

# Mathematics for Machine Learning - Questions and Solutions

Edwin Fennell

Note - I'm not going to write out the questions here since they are very, very inefficiently posed and no way am I going to TeX all of that.

**2.1** a. In order to show that this constitutes a group, we need to show four things:

- Closure - if  $a, b \in \mathbb{R}$  then clearly  $ab + a + b \in \mathbb{R}$ . Now, suppose that  $ab + a + b = -1$ . This rearranges to

$$a(b + 1) = -(b + 1)$$

or

$$(a + 1)(b + 1) = 0$$

Therefore if neither  $a$  nor  $b$  is equal to  $-1$ ,  $a * b$  also cannot be equal to  $-1$ , and therefore  $*$  is a valid group operation on  $\mathbb{R} \setminus \{-1\}$ .

- Identity - our identity is  $0$  since for any  $a \in \mathbb{R} \setminus \{-1\}$  we have

$$a * 0 = a \cdot 0 + 0 + a = a$$

- Inverse - given a fixed  $a \in \mathbb{R} \setminus \{-1\}$  we want to solve for  $x$  in the following:

$$a * x = ax + x + a = 0$$

we rearrange to get

$$x = \frac{-a}{a + 1}$$

Therefore all elements in  $\mathbb{R} \setminus \{-1\}$  have inverses under  $*$

- Associativity - we consider the respective values of  $(a * b) * c$  and  $a * (b * c)$  for arbitrary  $a, b, c \in \mathbb{R} \setminus \{-1\}$ :

$$(a * b) * c = (a * b)c + a * b + c = abc + ac + bc + ab + a + b + c$$

$$a * (b * c) = a(b * c) + b * c + a = abc + ac + bc + ab + a + b + c$$

and so we have associativity.

Now we need to show that the resulting group is Abelian, but this is clear from the definition of  $*$  being completely symmetric in its two operands.

□

b. Conveniently, from our proof of associativity we know immediately that

$$3 * x * x = 3x^2 + 3x + 3x + x^2 + x + x + 3 = 4x^2 + 8x + 3$$

Therefore we need to solve  $4x^2 + 8x + 3 = 15$ , or rather

$$4x^2 + 8x - 12 = 4(x^2 + 2x - 3) = 4(x + 3)(x - 1) + 0$$

From this we see that the solutions are exactly  $x = 1, x = -3$

**2.2** a. We need to show the four group axioms:

- Closure - By definition of  $\oplus$  the result of its application is a congruence class mod  $n$ . (Well-posedness is another matter but that isn't asked for here).

- Identity - the identity is  $\bar{0}$  since

$$\forall a \in \mathbb{Z}, \bar{a} \oplus \bar{0} = \overline{(a + 0)} = \bar{a}$$

- Inverses - the inverse of  $\bar{a}$  for any  $a \in \mathbb{Z}$  is  $\overline{-a}$ :

$$\forall a \in \mathbb{Z}, \bar{a} \oplus \overline{-a} = \overline{(a - a)} = \bar{0}$$

- Associativity - we have

$$\forall a, b, c \in \mathbb{Z}, (\bar{a} \oplus \bar{b}) \oplus \bar{c} = \overline{(a + b)} \oplus \bar{c} = \overline{a + b + c}$$

and also

$$\forall a, b, c \in \mathbb{Z}, \bar{a} \oplus (\bar{b} \oplus \bar{c}) = \bar{a} \oplus \overline{(b + c)} = \overline{a + b + c}$$

and so we have associativity. Assuming that the operator  $\oplus$  is well-defined, this more or less comes down to "addition is associative".

Therefore  $(\mathbb{Z}_n, \oplus)$  is indeed a group.

b. I'm not going to write out the multiplication table for  $\mathbb{Z}_5 \setminus \{\bar{0}\}$ . I will show that this is a group when I prove the general case in part d of this question. Assuming that it is a group, it is clearly Abelian from the symmetric nature of  $\otimes$ .

c. Again, I'll use the result from part d. 8 is composite so this is not a group.

d. Suppose that  $n$  is composite. Then  $\exists a, b$  s.t.  $1 < a, b < n$  and  $a \cdot b = n$ . Therefore we have

$$\bar{a} \otimes \bar{b} = \bar{n} = \bar{0}$$

Therefore  $\mathbb{Z}_n \setminus \{\bar{0}\}$  is not a group since it fails the requirement of closure.

