Vectors & Matrices

Solutions to Problem Sheet 9

1. (i) By direct computation, we find

$$\begin{aligned}(1-\sqrt[3]{2}+2\sqrt[3]{4})(a+b\sqrt[3]{2}+c\sqrt[3]{4}) &= a(1-\sqrt[3]{2}+2\sqrt[3]{4})+b(\sqrt[3]{2}-\sqrt[3]{2}\sqrt[3]{2}+2\sqrt[3]{2}\sqrt[3]{4})\\ &\quad +c(\sqrt[3]{4}-\sqrt[3]{2}\sqrt[3]{4}+2\sqrt[3]{4}\sqrt[3]{4})\\ &= a(1-\sqrt[3]{2}+2\sqrt[3]{4})+b(\sqrt[3]{2}-\sqrt[3]{4}+4)+c(\sqrt[3]{4}-2+4\sqrt[3]{2})\\ &= (a+4b-2c)+(-a+b+4c)\sqrt[3]{2}+(2a-b+c)\sqrt[3]{4}\,.\end{aligned}$$

Per the question, this value can be represented by the vector

$$\begin{pmatrix} a+4b-2c \\ -a+b+4c \\ 2a-b+c \end{pmatrix} \in \mathbb{Q}^3\,.$$

Since each component of this column vector has linear dependence on the values $a, b, c \in \mathbb{Q}$, there exists a 3×3 matrix A such that

$$A\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a+4b-2c \\ -a+b+4c \\ 2a-b+c \end{pmatrix}\,. \tag{1}$$

In particular, if we take

$$A = \begin{pmatrix} 1 & 4 & -2 \\ -1 & 1 & 4 \\ 2 & -1 & 1 \end{pmatrix}\,,$$

then by the definition of matrix multiplication, we have

$$\begin{pmatrix} 1 & 4 & -2 \\ -1 & 1 & 4 \\ 2 & -1 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a+4b-2c \\ -a+b+4c \\ 2a-b+c \end{pmatrix}\,.$$

Thus, A satisfies the equation (1), and therefore encodes the effect of multiplying a value $a+b\sqrt[3]{2}+c\sqrt[3]{4}$ by $1-\sqrt[3]{2}+2\sqrt[3]{4}$.

(ii) We construct the augmented matrix $(A|I)$ and perform the following elementary row operations:

$$\left(\begin{array}{ccc|ccc} 1 & 4 & -2 & 1 & 0 & 0 \\ -1 & 1 & 4 & 0 & 1 & 0 \\ 2 & -1 & 1 & 0 & 0 & 1 \end{array}\right) \sim R_2 + R_1 \left(\begin{array}{ccc|ccc} 1 & 4 & -2 & 1 & 0 & 0 \\ 0 & 5 & 2 & 1 & 1 & 0 \\ 2 & -1 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\sim {}^1\!/_5 R_2 \left(\begin{array}{ccc|ccc} 1 & 4 & -2 & 1 & 0 & 0 \\ 0 & 1 & {}^2\!/_5 & {}^1\!/_5 & {}^1\!/_5 & 0 \\ 2 & -1 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\sim \; R_3 - 2R_1 \left(\begin{array}{ccc|ccc} 1 & 4 & -2 & 1 & 0 & 0 \\ 0 & 1 & {}^2\!/_5 & {}^1\!/_5 & {}^1\!/_5 & 0 \\ 0 & -9 & 5 & -2 & 0 & 1 \end{array}\right)$$

$$\sim \; R_3 + 9R_2 \left(\begin{array}{ccc|ccc} 1 & 4 & -2 & 1 & 0 & 0 \\ 0 & 1 & {}^2\!/_5 & {}^1\!/_5 & {}^1\!/_5 & 0 \\ 0 & 0 & {}^{43}\!/_5 & -{}^1\!/_5 & {}^9\!/_5 & 1 \end{array}\right)$$

$$\sim \; {}^5\!/_{43} R_3 \left(\begin{array}{ccc|ccc} 1 & 4 & -2 & 1 & 0 & 0 \\ 0 & 1 & {}^2\!/_5 & {}^1\!/_5 & {}^1\!/_5 & 0 \\ 0 & 0 & 1 & -{}^1\!/_{43} & {}^9\!/_{43} & {}^5\!/_{43} \end{array}\right)$$

