HONOURS ALGEBRA
WORKSHOP 1 SOLUTIONS

The purpose of this workshop is to revise some notions from Introduction to Linear Algebra, or its accelerated entry equivalent, such as bases, dimension, subspaces. But it may all be from a slightly different viewpoint.

Exercise 1 (Discuss this with a partner)**.** In this exercise you should use your intuition about space around us, but instead of considering the ground field to be $F = \mathbb{R}$ we will take $F = \mathbb{F}_3$, the field with three elements.

(1) $\mathbb{F}_3$ is the set of integers modulo 3: it has a multiplication and addition defined by modular arithmetic. I'll write its elements as $\{\overline{0}, \overline{1}, \overline{2}\}$. It is a field: recall from the first lecture this means in particular that for each non-zero $\lambda \in \mathbb{F}_3$ there exists a (necessarily unique) $\lambda^{-1} \in \mathbb{F}_3$ such that $\lambda \cdot \lambda^{-1} = 1$. The questions is: what is λ^{-1} when $\lambda = \overline{1}$ and $\lambda = \overline{2}$?

(2) (a) Write out all the elements of the vector space $\mathbb{F}_3^2$ as column vectors.
(b) Find all the one-dimensional subspaces of $\mathbb{F}_3^2$. How many different bases does each of these subspaces have?
(c) Given a non-zero vector $\vec{v}_1 \in \mathbb{F}_3^2$, how many vectors $\vec{v}_2 \in \mathbb{F}_3^2$ are there such that $\vec{v}_1, \vec{v}_2$ is a linearly independent family?
(d) Count the number of indexed bases of $\mathbb{F}_3^2$. Forgetting indexing, how many bases are there?

(3) Now let V be an arbitrary two-dimensional vector space over $\mathbb{F}_3$. How many indexed bases are there for V?

(4) (a) Let $n \in \mathbb{N}$ with $n \geq 2$. How many non-zero vectors are there in $\mathbb{F}_3^n$?
(b) How many one-dimensional subspaces of $\mathbb{F}_3^n$ are there? Check it agrees with your answer for 2(d) when $n = 2$!
(c) How many two-dimensional subspaces of $\mathbb{F}_3^n$ are there? Check you get the correct answer when $n = 2$!

(5) Now replace the ground field $\mathbb{F}_3$ by $\mathbb{R}$. Can you still answer all the above questions? If not, why not? And could you retrieve some kind of constructive answer?

Solution 1. (1) We have that $\overline{1}^{-1} = \overline{1}$, and that $\overline{2}^{-1} = \overline{2}$, since the fact that $4 \equiv 1 \pmod 3$ implies that $\overline{2} \cdot \overline{2} = \overline{4} = \overline{1}$.

(2) (a) The elements are

$$\begin{pmatrix}\overline{0}\\ \overline{0}\end{pmatrix}, \begin{pmatrix}\overline{1}\\ \overline{0}\end{pmatrix}, \begin{pmatrix}\overline{2}\\ \overline{0}\end{pmatrix}, \begin{pmatrix}\overline{0}\\ \overline{1}\end{pmatrix}, \begin{pmatrix}\overline{1}\\ \overline{1}\end{pmatrix}, \begin{pmatrix}\overline{2}\\ \overline{1}\end{pmatrix}, \begin{pmatrix}\overline{0}\\ \overline{2}\end{pmatrix}, \begin{pmatrix}\overline{1}\\ \overline{2}\end{pmatrix}, \begin{pmatrix}\overline{2}\\ \overline{2}\end{pmatrix}.$$

(b) The one-dimensional subspaces are

$$\left\{\begin{pmatrix}a\\ \overline{0}\end{pmatrix} : a \in \mathbb{F}_3\right\}, \left\{\begin{pmatrix}\overline{0}\\ a\end{pmatrix} : a \in \mathbb{F}_3\right\}, \left\{\begin{pmatrix}a\\ a\end{pmatrix} : a \in \mathbb{F}_3\right\}, \text{ and } \left\{\begin{pmatrix}a\\ 2a\end{pmatrix} : a \in \mathbb{F}_3\right\}.$$

Each has two bases, e.g. the first subspace could have basis element $\begin{pmatrix}\overline{1}\\ \overline{0}\end{pmatrix}$ or $\begin{pmatrix}\overline{2}\\ \overline{0}\end{pmatrix}$.

(c) Given a non-zero $\vec{v}_1$, the set $\{\vec{v}_1, \vec{v}_2\}$ is linearly independent if and only if $\vec{v}_2$ is not a multiple of $\vec{v}_1$, i.e. if and only if it is not one of the three multiples of $\vec{v}_1$. E.g. if

Date: Thursday, 15th January, 2015.

$\vec{v}_1 = (\overline{1}, \overline{0})^\mathsf{T}$, then $\vec{v}_2$ cannot be $(\overline{0}, \overline{0})^\mathsf{T}$, $(\overline{1}, \overline{0})^\mathsf{T}$, or $(\overline{2}, \overline{0})^\mathsf{T}$. Thus for each vector $\vec{v}_1$ there are $9 - 3 = 6$ choices of $\vec{v}_2$ such that $\{\vec{v}_1, \vec{v}_2\}$ is linearly independent.

(d) A basis is precisely obtained by taking a non-zero vector $\vec{v}_1$ and choosing a vector $\vec{v}_2$ such that $\{\vec{v}_1, \vec{v}_2\}$ is linearly independent. We have $9 - 1 = 8$ choices for our vector $\vec{v}_1$, and for each one, as argued above, 6 choices for $\vec{v}_2$. Thus in total there are $8 \cdot 6 = 48$ indexed (i.e. ordered) bases.

Counting bases like this we count each set of vectors (i.e. unordered) twice. Thus the number of unindexed (i.e. forgetting the ordering) bases is $48/2 = 24$.

(3) Any two-dimensional vector space over $\mathbb{F}_3$ is isomorphic to $\mathbb{F}_3^2$, so the answers are the same as the previous part.

(4) (a) We have 3 choices for n components in each vector, so the total size of the space is 3^n. Therefore we have $3^n - 1$ non-zero vectors.

(b) Choose a non-zero vector. Then it determines a one-dimensional subspace. The only other vector which will generate the same subspace is the one non-zero non-identical multiple of it. Thus counting this way we count each subspace twice. So the total number of one-dimensional subspaces is half the number of non-zero vectors, i.e. $(3^n - 1)/2$. This gives $(9 - 1)/2 = 4$ in the case $n = 2$, as we would hope.

(c) To get a two-dimensional space we choose a non-zero vector and then choose a second vector which is not a multiple of the first. There are $(3^n - 1)$ choices for the first non-zero vector, and given each there are $(3^n - 3)$ choices for the second (i.e. all but the three multiples of our first vector). But we have previously calculated that each two-dimensional subspace has 48 ordered bases. I.e. following this method of counting will count each two-dimensional subspace 48 times. Thus we have in total $(3^n - 1)(3^n - 3)/48$ two-dimensional subspaces. For $n = 2$, we get $(9 - 1)(9 - 3)/48 = 8 \cdot 6/48 = 1$ as we would hope.

(5) We can no longer count so easily since $\mathbb{R}$ is infinite. But one could still imagine that the number of linearly independent sets of size 2 is in some sense a copy of $\mathbb{R} \setminus \{0\}$ times a copy of $\mathbb{R}^2 \setminus \mathbb{R}$. These considerations then begin to move into more geometric questions.

Exercise 2 (Group Work)**.** Don't take more than 15 minutes doing this: it is just a warm-up, to check you remember gaussian elimination and that you understood the point made in the first lecture about it.

Consider the three systems of linear equations:

$$\begin{aligned} x_1 + 2x_2 + x_3 + 4x_4 + 5x_5 &= A \\ x_1 + 2x_2 + 2x_3 + 6x_4 + 8x_5 &= B \\ 2x_1 + 4x_2 + 3x_3 + 11x_4 + 15x_5 &= C \end{aligned}$$

$$\begin{aligned} x_1 + 2x_2 + x_3 + 4x_4 + 5x_5 &= A \\ 2x_1 + 4x_2 + 3x_3 + 11x_4 + 15x_5 &= B \\ x_1 + 2x_2 + 2x_3 + 6x_4 + 8x_5 &= C \end{aligned}$$

$$\begin{aligned} x_1 + \ 4x_2 + \ x_3 + \ 2x_4 + \ 5x_5 &= A \\ 2x_1 + \ 6x_2 + \ x_3 + \ 2x_4 + \ 8x_5 &= B \\ 3x_1 + 11x_2 + 2x_3 + \ 4x_4 + 15x_5 &= C \end{aligned}$$

(1) Represent each by a suitable augmented coefficient matrix.
(2) Put each into echelon form. Compare the stair patterns you get. Do they look the same to you? Do they have anything in common?
(3) Now solve each of them, and write out the solutions as concisely as you can. How do the solutions compare?
(4) Let $A = B = C = 0$ so that the solution set is subspace of F^5. What is the dimension of the solution space?
(5) Let A, B, C be arbitrary. Write out the general solution by using the previous part and a specific solution.

Solution 2. (1) The associated coefficient matrices are, respectively,

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 4 & 5 & A \\ 1 & 2 & 2 & 6 & 8 & B \\ 2 & 4 & 3 & 11 & 15 & C \end{array}\right), \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 4 & 5 & A \\ 2 & 4 & 3 & 11 & 15 & B \\ 1 & 2 & 2 & 6 & 8 & C \end{array}\right), \text{ and } \left(\begin{array}{ccccc|c} 1 & 4 & 1 & 2 & 5 & A \\ 2 & 6 & 1 & 2 & 8 & B \\ 3 & 11 & 2 & 4 & 15 & C \end{array}\right).$$

(2) Performing Gaussian elimination on each matrix gives the following echelon form of each matrix respectively:

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 4 & 5 & A \\ 0 & 0 & 1 & 2 & 3 & B-A \\ 0 & 0 & 0 & 1 & 2 & C-A-B \end{array}\right),$$

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 4 & 5 & A \\ 0 & 0 & 1 & 3 & 5 & B-2A \\ 0 & 0 & 0 & -1 & -2 & C+A-B \end{array}\right), \text{ and}$$

$$\left(\begin{array}{ccccc|c} 1 & 4 & 1 & 2 & 5 & A \\ 0 & -2 & -1 & -2 & -2 & B-2A \\ 0 & 0 & -1/2 & -1 & 1 & C-2A-B/2 \end{array}\right).$$

The stair pattern of the third is different, but the length of the stairs is the same in each case: $(2+1+2 = 1+1+3)$.

(3) Reading off the equations represented by the reduced echelon form of the first system gives us:

$$\begin{aligned} x_4 &= C - A - B - 2x_5, \\ x_3 &= B - A - 2x_4 - 3x_5 = B - A - 2C + 2A + 2B + 4x_5 - 3x_5 = A + 3B - 2C + x_5, \text{ and} \\ x_1 &= A - 2x_2 - x_3 - 4x_4 - 5x_5 \\ &= A - 2x_2 - A - 3B + 2C - x_5 - 4C + 4A + 4B + 8x_5 - 5x_5 \\ &= 4A + B - 2C - 2x_2 + 2x_5. \end{aligned}$$

Setting $A = B = C = 0$ gives us the solution $(-2x_2+2x_5, x_2, x_5, -2x_5, x_5)$, which describes a two-dimensional subspace. The general solution is therefore

$$(4A + B - 2C, 0, A + 3B - 2C, C - A - B, 0) + (-2x_2 + 2x_5, x_2, x_5, -2x_5, x_5).$$

Similarly for the next system, we obtain

$$\begin{aligned}
-x_4 &= C + A - B + 2x_5, \\
x_3 &= B - 2A - 3x_4 - 5x_5 = B - 2A + 3C + 3A - 3B + 6x_5 - 5x_5 \\
&= A - 2B + 3C + x_5, \text{ and} \\
x_1 &= A - 2x_2 - x_3 - 4x_4 - 5x_5 \\
&= A - 2x_2 - A + 2B - 3C - x_5 + 4C + 4A - 4B + 8x_5 - 5x_5 \\
&= 4A - 2B + C - 2x_2 + 2x_5.
\end{aligned}$$

Setting $A = B = C = 0$ gives the solution $(-2x_2 + 2x_5, x_2, x_5, -2x_5, x_5)$, which describes the same two-dimensional space, and the general solution is therefore

$$(4A - 2B + C, 0, A - 2B + 3C, B - C - A, 0) + (-2x_2 + 2x_5, x_2, x_5, -2x_5, x_5).$$

For the last system, we obtain

$$\begin{aligned}
-x_3/2 &= C - 2A - B/2 + x_4 - x_5, \\
-2x_2 &= B - 2A + x_3 + 2x_4 + 2x_5 = B - 2A - 2C + 4A + B - 2x_4 + 2x_5 + 2x_4 + 2x_5 \\
&= 2B + 2A - 2C + 4x_5, \text{ and} \\
x_1 &= A - 4x_2 - x_3 - 2x_4 - 5x_5 \\
&= A + 4B + 4A - 4C + 8x_5 + 2C - 4A - B + 2x_4 - 2x_5 - 2x_4 - 5x_5 \\
&= A + 3B - 2C + x_5.
\end{aligned}$$

Setting $A = B = C = 0$ gives the solution $(x_5, -2x_5, -2x_4 + 2x_5, x_4, x_5)$, which describes a two-dimensional subspace, and the general solution is

$$(A + 3B - 2C, C - A - B, 4A + B - 2C, 0, 0) + (x_5, -2x_5, -2x_4 + 2x_5, x_4, x_5).$$

Exercise 3. A Challenge! Let $V = \mathbb{F}_3^n$. The general linear group $GL(V)$ of automorphisms $g : V \xrightarrow{\sim} V$ is defined on page 18 of the lecture notes. It is a finite group in this example, since V is a finite set. What is its order? [Hint 1: could you do this for $n = 2$? Hint 2: any automorphism $g : V \xrightarrow{\sim} V$ must send one basis to another.]

Solution 3. A challenge!

