $AA^{-1} = A^{-1}A = I$.

### Solution 9.1 COMPUTE ⋆

**Method.** Compute $ad - bc$ first. If it is 0, stop: the matrix is not invertible. Otherwise, swap $a$ and $d$, change the signs of $b$ and $c$, and divide by $ad - bc$.

(a) $ad - bc = 15 - 14 = 1$, so $\begin{bmatrix} 5 & 2 \\ 7 & 3 \end{bmatrix}^{-1} = \frac{1}{1}\begin{bmatrix} 3 & -2 \\ -7 & 5 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -7 & 5 \end{bmatrix}$.

**Check.** $\begin{bmatrix} 5 & 2 \\ 7 & 3 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -7 & 5 \end{bmatrix} = \begin{bmatrix} 15 - 14 & -10 + 10 \\ 21 - 21 & -14 + 15 \end{bmatrix} = I_2$ ✓.

(b) $ad - bc = 6 - 6 = 0$: **not invertible**. The second column $(6, 3)$ is 3 times the first $(2, 1)$, so the square collapses onto a line, as in Exercise 7.5.

(c) $ad - bc = 6 - (-4) = 10$, so $\begin{bmatrix} 3 & -1 \\ 4 & 2 \end{bmatrix}^{-1} = \frac{1}{10}\begin{bmatrix} 2 & 1 \\ -4 & 3 \end{bmatrix}$.

(d) $ad - bc = 0 - 1 = -1$, so the inverse is $\frac{1}{-1}\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$: **the matrix is its own inverse**. This makes sense: it is the reflection across $y = x$ (7.1(d)), and reflecting twice brings every point back.

### Solution 9.2 COMPUTE ⋆

**Method.** Write the system as $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ and multiply both sides **on the left** by $A^{-1}$: $A^{-1}A\vec{\mathbf{x}} = A^{-1}\vec{\mathbf{b}}$, so $\vec{\mathbf{x}} = A^{-1}\vec{\mathbf{b}}$.

$$\underbrace{\begin{bmatrix}5 & 2\\ 7 & 3\end{bmatrix}}_{A}\begin{bmatrix}x\\ y\end{bmatrix} = \begin{bmatrix}1\\ 2\end{bmatrix} \quad\Longrightarrow\quad \begin{bmatrix}x\\ y\end{bmatrix} = \begin{bmatrix}3 & -2\\ -7 & 5\end{bmatrix}\begin{bmatrix}1\\ 2\end{bmatrix} = \begin{bmatrix}3-4\\ -7+10\end{bmatrix} = \begin{bmatrix}-1\\ 3\end{bmatrix}.$$

**Check.** $5(-1)+2(3) = 1$ ✓ $\quad 7(-1)+3(3) = 2$ ✓. Since $A$ is invertible, this is the *only* solution.

## Solution 9.3 COMPUTE ⋆⋆

**Method.** The matrix is invertible exactly when $ad - bc \neq 0$. Compute $ad - bc$ as a function of $k$, factor it, and exclude its roots.

(a) $ad - bc = k \cdot k - 3 \cdot k = k^2 - 3k = k(k-3)$. This is zero for $k = 0$ or $k = 3$. The matrix is invertible **for every** $k$ **except** 0 **and** 3. (For $k = 3$: $\begin{bmatrix}3 & 3\\ 3 & 3\end{bmatrix}$ has equal columns. For $k = 0$: $\begin{bmatrix}0 & 3\\ 0 & 0\end{bmatrix}$ has a zero column.)

(b) $ad - bc = (1-k)(2-k) - 6 = 2 - 3k + k^2 - 6 = k^2 - 3k - 4 = (k-4)(k+1)$. The matrix is invertible **for every** $k$ **except** 4 **and** $-1$.

**Check.** For $k = 4$: $\begin{bmatrix}-3 & 2\\ 3 & -2\end{bmatrix}$; the second row is $-1$ times the first, so $ad - bc = 6 - 6 = 0$ ✓.