$$\sim R_2 - {}^2\!/_5 R_3 \left(\begin{array}{ccc|ccc} 1 & 4 & -2 & 1 & 0 & 0 \\ 0 & 1 & 0 & {}^9\!/_{43} & {}^5\!/_{43} & -{}^2\!/_{43} \\ 0 & 0 & 1 & -{}^1\!/_{43} & {}^9\!/_{43} & {}^5\!/_{43} \end{array}\right)$$

$$\sim \; R_1 + 2R_3 \left(\begin{array}{ccc|ccc} 1 & 4 & 0 & {}^{41}\!/_{43} & {}^{18}\!/_{43} & {}^{10}\!/_{43} \\ 0 & 1 & 0 & {}^9\!/_{43} & {}^5\!/_{43} & -{}^2\!/_{43} \\ 0 & 0 & 1 & -{}^1\!/_{43} & {}^9\!/_{43} & {}^5\!/_{43} \end{array}\right)$$

$$\sim \; R_1 - 4R_2 \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & {}^5\!/_{43} & -{}^2\!/_{43} & {}^{18}\!/_{43} \\ 0 & 1 & 0 & {}^9\!/_{43} & {}^5\!/_{43} & -{}^2\!/_{43} \\ 0 & 0 & 1 & -{}^1\!/_{43} & {}^9\!/_{43} & {}^5\!/_{43} \end{array}\right).$$

Therefore, the inverse A^{-1} is given by

$$A^{-1} = \begin{pmatrix} 5/43 & -2/43 & 18/43 \\ 9/43 & 5/43 & -2/43 \\ -1/43 & 9/43 & 5/43 \end{pmatrix} = \frac{1}{43} \begin{pmatrix} 5 & -2 & 18 \\ 9 & 5 & -2 \\ -1 & 9 & 5 \end{pmatrix} .$$

(iii) To express

$$\frac{3 + 6\sqrt[3]{2} + 7\sqrt[3]{4}}{1 - \sqrt[3]{2} + 2\sqrt[3]{4}}$$

in the form $a + b\sqrt[3]{2} + c\sqrt[3]{4}$ is equivalent to finding $a, b, c \in \mathbb{Q}$ such that

$$(1 - \sqrt[3]{2} + 2\sqrt[3]{4})(a + b\sqrt[3]{2} + c\sqrt[3]{4}) = 3 + 6\sqrt[3]{2} + 7\sqrt[3]{4} .$$

In part (i), we showed that this is equivalent to finding a vector

$$\mathbf{x} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

such that

$$A\mathbf{x} = \begin{pmatrix} 1 & 4 & -2 \\ -1 & 1 & 4 \\ 2 & -1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \\ 7 \end{pmatrix} .$$

Since we evaluated the inverse of A in part (ii), we can multiply both sides of the above equality by this A^{-1} to get

$$\begin{aligned} \mathbf{x} = A^{-1} \begin{pmatrix} 3 \\ 6 \\ 7 \end{pmatrix} = \frac{1}{43} \begin{pmatrix} 5 & -2 & 18 \\ 9 & 5 & -2 \\ -1 & 9 & 5 \end{pmatrix} \begin{pmatrix} 3 \\ 6 \\ 7 \end{pmatrix} &= \frac{1}{43} \begin{pmatrix} (5)(3) + (-2)(6) + (18)(7) \\ (9)(3) + (5)(6) + (-2)(7) \\ (-1)(3) + (9)(6) + (5)(7) \end{pmatrix} \\ &= \frac{1}{43} \begin{pmatrix} 129 \\ 43 \\ 86 \end{pmatrix} \\ &= \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} . \end{aligned}$$

Hence,

$$\frac{3+6\sqrt[3]{2}+7\sqrt[3]{4}}{1-\sqrt[3]{2}+2\sqrt[3]{4}} = 3+\sqrt[3]{2}+2\sqrt[3]{4}\,.$$

2. (i) First, we show that since the system $A\mathbf{x} = \mathbf{0}$ is homogeneous (that is, the right-hand side is equal to the zero vector), if the solution set can be written as $\mathbf{p} + \lambda\mathbf{u}$ for fixed vectors $\mathbf{p}$ and $\mathbf{u}$, then $\mathbf{p} = \mu\mathbf{u}$ for some scalar $\mu \in \mathbb{R}$. We note that $A\mathbf{0} = \mathbf{0}$ and so the zero vector is a solution to the system. Hence,

$$\mathbf{0} = \mathbf{p} + \lambda\mathbf{u}\,,$$

for some $\lambda \in \mathbb{R}$. If we define μ as $\mu = -\lambda$, then we have $\mathbf{p} = \mu\mathbf{u}$, as desired. In summary, since the system is homogeneous, $\mathbf{p}$ is a scalar multiple of $\mathbf{u}$, and so we can absorb it into the scaled term to write the solution set of the system $A\mathbf{x} = \mathbf{0}$ as $\{\lambda\mathbf{u} : \lambda \in \mathbb{R}\}$. We also note that this implies all non-zero solutions of the system are scalar multiples of each other.