HONOURS ALGEBRA
WORKSHOP 2 SOLUTIONS

The purpose of this workshop is to study examples of linear (in)dependence, spanning sets and bases, particularly for vector spaces you may not have used much before.

Exercise 1 (Group Work)**.** Split up into pairs at your table. Each pair take two of these problems so that their sum is 7. Answer these questions, with justification. Then, when you are all done, explain your answers to the others at the table.

(1) Let $V = \{f : \mathbb{R} \to \mathbb{R}\}$ be the vector space of functions that we have discussed in lectures.
   (a) Is the set $\{\cos(x), \sin(x), e^x\}$ linearly independent?
   (b) Is the set $\{\cos^2(x), \sin^2(x), 1\}$ linearly independent?

(2) Let $V = \{f : \mathbb{C} \to \mathbb{C}\}$ be the vector space of functions that we have discussed in lectures.
   (a) Is the set $\{\cos(z), \sin(z), e^z\}$ linearly independent?
   (b) Is the set $\{\cos(z), \sin(z), e^{\sqrt{-1}z}\}$ linearly independent?

(3) Let $S = \{\vec{u}_1, ..., \vec{u}_n\}$ and $T = \{\vec{u}_1, ..., \vec{u}_n, \vec{u}_{n+1}\}$ be subsets of a vector space V. Consider the statements:

   (A) If S is linearly independent then T is linearly independent
   (B) If T is linearly independent then S is linearly independent

   Which do you agree with?
   (a) A true, B false; (b) B true, A false; (c) Both true; (d) Both false.

(4) Let

$$S = \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \end{pmatrix} \right\} \subset F^4.$$

   Is S linearly independent?
   (a) Yes; (b) No; (c) Impossible to say.

(5) Let $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ belong to a vector space V. Suppose that

$$\langle \vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4 \rangle = \langle \vec{v}_1, \vec{v}_2, \vec{v}_3 \rangle.$$

   Which the following statements are necessarily true?
   (a) $\vec{v}_4 = \vec{0}$.
   (b) $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is a linearly independent subset of V.
   (c) $\vec{v}_4$ is an element of $\langle \vec{v}_1, \vec{v}_2, \vec{v}_3 \rangle$.
   (d) $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is a linearly dependent subset of V.

(6) Let $F = \mathbb{F}_3$, the field with three elements. Let

$$S = \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \end{pmatrix} \right\} \subset F^4.$$

   What is the dimension of the space spanned by S?

Solution 1. (1) (a) The set $\{\cos x, \sin x, e^x\}$ is linearly independent. Suppose

$$\alpha \cos x + \beta \sin x + \gamma e^x = 0,$$

for some α, β, and γ, i.e. holds for all $x \in \mathbb{R}$. Then evaluating this equation at $x = 0$ implies that $\alpha + \gamma = 0$, and furthermore evaluating at $x = \pi$ implies that $-\alpha + \gamma e^\pi = 0$, i.e. $\gamma(1 + e^\pi) = 0$, and thus $\gamma = 0$ and hence also $\alpha = 0$. Finally evaluating at $x = \pi/2$ implies that $\beta + \gamma e^{\pi/2} = 0$, and hence that $\beta = 0$.

(b) The set $\{\cos^2 x, \sin^2 x, 1\}$ is not linearly independent, since $\cos^2 x + \sin^2 x - 1 = 0$ for all x.

(2) (a) The set $\{\cos z, \sin z, e^z\}$ is linearly independent, by the same argument as for 1(a).

(b) The set $\{\cos z, \sin z, e^{\sqrt{-1}z}\}$ is not linearly independent, since $e^{\sqrt{-1}z} - \cos z - \sqrt{-1}\sin z = 0$ for all z.

(3) (b) is correct. To see that (A) is false, just consider $\vec{u}_{n+1} = \vec{0}$. Then T is linearly dependent whether S is or not, since T contains the zero vector.

To see that (B) is true, suppose that $T = \{\vec{u}_1, \ldots, \vec{u}_n, \vec{u}_{n+1}\}$ is linearly independent, and suppose that for some $\alpha_1, \ldots, \alpha_n \in F$ we have that

$$\alpha_1 \vec{u}_1 + \cdots + \alpha_n \vec{u}_n = \vec{0}.$$

Then we have in fact that

$$\alpha_1 \vec{u}_1 + \cdots + \alpha_n \vec{u}_n + 0\vec{u}_{n+1} = \vec{0},$$

from which the linear independence of T implies that $\alpha_1 = \cdots = \alpha_n = 0$, as required.

(4) (c) is correct: the situation is dependent on the underlying field. Suppose that $F = \mathbb{F}_3$. Then the set is linearly dependent, since

$$\begin{pmatrix}1\\2\\0\\2\end{pmatrix} + \begin{pmatrix}1\\0\\1\\2\end{pmatrix} + 2\begin{pmatrix}2\\2\\1\\1\end{pmatrix} = \begin{pmatrix}6\\6\\3\\6\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix}.$$

However, if $F = \mathbb{F}_2$, then the set is linearly independent, since

$$\begin{pmatrix}1\\2\\0\\2\end{pmatrix} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad \begin{pmatrix}1\\0\\1\\2\end{pmatrix} = \begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \text{ and } \quad \begin{pmatrix}2\\2\\1\\1\end{pmatrix} = \begin{pmatrix}0\\0\\1\\1\end{pmatrix}.$$

Then

$$\begin{pmatrix}0\\0\\0\\0\end{pmatrix} = \alpha\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + \beta\begin{pmatrix}1\\0\\1\\0\end{pmatrix} + \gamma\begin{pmatrix}0\\0\\1\\1\end{pmatrix} = \begin{pmatrix}\alpha+\beta\\0\\\beta+\gamma\\\gamma\end{pmatrix}$$

implies that $\alpha = \beta = \gamma = 0$.

(5) Only (c) is true. That (c) is indeed true follows since $\vec{v}_4 \in \langle \vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4 \rangle = \langle \vec{v}_1, \vec{v}_2, \vec{v}_3 \rangle$. Taking $V = \mathbb{R}^3$ and $\vec{v}_1 = (1,0,0)^\mathsf{T}$, $\vec{v}_2 = (0,1,0)^\mathsf{T}$, $\vec{v}_3 = (0,0,1)^\mathsf{T}$, and $\vec{v}_4 = (1,1,1)^\mathsf{T}$ we get counterexamples to (a) and (d). Taking $\vec{v}_1 = \vec{v}_2 = \vec{v}_3 = \vec{v}_4 = (1,0,0)^\mathsf{T}$, we get a counterexample to (a) and (b).

(6) The space is two-dimensional. We saw in (4) that the set is linearly dependent, and in fact we have

$$\begin{pmatrix}1\\2\\0\\2\end{pmatrix} = 2\begin{pmatrix}1\\0\\1\\2\end{pmatrix} + \begin{pmatrix}2\\2\\1\\1\end{pmatrix}.$$

Therefore

$$\langle S\rangle = \left\langle \begin{pmatrix}1\\0\\1\\2\end{pmatrix}, \begin{pmatrix}2\\2\\1\\1\end{pmatrix} \right\rangle.$$

This is a two-dimensional space since $(1,0,1,2)^\mathsf{T}$ and $(2,2,1,1)^\mathsf{T}$ are not multiples of each other.

Exercise 2 (Work on this in groups of more than one!)**.** Recall the vector space $\mathrm{Mat}(m\times n; F)$ from the end of Chapter 1, Section 2.

(1) What is the dimension of $\mathrm{Mat}(m\times n; F)$?
(2) Find a basis of this vector space.
(3) Let $p(z) \in F[z]$ be a polynomial whose coefficients belong to F. Given $A \in \mathrm{Mat}(n; F)$, let $p(A) \in \mathrm{Mat}(n; F)$ be the matrix you get by replacing each power of z in $p(z)$ by the corresponding power of A, i.e. if $p(z) = z^2 - z + 6$ then $p(A) = A^2 - A + 6I$. Show that: there exists a non-zero polynomial $p(z)$ such that $p(A)$ is the zero matrix.
(4) Let

$$A = \begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}.$$

Find an explicit non-zero polynomial $p(z)$ for which $p(A)$ is the zero matrix.
(5) Here is a fact, which you don't need to check,

$$B := \frac{1}{12}\begin{pmatrix}32&-12&8\\16&12&-8\\13&-15&28\end{pmatrix} = \begin{pmatrix}1&2&3\\3&2&1\\2&1&3\end{pmatrix}\cdot\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}\cdot\begin{pmatrix}1&2&3\\3&2&1\\2&1&3\end{pmatrix}^{-1}.$$

Find a non-zero polynomial $p(z)$ for which $p(B)$ is the zero matrix.

Solution 2. (1) It is mn-dimensional (see below).
(2) Let E_{ij} be the $(m\times n)$-matrix with a 1 in the (i,j)-th entry, and 0 elsewhere. Then we shall see that

$$\mathcal{B} = \{E_{ij} : 1 \le i \le m,\ 1 \le j \le n\}$$

is a basis. An arbitrary $(m\times n)$ matrix has the form

$$\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix} = \sum_{i,j=1}^{m,n} a_{ij}E_{ij},$$

so $\mathcal{B}$ spans the space. Furthermore if $\sum_{i,j=1}^{m,n} a_{ij}E_{ij} = \vec{0}$, then each $a_{ij} = 0$, so $\mathcal{B}$ is linearly independent.

(3) Consider the set $\{I, A, A^2, \ldots, A^N\}$, where $N = n^2$. Then this set has $n^2 + 1$ elements, all of which belong to $\mathrm{Mat}(n; F)$. Since this space is n^2-dimensional, this set must be linearly dependent. Thus there exist $\lambda_i \in F$, not all zero, such that

$$\lambda_0 A^0 + \lambda A_1 + \lambda_2 A^2 + \cdots + \lambda_N A^N = \vec{0}.$$

So the polynomial $p(z) = \lambda_0 + \lambda_1 z + \lambda_2 z^2 + \cdots + \lambda_N z^N$ is as required.

(4) From the previous part we know that we can find a polynomial of degree 9 satisfying the requirement, but we can do better. Since

$$A - I = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \; A - 2I = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ and } A - 3I = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

if we define $p(z) = (z-1)(z-2)(z-3)$, we see that since the product of the above three matrices is the zero matrix, we in fact have that $p(A) = (A - I)(A - 2I)(A - 3I) = \vec{0}$ as required.

(5) The point here is that B is much too complicated a matrix for us to calculate directly with, so instead we use the fact we are given, that there is a matrix Q such that $B = QAQ^{-1}$. Then we observe that, exploiting the associativity of matrix multiplication,

$$\begin{aligned}(QAQ^{-1})^n = (QAQ^{-1})(QAQ^{-1})\cdots(QAQ^{-1}) &= QA(Q^{-1}Q)A(Q^{-1}Q)\cdots(Q^{-1}Q)AQ^{-1} \\ &= QA\cdots AQ^{-1} \\ &= QA^nQ^{-1},\end{aligned}$$

and applying this to all appropriate exponents we see that for any polynomial $p(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_t z^t$, we have that

$$\begin{aligned}p(B) &= \alpha_0 + \alpha_1 B + \alpha_2 B^2 + \cdots + \alpha_t B^t \\ &= \alpha_0 + \alpha_1 QAQ^{-1} + \alpha_2 (QAQ^{-1})^2 + \cdots + \alpha_t (QAQ^{-1})^n \\ &= \alpha_0 + \alpha_1 QAQ^{-1} + \alpha_2 QA^2Q^{-1} + \cdots + \alpha_t QA^tQ^{-1} \\ &= Q\left(\alpha_0 + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_t A^t\right)Q^{-1} \\ &= Qp(A)Q^{-1}.\end{aligned}$$

So we see that if in fact $p(z) = (z-1)(z-2)(z-3)$, as before, we have that

$$p(B) = Qp(A)Q^{-1} = Q0Q^{-1} = 0.$$

HONOURS ALGEBRA
WORKSHOP 3 SOLUTIONS

The purpose of this workshop is to study how to describe linear mappings via different bases, and then to see that abstract algebra has utilitarian value beyond its natural beauty.

Exercise 1. First warm up. Do this in pairs: every pair should do (1) and then distribute (2), (3) and (4) amongst your pairs at the table.

The linear mapping $f : \mathbb{R}^3 \to \mathbb{R}^2$ is defined by

$$f(x_1, x_2, x_3) = (x_1 - x_2 + 2x_3, x_1 - x_3).$$

In $\mathbb{R}^2$, $\mathcal{A}$ is the basis $((1,1),(1,-1))$ and in $\mathbb{R}^3$, $\mathcal{B}$ is the basis $((1,1,0),(0,1,1),(1,0,1))$. Obtain:

(1) the matrix of f with respect to the standard bases of $\mathbb{R}^3$ and $\mathbb{R}^2$;
(2) the matrix of f with respect to the standard basis of $\mathbb{R}^3$ and the basis $\mathcal{A}$ of $\mathbb{R}^2$;
(3) the matrix of f with respect to the basis $\mathcal{B}$ of $\mathbb{R}^3$ and the standard basis of $\mathbb{R}^2$;
(4) the matrix of f with respect to the basis $\mathcal{B}$ of $\mathbb{R}^3$ and the basis $\mathcal{A}$ of $\mathbb{R}^2$.

And then stretch. I'd suggest you all work together now, combining your calculation skills with your geometric intuition for $\mathbb{R}^3$!

(5) Convince yourself (don't prove!) every rotation in $\mathbb{R}^3$ is determined by an **axis of rotation** and an **angle of rotation**. (Did you use the "right-hand rule"?)
(6) Show that: if the axis of rotation is the x-axis and you rotate by θ degrees, the matrix representing this linear transformation in standard coordinates is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

(7) Now prove, by a suitable of change of basis, my claim: with the standard basis, there is rotation in $\mathbb{R}^3$ with axis of rotation the line through the $\vec{0}$ and $(1,1,1)$, represented by

$$\begin{pmatrix} \frac{1+\sqrt{3}}{3} & \frac{1-\sqrt{3}}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1+\sqrt{3}}{3} & \frac{1-\sqrt{3}}{3} \\ \frac{1-\sqrt{3}}{3} & \frac{1}{3} & \frac{1+\sqrt{3}}{3} \end{pmatrix}$$

Can you tell me what the angle of rotation is for it?
(If you want a hint, go to the bottom of the next page!)