## Solution 9.4 THINK ⋆⋆ *Left or right?*

**Method.** To remove $A$, multiply by $A^{-1}$ on the **same side** as $A$, and do it on both sides of the equation.

First, $ad - bc = 2 - 1 = 1$, so $A^{-1} = \begin{bmatrix}1 & -1\\ -1 & 2\end{bmatrix}$.

(a) $AX = B$: $A$ is on the **left** of $X$, so multiply on the left:

$$A^{-1}(AX) = A^{-1}B \;\Rightarrow\; X = A^{-1}B = \begin{bmatrix}1 & -1\\ -1 & 2\end{bmatrix}\begin{bmatrix}1 & 0\\ 3 & 2\end{bmatrix} = \begin{bmatrix}-2 & -2\\ 5 & 4\end{bmatrix}.$$

(b) $XA = B$: $A$ is on the **right** of $X$, so multiply on the right:

$$(XA)A^{-1} = BA^{-1} \;\Rightarrow\; X = BA^{-1} = \begin{bmatrix}1 & 0\\ 3 & 2\end{bmatrix}\begin{bmatrix}1 & -1\\ -1 & 2\end{bmatrix} = \begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}.$$

**Check.** $AX = \begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}-2 & -2\\ 5 & 4\end{bmatrix} = \begin{bmatrix}1 & 0\\ 3 & 2\end{bmatrix} = B$ ✓ and $XA = \begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix} = \begin{bmatrix}1 & 0\\ 3 & 2\end{bmatrix} = B$ ✓.

(c) The two solutions are **different**. The notation $\frac{B}{A}$ would not say whether we mean $A^{-1}B$ or $BA^{-1}$, which is why the course forbids it.

**Common mistake**

Writing $AX = B \Rightarrow X = BA^{-1}$. Check the sizes and positions: $A^{-1}$ must end up *next to* $A$ in order to cancel it.

### Solution 9.5 THINK ⋆⋆ *An inverse without a formula*

**Method.** Isolate $I$ on one side, factor out $A$, then divide by the scalar.

(a) From $A^2 - 3A + 2I = O$:

$$A^2 - 3A = -2I \iff 3A - A^2 = 2I \iff A(3I - A) = 2I \iff A \cdot \tfrac{1}{2}(3I - A) = I.$$

(We used $3A - A^2 = A \cdot 3I - A \cdot A = A(3I - A)$.) In the same way, $(3I - A)A = 3A - A^2 = 2I$, so $\frac{1}{2}(3I - A) \cdot A = I$ as well. Both products equal $I$, so by definition $A$ is invertible and

$$A^{-1} = \tfrac{1}{2}(3I - A).$$

(b) $A^2 = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & 4 \end{bmatrix}$, so

$$A^2 - 3A + 2I = \begin{bmatrix} 1 & 3 \\ 0 & 4 \end{bmatrix} - \begin{bmatrix} 3 & 3 \\ 0 & 6 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \checkmark$$

Then $\frac{1}{2}(3I - A) = \frac{1}{2}\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$. The formula $(ad - bc = 2)$ gives $\frac{1}{2}\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$: the same matrix ✓.

### Solution 9.6 THINK ⋆⋆⋆

(a) (i) $ad - bc = 2(-2) - (-4)(1) = -4 + 4 = 0$, so $A$ is not invertible.

(ii) *Argument by contradiction.* Suppose $A^{-1}$ exists. Multiplying $A^2 = O$ on the left by $A^{-1}$:

$$A^{-1}A^2 = A^{-1}O \;\Rightarrow\; (A^{-1}A)A = O \;\Rightarrow\; IA = O \;\Rightarrow\; A = O.$$

But $A \neq O$: contradiction. So $A^{-1}$ does not exist. This argument shows more: *no* nonzero matrix with $A^2 = O$ is invertible.

