MA10209 Algebra 1A

Sheet 8 Problems and Solutions v1: GCS

27-xi-11

The course website is http://people.bath.ac.uk/masgcs/diary.html

Hand in work to your tutor by 13:00, Monday Nov 28.

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a 2 by 2 matrix with real entries. Define a map $f_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ by $f_A((x, y)) = (u, v)$ where

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

for all $(x, y) \in \mathbb{R}^2 = \mathbb{R} \oplus \mathbb{R}$. If $(x, y) \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$, we define $\lambda(x, y)$ to be $(\lambda x, \lambda y)$. We write $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ as I_2, and more generally, the matrix I_n is an $n \times n$ matrix with each entry 0, except that the entries on the leading diagonal are 1.

1. (a) Prove that $f_A = \mathrm{Id}_{\mathbb{R}^2}$ if, and only if, $A = I_2$.
Solution Suppose that $A = I_2$. Then $f_A((x, y)) = (x, y)$ for all $(x, y) \in \mathbb{R}^2$ because

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

Conversely if $f_A = \mathrm{Id}_{\mathbb{R}^2}$, then $f_A((1, 0)) = (1, 0)$ and $f_A((1, 0)) = (0, 1)$. These two matrix equations lead to four ordinary equations: $a = 1$, $b = 0$, $c = 0$ and $d = 1$ so $A = I_2$.

(b) Suppose that $u, v \in \mathbb{R}^2$ and $\lambda, \mu \in \mathbb{R}$. Prove that $f_A(\lambda u + \mu v) = \lambda f_A(u) + \mu f_A(v)$.
Solution This is a very important point. It means that the map f_A is, in some sense, compatible with the linear operations going on. Suppose that $u = (u_1, u_2)$ and $v = (v_1, v_2)$ then $\lambda u + \mu v = (\lambda u_1 + \mu v_1, \lambda u_2 + \mu v_2)$ so $f_A(\lambda u + \mu v) = (a(\lambda u_1 + \mu v_1) + b(\lambda u_2 + \mu v_2), c(\lambda u_1 + \mu v_1) + d(\lambda u_2 + \mu v_2))$. On the other hand $\lambda_1 f_A(u)) = \lambda_1(au_1 + bu_2, cu_1 + du_2)$ and $\lambda_1 f_A(v)) = \lambda_1(av_1 + bv_2, cv_1 + du_2)$. The required formula follows.

(c) Determine $f_A((0, 0))$, $f_A((1, 0))$, $f_A(1, 1))$ and $f_A((0, 1))$.
Solution $f_A((0, 0)) = (0, 0)$, $f_A((1, 0)) = (a, c)$, $f_A((0, 1)) = (b, d)$, $f_A((1, 1)) = (a + b, c + d)$, $f_A((0, 1)) = (b, d)$.

(d) Determine the area of the polygon which has the points in part (c) as vertices. There is a helpful formula involving the vector (cross) product which, if you know it, will shorten your calculation.

Solution We will use $\mathbf{i}, \mathbf{j}, \mathbf{k}$ for three mutually orthogonal unit vectors in $\mathbb{R}^3$ with $\mathbf{i} = (1,0,0)$, $\mathbf{j} = (0,1,0)$ and $\mathbf{k} = (0,0,1)$. We let $\times$ denote the (vector) cross product, so $\mathbf{i} \times \mathbf{j} = \mathbf{k}$ and we will take for granted the standard properties of the (vector) cross product. In particular, the (signed) area of the parallelogram defined by two vectors is their cross product, the unsigned area of a parallelogram being the product of the lengths of two adjacent sides, multiplied by the sine of the angle between them. In our case, the required area is therefore the signed length of the vector $(a,c) \times (b,d) = (a\mathbf{i} + d\mathbf{j}) \times (b\mathbf{i} + d\mathbf{j}) = (ad - bc)\mathbf{k}$ which is $ad - bc$. A more elementary method is to study the area of this triangle by means of the areas of lots of appropriate right-angled triangles using the co-ordinate system. It is reassuring that such a method exists, but it suffers from issues of diagram dependency.

(e) Describe $\{f_A((x,0)) \mid x \in \mathbb{R}\}$.
Solution This is $\{(ax, cx) \mid x \in \mathbb{R}\}$ which is also $\{x(a,c) \mid x \in \mathbb{R}\}$. If $(a,c) \neq (0,0)$, then this is the straight line through the origin and (a,c). However, if $(a,c) = (0,0)$, then this is the point $\{(0,0)\}$.