Now, if  $n$  is instead prime, then  $\mathbb{Z}_n \setminus \{\bar{0}\}$  is a group - we will show this by verifying the group axioms.

- Closure - suppose that  $a, b \in \mathbb{Z} \setminus \{\bar{0}\}$ . Now, suppose that

$$ab \equiv 0 \pmod{n}$$

Then  $ab = kn$  for some  $k \in \mathbb{Z}$ . Since  $n$  is prime,  $a$  and  $n$  are coprime, and therefore by Bezout's theorem,  $\exists u, v \in \mathbb{Z}$  s.t.

$$ua + vn = 1$$

Therefore

$$b = b \cdot 1 = b(ua + vn) = ab \cdot u + bvn = (uk + bv)n$$

and so we find that  $b$  is a multiple of  $n$ . This is a contradiction since  $b \in \mathbb{Z} \setminus \{\bar{0}\}$ . Therefore  $ab \not\equiv 0 \pmod{n}$  and we have

$$\bar{a}, \bar{b} \neq \bar{0} \implies \bar{a}\bar{b} \neq \bar{0}$$

and so we have closure

- Identity - the identity is trivially  $\bar{1}$
- Inverse - for any  $\bar{a} \neq \bar{0}$  we have that  $a$  and  $n$  are coprime. By Bezout's theorem we know that  $\exists u, v \in \mathbb{Z}$  s.t.

$$ua + vn = 1$$

Therefore

$$\bar{a} \otimes \bar{u} = \overline{au} = \overline{(1 - vn)} = \bar{1}$$

and so we have constructed an inverse for  $\bar{a}$

- Associativity - exactly the same proof as in part a. Essentially "multiplication is associative".

Therefore  $(\mathbb{Z}_n, \otimes)$  is indeed a group.

### 2.3 Let's check the four group requirements:

- Closure -  $\forall x_1, y_1, z_1, x_2, y_2, z_2 \in \mathbb{R}$  we have

$$\begin{pmatrix} 1 & x_1 & z_1 \\ 0 & 1 & y_1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & x_2 & z_2 \\ 0 & 1 & y_2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x_1 + x_2 & x_1y_2 + z_1 + z_2 \\ 0 & 1 & y_1 + y_2 \\ 0 & 0 & 1 \end{pmatrix}$$

so we have closure

- Identity - from the above calculation (or simply by knowing what the identity matrix is) we see that

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathcal{G}$$

and acts as the identity.

- Inverses - we see that  $\forall x, y, z \in \mathbb{R}$  we have

$$\begin{pmatrix} 1 & x_1 & z_1 \\ 0 & 1 & y_1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -x_1 & x_1 y_1 - z_1 \\ 0 & 1 & -y_1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- Associativity - we can just show this manually. Let

$$x_1, y_1, z_1, x_2, y_2, z_2, x_3, y_3, z_3 \in \mathbb{R}$$

Then

$$\begin{aligned} & \left( \begin{pmatrix} 1 & x_1 & z_1 \\ 0 & 1 & y_1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & x_2 & z_2 \\ 0 & 1 & y_2 \\ 0 & 0 & 1 \end{pmatrix} \right) \cdot \begin{pmatrix} 1 & x_3 & z_3 \\ 0 & 1 & y_3 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & x_1 + x_2 & x_1 y_2 + z_1 + z_2 \\ 0 & 1 & y_1 + y_2 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & x_3 & z_3 \\ 0 & 1 & y_3 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & x_1 + x_2 + x_3 & x_1 y_2 + x_1 y_3 + x_2 y_3 + z_1 + z_2 + z_3 \\ 0 & 1 & y_1 + y_2 + y_3 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

and

$$\begin{aligned} & \begin{pmatrix} 1 & x_1 & z_1 \\ 0 & 1 & y_1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \left( \begin{pmatrix} 1 & x_2 & z_2 \\ 0 & 1 & y_2 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & x_3 & z_3 \\ 0 & 1 & y_3 \\ 0 & 0 & 1 \end{pmatrix} \right) \\ &= \begin{pmatrix} 1 & x_1 & z_1 \\ 0 & 1 & y_1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & x_2 + x_3 & x_2 y_3 + z_2 + z_3 \\ 0 & 1 & y_2 + y_3 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & x_1 + x_2 + x_3 & x_1 y_2 + x_1 y_3 + x_2 y_3 + z_1 + z_2 + z_3 \\ 0 & 1 & y_1 + y_2 + y_3 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

and so we have associativity.