(b) Check the definition directly, using $A^2 = O$:

$$(I - A)(I + A) = I + A - A - A^2 = I - A^2 = I, \qquad (I + A)(I - A) = I - A + A - A^2 = I.$$

Both products equal $I$, so $I - A$ is invertible and $(I - A)^{-1} = I + A$. Explicitly, $I - A = \begin{bmatrix} -1 & 4 \\ -1 & 3 \end{bmatrix}$ and $I + A = \begin{bmatrix} 3 & -4 \\ 1 & -1 \end{bmatrix}$.

**Check.** With the formula: for $I - A$, $ad - bc = -3 + 4 = 1$, and the inverse is $\begin{bmatrix} 3 & -4 \\ 1 & -1 \end{bmatrix} = I + A$ ✓.

**Solution 9.7** TRUE / FALSE ⋆⋆ *$A$, $B$ square of the same order*

(a) **False.** $I$ and $-I$ are invertible, but $I + (-I) = O$ is not.

(b) **False.** $AB$ is invertible, but its inverse is $B^{-1}A^{-1}$ (socks and shoes). Indeed,

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I.$$

The formula $A^{-1}B^{-1}$ is wrong in general. With $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$:

$$A^{-1}B^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}, \qquad (AB)^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}.$$

(c) **False.** $A = B = I$: $(I + I)^{-1} = (2I)^{-1} = \frac{1}{2}I$, but $I^{-1} + I^{-1} = 2I$. (Already false for numbers: $\frac{1}{1+1} \neq \frac{1}{1} + \frac{1}{1}$.)

(d) **True.** Multiply $AB = AC$ on the left by $A^{-1}$: $A^{-1}AB = A^{-1}AC$, so $B = C$. We *can* simplify by an invertible matrix.

(e) **True.** $B = IB = (A^{-1}A)B = A^{-1}(AB) = A^{-1}O = O$.

(f) **True.** If the first row is zero ($a = b = 0$), then $ad - bc = 0 \cdot d - 0 \cdot c = 0$; the second row is similar.

(g) **True.** By the course, $(A^T)^{-1} = (A^{-1})^T$. Check: $A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I$, and $(A^{-1})^T A^T = (AA^{-1})^T = I$.

**Takeaway.** (d) and (e) are the cancellation rules that fail in Section 6. They come back as soon as $A$ is **invertible**. This is why invertible matrices are so important.

# 10 Synthesis

**What this section trains**

Using the whole chapter: finding the hidden assumption in a wrong argument, interpreting a product in a new context, and writing a short proof.

**Solution 10.1** THINK ⋆⋆ *Spot the error*

**Method.** For each step, ask which property of the course justifies it, then check whether that property really holds for matrices.

(a) **Error:** "$A(A - I) = O$, so $A = O$ or $A - I = O$". This uses "no zero divisors", which is false for matrices (6.1(c)). **Counterexample:** the projection $P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ of 7.4: $P^2 = P$, yet $P \neq O$ and $P \neq I$.

(b) **Error:** simplifying by $A$ means multiplying by $A^{-1}$, which is only possible if $A$ is invertible. **Counterexample:** $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$, $C = O$ (6.3(c)): $AB = AC = O$, but $B \neq C$.

(c) **Error:** an analogy is not a proof. The transpose of a *product* reverses the order: $(AB)^T = B^T A^T$. **Counterexample:** $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$: $(AB)^T = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}^T = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$, while $A^T B^T = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}$.

(d) **Error:** $ABAB = AABB$ swaps the middle factors $B$ and $A$, i.e. it assumes $BA = AB$. **Counterexample:** the same $A, B$: $AB = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}$, so $(AB)^2 = \begin{bmatrix} 11 & 8 \\ 4 & 3 \end{bmatrix}$, while $A^2 B^2 = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 9 & 4 \\ 2 & 1 \end{bmatrix}$.