Solution 1. (1)

$$\begin{pmatrix} 1 & -1 & 2 \\ 1 & 0 & -1 \end{pmatrix}$$

(2)

$$\begin{pmatrix} 1 & -1/2 & 1/2 \\ 0 & -1/2 & 3/2 \end{pmatrix}$$

(3)

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

Date: Thursday, 29th January, 2015.

(4)
$$\begin{pmatrix} 1/2 & 0 & 3/2 \\ -1/2 & 1 & 3/2 \end{pmatrix}$$

(5) It's intuitively clear. Pick a non-zero vector, and draw the line through the origin and your chosen vector to get an axis of rotation. Then rotate around this axis by the angle of rotation (clockwise, where clockwise is determined by the right-hand rule: stick your thumb in the direction of your vector and then rotate in the direction in which your middle finger points). Done rigorously, this is "Euler's Rotation Theorem."

(6) The x-axis is fixed, so we can think of rotation simply as happening on the (y, z)-plane, and the bottom-right square matrix then describes that rotation.

(7) Take the orthonormal basis proposed, viz $\mathcal{B} = \left(\frac{1}{\sqrt{3}}(1,1,1)^\mathsf{T}, \frac{1}{\sqrt{6}}(1,1,-2)^\mathsf{T}, \frac{1}{\sqrt{2}}(1,-1,0)^\mathsf{T}\right)$. We compare this with the x, y, and z as in (6). The matrix in question is ${}_{\mathcal{A}}[f]_{\mathcal{A}}$ for some $f\colon \mathbb{R}^3 \to \mathbb{R}^3$, where $\mathcal{A}$ is the standard basis $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$. Now I write it instead with respect to the basis $\mathcal{B}$, using the rule

$$ {}_{\mathcal{B}}[f]_{\mathcal{B}} = {}_{\mathcal{B}}[\mathrm{id}_{\mathbb{R}^3}]_{\mathcal{A}} \cdot {}_{\mathcal{A}}[f]_{\mathcal{A}} \cdot {}_{\mathcal{A}}[\mathrm{id}_{\mathbb{R}^3}]_{\mathcal{B}}. $$

Here ${}_{\mathcal{B}}[\mathrm{id}_{\mathbb{R}^3}]_{\mathcal{A}}$ $\left(= {}_{\mathcal{A}}[\mathrm{id}_{\mathbb{R}^3}]_{\mathcal{B}}^{-1}\right)$ and ${}_{\mathcal{A}}[\mathrm{id}_{\mathbb{R}^3}]_{\mathcal{B}}$ are the change of basis matrices. It is easy to see that

$$ {}_{\mathcal{A}}[\mathrm{id}_{\mathbb{R}^3}]_{\mathcal{B}} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{pmatrix}, $$

and I can also get (with thought or direct calculation) that

$$ {}_{\mathcal{A}}[\mathrm{id}_{\mathbb{R}^3}]_{\mathcal{B}}^{-1} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix}. $$

Now I calculate fo find

$$ {}_{\mathcal{B}}[f]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}, $$

and I deduce that I have a rotation by $\pi/6$ about the axis determined by $(1,1,1)^\mathsf{T}$.

Exercise 2 (Public Key Cryptography)**.** You need to work in groups of 3 for this: call yourself Alice, Bob and Eve. Alice and Bob want to communicate without Eve knowing, but Eve wants to find out what's going on. So whoever is Eve will have to play nasty.

- Eve picks a prime number p and a second number g and tells everyone (this number should have a special property; I'll discuss that later - for now pick a prime smaller than p). For instance $p = 29$ and $g = 2$.
- Alice and Bob pick numbers secretly for themselves: say a and b. I'll use 10 and 17.
- Alice calculates g^a modulo p and Bob calculates g^b modulo p. They tell everybody the answer of that calculation. In my example Alice gets 9 and Bob 21.
- Alice takes Bob's number, raises it to the power a and calculates that modulo p; Bob takes Alice's number and raises it to the power b and calculates modulo p. They don't tell anyone the outcome of this calculation, but they do get the same number. (Why?) In the example it is 4.

So Alice and Bob now have same piece of information - the number 4. But Eve doesn't have it! They can use that number now for encoding their messages and Eve is stuck.

Magic? Well, play the game and see what happens to Eve. Eve will find the only way she can work out the final number is to work backwards to get a (or b) from the outcome of the third step, i.e. by solving the problem: for which a is $2^a \equiv 9$ modulo 29. Doing this is called the **discrete logarithm problem**; there is no known efficient algorithm to solve it on a computer. Taking big prime numbers, of which there are infinitely many, then makes Alice and Bob's final number practically secret from Eve.

For this method to be effective, however, we are relying on a specific fact about finite fields whose context you will encounter within the next week in a more general setting. It is:

the non-zero elements of a finite field form a cyclic group under multiplication

That they form a group is trivial; it is the cyclic assertion that has content. You'll prove a special case now:

Let p be a prime number and let $\mathbb{F}_p = \{\overline{0}, \overline{1}, \ldots, \overline{p-1}\}$ be the set of integers modulo p. Just as in Workshop 1, where we used $\mathbb{F}_3$, this is a field and it has p elements. Your job is to prove that $\mathbb{F}_p^\times = \{\overline{1}, \ldots, \overline{p-1}\}$ is a cyclic group.

(1) Check first of all this is true for $\mathbb{F}_3, \mathbb{F}_5$ and $\mathbb{F}_7$.
(2) Let $N_p(d)$ denote the number of elements in $\mathbb{F}_p^\times$ that are of order d. Why is true that
$$\sum_{d|p-1} N_p(d) = p - 1?$$
(3) I need to remind you of Euler's function: $\phi : \mathbb{N} \to \mathbb{N}$ defined by $\phi(n)$ being the number of integers k in the range $1 \leq k \leq n$ for which $\gcd(k, n) = 1$. For instance $\phi(9) = 6$. Show that: if $d|p-1$ and $N_p(d) > 0$ then $N_p(d) = \phi(d)$. (It might help you to know that the polynomial $X^d - 1$ has at most d roots in $\mathbb{F}_p$: we'll a big generalization of that soon in lectures.)
(4) Let $n \in \mathbb{N}_{\geq 1}$. By counting how to express each of the fractions
$$\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, \frac{n}{n}$$
in reduced form, show that $\sum_{d|n} \phi(d) = n$.
(5) Deduce that $\mathbb{F}_p^\times$ is cyclic.

Solution 2. (1) $\mathbb{F}_3^\times = \{\overline{1}, \overline{2}\} = \langle\overline{2}\rangle$. $\mathbb{F}_5^\times = \langle\overline{2}\rangle$ since $\overline{1} = \overline{2}^0$, $\overline{2} = \overline{2}^1$, $\overline{3} = \overline{2}^3$, and $\overline{4} = \overline{2}^2$. $\mathbb{F}_7^\times = \langle\overline{3}\rangle$ (calculate it and see); it's not generated by $\overline{2}$ because $\overline{2}^3 = \overline{1}$.

(2) By Lagrange's theorem, each element of $\mathbb{F}_p^\times$ has order which divides the order of the group, i.e. divides $p - 1$. Hence
$$p - 1 = \left|\mathbb{F}_p^\times\right| = \sum_{d=1}^{p-1} N_p(d) = \sum_{d|p-1} N_p(d).$$

(3) Suppose that $a \in \mathbb{F}_p^\times$ has order d (so $N_p(d) > 0$). Then each of the d elements $1, a, a^2, \ldots, a^{d-1}$ satisfies the equation $X^d = \overline{1}$, and no two are equal since otherwise the order of a would be strictly less than d. But, by the hint, $X^d - \overline{1}$ can have at most d solutions in $\mathbb{F}_p$, and so it has exactly d solutions, which are those elements we just listed. Now, any element of $\mathbb{F}_p^\times$ having degree d must satisfy $X^d = \overline{1}$, and so must be of the form a^i for some i satisfying $0 \leq i \leq d - 1$. The order of a^i is, by last year's group theory, $d/\gcd(i, d)$. Thus the elements of order d are precisely those elements a^i with $\gcd(d, i) = 1$, proving the claim.

(4) We write i/n in reduced terms as a/d where $\gcd(a,d)=1$ and $1 \le a \le d$, where $d \mid n$. Thus the number of fractions is on the one hand n (this is evident from the list) and on the other hand is $\sum_{d|n} \phi(d)$.

(5) By (2) we have that $\sum_{d|p-1} N_p(d) = p-1$. By (4) we have that $\sum_{d|p-1} \phi(d) = p-1$. By (3) $N_p(d)$ is either 0 or $\phi(d)$ for $d \mid p-1$. It follows that $N_p(d) = \phi(d)$ for all $d \mid p-1$.

In particular, for $d = p-1$ we find that $N_p(p-1) = \phi(p-1) \ge 1$, i.e. there exists an element of $\mathbb{F}_p^\times$ with order $p-1$. This is a generator and therefore proves that $\mathbb{F}_p^\times$ is cyclic.

HINT for Exercise 1.7: I found it useful to know that $(\frac{1}{\sqrt{3}}(1,1,1), \frac{1}{\sqrt{6}}(1,1,-2), \frac{1}{\sqrt{2}}(1,-1,0))$ was an orthonormal basis, i.e. a basis of unit vectors for $\mathbb{R}^3$, all perpendicular to one another.

HONOURS ALGEBRA
WORKSHOP 4 SOLUTIONS

The purpose of this workshop is to think about linear mappings, practice change of basis, and to discover what you think is important and/or tricky, and to demonstrate that you are rhetors and the tutors are students.

Exercise 1. Let $\mathcal{S}(2) = (\vec{e}_1, \vec{e}_2)$ be the standard basis of $T = \mathbb{R}^2$ and let $\mathcal{B} = (\vec{v}_1 = -3\vec{e}_1 + 2\vec{e}_2, \vec{v}_2 = 2\vec{e}_1 - \vec{e}_2)$. Show that $\mathcal{B}$ is a basis of T. Now suppose that a linear mapping $f : T \to T$ is represented with respect to $\mathcal{S}(2)$ by the matrix

$$A = \begin{pmatrix} -6 & -9 \\ 4 & 6 \end{pmatrix}.$$

Find the matrix B that represents f with respect to $\mathcal{B}$.

Solution 1. If we do row reduction to $\begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}$ we get the identity matrix. Hence $\mathcal{B}$ is a basis.

Now $A(-3, 2)^T = (0, 0)^T$ and $A(2, -1)^T = (-3, 2)^T$ therefore the matrix with respect to $\mathcal{B}$ is

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Exercise 2. (1) Work out the matrix ${}_{\mathcal{B}}[f]_{\mathcal{A}}$ for the linear map

$$f : \mathbb{C}^3 \to \mathbb{C}^2; (x, y, z)^T \to (-x - y + 2z, 2x + 2y - 3z)^T$$

where $\mathcal{A} = ((0, 3, 2)^T, (1, 1, 1)^T, (1, 2, 2)^T)$ of $\mathbb{C}^3$ and $\mathcal{B} = ((1, 0)^T, (0, 1)^T)$ of $\mathbb{C}^2$. (You don't need to check either are bases: they are, you can assume it!)

(2) Write down a basis for the kernel of f.

Solution 2. (1)

$$f(0, 3, 2) = (1, 0)$$
$$f(1, 1, 1) = (0, 1)$$
$$f(1, 2, 2) = (1, 0).$$

Therefore the representing matrix is

$${}_{\mathcal{B}}[f]_{\mathcal{A}} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

(2) The first and last vectors from $\mathcal{A}$ are sent to the same vector, hence $(0, 3, 2)^T - (1, 2, 2)^T = (-1, 1, 0,)^T$ is in the kernel of f. Since the rank of the representing matrix is 2, the rank-nullity theorem implies that the kernel of f has dimension $3 - 2 = 1$ and so the kernel is 1 dimensional and this vector is a basis vector.

Date: Thursday, 8th February, 2018.

HONOURS ALGEBRA
SOLUTIONS FOR WORKSHOP 5

Exercise 1. Criticise the following argument in under 1 minute: The ring of integers $\mathbb{Z}$ is a field because every nonzero element has a multiplicative inverse (for example, 6 has inverse $1/6$).

Solution. *Although it is true that nonzero elements of $\mathbb{Z}$ have inverses* **in** $\mathbb{Q}$*, it is not true that they have inverses* **in** $\mathbb{Z}$*, and this is what is required if $\mathbb{Z}$ is to be a field. This is relevant to your answer in the next part. Not only do you have to find an inverse, but you have to show that this inverse is in F.*

Exercise 2. [This is a warm-up, to be studied in pairs, one person taking on the first part, the other taking on the second part, the third part together reuniting you. Don't let this take longer than 10 minutes.]

(1) By using the test for a subring plus something, or otherwise, show that the following subset of $\mathrm{Mat}(2;\mathbb{R})$ is a field:
$$R := \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} : a, b \in \mathbb{R} \right\}$$
(2) Construct a ring homomorphism from $\mathbb{R}[X]$ to $\mathbb{C}$ that is surjective. Calculate its kernel.
(3) What do the constructions above have anything in common?

Solution. *(1) To show that R is a ring we need to check that if $A, B \in R$ then $A + B \in R, -A \in R, 0 \in R, A \cdot B \in R$ and $\mathrm{id} \in R$. The laws of associativity and distributivity, as well as knowing 0 and id do what we expect, all hold because they hold in $\mathrm{Mat}(2;\mathbb{R})$.*

It's obvious that $0 \in R$. Let $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ and $B = \begin{pmatrix} c & d \\ -d & c \end{pmatrix}$ be two elements of R, so $a, b, c, d \in \mathbb{R}$. Then I can check both $A + B \in R$ and $-A \in R$ by showing that $A - B \in R$:
$$A - B = \begin{pmatrix} a-c & b-d \\ -b+d & a-c \end{pmatrix} = \begin{pmatrix} a-c & b-d \\ -(b-d) & a-c \end{pmatrix} \in R,$$
and
$$A \cdot B = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \cdot \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{pmatrix} \in R.$$
That $\mathrm{id} \in R$ is clear. Thus R is a ring.