Therefore  $(\mathcal{G}, \cdot)$  is a group. It is not Abelian though. We can see this from observing that

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

and

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

**2.4** The first product is not computable since the column count of the first matrix is not equal to the row count of the second. All the other products are valid. I will not compute them here.

**2.5** We could just ask a linear equation solver to do this for us, but I think that the point of this question is to understand the solving process. In each case we use row operations to put the "query matrix" into echelon row form. Note that we can combine the ability to multiply rows by non-zero constants and adding a row to another row to simply add a non-zero multiple of a row to a different row.

a. We need to solve

$$\begin{pmatrix} 1 & 1 & -1 & -1 \\ 2 & 5 & -7 & -5 \\ 2 & -1 & 1 & 3 \\ 5 & 2 & -4 & 2 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ -2 \\ 4 \\ 6 \end{pmatrix}$$

We add multiples of row 1 to all other rows in order to force as many zeros as possible in the first column. giving

$$\begin{pmatrix} 1 & 1 & -1 & -1 \\ 0 & 3 & -5 & -3 \\ 0 & -3 & 3 & 5 \\ 0 & -3 & 1 & 7 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ -4 \\ 2 \\ 1 \end{pmatrix}$$

We repeat for row 2 and the second column:

$$\begin{pmatrix} 1 & 1 & -1 & -1 \\ 0 & 3 & -5 & -3 \\ 0 & 0 & -2 & 2 \\ 0 & 0 & -4 & 4 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ -4 \\ -2 \\ -3 \end{pmatrix}$$

and then row 3 and the third column:

$$\begin{pmatrix} 1 & 1 & -1 & -1 \\ 0 & 3 & -5 & -3 \\ 0 & 0 & -2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ -4 \\ -2 \\ 1 \end{pmatrix}$$

Understandably,  $0 = 1$  has no solutions, so there are no solutions to this equation.

b. We need to solve

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 1 \\ 1 & 1 & 0 & -3 & 0 \\ 2 & -1 & 0 & 1 & -1 \\ -1 & 2 & 0 & -2 & -1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 3 \\ 6 \\ 5 \\ 1 \end{pmatrix}$$

Again we do the whole "row reduction" thing. We add multiples of the first row the other rows in order to make a bunch of zeroes in the first column:

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 1 \\ 0 & 2 & 0 & -3 & -1 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 1 & 0 & -2 & 0 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 3 \\ 3 \\ -1 \\ 4 \end{pmatrix}$$

Then we swap the second and third rows:

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 2 & 0 & -3 & -1 \\ 0 & 1 & 0 & -2 & 0 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 3 \\ -1 \\ 3 \\ 4 \end{pmatrix}$$

Then we use the second row to introduce zeros into the second column:

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 0 & 0 & -5 & 5 \\ 0 & 0 & 0 & -3 & 3 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 3 \\ -1 \\ 5 \\ 5 \end{pmatrix}$$

Multiply the third row by 3 and the last row by 5:

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 0 & 0 & -15 & 15 \\ 0 & 0 & 0 & -15 & 15 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 3 \\ -1 \\ 15 \\ 25 \end{pmatrix}$$

and then subtract the third row from the 4th to give

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 0 & 0 & -15 & 15 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 3 \\ -1 \\ 15 \\ 10 \end{pmatrix}$$

Therefore any solution to this equation would require  $10 = 0$ , and thus there are no solutions.

## 2.6 We need to solve

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

We can pretty straightforwardly subtract the first row from the third, giving

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

Dealing with inhomogeneity is annoying though, so a simpler practice would be to solve the homogeneous

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

It doesn't take too much imagination here to see that the vectors

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ -1 \\ 1 \\ 1 \end{pmatrix}$$

are linearly independent and lie in the kernel of this matrix. It also doesn't take too much imagination to see that

$$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

are linearly independent, and that the span of their images under  $\mathbf{A}$  is all of  $\mathbb{R}^3$ . Therefore by rank-nullity theorem the kernel of  $A$  is exactly

$$\text{span} \left( \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ -1 \\ 1 \\ 1 \end{pmatrix} \right)$$

Now that we have solved the homogenous equation, we need a specific solution to

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

in order to get our general solution. We can see from our construction of a basis of  $\text{Im}(\mathbf{A})$  that

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

is such a solution. Therefore we express the general solution as

$$\left\{ \begin{pmatrix} 0 \\ 2 \\ 0 \\ -1 \\ 0 \\ -1 \end{pmatrix} + \mathbf{y}, \mathbf{y} \in \text{span} \left( \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ -1 \\ 1 \\ 1 \end{pmatrix} \right) \right\}$$

**2.7** Note that both of the equations posed are linear, so we can combine them into a single linear system describing our three unknowns. This is a little fiddly so I will do so in several separate steps.