(f) Describe $\{f_A((0,y)) \mid y \in \mathbb{R}\}$.
Solution This is $\{(by, dy) \mid y \in \mathbb{R}\}$which is also $\{y(b,d) \mid y \in \mathbb{R}\}$. If $(b,d) \neq (0,0)$, then this is the straight line through the origin and (b,d). However, if $(b,d) = (0,0)$, then this is the point $\{(0,0)\}$.

2. Let $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ and $C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$ be 2×2 matrices with real entries. Prove that $f_B \circ f_C = f_{BC}$ where BC denotes the product of the matrices B and C.
Solution Suppose that $(x,y) \in \mathbb{R}^2$. We have $f_C((x,y)) = (u,v)$ where

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c_{11}x + c_{12}y \\ c_{21}x + c_{22}y \end{pmatrix}.$$

Now $f_B(f_C((x,y)) = (r,s)$ where

$$\begin{pmatrix} r \\ s \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} b_{11}u + b_{12}v \\ b_{21}u + b_{22}v \end{pmatrix} =$$

$$\begin{pmatrix} b_{11}(c_{11}x + c_{12}y) + b_{12}(c_{21}x + c_{22}y) \\ b_{21}(c_{11}x + c_{12}y) + b_{22}(c_{21}x + c_{22}y) \end{pmatrix} = \begin{pmatrix} (b_{11}c_{11} + b_{12}c_{21})x + (b_{11}c_{12} + b_{12}c_{22})y \\ (b_{21}c_{11} + b_{22}c_{21})x + (b_{21}c_{12} + b_{22}c_{22})y \end{pmatrix}$$

and this is $f_{BC}((x,y))$. The maps f_{BC} and $f_B \circ f_C$ have the same domain, $\mathbb{R}^2$, the same codomain, $\mathbb{R}^2$, and act in the same way on an each element, so $f_{BC} = f_B \circ f_C$.

3. Suppose that co-ordinates $u, v \in \mathbb{R}^2$ and $u \neq v$. The straight line through u and v is $\{\lambda u + (1-\lambda)v \mid \lambda \in \mathbb{R}\}$.
Comment We elaborate on this. Suppose that geometric points U and V have co-ordinates $u, v \in \mathbb{R}^2$. A point Z with co-ordinates $z \in \mathbb{R}^2$ will be on the line through U and V if, and only if, there is $\lambda \in \mathbb{R}$ such that $z = v + \lambda(u - v) = \lambda u + (1-\lambda)v$. Note that Z divides VU in the ratio $\lambda : 1 - \lambda$ but that is in some sense the wrong way round. Let us turn it round. Put another way, Z divides UV in the ratio $1 - \lambda : \lambda$, and viewed as a 'number' this is $(1-\lambda)/\lambda$. Notice that this ratio is positive if, and only if, Z is strictly between U and V. It is 0 when $Z = U$ so $\lambda = 1$. It is undefined, or infinite if you

prefer, when $Z = V$ so $\lambda = 0$. The point Z is on UV produced beyond U when $\lambda > 1$ and the ratio is in the open interval $(-1, 0)$. The point Z is on UV produced beyond V when $\lambda < 0$, which corresponds the the ratio lying in the open interval $(-\infty, -1)$.

The description of the co-ordinates of a point on a line as $z = \lambda u + (1-\lambda)v$ can be made more symmetric. You can write it as $z = \lambda u + \mu v$ where λ, μ range over $\mathbb{R}$ subject to the condition $\lambda + \mu = 1$. Put in these terms, points on the line between U and V are those where the description of their co-ordinates has $0 < \lambda < 1$ (or equivalently $0 < \lambda < 1$).

(a) Suppose that f_A is injective. Prove that if $L \subseteq \mathbb{R}^2$ is a straight line, then $\{f_A(w) \mid w \in L\}$ is a straight line.
Solution If $w \in L$, then there is $\lambda \in \mathbb{R}$ such that $w = \lambda u + (1-\lambda)v$. Now $f_A(w) = f_A(\lambda u + (1-\lambda)v) = \lambda f_A(u) + (1-\lambda)f_A(v)$ by the result of Problem 1(b). If f_A is injective, then $f_A(u) \neq f_A(v)$ so as λ varies, the point $f_A(w)$ ranges over a straight line.