To see that is even a field I need to check that it is commutative and that each non-zero element has an inverse. Commutativity is clear:
$$A \cdot B = \begin{pmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{pmatrix} = \begin{pmatrix} ca-db & cb+da \\ -(cb+da) & ca-db \end{pmatrix} = B \cdot A.$$
Finally, $\det(A) = a^2 + b^2$ and this is zero if and only if both a and b are zero. Assuming the $A \neq 0$ then we find
$$A^{-1} = \frac{1}{a^2+b^2} \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} \frac{a}{a^2+b^2} & -\frac{-b}{a^2+b^2} \\ -\frac{-b}{a^2+b^2} & \frac{a}{a^2+b^2} \end{pmatrix}$$
and this is an element of R. Hence R is a field.

(2) I would take $f : \mathbb{R}[X] \to \mathbb{C}$ to be evalutation at $\sqrt{-1}$, which we have already said in the notes is a ring homomorphism. The explanation of surjectivity and what its kernel is can be found in Example 3.6.10 in the notes.

Date: Thursday, 17th October, 2013.

(3) Both answers describe $\mathbb{C}$. The field R and the factor ring $\mathbb{R}[X]/\ker f$ are isomorphic to $\mathbb{C}$. For the isomorphism between $\mathbb{C}$ and R simply send $a+\sqrt{-1}b$ to $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$. The isomorphism between $\mathbb{R}[X]/\ker f$ and $\mathbb{C}$ is explained in Example 3.6.10. In both cases, we have constructed a (well-known) field from a ring.

Exercise 3. Answer this quickly! Does the polynomial $X^2+3 \in \mathbb{F}_5[X]$ have a root in $\mathbb{F}_5$?

Solution. *No! Evaluatation at $0,1,2,3,4$ is never 0, as it yields $3,4,2,2,4$ respectively.*

Exercise 4. [I think, for variety, it's good that you do this in threes.] Let

$$F := \left\{ \begin{pmatrix} a & b \\ 2b & a \end{pmatrix} \mid a,b \in \mathbb{F}_5 \right\}.$$

(1) Show that: F is a ring with addition and multiplication given by matrix multiplication in $\mathrm{Mat}(2;\mathbb{F}_5)$.
(2) Show that: F is a commutative ring.
(3) Show that: if β is a nonzero element of F then the determinant $\det(\beta)$ is nonzero.
(4) Show that: F is a field by finding an inverse (in F) for each nonzero element of F.
(5) How many elements does F have?
(6) Identify the element $a \in \mathbb{F}_5$ with the element $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ in F, so that we may think of $\mathbb{F}_5 \subset F$. Set $\alpha := \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} \in F$. Show that: every element in F can be written as $a+b\alpha$, with $a,b \in \mathbb{F}_5$.
(7) Calculate α^2 and hence show that α is a root of a quadratic polynomial over $\mathbb{F}_5$.
(8) Explain how to do multiplications in F when you write elements in the form $a+b\alpha$.
(9) Construct a ring homomorphism from $\mathbb{F}_5[X]$ to F that is surjective. Calculate its kernel.

Solution. *(1) To show that F is a ring we need to check that if $A,B \in F$ then $A+B \in F, -A \in F, 0 \in F, A \cdot B \in F$ and $\mathrm{id} \in F$. The laws of associativity and distributivity, as well as knowing 0 and id do what we expect, all hold because they hold in $\mathrm{Mat}(2;\mathbb{F}_5)$.*

It's obvious that $0 \in F$. Let $A = \begin{pmatrix} a & b \\ 2b & a \end{pmatrix}$ and $B = \begin{pmatrix} c & d \\ 2d & c \end{pmatrix}$ be two elements of F, so $a,b,c,d \in \mathbb{F}_5$. Then I can check both $A+B \in F$ and $-A \in F$ by showing that $A-B \in F$:

$$A-B = \begin{pmatrix} a-c & b-d \\ 2b-2d & a-c \end{pmatrix} = \begin{pmatrix} a-c & b-d \\ 2(b-d) & a-c \end{pmatrix} \in F,$$

and

$$A \cdot B = \begin{pmatrix} a & b \\ 2b & a \end{pmatrix} \cdot \begin{pmatrix} c & d \\ 2d & c \end{pmatrix} = \begin{pmatrix} ac+2bd & ad+bc \\ 2(ad+bc) & ac+2bd \end{pmatrix} \in F.$$

That $\mathrm{id} \in F$ is clear. Thus F is a ring.

(2) To show that F is a commutative ring, just check that $A \cdot B = B \cdot A$ for $A,B \in F$.

(3) Let $0 \neq \beta = \begin{pmatrix} a & b \\ 2b & a \end{pmatrix} \in F$. Note that $\det(\beta) = a^2 - 2b^2 \in \mathbb{Z}_5$. We show that $\det(\beta) \neq 0$. Suppose first that $b = 0$. Then $a \neq 0$ and so $\det(\beta) = a^2 \neq 0$. Next, suppose that $b \neq 0$. Then set $c := ab^{-1} \in \mathbb{F}_5$. If $\det(\beta) = 0$ then $c^2 - 2 = 0$. However, you can check that there is no such element in $\mathbb{F}_5$. Thus, $\det(\beta) \neq 0$. Hence, in all cases, $\Delta := \det(\beta) \neq 0$; so that Δ^{-1} exists in $\mathbb{F}_5$.

(4) Thus β is invertible in $M_2(\mathbb{F}_5)$, and we can calculate that

$$\beta^{-1} = \begin{pmatrix} a\Delta^{-1} & -b\Delta^{-1} \\ 2(-b\Delta^{-1}) & a\Delta^{-1} \end{pmatrix}$$

and note that this is an element of F, since $a\Delta^{-1}, -b\Delta^{-1}$ are elements of $\mathbb{F}_5$. Thus every nonzero element of the commutative ring F has an inverse in F; so F is a field.

(5) There are 5 choices for a and 5 for b. Thus, there are $5^2 = 25$ elements in F.

(6) For any element in F, we have

$$\begin{pmatrix} a & b \\ 2b & a \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} 0 & b \\ 2b & 0 \end{pmatrix} = a + b\alpha.$$

Thus every element in F can be written as $a + b\alpha$ with $a, b \in \mathbb{Z}_5$.

(7) Calculate $\alpha^2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 2$.

(8) Thus, α is a root of the polynomial $X^2 - 2 \in \mathbb{F}_5[X]$. (Note that earlier in the exercise we have checked that this polynomial has no root in $\mathbb{Z}_5$.) To do calculations with elements written in this form, we do ordinary addition and multiplication, and then replace every occurence of α^2 by 2, since $\alpha^2 - 2 = 0$, and remember that we are working over $\mathbb{F}_5$; so we always reduce modulo 5. This is just like doing complex multiplications with $a + \sqrt{-1}b$, where you replace $(\sqrt{-1})^2$ by -1 because i is a root of $X^2 + 1 = 0$.

As an example,

$$(3 + 2\alpha) \cdot (4 - \alpha) = 12 - 3\alpha + 8\alpha - 2\alpha^2 = 12 + 5\alpha - 4 = 8 - 5\alpha = 3$$

where the last equality comes because $5 = 0$ in $\mathbb{F}_5$.

Note that $3 + 2\alpha = \begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix}$ and that $4 - \alpha = \begin{pmatrix} 4 & -1 \\ -2 & 4 \end{pmatrix}$, and check that you get the same answer by matrix multiplication.

(9) $f : \mathbb{F}_5[X] \to F$ can be defined by $f(a_m X^m + \cdots + a_1 X + a_0) = a_m\alpha^m + \cdots + a_1\alpha + a_0$ (i.e. evaluation at α). It is surjective by (6), and by (7) the kernel contains the principal ideal generated by the polynomial $X^2 + 3$. Copy the argument in Example 3.6.10 to show that actually $\ker f = {}_{\mathbb{F}_5[X]}\langle X^2 + 3\rangle$.

Exercise 5. Try to repeat Exercise 2 with $\mathbb{F}_7$ instead of $\mathbb{F}_5$. Does it all work? If not, what does go through and what fails?

Solution *We now work over $\mathbb{F}_7$ instead of $\mathbb{F}_5$. You can check as above that F is a commutative ring. However, the proof that we have a field breaks down. This part of the proof relied on the fact that every nonzero element of F had nonzero determinant, and this, in turn, relied on the fact that there was no element in $\mathbb{F}_5$ for whose square was equal to 2. This breaks down over $\mathbb{F}_7$; for example, $3^2 = 9 = 2 \in \mathbb{F}_7$. Having observed this, it is easy to see that, for example, $A := \begin{pmatrix} 3 & 1 \\ 2 & 3 \end{pmatrix}$ has no inverse in F; so that F is not a field. You know therefore that F is not even a domain: to see this find a nonzero matrix $B \in F$ with $AB = 0$.*

HONOURS ALGEBRA
WORKSHOP 6 SOLUTIONS

This workshop will help you to understand factor modules and factor rings. In order for you to have confidence working with factor things you need to work through examples: then you'll get used to what you can and cannot do, and you'll also appreciate better the definitions and theorems in the notes.

Every exercise has more than one correct answer, so I want you to work in duos or trios, and to compare your complete answers to each exercise with the other people at your table.

Exercise 1. Let $V = \mathbb{R}^4$ and $U = \{\vec{v} = (v_1, v_2, v_3, v_4)^\mathsf{T} \in \mathbb{R}^4 : v_1 = v_2, v_3 = v_4\}$.

(1) Show that U is a subspace of V and find a basis for U.
(2) Extend this basis of U to a basis of V. Write out $\vec{v} \in V$ in terms of this basis.
(3) Describe the elements of the factor module V/U. Write down a basis for V/U and prove that it is a basis.
(4) Write down the matrix that represents the canonical mapping $\mathsf{can} : V \to V/U$ sending $\vec{v}$ to $\vec{v} + U$ in terms of the (ordered) basis $\{\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4\}$ of V and the one you chose in (3) for V/U. Also write the matrix representing this mapping, this time in terms of the (ordered) bases you chose for (2) and (3).

Solution 1. (1) Evidently $\vec{0} \in U$. Suppose $\vec{v}, \vec{w} \in U$. Then $(\vec{v}+\vec{w})_i = v_i + w_i$, so $v_1 + w_1 = v_2 + w_2$ and $v_3 + w_3 = v_4 + w_4$, since $v_1 = v_2$, $v_3 = v_4$, $w_1 = w_2$, and $w_3 = w_4$. Finally, for $\lambda \in \mathbb{R}$ we have that $(\lambda\vec{v})_1 = \lambda v_1 = \lambda v_2 = (\lambda\vec{v})_2$, and $(\lambda\vec{v})_3 = \lambda v_3 = \lambda v_4 = (\lambda\vec{v})_4$. Therefore U is a subspace.

Take $\vec{b}_1 = (1, 1, 0, 0)^\mathsf{T}$ and $\vec{b}_2 = (0, 0, 1, 1)^\mathsf{T}$. These are obviously linearly independent as they are not multiples of each other. Now let $\vec{v} \in U$. Then $\vec{v} = (v_1, v_1, v_3, v_3)^\mathsf{T} = v_1(1, 1, 0, 0)^\mathsf{T} + v_3(0, 0, 1, 1)^\mathsf{T}$, so $\langle \vec{b}_1, \vec{b}_2 \rangle = U$. Hence $\{\vec{b}_1, \vec{b}_2\}$ is a basis.

(2) Let $\vec{b}_3 = (1, -1, 0, 0)^\mathsf{T}$ and $\vec{b}_4 = (0, 0, 1, -1)^\mathsf{T}$. Then

$$\vec{v} = \frac{v_1 + v_2}{2}\vec{b}_1 + \frac{v_1 - v_2}{2}\vec{b}_3 + \frac{v_3 + v_4}{2}\vec{b}_2 + \frac{v_3 - v_4}{2}\vec{b}_4.$$

(3) We have that $(v_1, v_2, v_3, v_4)^\mathsf{T} + U = (w_1, w_2, w_3, w_4)^\mathsf{T} + U$ if and only if $(v_1 - w_1, v_2 - w_2, v_3 - w_3, v_4 - w_4)^\mathsf{T} \in U$, which holds if and only if $v_1 - w_1 = v_2 - w_2$ and $v_3 - w_3 = v_4 - w_4$.

But it is better to consider that $\alpha_1\vec{b}_1 + \alpha_2\vec{b}_2 + \alpha_3\vec{b}_3 + \alpha_4\vec{b}_4 + U = \beta_1\vec{b}_1 + \beta_2\vec{b}_2 + \beta_3\vec{b}_3 + \beta_4\vec{b}_4 + U$ if and only if $(\alpha_1 - \beta_1)\vec{b}_1 + (\alpha_2 - \beta_2)\vec{b}_2 + (\alpha_3 - \beta_3)\vec{b}_3 + (\alpha_4 - \beta_4)\vec{b}_4 \in U$, which holds if and only if $\alpha_3 = \beta_3$ and $\alpha_4 = \beta_4$.

Then we claim $\{\vec{b}_3 + U, \vec{b}_4 + U\}$ is a basis for V/U. By the above argument it spans V/U. It is linearly independent since

$$\mu(\vec{b}_3 + U) + \tau(\vec{b}_4 + U) = \mu\vec{b}_3 + \tau\vec{b}_4 + U = \vec{0} + U$$

only if $\mu = 0 = \tau$ by above.

(4) The matrices are respectively

$$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}, \text{ and } \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

Date: Thursday, 26th February, 2015.

(For example, we see from part (2) that $\vec{e}_2 = \frac{1}{2}\vec{b}_1 - \frac{1}{2}\vec{b}_3$, so $\vec{e}_2 + U = -\frac{1}{2}(\vec{b}_3 + U)$, hence the second column of the first matrix is $(-\frac{1}{2}, 0)^\mathsf{T}$.)

Exercise 2. Let V be the real vector space of polynomials $\mathbb{R}[X]_{<4}$ of degree less than 4. Let $U = \{P \in V : P(3) = 0\}$.