We see that

$$\mathbf{A}\mathbf{x} = 12\mathbf{x}$$

rearranges to

$$(\mathbf{A} - 12\mathbf{I})\mathbf{x} = \mathbf{0}$$

Our second requirement can be restated as

$$\begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \mathbf{x} = 1$$

We can combine these into a single linear system, giving

$$\begin{pmatrix} 1 & 1 & 1 \\ -6 & 4 & 3 \\ 6 & -12 & 9 \\ 0 & 8 & -12 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

We will use Gaussian elimination to solve this. Reducing the first column:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 10 & 9 \\ 0 & -18 & 3 \\ 0 & 8 & -12 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 6 \\ -6 \\ 0 \end{pmatrix}$$

Swapping rows 2 and 4, then dividing row 2 through by 4:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & -3 \\ 0 & -18 & 3 \\ 0 & 10 & 9 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ -6 \\ 6 \end{pmatrix}$$

Now we reduce column 2:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & -3 \\ 0 & 0 & -24 \\ 0 & 0 & 24 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ -6 \\ 6 \end{pmatrix}$$



And then add the 3rd row to the 4th to remove it altogether:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & -3 \\ 0 & 0 & -24 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ -6 \\ 0 \end{pmatrix}$$

Or more succinctly:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & -3 \\ 0 & 0 & -24 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ -6 \end{pmatrix}$$

The matrix on the left is upper triangular, and so we can clearly see that the determinant is non-zero, meaning that the matrix is invertible. Computing the inverse is not mathematically interesting so I will let the computer do the work, revealing the unique solution to our equation as

$$\mathbf{x} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & -3 \\ 0 & 0 & -24 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \\ -6 \end{pmatrix} = \frac{1}{8} \cdot \begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix}$$

**2.8** a. The rows of this matrix are not independent - the second row is the arithmetic average of the other two. Therefore the matrix is non-invertible.

b. This one is not quite so obvious from the outset. We could easily find out whether or not the determinant is zero by plugging it into Wolfram Alpha, but let's do this manually instead. Row operations do not change the row space of the matrix, and we will use this to our advantage.

We start by adding multiples of the first row to everything to "zero out" the first column:

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

Now we do the same for the second row and column:

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

Now we switch rows 3 and 4:

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

And now we reduce the third column using the third row:

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

The determinant of this upper triangular matrix is -1, so it is invertible. Since the row space is unchanged, the original matrix was also invertible.

At this point I have no qualms about getting Wolfram Alpha to compute the inverse for me, since I know that it exists. Plugging constants into the lengthy formula for a 4x4 inverse is not an instructive or enjoyable exercise.

Thus we find the inverse to be

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

- 2.9** a. Despite appearances, this is a vector space. Note that  $f : \mathbb{R} \rightarrow \mathbb{R}$  given by  $f(x) = f(x^3)$  is a surjection. Therefore we can rewrite

$$A = \{(x, x+y, x-y) | x, y \in \mathbb{R}\}$$

or more succinctly:

$$A = \left\{ x \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \mid x, y \in \mathbb{R} \right\}$$

which is a 2D plane through the origin and thus clearly a subspace of  $\mathbb{R}^3$

b. This is not a vector space.  $(1, -1, 0) \in B$  but  $-1 * (1, -1, 0) = (-1, 1, 0) \notin B$

c. This is a subspace iff  $\gamma = 0$ . If  $\gamma \neq 0$  then  $(0, 0, 0) \notin C$ , and having an additive identity is a requirement of a vector space.

If  $\gamma = 0$  then we can pick  $\xi_2$  and  $\xi_3$  freely, at which point  $\xi_1$  is uniquely determined as  $2\xi_2 - 3\xi_3$ . Thus we can write  $C$  as

$$A = \left\{ x \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \mid x, y \in \mathbb{R} \right\}$$

which is a plane through the origin.  $C$  is a subspace in this case.

d.  $D$  is not a subspace since it is not closed under scalar multiplication in  $\mathbb{R}$ .  $(1, 1, 1) \in D$ , but  $\frac{1}{2}(1, 1, 1) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}) \notin D$

**2.10** a. The vectors are linearly dependent:  $x_3 - 2x_1 = x_2$  b. These vectors are linearly independent. Suppose that we have

$$a_1x_1 + a_2x_2 + a_3x_3 = 0$$

Looking at dimension 3, we see that necessarily  $a_1 = 0$ . Then, looking at dimension 2, we see that this forces  $a_2 = 0$ , which also forces  $a_3 = 0$ .

