STEP Support Programme

STEP 2 Matrices Hints

Question 1

The complex number $x + \mathrm{i}y$ is mapped into the complex number $X + \mathrm{i}Y$ where X and Y are given by the equation

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} X \\ Y \end{pmatrix}.$$

Which numbers are invariant under the mapping?

This question may look somewhat daunting, as it involves both complex numbers and matrices. But looks can be deceptive. Just give it a go! What would it mean algebraically for a number to be invariant under this mapping?

Question 2

The simultaneous equations

$$\begin{aligned} x + 2y &= 4, \\ 2x - y &= 0, \\ 3x + y &= 5 \end{aligned}$$

may be written in matrix form as

$$\begin{pmatrix} 1 & 2 \\ 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 5 \end{pmatrix}, \quad \text{or} \quad \mathbf{AX} = \mathbf{B}.$$

Carry out numerically the procedure of the three following steps:

(1) $\mathbf{A}^{\mathrm{T}}\mathbf{AX} = \mathbf{A}^{\mathrm{T}}\mathbf{B}$;

(2) $(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{AX} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{B}$;

(3) $\mathbf{IX} = \begin{pmatrix} x \\ y \end{pmatrix} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{B}$.

You will need to ensure you are clear about how to multiply rectangular matrices for this question. Once you are, this numerical calculation is reasonably straightforward if you write down all of the matrices involved explicitly. (Leave the vector $\mathbf{X}$ as it is.) It turns out that $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ is a square matrix, so you can find its inverse.

Verify that the values of x, y so found do not satisfy all the original three equations. Suggest a reason for this.

Once you have shown that the x and y found do not satisfy the original equations, you could try to find values of x and y which *do* satisfy them. What do you notice?

Under what circumstances will the procedure given above, when applied to a set of three simultaneous equations in two variables, result in values which satisfy the equations?

This means: take any three simultaneous equations in two variables and apply the above procedure. Does it result in values which satisfy all three equations? What can you say about the simultaneous equations in those cases where the procedure *does* result in values (x and y) which satisfy the equations?

Given the previous part, you should be able to make a reasonable guess as to when the procedure might work. Proving this is indeed the case is a little more subtle. It is helpful to assume that there is a solution $\mathbf{X}$, and then work through the procedure under this assumption.

There is a further subtlety which you could explore if you want a challenge, which is determining the conditions under which $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ is invertible.

Question 3

> Let $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ be real 2×2 matrices and write
>
> $$[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}, \text{ etc.}$$
>
> Prove that:
>
> (i) $[\mathbf{A}, \mathbf{A}] = \mathbf{O}$, where $\mathbf{O}$ is the zero matrix,
>
> (ii) $[[\mathbf{A}, \mathbf{B}], \mathbf{C}] + [[\mathbf{B}, \mathbf{C}], \mathbf{A}] + [[\mathbf{C}, \mathbf{A}], \mathbf{B}] = \mathbf{O}$,
>
> (iii) if $[\mathbf{A}, \mathbf{B}] = \mathbf{I}$, then $[\mathbf{A}, \mathbf{B}^m] = m\mathbf{B}^{m-1}$ for all positive integers m.
>
> At each step you should state clearly any properties of matrices which you use.

The key property of matrices we need here is that $(\mathbf{PQ})\mathbf{R} = \mathbf{P}(\mathbf{QR})$ for any conformable matrices $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{R}$. This is known as the *associativity* of matrix multiplication, or that matrix multiplication is *associative*.

For part (iii), "for all positive integers m" should suggest using induction. The algebra is somewhat tricky, so may take you a few attempts: if you don't find it the first time round, have another go, and then another.

> The *trace*, $\text{Tr}(\mathbf{A})$, of a matrix
>
> $$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$
>
> is defined by
>
> $$\text{Tr}(\mathbf{A}) = a_{11} + a_{22}.$$
>
> Prove that:
>
> (iv) $\text{Tr}(\mathbf{A} + \mathbf{B}) = \text{Tr}(\mathbf{A}) + \text{Tr}(\mathbf{B})$,
>
> (v) $\text{Tr}(\mathbf{AB}) = \text{Tr}(\mathbf{BA})$,
>
> (vi) $\text{Tr}(\mathbf{I}) = 2$.

Though this may look somewhat unfamiliar, writing out the calculations explicitly for (iv) and (v) will allow you to prove these. Part (vi) is a very quick calculation: don't be put off!

> Deduce that there are no matrices satisfying $[\mathbf{A}, \mathbf{B}] = \mathbf{I}$. Does this in any way invalidate the statement in (iii)?