(1) Show that U is a subspace of V.
(2) Find a basis for U and extend it to a basis for V. Express $P \in V$ explicitly in terms of this basis.
(3) Write down a basis for V/U.
(4) Write down the matrix that represents the canonical mapping $\mathsf{can} : V \to V/U$ sending P to $P + U$ in terms of the (ordered) basis $\{X^3, X^2, X^1, 1\}$ of V and the one you chose in (3) for V/U.

Solution 2. (1) $P = \vec{0} \in U$. If $P, Q \in U$, then $(P+Q)(3) = P(3) + Q(3) = 0$, and $(\lambda P)(3) = \lambda P(3) = 0$, so it is a subspace.

(2) $\{X - 3, X^2 - 3X, X^3 - 3X^2\}$ is a basis of U. It is obviously linearly independent. Now suppose $P(3) = 0$. Then $(X - 3) \mid P$, i.e. there exists $Q = aX^2 + bX + c$ for some a, b, c such that $P = Q(X - 3)$. Then

$$P = (aX^2 + bX + c)(X - 3) = a(X^3 - 3X^2) + b(X^2 - 3X) + c(X - 3),$$

so it spans.

Extend this basis with the element $\{1\}$. Then

$$\begin{aligned} P &= a_3X^3 + a_2X^2 + a_1X + a_0 \\ &= a_3(X^3 - 3X^2) + a_2X^2 + 3a_3X^2 + a_1X + a_0 \\ &= a_3(X_3 - 3X^2) + (a_2 + 3a_3)(X^2 - 3X) + a_1X + 3(a_2 + 3a_3)X + a_0 \\ &= a_3(X^3 - 3X^2) + (a_2 + 3a_3)(X^2 - 3X) + (a_1 + 3a_2 + 9a_3)(X - 3) + (a_0 + 3a_1 + 9a_2 + 27a_3). \end{aligned}$$

(3) Use $\{1 + U\}$.
(4) We can just read this off from the explicit representation of a polynomial P in part (2): the matrix is $\begin{pmatrix} 27 & 9 & 3 & 1 \end{pmatrix}$.

Exercise 3. Let $R = \mathbb{F}_3[X]$.

(1) Find a cubic polynomial $P \in R$ that has no root.
(2) Let $I = {}_R\langle P \rangle$, the ideal generated by P. Write down the elements of R/I. How many are there?
(3) Prove that R/I is a field.
(4) The group of units $(R/I)^\times = R/I \setminus \{0_{R/I}\}$ is a cyclic group. Find a generator!

Solution 3. (1) We just run through a, b, c in $X^3 + aX^2 + bX + c$ discounting things that obviously have roots, e.g. when $c = 0$. $X^3 + 2X^2 + 1$ is such a polynomial.

(2) Let $S = \mathbb{F}_3[X]/I$, where $I = {}_{\mathbb{F}_3[X]}\langle X^3 + 2X^2 + 1 \rangle$. Then for any $P \in \mathbb{F}_3[X]$ we can find Q and R such that $P = (X^3 + 2X^2 + 1)Q + R$, where $\deg(R) \leq 2$. So $R = r_2X^2 + r_1X + r_0$, for some r_0, r_1, r_2. I.e. elements of S have the form $R + I$ were R has $\deg(R) \leq 2$.

Suppose $R + I = R' + I$. Then $R - R' \in I$, which implies that $R - R' = 0$ by considering the degrees of the polynomials. So $S = \{R + I : \deg(R) \leq 2\}$ and then $|S| = 27$, since there are three independent choices fo r_0, r_1, and r_2.

(3) S is a ring by general theory, and is obviously commutative since $\mathbb{F}_3[X]$ is:

$$(P + I)(Q + I) = PQ + I = QP + I = (Q + I)(P + I).$$

So we need to check that any non-zero element $R + I$ is invertible. There is a greatest common divisor for polynomials just as there is for integers. We claim that $\gcd(X^3 + 2X^2 + 1, R) = 1$. Suppose $P \mid (X^3 + 2X^2 + 1)$. Then $\deg(P) \neq 1$, since $X^3 + 2X^2 + 1$ has no roots and therefore no linear factors. Similarly, if $\deg(P) = 2$ then $(X^3 + 2X^2 + 1) = PQ$, and consideration of degree shows that $Q \mid X^3 + 2X^2 + 1$ would have $\deg(Q) = 1$, which is again a contradiction with the non-existence of roots of $X^3 + 2X^2 + 1$. Thus since the degree of $\gcd(X^3 + 2X^2 + 1, R)$ is at most 2, and is in fact neither 1 nor 2, the degree must be 0, i.e. it is the constant polynomial, and in fact it must be the constant polynomial 1.

Now, the extended Euclidean algorithm implies that there exist $A, B \in \mathbb{F}_3$ such that $(X^3 + 2X^2 + 1)A + RB = 1$, which implies that $(R + I)(B + I) = RB + I = 1 + I$, i.e. $B + I$ is the inverse of $R + I$, as required.

(4) $|(R/I)^\times| = 26$, so the only possible orders of non-trivial proper subgroups are 2 or 13. Let's try $X + I$ as a generator. Obviously

$$(X + I)^2 = X^2 + I \neq 1 + I.$$

Further,

$$X^3 + I = (X^3 + 2X^2 + 1 - 2X^2 - 1) + I = (X^2 + 2) + I,$$

which implies that

$$\begin{aligned} X^6 + I = (X^3 + I)^2 = (X^2 + 2)^2 + I = (X^4 + 4X^2 + 4) + I \\ = X^2 + 2X + 2 + 4X^2 + 4 + I \\ = 2X^2 + 2X + I, \end{aligned}$$

because $X^4 + I = XX^3 + I = X(X^2 + 2) + I = X^3 + 2X + I = X^2 + 2X + 2 + I$. Using this again, we see that

$$\begin{aligned} X^{12} + I = (X^6 + I)^2 = (2X^2 + 2X)^2 + I = 4X^2(X + 1)^2 + I \\ = X^2(X^2 + 2X + 1) + I \\ = X^4 + 2X^3 + X^2 + I \\ = (X^2 + 2X + 2) + 2(X^2 + 2) + X^2 + I \\ = X^2 + 2X + I. \end{aligned}$$

Finally then

$$X^{13} + I = X(X^2 + 2X) + I = X^3 + 2X^2 + I = X^2 + 2 + 2X^2 + I = 2 + I \neq 1 + I.$$

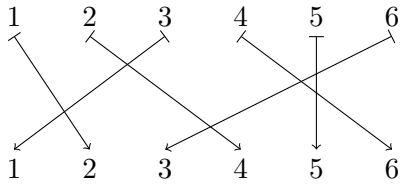
So the order of the subgroup generated by $X + I$ is greater than 13, therefore it must be 26, i.e. the subgroup is the whole group and $X + I$ is a generator.

HONOURS ALGEBRA
WORKSHOP 7 SOLUTIONS

This workshop will help to unpack the definition of the determinant given in the course. For that, you need to think about permutations and about a few explicit examples.

Exercise 1 (The sign of a permutation)**.** Recall the following from the notes or from Fundamentals of Pure Mathematics, all illustrated by the permutation $(1\,2\,4\,6\,3) \in \mathfrak{S}_6$:

- An inversion of a permutation $\sigma \in \mathfrak{S}_n$ is a pair (i, j) such that $1 \leq i < j \leq n$ and $\sigma(i) > \sigma(j)$. The number of inversions of the permutation σ is called the **length of** σ and written $\ell(\sigma)$. In formulas:
$$\ell(\sigma) = |\{(i, j) : i < j \text{ but } \sigma(i) > \sigma(j)\}|$$
So $\ell((1\,2\,4\,6\,3)) = 6$: the inversions are $(1, 3), (2, 3), (2, 6), (4, 5), (4, 6)$ and $(5, 6)$.
- A permutation can be represented by a diagram, from which I get a **crossing number** for σ as the number of times lines cross:



- Each element can be written as a product of transpositions from which I get a **transposition number** of σ as the number of transpositions I use to write out σ:
$$(1\,2\,4\,6\,3) = (1\,2)(3\,4)(4\,6)(2\,3)$$

(1) Calculate the length, crossing number and transposition number of the following elements of $\mathfrak{S}_4$: $\mathrm{id}, (1\,3), (1\,2\,3), (1\,2\,3\,4), (1\,2)(3\,4)$.
(2) What do you notice about these numbers? What does it have to do with sgn, the sign of a permutation?
(3) With which of three numbers (length, crossing, transposition) is the following statement quite obvious? And with which is not quite obvious?
   (a) $\mathsf{sgn}(\sigma) = \mathsf{sgn}(\sigma^{-1})$ for $\sigma \in \mathfrak{S}_n$
   (b) $\mathsf{sgn}((a\,b)) = -1$ for $(a\,b) \in \mathfrak{S}_n$ a transposition
   (c) $\mathsf{sgn}(\sigma) = \mathsf{sgn}(\tau^{-1}\sigma\tau)$ for $\sigma, \tau \in \mathfrak{S}_n$
(4) Which of the numbers (length, crossing, transposition) is unambiguously defined?
(5) Why does your answer to (1) and (3c) determine the sign of all elements in $\mathfrak{S}_4$?

Solution 1.

(1)

Element	Length	Crossing	Transposition
id	0: no inversions	0	0
$(1\,3)$	3: $(1,2)$, $(1,3)$, and $(2,3)$ are inversions	3	1: $(1\,3) = (1\,3)$
$(1\,2\,3)$	2: $(1,3)$, $(2,3)$ are inversions	2	2: $(1\,2\,3) = (1\,2)(2\,3)$
$(1\,2\,3\,4)$	3: $(1,4)$, $(2,4)$, $(3,4)$ are inversions	3	3: $(1\,2\,3\,4) = (1\,2)(2\,3)(3\,4)$
$(1\,2)(3\,4)$	2: $(1,2)$, $(3,4)$ are inversions	2	2: $(1\,2)(3\,4) = (1\,2)(3\,4)$

Date: Thursday, 5th March, 2015.

(2) Along any line in the above table the parity is the same, i.e. the numbers work out to be all even or all odd. This relates to the sign of the permutation: it can be proved that the numbers are all even exactly when the sign is 1, and all odd when it it -1.

(3) (a) Not obvious for length; obvious for crossings (turn the diagram upside down); obvious for transpositions (reverse the order of multiplication, e.g. if $(1\,2\,3) = (1\,2)(2\,3)$ then $(1\,2\,3)^{-1} = (2\,3)(1\,2)$).
(b) Not obvious for length; requires thought for crossing; obvious for transpositions.
(c) Not obvious for length; obvious for crossing (put the diagram for τ above that for σ, and that for τ^{-1}: then this gives $\tau^{-1}\sigma\tau$ and the number of crossings increases by an even number); more obvious for transposition (e.g. $(1\,2\,3\,4) = (1\,2)(2\,3)(3\,4)$ so $(1\,2\,3\,4)^{-1}(1\,2\,3)(1\,2\,3\,4) = (3\,4)(2\,3)(1\,2)(1\,2\,3)(1\,2)(1\,3)(3\,4)$, where each transposition appears twice).

(4) Only the length is unambiguous. The crossing number depends on how one draws the diagram, e.g. how straight one draws ones lines. For transpositions, e.g. $(1\,3) = (1\,2)(2\,3)(1\,2)$.

(5) A conjugacy class of permutations is determined by its "cycle type". (3c) says that sgn is constant on conjugacy classes. (1) gives a representative of each conjugacy class in $\mathfrak{S}_4$.

Exercise 2 (The definition of the determinant)**.** Let $A = (a_{ij}) \in \mathrm{Mat}(3; R)$ be a (3×3)-matrix with entries in a commutative ring R. Write out the determinant of A using only Definition 4.2.1 in the notes. If I asked you to write it out when $A \in \mathrm{Mat}(4; R)$ could you do it?

Solution 2. We can now calculate confidently that $\mathsf{sgn}(e) = \mathsf{sgn}((1\,2\,3)) = \mathsf{sgn}((1\,3\,2)) = 1$, and that $\mathsf{sgn}((1\,2)) = \mathsf{sgn}((1\,3)) = \mathsf{sgn}((2\,3)) = -1$. So

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}.$$

We could do it for $\mathfrak{S}_4$, but it would be too boring.

Exercise 3 (Calculating determinants for matrices in standard forms)**.** (1) Here is a procedure to produce matrices from permutations: $M\colon \mathfrak{S}_n \to \mathrm{Mat}(n; R)$. Given $\sigma \in \mathfrak{S}_n$ let $M(\sigma)$ be the matrix whose (i, j) entry is

$$M(\sigma)_{i,j} = \begin{cases} 1 & i = \sigma(j) \\ 0 & \text{otherwise.} \end{cases}$$

Take the elements of $\mathfrak{S}_4$ that you use in Exercise 1(1) and calculate $\mathsf{det}(M(\sigma))$ for each of them using the *old* fact you know that you can swap two columns at the cost of multiplying the determinant by -1. Does it look familiar? What happens if you calculate with Definition 4.2.1? Why is this fact critical for all determinant calculations?

(2) (This is Example 4.2.4 in the notes.) Show that: The determinant of an upper triangular matrix is the product of the entries along the diagonal.

(3) (This is Exercise 63 in the notes.) Show that: the determinant of a block-upper triangular matrix with square blocks along the diagonal is the product of the determinants of the blocks along the diagonal

$$\mathsf{det}\left(\begin{array}{c|c|c|c} A_1 & * & * & * \\ \hline 0 & A_2 & * & * \\ \hline 0 & 0 & \ddots & * \\ \hline 0 & 0 & 0 & A_t \end{array}\right) = \mathsf{det}(A_1)\mathsf{det}(A_2)\cdots\mathsf{det}(A_t).$$

Solution 3. (1) The relevant matrices are

$$\mathrm{id} \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \ (1\,3) \mapsto \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \ (1\,2\,3) \mapsto \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$(1\,2\,3\,4) \mapsto \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \ (1\,2)(3\,4) \mapsto \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

The determinants of these can be calculated by swapping columns until we get to the identity matrix. Then $\mathsf{det}(M(\mathrm{id})) = 1$, $\mathsf{det}(M((1\,3))) = -1$, $\mathsf{det}(M((1\,2\,3))) = 1$, $\mathsf{det}(M((1\,2\,3\,4))) = -1$, and $\mathsf{det}(M((1\,2)(3\,4))) = 1$. The column swaps we do are those in writing out σ as a product of transpositions, but backwards, i.e. $\mathsf{det}(\sigma) = \mathsf{sgn}(\sigma^{-1})$ $(= \mathsf{sgn}(\sigma)$ by exercise 1 (3a)).