**2.11** We want to solve

$$a_1x_1 + a_2x_2 + a_3x_3 = y$$

which we can rewrite as

$$\begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & -1 \\ 1 & 3 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = y$$

If we check we can see that the matrix on the left is invertible. We can therefore compute its inverse and write

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 5 & 5 & -5 \\ -2 & -1 & 3 \\ 1 & -2 & 1 \end{pmatrix} y = \begin{pmatrix} -6 \\ 3 \\ 2 \end{pmatrix}$$

Therefore we have

$$-6x_1 + 3x_2 + 2x_3 = y$$

**2.12** Suppose we want to find the generic form of a point  $x \in U_1 \cap U_2$ . Then we can write

$$x = \begin{pmatrix} 1 & 2 & -1 \\ 1 & -1 & 1 \\ -3 & 0 & -1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} -1 & 2 & -3 \\ -2 & -2 & 6 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

for some  $a_1, a_2, a_3, b_1, b_2, b_3$ . We are not however free to pick these 6 values independently however we want - we do actually need the above equality to hold.

So we need

$$\begin{pmatrix} 1 & 2 & -1 \\ 1 & -1 & 1 \\ -3 & 0 & -1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} -1 & 2 & -3 \\ -2 & -2 & 6 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

which we can rewrite as

$$\begin{pmatrix} 1 & 2 & -1 & -1 & 2 & -3 \\ 1 & -1 & 1 & -2 & -2 & 6 \\ -3 & 0 & -1 & 2 & 0 & -2 \\ 1 & -1 & 1 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ b_1 \\ b_2 \\ b_3 \end{pmatrix} = 0$$

We can now solve this by Gaussian elimination. We reduce column 1 with multiples of the first row:

$$\begin{pmatrix} 1 & 2 & -1 & -1 & 2 & -3 \\ 0 & -3 & 2 & -1 & -4 & 9 \\ 0 & 6 & -4 & -1 & 6 & -11 \\ 0 & -3 & 2 & 2 & -2 & 2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ b_1 \\ b_2 \\ b_3 \end{pmatrix} = 0$$

and now we reduce column 2 with multiples of the second row:

$$\begin{pmatrix} 1 & 2 & -1 & -1 & 2 & -3 \\ 0 & -3 & 2 & -1 & -4 & 9 \\ 0 & 0 & 0 & -3 & -2 & 7 \\ 0 & 0 & 0 & 3 & 2 & -7 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ b_1 \\ b_2 \\ b_3 \end{pmatrix} = 0$$

and now we reduce column 4 with multiples of the third row:

$$\begin{pmatrix} 1 & 2 & -1 & -1 & 2 & -3 \\ 0 & -3 & 2 & -1 & -4 & 9 \\ 0 & 0 & 0 & -3 & -2 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ b_1 \\ b_2 \\ b_3 \end{pmatrix} = 0$$

We can now work out the full set of solutions from this echelon row form. There is a general methodology here that I will lay out rigorously at the end of this chapter, but for now I will just execute this methodology.

Actually, in this case we don't even have to solve this entire set of linear equations. As soon as we solve for  $b_1, b_2, b_3$  we have our completely general expression for  $x$ .

We start from the last non-zero row of the matrix and resolve each row in turn:

- We have  $-3b_1 - 2b_2 + 7b_3 = 0$ , which allows us to pick  $b_1, b_2$  freely, and then this forces  $b_3 = \frac{3b_1 + 2b_2}{7}$

We can now stop since we have determined all possible value combinations for  $b_1, b_2, b_3$ , and so we have the general form for  $x$  as

$$\begin{pmatrix} -1 & 2 & -3 \\ -2 & -2 & 6 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \frac{3b_1 + 2b_2}{7} \end{pmatrix} = \frac{1}{7} \begin{pmatrix} -16b_1 + 8b_2 \\ 4b_1 - 2b_2 \\ 8b_1 - 4b_2 \\ 4b_1 - 2b_2 \end{pmatrix} = (2b_1 - b_2) \begin{pmatrix} -8 \\ 2 \\ 4 \\ 2 \end{pmatrix}$$

We can select  $b_1, b_2$  to force  $2b_1 - b_2$  to take any value we like. Therefore the singleton set containing the vector

$$\begin{pmatrix} -8 \\ 2 \\ 4 \\ 2 \end{pmatrix}$$

is a basis for  $U_1 \cap U_2$ .