Next, we define U to be the 3×3 matrix formed by performing elementary row operations on A until it is in reduced row echelon form (by Theorem 6.2.7, such a sequence of elementary row operations always exists). This matrix U must, by definition, have some number of leading ones.

We know that U cannot have three leading ones, or else it would have the same number of leading ones as it has rows, reducing it to the identity matrix I_3. By the Invertible Matrix Theorem, if A was row equivalent to I_3, then the only solution of $A\mathbf{x} = \mathbf{0}$ would be the trivial solution $\mathbf{x} = \mathbf{0}$, but we have assumed this to not be true ($\mathbf{u}$ was specified as being non-zero in the question).

Consider the case where U has two leading ones. These would have to be found in two distinct columns of A, giving us three possibilities:

- The leading ones are found in columns 1 and 2,

$$\begin{pmatrix} 1 & 0 & \alpha \\ 0 & 1 & \beta \\ 0 & 0 & 0 \end{pmatrix}\,,$$

  for some $\alpha, \beta \in \mathbb{R}$.

– The leading ones are found in columns 1 and 3,

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

for some $\alpha \in \mathbb{R}$.

– The leading ones are found in columns 2 and 3,

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

This final possibility cannot occur in this case however, as any sequence of elementary row operations we perform on this matrix would still give only zero entries in the first column, whereas the question specified at least one non-zero entry exists there.

Finally, we show that U cannot have fewer than two leading ones. Consider that if U had fewer than two leading ones, it would have at least two zero rows (that is, at least two rows containing only zero entries). This would give it the form

$$\begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

with each $u_{ij} \in \mathbb{R}$. Let $\mathbf{x}$ be a solution of the system $U\mathbf{x} = \mathbf{0}$, then by definition

$$U\mathbf{x} = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} u_{11}x_1 + u_{12}x_2 + u_{13}x_3 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Therefore, any vector $\mathbf{x}$ satisfying the constraint $u_{11}x_1 + u_{12}x_2 + u_{13}x_3$ is a solution to the system $U\mathbf{x} = \mathbf{0}$. We know that u_{11} is non-zero as, per the argument above, if it were zero then we would have no non-zero elements in the first column of A, contradicting the statement given in the question.

Thus, the vectors

$$\begin{pmatrix} u_{12} \\ -u_{11} \\ 0 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} u_{13} \\ 0 \\ -u_{11} \end{pmatrix}$$

would both be non-zero solutions to the system $U\mathbf{x} = \mathbf{0}$. Furthermore, it is clear that these two vectors do not lie on the same line $\lambda\mathbf{u}$, since any multiple of the first vector would have a zero entry in the third row, whereas the second vector has the non-zero value $-u_{11}$ in its third row.

We have therefore found two non-zero solutions to $U\mathbf{x} = \mathbf{0}$ that are not scalar multiples of each other. If we let $M = E_k E_{k-1} \ldots E_1$ be defined as the matrix product of the elementary matrices that perform the row operations required to get the matrix U back to the original matrix A, then we can use Lemma 7.6.1 to deduce that the systems $U\mathbf{x} = \mathbf{0}$ and $A\mathbf{x} = MU\mathbf{x} = M\mathbf{0} = \mathbf{0}$ will have equal solution sets.

Hence, $A\mathbf{x} = \mathbf{0}$ has two solutions that are not scalar multiples of each other, a contradiction. This tells us that U cannot have fewer than two leading ones.

In summary, we have shown that if we choose elementary row operations that reduce A to reduced row echelon form, then the resulting matrix will have exactly two leading ones, one of which will be in the first column. A will therefore be row equivalent to a matrix of the form

$$\begin{pmatrix} 1 & 0 & \alpha \\ 0 & 1 & \beta \\ 0 & 0 & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 1 & \alpha & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

for some $\alpha, \beta \in \mathbb{R}$.

(ii) Similarly to part (i), we take U to be the matrix formed by performing the elementary row operations on A that bring it to reduced row echelon form. As above, U cannot have three leading ones, as this would make it the identity matrix, contradicting the Invertible Matrix Theorem.