The first part invites you to put together all the different parts of this question in a useful way. The second part is an interesting question of mathematical logic: when is an "if ... then" statement true? We have the following standard definition of the meaning of the statement "if A then B" in mathematics (where A and B are some sort of statement):

A	B	if A then B
true	true	true
true	false	false
false	true	true
false	false	true

This may seem somewhat confusing at first: the third column specifies whether the *whole statement* "if A then B" is true or false. Once that makes a little sense, the last two lines are quite surprising: if A is false, then we regard the whole "if A then B" statement as *true*. This is frequently different from the logic of ordinary conversation, where a statement such as "if it's raining, then I'll carry an umbrella" is usually understood to also mean "and if it's not raining, then I won't carry an umbrella". But in mathematical logic, it is extremely useful to understand the statement "if A then B in the way described above. One way of thinking about this is to ask: In which situation would we know that "if A then B" is *false*? We would only know that it is false if A itself were true, and yet B were false. And our convention is that we take the statement to be true otherwise.

Question 4

Matrices $\mathbf{P}$ and $\mathbf{Q}$ are given by

$$\mathbf{P} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \mathbf{Q} = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}$$

(where $\mathrm{i}^2 = -1$). Show that $\mathbf{P}^2 = \mathbf{Q}^2$, $\mathbf{PQP} = \mathbf{Q}$ and $\mathbf{P}^4 = \mathbf{I}$, the identity matrix.

Don't be put off by the complex numbers appearing in these matrices! The rules for multiplying matrices are still the usual ones.

Deduce that, for all positive integers n, $\mathbf{P}^n\mathbf{Q}\mathbf{P}^n = \mathbf{Q}$.

This sounds like a proof by induction question. Alternatively, since $\mathbf{P}^4 = \mathbf{I}$, we could prove it by showing that a few cases are true, though that is more work.

Hence, or otherwise, show that if $\mathbf{X}$ and $\mathbf{Y}$ are each matrices of the form

$$\mathbf{P}^m\mathbf{Q}^n, \quad m = 1, 2, 3, 4;\ n = 1, 2$$

then $\mathbf{XY}$ has the same form.

This is somewhat fiddly. There are exactly 8 matrices of this form, and your argument must show that the product of any pair of them can be written in the same form. You have all of the earlier results at your disposal. One particularly useful thing to bear in mind is that we can add 4 to a power of $\mathbf{P}$ without changing it (and likewise subtract 4).

Question 5

(a) Show that if $\mathbf{A} = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$, then

$$\mathbf{A}^2 - (p+s)\mathbf{A} + (ps - qr)\mathbf{I} = \mathbf{O},$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{O}$ is the zero matrix.

You are only being asked to show that this equation is true, not asked to derive it. So it is enough just to calculate the left hand side and show that it gives the zero matrix.

(b) Given that $\mathbf{X} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and that $\mathbf{X}^2 = \mathbf{O}$, show that $\mathbf{X}$ can be written either in terms of a and b only or in terms of c only, or of b only.

One way to do this is to work out $\mathbf{X}^2$ and then equate each coefficient to zero. This might well be quite hard work.

A more straightforward approach is to use part (a) of the question, replacing $\mathbf{A}$ by $\mathbf{X}$. What does this tell us about $\mathbf{X}$?

The question suggests that there are several different cases: one in which $\mathbf{X}$ can be written in terms of a and b only, one in which it can be written in terms of c only and one in which it can be written in terms of b only. So the solution will need to consider several different cases.

(b) (continued)

Show that when $\mathbf{X}$ is written in terms of c only, the solution can be written in the form:

$$\mathbf{X} = c \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

and interpret this result in terms of transformations of the plane represented by these matrices, relating your answer to the fact that $\mathbf{X}^2 = \mathbf{O}$.

This provides a useful check on the answer to the first part: this expression should expand to one of the possible forms for $\mathbf{X}$. In what order are the three transformations applied, and what are these transformations?

Question 6

A mapping $(x, y) \to (u, v)$ is given by

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -8 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Show briefly that this mapping is not one to one.

Can you find two points mapping to the same (u, v)? It might be helpful to write expressions for u and v in terms of x and y.

Find the locus, L, of all points which map to $(1, -4)$.

That is, find the set of all such points.

Describe the locus of (u, v) as (x, y) is allowed to vary throughout the plane.

Show that any given point, P, on this locus is the image of just one point on the y-axis, and describe how the set of all points with image P is related to the locus L.

Work with explicit coordinates: what are the coordinates of P on this locus?

Question 7

You are given that $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{R}$ are 2×2 matrices, $\mathbf{I}$ is the identity matrix and $\mathbf{P}^{-1}$ exists.