- 2.13** a. The easiest way to work this out is to look at the column space, which is exactly the image space. Then via rank-nullity theorem

$$\dim(\ker(A_1)) + \dim(\text{Im}(A_1)) = \dim(\text{Preim}(A_1)) = 3$$

For  $A_1$  we see that the third column is the sum of the first two, and that the first two columns are not linearly dependent. Therefore  $\dim(\text{Im}(A_1)) = 2$ , and  $\dim(\ker(A_1)) = \dim(U_1) = 1$ .

For  $A_2$ , we also see that the third column is the sum of the first two, and that the first two columns are not linearly dependent. Therefore  $\dim(\text{Im}(A_2)) = 2$ , and  $\dim(\ker(A_2)) = \dim(U_2) = 1$ .

- b. From our observations before, we see that

$$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

lies in both  $U_1$  and  $U_2$ . The dimension of each of these spaces is 1, so therefore the singleton set containing only this vector acts as a basis for both  $U_1$  and  $U_2$ .

c. We have  $U_1 = U_2 = U_1 \cap U_2$ , so

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\}$$

is also a basis for  $U_1 \cap U_2$

**2.14** a. We already know that the column space has dimension 2 for each of the above matrices:  $\dim(U_1) = \dim(U_2) = 2$  b. We also know from the previous question that the first two columns of each matrix generate its column space. So

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} \right\}$$

is a basis of  $U_1$  and

$$\left\{ \begin{pmatrix} 3 \\ 1 \\ 7 \\ 3 \end{pmatrix}, \begin{pmatrix} -3 \\ 2 \\ -5 \\ -1 \end{pmatrix} \right\}$$

is a basis of  $U_2$ .

c. We note that

$$\begin{pmatrix} 3 \\ 1 \\ 7 \\ 3 \end{pmatrix}$$

lies in  $U_1$  but that

$$\begin{pmatrix} -3 \\ 2 \\ -5 \\ -1 \end{pmatrix}$$

does not. Therefore the dimension of  $U_1 \cap U_2$  is at least 1 but less than 2, since it does not contain all of  $U_2$ . Therefore the dimension is exactly 1 and we can write our basis of  $U_1 \cap U_2$  as

$$\left\{ \begin{pmatrix} 3 \\ 1 \\ 7 \\ 3 \end{pmatrix} \right\}$$

**2.15** a. In order to see whether a set is a subspace of a space, we need to check closure under addition and scalar multiplication - all the other requirements for being a vector space are inherited from the parent space.

For  $F$  we can check these two properties directly:

- Closure under addition - suppose that we have

$$(x_1, y_1, z_1), (x_2, y_2, z_2) \in F$$

Then we also have

$$(x_1 + x_2, y_1 + y_2, z_1 + z_2) \in F$$

since

$$(x_1 + x_2) + (y_1 + y_2) - (z_1 + z_2) = (x_1 + y_1 - z_1) + (x_2 + y_2 - z_2) = 0 + 0 = 0$$

- Closure under scalar multiplication - suppose that we have

$$(x, y, z) \in F$$

Then for any  $\lambda \in \mathbb{R}$  we have

$$(\lambda x, \lambda y, \lambda z) \in F$$

since

$$(\lambda x + \lambda y - \lambda z) = \lambda(x + y - z) = \lambda \cdot 0 = 0$$

With  $G$  it is a lot easier to see from the construction that it is a 2-dimensional subspace and work out a fairly obvious basis, but we'll handle it in the same way we did  $F$  for the sake of practice.

- Closure under addition - suppose that we have (completely generally)

$$(a_1 - b_1, a_1 + b_1, a_1 - 3b_1), (a_2 - b_2, a_2 + b_2, a_2 - 3b_2) \in G$$

then the sum of these two elements, which we call  $v$ , is also in  $G$ . We can see this since if we set  $A = a_1 + a_2$  and  $B = b_1 + b_2$ , then we have

$$v = (A - B, A + B, A - 3B)$$

and so  $v \in G$ .