(b) Suppose that f_A is injective. Choose three distinct points $u, v, w \in \mathbb{R}^2$ which are not collinear. Prove that $f_A(u), f_A(v)$ and $f_A(w)$ are not collinear.
Solution Suppose, for contradiction, that $f_A(u), f_A(v)$ and $f_A(w)$ are collinear. They must be distinct since f_A is injective. Therefore $f_A(w) = \lambda f_A(u) + (1-\lambda)f_A(v)$ for some real number λ. Now $f_A(w) = f_A(\lambda u + (1-\lambda)v)$ by Problem 1(b), but w is not on the line through u and v whereas $\lambda u + (1-\lambda)v$ is on that line. Therefore $w \neq \lambda u + (1-\lambda)v$ and this violates the injectivity of f_A.

(c) Suppose that f_A is injective. Prove that it is bijective.
Solution Choose three points u, v and w as in the previous part. Choose any $t \in \mathbb{R}^2$. Draw a line L through t which intersects both the line L_1 through $f_A(u)$ and $f_A(v)$, and the line L_2 through $f_A(u)$ and $f_A(w)$ at different points. Now points on the lines L_1 and L_2 are in the image of f_A, so there are $r, s \in \mathbb{R}^2$ such that $L \cap L_1 = \{f_A(r)\}$ and $L \cap L_2 = \{f_A(s)\}$. Now t is on the line through $f_A(r)$ and $f_A(s)$, so there is $\theta \in \mathbb{R}$ such that

$$t = \theta f_A(r) + (1-\theta)f_A(s) = f_A(\theta r + (1-\theta)s)$$

by Problem 1(b). Therefore f_A is surjective.

(d) Prove that if $ad \neq bc$, then f is bijective.
Solution Since $ad \neq bc$, Problem 1(d) tells us that the parallelogram formed by the images under f_A of $(0,0), (0,1), (1,1)$ and $(1,0)$ is not flat, so $(0,0), (0,1)$ and $(1,0)$ do not map to collinear points. Suppose that $f_A(x_1, y_1) = f_A(x_2, y_2)$, so by 1(b) $x_1 f_A((1,0)) + y_1 f_A((0,1)) = x_2 f_A((1,0)) + y_2 f_A((0,1))$, then $(x_1 - x_2)f_A((1,0)) + (y_1 - y_2)f_A((0,1)) = (0,0)$ so the parallelogram would be flat (degenerate) which it is not. Therefore f_A is injective, and by part(c), it is bijective.

(e) Suppose that f_A is surjective. Prove that $ad - bc \neq 0$.
Solution If $f_A((0,0) = (0,0)$, $f_A((1,0)$ and $f_A((0,1))$ are collinear, then they must all lie on a line through the origin $(0,0)$. For each $(x,y) \in \mathbb{R}^2$, we then have $f_A((x,y)) = f_A(x(1,0) + y(0,1)) = xf_A((1,0)) + yf_A((0,1))$ is also on this line, but f_A is surjective, so this does not happen. Therefore the parallegram $(0,0) = f_A(0,0)), f_A((1,0)), f_A((1,1)), f_A((0,1))$ is not flat, so $ad - bc \neq 0$.

4. Prove that the following are equivalent. (a) f_A is injective, (b) f_A is surjective, (c) $ad - bc \neq 0$, (d) f_A is bijective.
Solution (a) $\Rightarrow$ (b) by 3(c), (b) $\Rightarrow$ (c) by 3(e), (c) $\Rightarrow$ (d) by 3(d), and (d) $\Rightarrow$ (a) is a formality.

5. Suppose that f_A is bijective, that $\begin{pmatrix} 1 \\ 0 \end{pmatrix} = A \begin{pmatrix} p \\ q \end{pmatrix}$, and that $\begin{pmatrix} 0 \\ 1 \end{pmatrix} = A \begin{pmatrix} r \\ s \end{pmatrix}$.

