

# 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

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