- Closure under scalar multiplication - suppose that we have

$$(a - b, a + b, a - 3b) \in G$$

and some  $\lambda \in \mathbb{R}$ . Set  $A = \lambda a$ ,  $B = \lambda b$  and we see that

$$(\lambda(a - b), \lambda(a + b), \lambda(a - 3b)) = (A - B, A + B, A - 3B) \in G$$

Therefore both  $F$  and  $G$  are subspaces of  $\mathbb{R}^3$ .

b. Pick a completely general element of  $G$ , namely  $(a - b, a + b, a - 3b)$  for some  $a, b \in \mathbb{R}$ . This element also lies in  $F$ , and therefore in  $F \cap G$ , exactly when  $a - b + a + b - a + 3b = 0$ , or rather  $a = -3b$ . Therefore the general element of  $F \cap G$  can be written as  $\{(-4b, -2b, -6b) | b \in \mathbb{R}\}$ .

c. Really? Ok then ...

We can see that  $F$  is a 2D space. It contains the mutually linearly independent vectors  $(1, 0, 1)$  and  $(1, 1, 2)$ , and doesn't include all of  $\mathbb{R}^3$ . Therefore  $\{(1, 0, 1), (1, 1, 2)\}$  is in fact a basis of  $F$ .

The way that  $G$  is defined makes it very clear that  $\{(1, 1, 1), (-1, 1, -3)\}$  is a basis of  $G$ .

Now that we have a basis for  $F$  and  $G$  we can say that the general element  $v$  of  $F \cap G$  satisfies

$$v = a \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + b \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = c \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + d \begin{pmatrix} -1 \\ 1 \\ -3 \end{pmatrix}$$

We want to find all possible values for  $a, b, c, d$ . To do this we rewrite the above equation as

$$\begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & -1 & -1 \\ 1 & 2 & -1 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = 0$$

Now we use row operations to put this into echelon row form. We reduce the first column:

$$\begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = 0$$

Then the second:

$$\begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = 0$$

From this we see that we have completely free choice over the value of  $d$ , at which point  $a, b, c$  are completely determined. In particular we see that  $c = -3d$ , and so our general form for  $v$  is

$$v = c \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + d \begin{pmatrix} -1 \\ 1 \\ -3 \end{pmatrix} = -3d \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + d \begin{pmatrix} -1 \\ 1 \\ -3 \end{pmatrix} = d \begin{pmatrix} -4 \\ -2 \\ -6 \end{pmatrix}$$

which is exactly the answer we got in part b.

- 2.16** a. Yes - integration distributes over addition and commutes with scalar multiplication  
b. Yes - differentiation distributes over addition and commutes with scalar multiplication



- c. No -  $\cos(0) \neq 0$  so  $\forall x, \cos(x+0) \neq \cos(x) + \cos(0)$
- d. Yes - matrix multiplication distributes over addition and commutes with scalar multiplication
- e. No -  $\Phi(x) \neq 0$  (where in this case 0 is notation abuse meaning the 2x2 zero matrix) so we have the same problem as in part c.

**2.17** a. The transformation matrix is

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

b. The columns of this matrix form a linearly independent set, so the rank of the transformation is 3.

c. We already know that

$$\left\{ \begin{pmatrix} 3 \\ 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -3 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

is a basis for our image, which has dimension 3. By rank-nullity we see that the dimension of the kernel must be 0, and therefore is the trivial subspace.

**2.18**  $f$  and  $g$  are both automorphisms so they both have trivial kernel and full image, as do  $f \circ g$  and  $g \circ f$ . Therefore all the claimed statements are vacuously true.

**2.19** a. This square matrix has determinant  $-2 \neq 0$ , and therefore has full rank and trivial kernel.

b. Here I am going to go out of my way and explain the whole "basis change" thing in detail in order to achieve a clear understanding.

In essence a matrix represents a linear map between two vector spaces. But it's a little more strict than that - a matrix defines a linear map between two vector spaces using a basis of each space. By specifying where each of the basis vectors of the domain space go in terms of our image basis, we define our transformation matrix.

Let's be more specific to this case - we have the same basis for both our domain and image, which we can write as  $\{v_1, v_2, v_3\}$ . Pick an arbitrary vector in our domain  $x = a_1v_1 + a_2v_2 + a_3v_3$ . Then we find unique  $b_1, b_2, b_3$  s.t.

$$A_\Phi x = A_\Phi(a_1v_1 + a_2v_2 + a_3v_3) = b_1v_1 + b_2v_2 + b_3v_3$$

Now, let  $B$  denote the invertible square matrix with  $v_1, v_2, v_3$  as its columns. We can use this to immediately rewrite the above equation as

$$A_{\Phi}B \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = B \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$B$  is invertible so we can pre-multiply by  $B^{-1}$  to get

$$B^{-1}A_{\Phi}B \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

The above construction is completely general, and thus we have found our target matrix  $\bar{A}_{\Phi} = B^{-1}A_{\Phi}B$ , since it defines where every vector in the domain, expressed in terms of our chosen basis, is mapped to in the image, also expressed in terms of our chosen basis.