It could also not have two leading ones, as (per the derivation given in part (i)) this would mean that it takes the form

$$\begin{pmatrix} 1 & 0 & \alpha \\ 0 & 1 & \beta \\ 0 & 0 & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 1 & \alpha & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

for some $\alpha, \beta \in \mathbb{R}$. These cases are solved by $\lambda\mathbf{u_1}$ and $\lambda\mathbf{u_2}$ respectively, where

$$\mathbf{u_1} = \begin{pmatrix} \alpha \\ \beta \\ -1 \end{pmatrix}, \quad \text{and} \quad \mathbf{u_2} = \begin{pmatrix} \alpha \\ -1 \\ 0 \end{pmatrix}.$$

Hence, both systems have solutions sets of the form $\{\lambda\mathbf{u} : \lambda \in \mathbb{R}\}$, meaning that all non-zero solutions are scalar multiples of each other. As per the argument in part (i), this would equally hold for the original system $A\mathbf{x} = \mathbf{0}$, which has a solution set given by the plane $\mathbf{x} \cdot \mathbf{n} = d$.

We proceed by showing that this is impossible, i.e. that the plane given by $\mathbf{x} \cdot \mathbf{n} = d$ cannot be contained in the line given by $\lambda\mathbf{u}$. We can evaluate d by noting that since the zero vector solves the system $A\mathbf{x} = \mathbf{0}$, it must also lie in the plane $\mathbf{x} \cdot \mathbf{n} = d$, and so $d = \mathbf{0} \cdot \mathbf{n} = 0$.

The concluding argument follows as before. The vector $\mathbf{n}$ is non-zero, and without loss of generality, we assume $n_1 1 \neq 0$. Hence the vectors

$$\begin{pmatrix} n_{12} \\ -n_{11} \\ 0 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} n_{13} \\ 0 \\ -n_{11} \end{pmatrix}$$

are non-zero, do not lie in the same line, and yet both lie in the plane $\mathbf{x} \cdot \mathbf{n} = 0$. Thus, the matrix U cannot have two leading ones. If the matrix U had a single leading one, it would have to be of the form

$$\begin{pmatrix} 1 & \alpha & \beta \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & 1 & \alpha \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The condition that there should be at least one non-zero entry in the first column prohibits the latter two cases, and so the only possible case is that

$$U = \begin{pmatrix} 1 & \alpha & \beta \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} .$$

Hence, there exist a sequence of elementary row operations that brings A to this reduced row echelon form U, i.e. there are elementary matrices $E_1, \dots, E_k$ such that $A = E_k E_{k-1} \dots E_1 U$, and thus, A is row equivalent to this U.

As a final note, we mention that it would be impossible for U to have no leading ones, as this would mean that A would be row equivalent to the zero matrix, and so

$$A = E_k E_{k-1} \dots E_1 O_{3\times 3} = O_{3\times 3} ,$$

showing that the only matrix row equivalent to the 3×3 zero matrix is the zero matrix itself. If A were the zero matrix however, the system $A\mathbf{x} = \mathbf{0}$ would be solved by all vectors in $\mathbb{R}^3$, not just in the plane $\mathbf{x} \cdot \mathbf{n} = 0$.

3. If A is equal to its own inverse, then $A = A^{-1}$, and moreover

$$A^2 = AA = AA^{-1} = I ,$$

where I is the identity matrix. We therefore have

$$\begin{aligned} (A - I)(A + I) &= A(A + I) - I(A + I) \\ &= A^2 + AI - IA - I^2 \\ &= I + A - A - I \\ &= O , \end{aligned}$$

where O is the zero matrix. Let $\mathbf{u}$ be some (appropriately-sized) non-zero vector. If we assume $A + I$ is invertible, then $\mathbf{v} = (A + I)\mathbf{u}$ will also be non-zero (since, if $A + I$ is invertible the only solution to $(A + I)\mathbf{x} = \mathbf{0}$ should be the zero vector itself). Hence,

$$(A - I)\mathbf{v} = (A - I)(A + I)\mathbf{u} = O\mathbf{u} = \mathbf{0} ,$$

giving $A\mathbf{v} - I\mathbf{v} = \mathbf{0}$, or equivalently, $A\mathbf{v} = \mathbf{v}$. Note that this analysis assumed that $A + I$ was invertible. If this were not true, then by the Invertible Matrix Theorem, there would exist some non-zero vector v such that

$$(A + I)\mathbf{v} = \mathbf{0} \ ,$$

and hence, $A\mathbf{v} = -\mathbf{v}$.