If we calculate using the definition in the notes, we note first that

$$M(\sigma)_{i,\tau(i)} = \begin{cases} 1 & i = \sigma(\tau(i)), \\ 0 & \text{otherwise.} \end{cases}$$

This means that for $\tau \in \mathfrak{S}_n$,

$$M(\sigma)_{1,\tau(1)} \cdots M(\sigma)_{n,\tau(n)} = \begin{cases} 1 & i = \sigma(\tau(i)) \text{ for all } i, \text{ i.e. } \tau = \sigma^{-1}, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore

$$\mathsf{det}(M(\sigma)) = \sum_{\tau \in \mathfrak{S}_n} \mathsf{sgn}(\tau) M(\sigma)_{1,\tau(1)} \cdots M(\sigma)_{n,\tau(n)} = \mathsf{sgn}(\sigma^{-1}).$$

This is critical because the multilinearity and alternating properties of the determinant reduce every calculation to a calculation of the determinant of one of the $M(\sigma)$, see theorem 4.3.6.

(2) The reason is given in the notes.

(3) Suppose the matrix is in the block form described. Suppose A_i has dimensions $n_i \times n_i$. Split the set $\{1, \ldots, n\}$ into

$$\{1, \ldots, n_1\} \cup \{n_1 + 1, \ldots, n_1 + n_2\} \cup \cdots \cup \{n_1 + \cdots + n_{t-1} + 1, \ldots, n_1 + \cdots + n_{t-1} + n_t = n\}.$$

So the ith set in this union denotes those rows and columns where we find A_i. Call these subsets of the indices $J_1, \ldots, J_t$. Observe that if $\sigma \in \mathfrak{S}_n$ has the property that σ does *not* send each J_i to J_i, then there must be at least one element p in some J_i that gets sent to an element of J_k with $k > j$. Then

$$a_{1\sigma(1)} \cdots a_{n\sigma(n)} = a_{1\sigma(1)} \cdots a_{p\sigma(p)} \cdots a_{n\sigma(n)} = 0,$$

since $a_{p\sigma(p)}$ is an entry directly below A_i and is therefore zero.

Therefore the only permutations that contribute to the determinant are those $\sigma \in \mathfrak{S}_n$ such that σ sends each J_i to J_i. (For understanding, look at two extreme cases: $t = 1$, i.e. only one (big) block, then we ask that σ sends $\{1, \ldots, n\}$ to $\{1, \ldots, n\}$, which is no restriction at all; and $t = n$, then $J_i = \{i\}$, and this is an upper triangular matrix, and σ

must send each i to i, i.e. be the identity element—this is what appears in the explanation of the previous question.) So

$$\mathsf{det}(A) = \sum_{\sigma \in \mathrm{Sym}(J_1) \times \cdots \times \mathrm{Sym}(J_t)} \mathsf{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

Write such an element $\sigma \in \mathrm{Sym}(J_1) \times \cdots \times \mathrm{Sym}(J_t)$ as a t-tuple $(\sigma_1, \ldots, \sigma_t)$, where σ_i describes the permutation of the set J_i that σ produces. By thinking of diagrams, lengths, or transpositions, we see that $\mathsf{sgn}(\sigma) = \mathsf{sgn}(\sigma_1) \cdots \mathsf{sgn}(\sigma_t)$. So

$$\begin{aligned}
&\mathsf{det}(A) \\
&= \sum_{\sigma_1 \in \mathrm{Sym}(J_1), \ldots, \sigma_t \in \mathrm{Sym}(J_t)} \mathsf{sgn}(\sigma_1) \cdots \mathsf{sgn}(\sigma_t) a_{1\sigma_1(1)} \cdots a_{n_1\sigma_1(n_1)} a_{(n_1+1)\sigma_2(n_1+1)} \cdots a_{(n_1+n_2)\sigma_2(n_1+n_2)} \cdots a_{n\sigma_t(n)} \\
&= \mathsf{det}(A_1) \cdots \mathsf{det}(A_t),
\end{aligned}$$

as required.

HONOURS ALGEBRA
WORKSHOP 8 SOLUTIONS

In this workshop we will see that matrices describe many more things than linear mappings, and as an adjunct that their eigenvalues are obviously useful.

Exercise 1. There are 91 students registered for Honours Algebra. Let $M \in \mathrm{Mat}(91;\mathbb{R})$ be the matrix whose (i,j) entry is 1 if student i and student j have met each other and 0 if they have not. (I will assume that $M_{ii} = 0$.)

(1) Let $\vec{u} \in \mathbb{R}^{91}$ be the vector each of whose entries if 1. What does the vector $M\vec{u}$ represent?
(2) What information is contained in M^2?

Solution 1. (1) The ith entry of $M\vec{u}$ is the number of people registered for Honours Algebra that student i has met.

(2) $(M^2)_{ij}$ is the number of students registered for Honours Algebra who have met both student i and student j, when $i \neq j$. To see this we see that

$$(M^2)_{ij} = \sum_{k=1}^{91} M_{ik}M_{kj},$$

and $M_{ik}M_{kj} \neq 0$ if and only if $M_{ik} = 1 = M_{kj}$ if and only if i and k and k and j know each other.

For $i = j$, $(M^2)_{ii}$ is the number of students registered for Honours Algebra whom student i has met. To see this we see that

$$(M^2)_{ii} = \sum_{k=1}^{91} M_{ik}M_{ki} = \sum_{k=1}^{91} M_{ik}M_{ik},$$

where $M_{ik}M_{ik} \neq 0$ if and only if $M_{ik} = 1$ if and only if i has met k.

Exercise 2. Let $A = \begin{pmatrix} 7 & 2 \\ 1 & 6 \end{pmatrix}$.

(1) Find the two eigenvalues of the matrix A.
(2) Find eigenvectors for both eigenvalues.
(3) Find an invertible matrix P such that $P^{-1}AP$ is diagonal.

Solution 2. (1)

$$\chi_A(x) = \det \begin{pmatrix} 7-x & 2 \\ 1 & 6-x \end{pmatrix} = 42 - 13x + x^2 - 2 = (x-8)(x-5).$$

So the eigenvalues are 5 and 8.

(2)

$$A - 8I = \begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix}$$

and $\vec{v} = \binom{2}{1}$ is in the nullspace of this matrix, thus $A\binom{2}{1} = 8\binom{2}{1}$.

Date: Thursday, 7th November, 2013.

$$A - 5I = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix},$$

and $\vec{w} = \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ is in the nullspace, thus $A\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right) = 5\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$.

Therefore $\left(\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right)$ is an eigenvector with eigenvalue 8, and $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ is an eigenvector with eigenvalue 5.

(3) We know that $\mathcal{B} = \left\{\left(\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)\right\}$ is an ordered basis of $\mathbb{R}^2$ consisting of eigenvectors of the linear transformation $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \mapsto A\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$. I.e.

$$_{\mathcal{B}}[f]_{\mathcal{B}} = \begin{pmatrix} 8 & 0 \\ 0 & 5 \end{pmatrix}.$$

Thus

$$_{\mathcal{B}}[f]_{\mathcal{B}} = {}_{\mathcal{B}}[\mathrm{id}]_{\mathcal{S}(2)} \cdot {}_{\mathcal{S}(2)}[f]_{\mathcal{S}(2)} \cdot {}_{\mathcal{S}(2)}[\mathrm{id}]_{\mathcal{B}},$$

where $\mathcal{S}(2)$ is the standard basis of $\mathbb{R}^2$, and hence ${}_{\mathcal{S}(2)}[f]_{\mathcal{S}(2)} = A$. Let $P = {}_{\mathcal{S}(2)}[\mathrm{id}]_{\mathcal{B}} = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$. Then $P^{-1}AP = \begin{pmatrix} 8 & 0 \\ 0 & 5 \end{pmatrix}$.

Exercise 3. Each year $1/4$ of the haggises[1] outside Scotland move in and $1/8$ of the haggises inside Scotland move out. Let h_k be the number of haggis in Scotland in Year k and let g_k be the number of haggis outside of Scotland in Year k ($k \geq 0$).

(1) Write down a matrix equation that describes the number of haggises inside and outside in Year 1, in terms of h_0 and g_0. The matrix you write down should have the following two properties:
- each column sums to 1
- each entry is non-negative.

Why?

(2) Write down a matrix equation that describes the number of haggises inside and outside in Year k, in terms of h_0 and g_0. The matrix you write down should have the following two properties:
- each column sums to 1
- each entry is non-negative.

Why?

(3) You can now use a variation on your solution to Exercise 2 to solve this. Do it!

(4) What do you expect the proportion of haggises inside and outside Scotland to be in the long run?

(5) Analyse mathematically what you mean by "the long run".

Solution 3. (1) The appropriate equation is

$$\begin{pmatrix} 7/8 & 1/4 \\ 1/8 & 3/4 \end{pmatrix} \begin{pmatrix} h_0 \\ g_0 \end{pmatrix} = \begin{pmatrix} h_1 \\ g_1 \end{pmatrix},$$

i.e. $\frac{7}{8}h_0 + \frac{1}{4}g_0 = h_1$ ($7/8$ of haggises in Scotland stay and $1/4$ of haggises outside Scotland come) and $\frac{1}{8}h_0 + \frac{3}{4}g_0 = g_1$. Each column sums to one because a column is the proportion of haggises staying or going, which must total 1 as long as there is no birth or death—this is also why the numbers are non-negative.

[1] If you think it is disrespectful to mix haggises and mathematics, think again.

(2) Let $B = \begin{pmatrix} 7/8 & 1/4 \\ 1/8 & 3/4 \end{pmatrix}$. Then

$$B^k \begin{pmatrix} h_0 \\ g_0 \end{pmatrix} = \begin{pmatrix} h_k \\ g_k \end{pmatrix}$$

is the formula we want. The reasons for summing to one are the same as before: the columns are proportions staying or going after k years.

(3) $B = \frac{1}{8}A$, where A is as in exercise 2, so $B^k = \frac{1}{8^k}A^k$. In particular, $B = \frac{1}{8}A = \frac{1}{8}PD'P^{-1}$, where $D' = \left(\begin{smallmatrix} 8 & 0 \\ 0 & 5 \end{smallmatrix}\right)$, i.e. $B = PDP^{-1}$ where $D = \begin{pmatrix} 1 & 0 \\ 0 & 5/8 \end{pmatrix}$. This means that $B^k = (PDP^{-1})^k = PD^kP^{-1}$, and so

$$\begin{aligned} B^k \begin{pmatrix} h_0 \\ g_0 \end{pmatrix} &= PD^kP^{-1} \begin{pmatrix} h_0 \\ g_0 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & (5/8)^k \end{pmatrix} \cdot \frac{1}{3} \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} h_0 \\ g_0 \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} h_0 + g_0 \\ (5/8)^k(h_0 - 2g_0) \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 2(h_0 + g_0) + (5/8)^k(h_0 - 2g_0) \\ (h_0 + g_0) - (5/8)^k(h_0 - 2g_0) \end{pmatrix} \\ &= \begin{pmatrix} h_k \\ g_k \end{pmatrix}, \end{aligned}$$

i.e.

$$\begin{aligned} h_k &= \left(\frac{2}{3} + \frac{1}{3} \left(\frac{5}{8} \right)^k \right) h_0 + \left(\frac{2}{3} - \frac{2}{3} \left(\frac{5}{8} \right)^k \right) g_0, \\ g_k &= \left(\frac{1}{3} - \frac{1}{3} \left(\frac{5}{8} \right)^k \right) h_0 + \left(\frac{1}{3} + \frac{2}{3} \left(\frac{5}{8} \right)^k \right) g_0. \end{aligned}$$

(4) In the long run, as $k \to \infty$, $h_k \to \frac{2}{3}(h_0 + g_0)$ and $g_k \to \frac{1}{3}(h_0 + g_0)$. But $h_0 + g_0$ is the total population. So $2/3$ of the haggises will be in Scotland, and $1/3$ will not be.

(5) Let $h_\infty = \frac{2}{3}(h_0 + g_0)$ and $g_\infty = \frac{1}{3}(h_0 + g_0)$. This is what we expect to happen. Then

$$\begin{pmatrix} h_k \\ g_k \end{pmatrix} - \begin{pmatrix} h_\infty \\ g_\infty \end{pmatrix} = \frac{1}{3} \left(\frac{5}{8} \right)^k (h_0 - 2g_0) \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

So the rate of convergence is determined by $(5/8)^k$, which is the second eigenvalue to the power k. So if we want to get a good approximation to the stable (i.e. final) state, given h_0 and g_0, we just need to find k big enough depending on $(5/8)$.

Exercise 4. If that was all too easy, ask me for another problem!

HONOURS ALGEBRA
WORKSHOP 9 SOLUTIONS

In this workshop we will play with Cauchy's inequality and it's consequences. This is supposed to be an illustration of one way in which you do mathematics. You keep asking questions. The questions here follow the book "The Cauchy-Schwarz Master Class" by Steele which I like. It has lots more like this, together with good explanations of what you're doing when you do mathematics.

There are 5 exercises: do 1 in two groups (one part each), then work together on the rest.

Exercise 1 (We shall not cease from exploration)**.** Cauchy's inequality asserts

$$x_1y_1 + \cdots + x_ny_n \leq \sqrt{x_1^2 + \cdots + x_n^2}\sqrt{y_1^2 + \cdots + y_n^2}$$

for $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$.