**2.20** a. Each pair of vectors is clearly linearly independent and thus span  $\mathbb{R}^2$ . I'm not going to draw them because I can't be bothered with TeX diagrams.

b. I could just write the answer here, but this is again an opportunity for detail and clarity. Suppose we have two bases  $B$  and  $B'$  of a space  $A$ , and we pick an arbitrary point  $v \in A$ . Then let  $x_1, \dots, x_n$  be the coefficients of  $v$  in  $B$  and  $y_1, \dots, y_n$  be the coefficients of  $v$  in  $B'$ . Then if we define some  $M_i$  as the square matrices with their columns as the basis vectors of  $B_i$ , we can see that

$$M_1 \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} = M_2 \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix}$$

and therefore

$$\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} = M_1^{-1}M_2 \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix}$$

Our  $v$  was arbitrary so this holds completely generally, and thus our "basis change" matrix from  $B'$  to  $B$  is  $M_1^{-1}M_2$ .

In our specific case we have

$$M_1 = \begin{pmatrix} 2 & -1 \\ 1 & -1 \end{pmatrix}, M_2 = \begin{pmatrix} 2 & 1 \\ -2 & 1 \end{pmatrix}$$

and so we compute our basis change matrix as

$$M_1^{-1}M_2 = \begin{pmatrix} 4 & 0 \\ 6 & -1 \end{pmatrix}$$

- c. i. I'm not doing this yet again. It's a pretty easy computation.  
 c. ii. Here we get payoff from working out the completely general result in part b. Our specific values for  $M_1$  and  $M_2$  in this case are

$$M_1 = I, M_2 = \begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \\ -1 & 2 & -1 \end{pmatrix}$$

and so our basis change matrix is just  $M_2$ . (In fact the basis change matrix from any arbitrary basis to the standard basis is just the matrix with the vectors of the original basis making up the columns)

- d. To work out  $A_\Phi$  w.r.t  $B$  and  $C$ , we should work out where  $\Phi$  actually sends all the basis vectors of  $B$ . Since  $\Phi$  is linear we see that

$$\Phi(b_1) = \frac{1}{2}(\Phi(b_1 + b_2) + \Phi(b_1 - b_2)) = c_1 + 2c_3$$

and

$$\Phi(b_2) = \frac{1}{2}(\Phi(b_1 + b_2) - \Phi(b_1 - b_2)) = -c_1 + c_2 - c_3$$

Therefore the transformation matrix  $A_\Phi$  w.r.t  $B$  as our domain basis and  $C$  as our image basis is

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

- e. We can think of the linear map in the required bases as the composition of several transforms that we have already computed. First convert our domain representation from basis  $B'$  to basis  $B$ , then apply the transform  $\Phi$ , and then convert our image representation from basis  $C$  to basis  $C'$ . We have already computed all three of these matrices and so we can just compose them in order to get the required transformation matrix as

$$\begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \\ -1 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 6 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -10 & 3 \\ 12 & -4 \end{pmatrix}$$

- f. I'm not going through this rigmarole. The first three parts just involve applying the three constituent transformations in turn. The fourth is just applying the combined transform all at once.

## Key points for Chapter 2

- We find it beneficial to define the concept of a vector space. This is a bunch of "things" with an addition function and scalar multiplication (over some field), with closure under each. This is really abstract and general but whatever - this is just the maths that we're developing to use later.

- We can introduce the concept of dimension - what is the smallest number of "vectors" in this space that we need to be able to generate the entire space with linear combinations thereof. There are absolutely infinite-dimensional vector spaces (consider the field  $\mathbb{Q}$  over the vector space  $\mathbb{R}$ ) but we don't care about them. We call any of these sets of generating vectors a "basis".
- In maths we don't just care about things - we care about maps between things. We can define a linear map as something that "respects linearity" - i.e. we map from one vector space to another, both over the same field, with a map that distributes over addition and commutes with scalar multiplication.
- Suppose we have