(i) Prove, by expanding both sides, that

$$\det(\mathbf{PQ}) = \det \mathbf{P} \det \mathbf{Q}.$$

Deduce that

$$\det(\mathbf{P}^{-1}\mathbf{Q} + \mathbf{I}) = \det(\mathbf{Q}\mathbf{P}^{-1} + \mathbf{I}).$$

The first part is a straightforward calculation using the definitions of matrix multiplication and determinant.

For the second part, how can you use the first part to help you?

(ii) If $\mathbf{PX} = \mathbf{XP}$ for *every* 2×2 matrix $\mathbf{X}$, prove that $\mathbf{P} = \lambda\mathbf{I}$, where λ is a constant.

This is not related to the first part of the question!

What choices of matrix $\mathbf{X}$ would be helpful to use here? It may take more than one such matrix to reach the desired conclusion.

(iii) If $\mathbf{RQ} = \mathbf{QR}$, prove that

$$\mathbf{RQ}^n = \mathbf{Q}^n\mathbf{R} \quad \text{and} \quad \mathbf{R}^n\mathbf{Q}^n = \mathbf{Q}^n\mathbf{R}^n$$

for any positive integer n.

The first part looks like a standard induction argument.

The second part has lots of ns in it. One way to approach it is by induction on n as normal. Another approach is to think about the powers of $\mathbf{R}$ and $\mathbf{Q}$ as two different numbers, and instead attempt to prove the more general statement $\mathbf{R}^m\mathbf{Q}^n = \mathbf{Q}^n\mathbf{R}^m$; the case $m = 1$ is the first part. For the second approach, it is important to be clear what values m and n are allowed to take at each stage of the argument.

Question 8

The real 3×3 matrix $\mathbf{A}$ is such that $\mathbf{A}^2 = \mathbf{A}$.

(i) Prove that $(\mathbf{I} - \mathbf{A})^2 = \mathbf{I} - \mathbf{A}$.

It is not necessary to write out the matrix $\mathbf{A}$ explicitly to answer this question. Is it always true that $(\mathbf{Q} - \mathbf{R})^2 = \mathbf{Q}^2 - 2\mathbf{QR} + \mathbf{R}^2$ for matrices as it is with real numbers? So you will need to take some care with the algebra here!

(ii) Express $(\mathbf{I} - \mathbf{A})^3$ in the form $\mathbf{I} + k\mathbf{A}$, where k is a number to be determined.

(iii) Prove that, for all real constants λ and all positive integers n,

$$(\mathbf{I} + \lambda\mathbf{A})^n = \mathbf{I} + \big((\lambda + 1)^n - 1\big)\mathbf{A}.$$

Use this result to verify your answer to (ii).

This seems like a standard application of induction: fix λ and induct on n.

Acknowledgements

The exam questions are reproduced by kind permission of Cambridge Assessment Group Archives. Unless otherwise noted, the questions are reproduced verbatim, except for some differences in convention between the papers as printed and STEP, specifically: the use of j to represent $\sqrt{-1}$ has been replaced by i; variables representing matrices are written in boldface type rather than italics; transpose is denoted $\mathbf{A}^{\mathrm{T}}$ rather that A'.

In the list of sources below, the following abbreviations are used:

O&C	Oxford and Cambridge Schools Examination Board
SMP	School Mathematics Project
MEI	Mathematics in Education and Industry
QP	Question paper
Q	Question

1 O&C, A level Mathematics (SMP), 1966, QP Mathematics II, Q A3

2 O&C, A level Mathematics (SMP), 1967, QP Mathematics II, Q B22

3 O&C, A level Mathematics (MEI), 1968, QP MEI 20, Pure Mathematics III (Special Paper), Q 3; editorial changes here: the definition of **O** is inserted, the implication symbol is written in words, and the reference to the matrix ring is removed

4 O&C, A level Mathematics (MEI), 1968, QP MEI 143*, Pure Mathematics I, Q 6

5 O&C, A level Mathematics (MEI), 1980, QP 9655/1, Pure Mathematics 1, Q 2; editorial changes here: use **O** rather than **0** for the zero matrix, and define the notation.

6 O&C, A level Mathematics (MEI), 1981, QP 9655/1, Pure Mathematics 1, Q 6(b)

7 O&C, A level Mathematics (MEI), 1986, QP 9657/0, Mathematics 0 (Special Paper), Q 2; editorial change here: **I** is called the identity matrix rather than the unit matrix

8 O&C, A level Mathematics (MEI), 1987, QP 9650/2, Mathematics 2, Q 16