(1) Prove this by induction on n. The case $n = 1$ is trivial, so figure the $n = 2$ case directly before trying in general.
(2) Let t be a variable and consider the quadratic polynomial in t:

$$\sum_{i=1}^{n}(x_it + y_i)^2$$

This is always positive, right? So, use that to prove Cauchy's inequality.

Do you like one version more than the other? [By the way, I will prove this in lectures by another method, just to illustrate something else. And you are going to meet another (stronger?) proof, below.]

Solution 1. (1) For $n = 2$ the inequality states that

$$(x_1y_1 + x_2y_2)^2 \leq (x_1^2 + x_2^2)(y_1^2 + y_2^2).$$

Expanding and bringing all to one side we see that this is equivalent to

$$0 \leq (x_1y_2)^2 - 2(x_1y_2)(x_2y_1) + (x_2y_1)^2.$$

The RHS factorizes as $(x_1y_2 - x_2y_1)^2$, which is certainly non-negative.

For the inductive step, we assume the inequality holds for n terms. Then for $n+1$ terms, we combine the cases for n and $n = 2$:

$$\begin{aligned} x_1y_1 + \cdots + x_ny_n + x_{n+1}y_{n+1} &= (x_1y_1 + \cdots + x_ny_n) + x_{n+1}y_{n+1} \\ &\leq (x_1^2 + \cdots + x_n^2)^{1/2}(y_1^2 + \cdots + y_n^2)^{1/2} + x_{n+1}y_{n+1} \\ &\leq \left((x_1^2 + \cdots + x_n^2) + x_{n+1}^2\right)^{1/2}\left((y_1^2 + \cdots + y_n^2) + y_{n+1}^2\right)^{1/2}, \end{aligned}$$

where the last line follows from the $n = 2$ case applied to $(x_1^2 + \cdots + x_n^2)^{1/2}$ and x_{n+1}, and $(y_1^2 + \cdots + y_n^2)^{1/2}$ and y_{n+1}.

(2) Write $\sum_{i=1}^{n}(x_it + y_i)^2$ as $\alpha t^2 + \beta t + \gamma$ for some α, β, γ depending on x_i and y_i. Since it is a sum of squares, we know that $\alpha t^2 + \beta t + \gamma \geq 0$ for any choice of t, and so the quadratic has

Date: Thursday, 14th November, 2013.

at most one real solution. Therefore the discriminant satisfies $\beta^2 - 4\alpha\gamma \leq 0$, i.e. $\beta^2 \leq 4\alpha\gamma$. But now, what are α, β, γ in terms of x_i and y_i precisely? We have that

$$\alpha = \sum_{i=1}^{n} x_i^2,\ \beta = 2\sum_{i=1} x_i y_i,\ \text{and}\ \gamma = \sum_{i=1}^{n} y_i^2,$$

so we have that $4\left(\sum_{i=1}^n x_i y_i\right)^2 \leq 4\left(\sum_{i=1}^n x_i^2\right)\left(\sum_{i=1}^n y_i^2\right)$, which after dividing by 4 and taking square roots gives the required inequality.

Exercise 2 (And the end of all our exploring)**.** Begin by making sure that you can see how to deduce from Cauchy's inequality the fact that

$$\sum_{k=1}^{\infty} x_k^2 < \infty \text{ and } \sum_{k=1}^{\infty} y_k^2 < \infty \text{ together imply that } \sum_{k=1}^{\infty} |x_k y_k| < \infty.$$

OK? Now I want you to prove this fact without using Cauchy's inequality. Here's an approach:

The statement you want suggests that you'll need to show that the product $x_k y_k$ is small x_k^2 and y_k^2 are small.

(1) Can you find a C (as efficiently as possible) such that $xy \leq C(x^2 + y^2)$ for all real x, y?
(2) Now apply what you've just done to $x = |x_k|$ and $y = |y_k|$ and sum over all k. Do you see a new inequality, looking different from Cauchy's inequality?[1]

Solution 2. By hypothesis there exist N_1, N_2 such that $\sum_{k=1}^{\infty} x_k^2 \leq N_1$ and $\sum_{k=1}^{\infty} y_k^2 \leq N_2$. Changing y_k to $-y_k$ if necessary we may assume that $|x_k y_k| = x_k y_k$. Then Cauchy-Schwarz states that for any t,

$$\sum_{k=1}^{t} x_k y_k = \sum_{k=1}^{t} |x_k y_k| \leq \left(\sum_{k=1}^{t} x_k^2\right)^{1/2} \left(\sum_{k=1}^{t} y_k^2\right)^{1/2} \leq N_1^{1/2} N_2^{1/2}.$$

Therefore $\sum_{k=1}^{\infty} |x_k y_k| \leq N_1^{1/2} N_2^{1/2}$ and this sum is therefore finite.

(1) $C = 1/2$ works:

$$xy \leq \frac{1}{2}x^2 + \frac{1}{2}y^2 \text{ if and only if } 0 \leq \frac{1}{2}x^2 - xy + \frac{1}{2}y^2 = \frac{1}{2}(x-y)^2.$$

(2) We get

$$\sum_{k=1}^{\infty} |x_k y_k| \leq \frac{1}{2}\sum_{k=1}^{\infty} x_k^2 + \frac{1}{2}\sum_{k=1}^{\infty} y_k^2. \tag{1}$$

Exercise 3 (Will be to arrive where we started)**.** Calculate the inequality you just produced with normalised variables

$$\hat{x}_j = x_j / (\sum_{k=1}^{\infty} x_k^2)^{1/2} \text{ and } \hat{y}_j = y_j / (\sum_{k=1}^{\infty} y_k^2)^{1/2}$$

What do I mean by normalised? Can you find a new interesting inequality out of this?

Solution 3. Normalization means

$$\sum_{j=1}^{\infty} \hat{x}_j = \sum_{j=1}^{\infty} \left(x_j / \sum_{k=1}^{\infty} x_k^2 \right) = 1,$$

[1] Go to the next page to see what I am thinking of; but only go there to check your answer, not to be a lazybones

and similarly for $\hat{y}_j$. Thus we find from (1) that

$$\sum_{j=1}^{\infty}\left(x_j/\sum_{k=1}^{\infty}x_k^2\right)\left(y_j/\sum_{k=1}^{\infty}y_k^2\right)=\sum_{j=1}^{\infty}\hat{x}_j\hat{y}_j\leq\frac{1}{2}\sum_{j=1}^{\infty}\hat{x}_j^2+\frac{1}{2}\sum_{j=1}^{\infty}\hat{y}_j^2=\frac{1}{2}\cdot 1+\frac{1}{2}\cdot 1=1.$$

Multiplying by the common denominator on the LHS and switching to the same indices gives us

$$\sum_{j=1}^{\infty}x_jy_j\leq\left(\sum_{j=1}^{\infty}x_j^2\right)\left(\sum_{j=1}^{\infty}y_j^2\right),$$

i.e. Cauchy's inequality with infinite sums.

Exercise 4 (No inequality)**.** Let's go back to Cauchy's inequality. Precisely when do we get an equality between the two sides of the inequality?

Solution 4. By tracing back up the previous solution, we get equality precisely when each term

$$\hat{x}_j\hat{y}_j\leq\frac{1}{2}\hat{x}_j^2+\frac{1}{2}\hat{y}_j^2$$

is an equality, i.e. precisely when $\frac{1}{2}(\hat{x}_j-\hat{y}_j)^2=0$, which holds precisely when $\hat{x}_j=\hat{y}_j$, i.e. when $x_j=\lambda y_j$, where $\lambda=\sum_{k=1}^{\infty}x_k^2/\sum_{k=1}^{\infty}y_k^2$—i.e. when the sequences are linear multiples of each other.

Exercise 5 (Notation)**.** Has the notation that you've used been a pain. Lots of sums, infinities, and so on? Can you find a concise way to write key statements down? What are the important properties of numbers in the notation that you are using? Have you just invented an axiomatic system?

Solution 5. Let $\sum x_iy_i=x\cdot y$, $\sum x_i^2=x\cdot x$ and $\sum y_i^2=y\cdot y$. Then Cauchy says that

$$x\cdot y\leq(x\cdot x)^{1/2}(y\cdot y)^{1/2}.$$

For this to be absolutely sensible we'd want $x\cdot y=y\cdot x$ and that $(x\cdot x)\geq 0$ with equality only for $x=0$. This is the start of the definition of an inner product space.

Here's the inequality I am thinking of:

$$\sum_{k=1}^{\infty}|x_ky_k|\leq\frac{1}{2}\sum_{k=1}^{\infty}x_k^2+\frac{1}{2}\sum_{k=1}^{\infty}y_k^2.$$

HONOURS ALGEBRA
WORKSHOP 10 SOLUTIONS

You will have noticed in the notes and lectures on inner product spaces that, given a vector $\vec{v}$, I have used the mapping $(-, \vec{v})$ several times (Proofs of Theorems 5.1.10 and 5.3.4). In this workshop we will study such mappings more generally: this the theory of dual vector spaces. These spaces are important in analysis, in algebra, in geometry, in physics, basically in you name it.

Throughout this workshop, V will be a vector space over a field F. I define the **dual vector space** V^* to be the space of linear mappings

$$V^* = \mathrm{Hom}_F(V, F)$$

Recall from Exercise 15 in the notes that this is a vector space: if $\theta, \phi \in V^*$, $\vec{v} \in V$ and $\lambda \in F$ then

$$\begin{aligned}(\theta + \phi)(\vec{v}) &= \theta(\vec{v}) + \phi(\vec{v}) \\ (\lambda\theta)(\vec{v}) &= \lambda(\theta(\vec{v}))\end{aligned}$$

You do not need to check this. You can look at the solutions to Exercise 15 if you want a crutch.

Exercise 1. Assume that V is an inner product space.

(1) Let $\vec{v} \in V$. Show that: the mapping

$$(-, \vec{v}) : V \to F, \quad \vec{w} \mapsto (\vec{w}, \vec{v})$$

is an element of V^*. Call it $\epsilon_{\vec{v}}$.

(2) Show that the mapping

$$\Delta : V \to V^*, \quad \vec{v} \to \epsilon_{\vec{v}}$$

is injective.

(3) Is Δ linear?

Solution 1. (1) We are required to show that $\epsilon_{\vec{v}}$ is linear. This is the case:

$$\epsilon_{\vec{v}}(\alpha\vec{w}) = (\alpha\vec{w}, \vec{v}) = \alpha(\vec{w}, \vec{v}) = \alpha(\epsilon_{\vec{v}}(\vec{w})),$$
$$\epsilon_{\vec{v}}(\vec{w}_1 + \vec{w}_2) = (\vec{w}_1 + \vec{w}_2, \vec{v}) = (\vec{w}_1, \vec{v}) + (\vec{w}_2, \vec{v}) = \epsilon_{\vec{v}}(\vec{w}_1) + \epsilon_{\vec{v}}(\vec{w}_2).$$

(2) Suppose $\epsilon_{\vec{v}_1} = \epsilon_{\vec{v}_2}$. Then for all $\vec{w} \in V$, we have that

$$(\vec{w}, \vec{v}_1) = \epsilon_{\vec{v}_1}(\vec{w}) = \epsilon_{\vec{v}_2}(\vec{w}) = (\vec{w}, \vec{v}_2),$$

i.e. for all $\vec{w} \in V$ we have that $(\vec{w}, \vec{v}_1 - \vec{v}_2) = 0$. Applying this to $\vec{w} = \vec{v}_1 - \vec{v}_2$, we see that $(\vec{v}_1 - \vec{v}_2, \vec{v}_1 - \vec{v}_2) = 0$, which implies that $\vec{v}_1 - \vec{v}_2 = \vec{0}$, i.e. that $\vec{v}_1 = \vec{v}_2$. Therefore Δ is injective.

(3) This depends on whether the underlying scalar field is real or complex: the answer is yes for $\mathbb{R}$, no for $\mathbb{C}$. We have that

$$\epsilon_{\vec{v}_1+\vec{v}_2}(\vec{w}) = (\vec{w}, \vec{v}_1 + \vec{v}_2) = (\vec{w}, \vec{v}_1) + (\vec{w}, \vec{v}_2) = \epsilon_{\vec{v}_1}(\vec{w}) + \epsilon_{\vec{v}_2}(\vec{w}) = (\epsilon_{\vec{v}_1} + \epsilon_{\vec{v}_2})(\vec{w})$$

for all $\vec{w}$, hence

$$\Delta(\vec{v}_1 + \vec{v}_2) = \epsilon_{\vec{v}_1+\vec{v}_2} = \epsilon_{\vec{v}_1} + \epsilon_{\vec{v}_2} = \Delta(\vec{v}_1) + \Delta(\vec{v}_2).$$

But

$$\epsilon_{\lambda\vec{v}_1}(\vec{w}) = (\vec{w}, \lambda\vec{v}_1) = \overline{\lambda}(\vec{w}, \vec{v}_1) = \overline{\lambda}\epsilon_{\vec{v}_1}(\vec{w})$$

for all $\vec{w}$, so

$$\Delta(\lambda\vec{v}_1) = \epsilon_{\lambda\vec{v}_1} = \overline{\lambda}\epsilon_{\vec{v}_1} = \overline{\lambda}\Delta(\vec{v}_1).$$

Date: Thursday, 21st November, 2013.

Exercise 2. Assume that V is a finite dimensional F-vector space.

(1) Let $\mathcal{A} = (\vec{v}_1, \ldots, \vec{v}_n)$ be an ordered basis of V. Show that: $(\theta_1, \ldots, \theta_n)$ is an ordered basis of V^* where I define $\theta_i : V \to F$ by the rule

$$\theta_i(\sum_{j=1}^{n} \lambda_j \vec{v}_j) = \lambda_i$$

This is called the **dual basis** to $\mathcal{A}$.

(2) Let W be a finite dimensional F-vector space and $f : V \to W$ a linear mapping. Show that: the mapping $f^* : W^* \to V^*$ defined by

$$f^*(\theta)(\vec{v}) = \theta(f(\vec{v})), \quad \text{for } \theta \in W^*, \vec{v} \in V$$

is linear. This is called the **dual mapping** to f.