**Takeaway.** All four errors come from the same habit: treating matrices like numbers. Before each step, ask: *does this step need $AB = BA$? Does it need $A$ to be invertible?*

## Solution 10.2 COMPUTE ⋆⋆ *Counting flights*

(a) Row $i$ lists the flights *leaving* city $i$:

$$M = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \qquad \begin{array}{l} \text{from 1: to 2 and 3} \\ \text{from 2: to 3} \\ \text{from 3: to 1} \end{array}$$

$M$ is **not symmetric**: $m_{12} = 1$ (flight $1 \to 2$) but $m_{21} = 0$ (no flight $2 \to 1$). A symmetric $M$ would mean that every flight also goes back the other way.

(b)

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

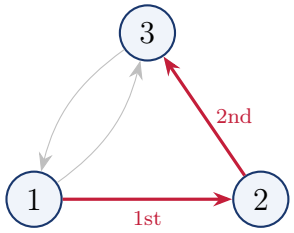
**Why it counts routes.** The term $m_{ik} m_{kj}$ is 1 if both $i \to k$ and $k \to j$ are flights (a route $i \to k \to j$ with a stop at $k$), and 0 otherwise. Summing over the three possible stops $k = 1, 2, 3$ counts all the two-flight routes from $i$ to $j$.

$$(M^2)_{13} = \underbrace{m_{11} m_{13}}_{0 \cdot 1} + \underbrace{m_{12} m_{23}}_{1 \cdot 1} + \underbrace{m_{13} m_{33}}_{1 \cdot 0} = 1 : \qquad \text{the only route is } 1 \to 2 \to 3.$$

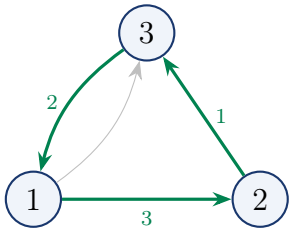
(c)

$$M + M^2 = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Entry $(i, j)$ counts the routes from $i$ to $j$ with *one or two* flights. Row 2 is $(1, 0, 1)$: from city 2 you can reach **city 1** ($2 \to 3 \to 1$) and **city 3** (direct), but $(M + M^2)_{22} = 0$: **you cannot come back to 2** with at most two flights.




$(M^2)_{13} = 1$: $1 \to 2 \to 3$




back to 2 needs 3 flights

To come back to 2, you need three flights: $2 \to 3 \to 1 \to 2$. Indeed, $(M^3)_{22} = 1$.

### Solution 10.3 PROVE ⋆⋆⋆

**Method.** An "if and only if" statement needs two directions. Here a single chain of equivalences gives both at once.

Since $A^T = A$ and $B^T = B$:

$$(AB)^T = B^T A^T = BA.$$

So

$$AB \text{ symmetric} \iff (AB)^T = AB \iff BA = AB.$$

■

**Illustration.** $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ are symmetric but do not commute, and indeed $AB = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$ is not symmetric. A product of symmetric matrices is not symmetric in general.

### Solution 10.4 PROVE ⋆⋆⋆ *Matrices that commute with everything*

**Method.** If $A$ commutes with *every* matrix, it commutes in particular with well-chosen simple matrices. Each choice gives conditions on $a, b, c, d$.

(a)
$$AE = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix}, \qquad EA = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}.$$
$AE = EA$ gives $b = 0$ (entry $(1,2)$) and $c = 0$ (entry $(2,1)$).

(b) Now $A = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$:
$$AJ = \begin{bmatrix} 0 & a \\ d & 0 \end{bmatrix}, \qquad JA = \begin{bmatrix} 0 & d \\ a & 0 \end{bmatrix}.$$
$AJ = JA$ gives $a = d$.

(c) So $A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} = aI_2$, a **scalar matrix**. Conversely, every scalar matrix commutes with every matrix (6.2(b)). **Conclusion:** the matrices that commute with all $2 \times 2$ matrices are exactly the scalar matrices $kI_2$.

**Takeaway.** The scalar matrices are the only ones that behave like numbers with respect to the product. Every other matrix fails to commute with at least one matrix.