   (a) Determine p, q, r and s in terms of a, b, c and d.
   Solution It is helpful to introduce some notation: $\Delta = ad - bc$. By Problem 4, $\Delta \neq 0$. We have $ap + bq = 1$ and $cp + dq = 0$. Multiply the first equation by d, the second by b and subtract to discover that $\Delta p = d$. Instead multiply the first by c and the second by a and subtract to learn that $\Delta q = -c$. . A similar analysis applied to the second equation yields that $\Delta r = -b$ and $\Delta s = a$. Therefore

$$p = d/\Delta, q = -c/\Delta, r = -b/\Delta \text{ and } s = a/\Delta.$$

   (b) Deduce that if f_A is bijective, there is a unique matrix B such that $AB = I_2$, and moreover this matrix B is the unique matrix B such that $BA = I_2$.
   Solution Let $B = \begin{pmatrix} d/\Delta & -c/\Delta \\ -b/\Delta & a/\Delta \end{pmatrix}$ so

$$B \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix}$$

   and

$$B \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} r \\ s \end{pmatrix}$$

   and therefore

$$AB \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

   and

$$AB \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

   By problem 1(a), the maps f_A and f_B are mutually inverse so $AB = I_2$ (or by direct calculation). The mutual inverse property holds 'the other way round' too, and therefore $f_{BA} = f_B \circ f_A = \mathrm{Id}_{\mathbb{R}^2}$ so $BA = I_2$. The first equality holds because of Problem 2.

6. Suppose that ε and θ are positive real numbers. Let $e_1 = (1,0)$, $e_2 = (0,1)$, and $v \in \mathbb{R}^2$ be an arbitrary vector. Compare the areas of the rectangle with vertices $v, v + \varepsilon e_1, v + \varepsilon e_1 + \theta e_2, v + \theta e_2$, the region obtained by applying f_A to each point on, or in the interior of, this rectangle.
Solution This rectangle has area $\varepsilon\theta$. The image of this rectangle is a parallelogram with area $\varepsilon(a,c) \times \theta(b,d) = \varepsilon\theta(a\mathbf{i} + d\mathbf{j}) \times (b\mathbf{i} + d\mathbf{j}) = \varepsilon\theta(ad - bc)\mathbf{k}$. Thus the effect on a rectangle with sides aligned to the co-ordinate axes is to produce a parallelogram with signed area changed by a factor of $ad - bc$.

7. Show that if $S \subseteq \mathbb{R}^2$ is a region with area $|S|$, then the region obtained by applying f_A to each element of S has area $|ad - bc| \cdot |S|$. *The punctilious may confine themselves to regions S bounded by horizontal and vertical line segments, but students with a relaxed attitude to analysis should be able to see that the same result applies to any "reasonable" region of $\mathbb{R}^2$: to circular disks and to other regions bounded by calm and reassuringly dull curves.*

   Solution Assume that the boundary of S consists of finitely many line segments, each of which is either a horizontal or vertical line segment. The region may be decomposed into finitely many rectangles by producing all line segments in the diagram. If f_A is not bijective, then the image of f_A will be a subset of a line through the origin (see the proof, not the statement, of 3(e)), and so the image of S will be flat, and will have area 0.

   More interestingly, if f_A is bijective, then $ad - bc \neq 0$, and f maps the interiors of disjoint rectangles to the interiors of disjoint parallelograms by Problem 6. This process multiplies the signed area of each rectangle by $ad - bc$, again by Problem 6. Therefore the application of f_A has the effect of producing a region with (signed) area $ad - bc$ times the area of S.

   A more delicate and rigorous theory of area (measure theory) is needed make the glib statement "f_A multiplies area by $ad - bc$" into a precise and correct mathematical statement.

8. Suppose that U, V, W are distinct collinear points in the plane. Identify the plane with $\mathbb{R}^2$, and suppose that the co-ordinates of U, V and W are u, v and w, respectively. Suppose that A is a 2×2 matrix with real entries, and that the map f_A is bijective. Let the geometric points U', V' and W' have co-ordinates $f_A(u), f_A(v)$ and $f_A(w)$. Show that $UV : VW = U'V' : V'W'$, in other words that V divides the interval UW in the same ratio that V' divides the interval $U'W'$.

   Solution Recall that we have parameterization of the line UW as the set of points with co-ordinates $\lambda u + (1 - \lambda)w$ and f_A will carry a point with a particular value of λ to $\lambda f_A(u) + (1 - \lambda) f_A(w)$. However, the first statement means that V divides UW in the ratio $1 - \lambda : \lambda$, and the second asserts that V' divides $U'W'$ in the ratio $1 - \lambda : \lambda$ as required. Note that these ratios are negative when V is not in the interior of the interval UW, as discussed in the Comment before the solutions to Problem 3.