(3) Let $\mathcal{A} = (\vec{v}_1, \ldots, \vec{v}_n)$ and $\mathcal{B} = (\vec{w}_1, \ldots, \vec{w}_m)$ be ordered bases of V and W respectively. Let $\mathcal{A}^*$ and $\mathcal{B}^*$ be the dual bases of V^* and W^* obtained from applying (1). Show that:

$${}_{\mathcal{A}^*}[f^*]_{\mathcal{B}^*} = ({}_{\mathcal{B}}[f]_{\mathcal{A}})^{\mathsf{T}}$$

In other words, the transpose of a matrix describes the dual mapping!

Solution 2. (1) Let $\sum_{i=1}^{n} \alpha_i \theta_i = 0$ in V^*, i.e. $(\sum_{i=1}^{n} \alpha_i \theta_i)(\vec{v}) = 0$ for all $\vec{v} \in V$. Then by definition of the addition in V^*, we have that

$$\sum_{i=1}^{n} \alpha_i \theta_i(\vec{v}) = 0.$$

Apply this to $\vec{v} = \vec{v}_j$ to see that

$$\alpha_j = \sum_{i=1}^{n} \alpha_i \theta_i(\vec{v}_j) = 0,$$

by definition of θ_i. Therefore $\{\theta_1, \ldots, \theta_n\}$ is linearly independent.

Let $\theta \in V^*$ be arbitrary. We claim that $\theta = \sum_{i=1}^{n} \lambda_i \theta_i$, where $\lambda_i = \theta(\vec{v}_i)$. To see this, let $\vec{v} = \sum_{j=1}^{n} \alpha_j \vec{v}_j$ be arbitrary. Then linearity of θ implies that

$$\theta(\vec{v}) = \theta\left(\sum_{j=1}^{n} \alpha_j \vec{v}_j\right) = \sum_{j=1}^{n} \alpha_j \theta(\vec{v}_j),$$

while

$$\sum_{i=1}^{n} \lambda_i \theta_i(\vec{v}) = \sum_{i=1}^{n} \lambda_i \theta_i \left(\sum_{j=1}^{n} \alpha_j \vec{v}_j\right) = \sum_{i=1}^{n} \lambda_i \left(\sum_{j=1}^{n} \alpha_j \theta_i(\vec{v}_j)\right) = \sum_{i=1}^{n} \lambda_i \alpha_i.$$

But $\sum_{j=1}^{n} \alpha_j \theta(\vec{v}_j) = \sum_{j=1}^{n} \alpha_j \lambda_j$ by definition, so

$$\theta(\vec{v}) = \sum_{i=1}^{n} \lambda_i \theta_i(\vec{v})$$

for all $\vec{v} \in V$, so indeed $\theta = \sum_{i=1}^{n} \lambda_i \theta_i$. Thus $\{\theta_1, \ldots, \theta_n\}$ spans V^*.

(2) It is clear that $f^*(\theta)$ is a function from V to F: this is what the formula

$$f^*(\theta)(\vec{v}) = \theta(f(\vec{v})) \in F$$

shows. We need to check that it is linear. So let $\theta, \psi \in W^*$ and $\lambda \in F$. We first show that $f^*(\theta + \psi) = f^*(\theta) + f^*(\psi)$: for all $\vec{v} \in V$, we have

$$f^*(\theta+\psi)(\vec{v}) = (\theta+\psi)(f(\vec{v})) = \theta(f(\vec{v})) + \psi(f(\vec{v})) = f^*(\theta)(\vec{v}) + f^*(\psi)(\vec{v}) = (f^*(\theta) + f^*(\psi))(\vec{v}),$$

as required. We now show that $f^*(\lambda\theta) = \lambda f^*(\theta)$: for all $\vec{v} \in V$, we have

$$f^*(\lambda\theta)(\vec{v}) = (\lambda\theta)(f(\vec{v})) = \lambda(\theta(f(\vec{v}))) = \lambda(f^*(\theta))(\vec{v}) = (\lambda f^*(\theta))(\vec{v}),$$

as required.

If we were thinking straight, we would observe that $f^*(\theta)$ is the composition $\theta \circ f \colon V \to F$ of the two linear mappings $f \colon V \to W$ and $\theta \colon W \to F$, and is therefore itself linear.

(3) Let $\mathcal{A}^* = (\theta_1, \ldots, \theta_n)$ and $\mathcal{B}^* = (\psi_1, \ldots, \psi_m)$ be the corresponding ordered dual bases for V^* and W^* respectively. Then $C = {}_{\mathcal{A}^*}[f^*]_{\mathcal{B}^*}$ has as its jth column $(c_{ij})_{i=1}^n$ where

$$f^*(\psi_j) = \sum_{i=1}^{n} c_{ij}\theta_i.$$

How do we calculate these c_{ij}? Well, if $D = {}_{\mathcal{B}}[f]_{\mathcal{A}}$, then

$$d_{jk} = \psi_j\left(\sum_{t=1}^{m} d_{tk}\vec{w}_t\right) = \psi_j(f(\vec{v}_k)) = f^*(\psi_j)(\vec{v}_k) = \left(\sum_{i=1}^{n} c_{ij}\theta_i\right)(\vec{v}_k) = \sum_{i=1}^{n} c_{ij}\theta_i(\vec{v}_k) = c_{kj}.$$

So indeed $C = D^{\mathsf{T}}$.

Exercise 3. (Health Warning: this exercise will twist your brain; skip to 4 if you are worried. And you need to do this with a partner).

(1) Let V be a one-dimensional vector space. Pick a non-zero vector (make sure your partner picks a different one). Let's call your vector $\vec{v}$. This gives a basis for V and so you can apply Exercise 2(1) to get a dual basis vector, say $\theta \in V^*$. Show that: the mapping $V \to V^*$ that sends $\lambda\vec{v}$ to $\lambda\theta$ is an isomorphism.
(2) Is your isomorphism exactly the same as your partner's?
(3) Now let V be an arbitrary finite dimensional vector space. Without ever picking a basis, construct an explicit isomorphism $V \xrightarrow{\sim} (V^*)^*$.

Solution 3. (1) We'll do this by example. Consider $V = \mathbb{R}$, which is a one-dimensional real vector space.

Choose $\vec{v} = 1 \in \mathbb{R}$. Then its dual basis consists of $\theta \in V^*$ such that, by linearity of θ,

$$\theta(c) = \theta(c \cdot 1) = c\theta(1) = c,$$

i.e. is the identity mapping.

If I take $\vec{v}' = 2 \in \mathbb{R}$, then its dual basis consists of $\theta' \in V^*$ where, by linearity of θ',

$$\theta'(c) = \theta'(\frac{c}{2} \cdot 2) = \frac{c}{2}\theta'(2) = \frac{c}{2},$$

i.e. the operation of halving.

So we get maps from $V = \mathbb{R}$ to $V^* = \mathbb{R}$ which are respectively multiplying by 1 and $1/2$.

(2) We have just shown that $\theta = \mathrm{id}$ whereas $\theta' = \frac{1}{2}\mathrm{id}$, i.e. they are not the same.
(3) We will construct $\chi \colon V \to (V^*)^*$ by $\vec{v} \mapsto \chi_{\vec{v}}$ where $\chi_{\vec{v}} \colon V^* \to F$ is defined by $\chi_{\vec{v}}(\theta) = \theta(\vec{v})$.

We first prove that this is injective. $\chi_{\vec{v}} = \chi_{\vec{w}}$ if and only if $\theta(\vec{v}) = \theta(\vec{w})$ for all $\theta \in V^*$, if and only if $\theta(\vec{v} - \vec{w}) = 0$ for all $\theta \in V^*$. Is there a linear mapping $\theta \colon V \to F$ such that $\theta(\vec{v} - \vec{w}) \neq 0$? Yes, unless $\vec{v} - \vec{w} = \vec{0}$: just take $\vec{v} - \vec{w}$ to be the first element of a basis and use the dual basis construction. Therefore $\chi_{\vec{v}} = \chi_{\vec{w}}$ if and only if $\vec{v} = \vec{w}$.

We prove that χ is linear:

$$\chi_{\vec{v}+\vec{w}}(\theta) = \theta(\vec{w}+\vec{w}) = \theta(\vec{v}) + \theta(\vec{w}) = \chi_{\vec{v}}(\theta) + \chi_{\vec{w}}(\theta) = (\chi_{\vec{v}} + \chi_{\vec{w}})(\theta),$$

and

$$\chi_{\lambda\vec{v}}(\theta) = \theta(\lambda\vec{v}) = \lambda\theta(\vec{v}) = \lambda\chi_{\vec{v}}(\theta).$$

Therefore χ is an injective linear mapping between two vector spaces of the same dimension (since by exercise 2(1) we have that $\dim(V) = \dim(V^*) = \dim((V^*)^*)$). Thus it is an isomorphism.

Exercise 4. Let U be a subspace of a finite dimensional F-vector space V.

(1) Show that:
$$\mathrm{Ann}(U) \coloneqq \{\theta \in V^* : \theta(\vec{u}) = 0 \text{ for all } \vec{u} \in U\}$$
is a subspace of V^*.
(2) What is $\mathrm{Ann}(\{\vec{0}\})$? What is $\mathrm{Ann}(V)$?
(3) Let $f\colon V \to W$ be a linear mapping. Show that: $\ker(f^*) = \mathrm{Ann}(\mathrm{im} f)$.
(4) Show that: U^* is isomorphic to the quotient vector space $V^*/\mathrm{Ann}(U)$.

Solution 4. (1) The zero function is in $\mathrm{Ann}(U)$. Let $\theta, \psi \in \mathrm{Ann}(U)$. Then for any $\vec{u} \in U$,

$$(\theta + \psi)(\vec{u}) = \theta(\vec{u}) + \psi(\vec{u}) = 0 + 0 = 0,$$

so $\theta + \psi \in \mathrm{Ann}(U)$. Let $\lambda \in F$. Then again, for any $\vec{u} \in U$,

$$(\lambda\theta)(\vec{u}) = \lambda\theta(\vec{u}) = \lambda \cdot 0 = 0,$$

so $\lambda\theta \in \mathrm{Ann}(U)$. Therefore $\mathrm{Ann}(U)$ is a subspace.

(2) $\mathrm{Ann}(\{\vec{0}\}) = V^*$, since every linear mapping sends $\vec{0}$ to 0. On the other hand $\mathrm{Ann}(V) = \{0\}$, i.e. consists only of the zero mapping.

(3) $\theta \in W^*$ is in $\ker f^*$ if and only if $f^*(\theta) \in V^*$ is the zero mapping, i.e if and only if $f^*(\theta)(\vec{v}) = \theta(f(\vec{v})) = 0$ for all $\vec{v} \in V$, if and only if $\theta \in \mathrm{Ann}(\mathrm{im} f)$.

(4) Let $f\colon U \to V$ be the inclusion $f(\vec{u}) = \vec{u}$. Then $f^*\colon V^* \to U^*$ has kernel $\ker f^* = \mathrm{Ann}(U)$. So by the first isomorphism theorem, it remains to show that f^* is surjective. To do this, we pick a basis $\{\vec{u}_1, \ldots, \vec{u}_k\}$ of U and extend it by $\vec{v}_{k+1}, \ldots, \vec{v}_n$ to a basis of V. Let $\{\theta_1, \ldots, \theta_n\}$ be the corresponding dual basis for V^* and $\{\psi_1, \ldots, \psi_k\}$ be the corresponding dual basis for U^*. Then $f^*(\theta_j) = \psi_j$ for $j \leq k$, so surjectivity follows.

HONOURS ALGEBRA
WORKSHOP 11 SOLUTIONS

The purpose of this workshop is to perform two different types of Jordan normal form calculations!

Exercise 1. The matrix with entries in $\mathbb{C}$

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 4 & -3 \\ 4 & 8 & -6 \end{pmatrix}$$

has characteristic polynomial $\chi_A(x) = x^2(-1-x)$. Find by hand an invertible matrix P such that $P^{-1}AP$ is in Jordan Normal Form.

Solution 1. A has nullspace spanned by $(1,1,2)^{\mathsf{T}}$. We calculate that

$$A^2 = \begin{pmatrix} -1 & -3 & 2 \\ -2 & -6 & 4 \\ -4 & -12 & 8 \end{pmatrix}$$

has nullspace spanned by $(2,0,1)^{\mathsf{T}}$ and $(1,1,2)^{\mathsf{T}}$. These two vectors therefore span $E^{\mathrm{gen}}(0,A)$.

$$A + I = \begin{pmatrix} 2 & 1 & -1 \\ 2 & 5 & -3 \\ 4 & 8 & -5 \end{pmatrix}$$

has nullspace spanned by $(1,2,4)^{\mathsf{T}}$, which therefore spans $E^{\mathrm{gen}}(-1,A)$.

So pick the basis $\mathcal{B} = \mathcal{B}_0 \cup \mathcal{B}_{-1}$, where

$$\mathcal{B}_{-1} = \left\{ \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} \right\},\ \mathcal{B}_0 = \left\{ \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, A \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\}.$$

Then the JNF is

$$P^{-1}AP = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

where

$$P = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 2 \\ 2 & 1 & 4 \end{pmatrix}.$$

Exercise 2. A (5×5)-matrix with entries in $\mathbb{C}$ has two eigenvalues: 0 and 1. How many possible Jordan Normal Forms (up to re-ordering) does the matrix have?

Solution 2. Along the diagonal of the JNF we must have one of the following four sequences:

(1) $(0,0,0,0,1)$,
(2) $(0,0,0,1,1)$,
(3) $(0,0,1,1,1)$, or
(4) $(0,1,1,1,1)$.

Date: Thursday, 28th November, 2013.

In case (1), the Jordan blocks corresponding to zero could have sizes 4; $3, 1$; $2, 2$; $2, 1, 1$; or $1, 1, 1, 1$.

For (2) the Jordan blocks corresponding to zero could have sizes 3; $2, 1$; or $1, 1, 1$, and the blocks corresponding the one could have sizes 2 or $1, 1$.

Case (3) is analogous to (2) with the roles of zero and one swapped.

Case (4) is analogous to (1) with the roles of zero and one swapped.

So the total number is $5 + (3 \times 2) + (3 \times 2) + 5 = 22$.