9. Suppose that XYZ is a triangle in the plane. Show that we may endow the plane with co-ordinates to identify it with $\mathbb{R}^2$, and produce a 2×2 matrix A such that when we apply f_A to the region bounded by the triangle, the result is a region bounded by an equilateral triangle.

   Solution The Comment at the start of the solutions to Problem 3 (alternatively the result of Problem 8) force any line segment to map to a line segment. Therefore the boundary of a triangle is carried to the boundary of a triangle. Now any point in the interior of a triangle is in the interior of a line segment joining two points on its boundary. Thus when a linear maps carries a triangle (the union of three line segments) to a triangle, then it must carry the points in the interior of the first triangle to (all) the points in the interior of the image triangle.

   As for the main problem, it has many solutions. We select a straightforward route. We choose co-ordinates and units so that Y is at $(0,0)$, Z is at $(1,0)$ and X is at (v,w). We choose A so that f_A fixes $(1,0)$ and sends (v,w) to $(\frac{1}{2},\frac{\sqrt{3}}{2})$. We need to find a,b,c and d such that

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

and

$$\begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix}.$$

Multiplying out, and noticing that $w \neq 0$ we can use linear algebra to find the unique solution: $a = 1$, $d = \sqrt{3}/(2w)$, $c = 0$ and $b = (1/2 - v)/w$.

This works, and only the matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ does the specified job.

10. (Tutor pacifier) Let $ABCDEF$ be a convex hexagon whch has parallel opposite sides and area 1. The lines AB, CD and EF meet in pairs to determine the vertices of a triangle. Similarly, the lines BC, DE and FA meet in pairs to determine the vertices of another triangle. Show that the area of at least one of these two triangles is at least $3/2$. *An classical proof is possible, but you may find some of the ideas developed in this problem sheet will assist you to construct a conceptual proof (i.e. a proof consisting of a sequence of ideas, rather than algebraic or trigonometric manipulation).*

Solution Not yet, but here is a solution to Problem 10 of Sheet 6, which first appeared in the Anglo-French Canadian Problems magazine *Crux Mathematicorum* in 2000.

"A mathematical tree (i.e. a vertical unit interval) grows at each point of an infinite plane with integral co-ordinates except for the origin $(0,0)$ where an observer, of height 1, stands. Many trees are visible, including those at $(1,0),(7,8)$ and $(45,-7)$. Other trees are invisible, because the view of them from the origin is obstructed by other trees. For example, the view of the tree at $(-14,91)$ is obstructed by the tree at $(-2,13)$.

Show that it is possible for a *Tunguska event* of diameter 10^{10} to happen, yet be unknown to the observer. In other words, show that there is a circle in the plane of diameter 10^{10} which has only invisible trees in its interior."

This solution is based on the Chinese Remainder Theorem. A tree at (k,l) is visible from the origin if, and only if, $\gcd(k,l) = 1$. Let $n = 10^{10} + 1$. By Problem 9 of sheet 5, there are positive integers $a_1,\ldots,a_n$ and $b_1,\ldots,b_n$ with the a_i pairwise coprime, the b_j pairwise coprime, but each a_i and b_j not coprime. Use the CRT to find a positive integer u such that $u \equiv -i \bmod a_i$ for each i. You use CRT to find an integer of the required form, then add a positive multiple of the product of the a_i so that you get a positive solution. Similarly you use the CRT find a positive integer v such that $v \equiv -j \bmod b_j$ for each j. Now for each pair i,j of indices in the range 1 to n, we have a_i divides $u+j$ and b_j divides $v+j$. Let g_{ij} denote the gcd of a_i and b_j, so, by design, each g_{ij} is bigger than 1. The tree at $(u+i,v+j)$ is therefore invisible from the origin for each i and j in the given range. We have found a square of invisible trees of side 10^{10}. Each tree is obstructed from view by a visible tree. We can put the centre of the Tunguska event at the centre of this square, and only invisible trees will be destroyed, and none of the visible "blocking" trees will be destroyed, so the observer will be unaware of the Tunguska event.