Vectors & Matrices

Solutions to Problem Sheet 8

1. (i) We let $\mathbf{x}$ represent our column vector of variables. Each Cartesian equation has three variables, x, y, z, representing the distance of a point along the x, y and z axes respectively. This tells us that we should take vector $\mathbf{x}$ to be of the form

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} .$$

Next, we need a matrix A that maps this three dimensional vector to the column vector

$$\begin{pmatrix} 3x - y + 7z \\ -2x - 3y + 5z \\ x + y - 2z \end{pmatrix} ,$$

in order that we retrieve the left-hand sides of each of the Cartesian equations given in the question. We can see that

$$\begin{pmatrix} 3 & -1 & 7 \\ -2 & -3 & 5 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3x - y + 7z \\ -2x - 3y + 5z \\ x + y - 2z \end{pmatrix} ,$$

and so we shall define A to be the matrix on the left-hand of this equation. Finally, the Cartesian equations in the question show that we require values x, y, z such that the right-hand vector above becomes

$$\begin{pmatrix} 8 \\ 3 \\ -2 \end{pmatrix} ,$$

and so we call this vector $\mathbf{b}$. We have shown that our linear system can now be expressed as the matrix equation $A\mathbf{x} = \mathbf{b}$, for our matrix A, and vectors $\mathbf{x}$ and $\mathbf{b}$.

(ii) We have

$$A = \begin{pmatrix} 3 & -1 & 7 \\ -2 & -3 & 5 \\ 1 & 1 & -2 \end{pmatrix}.$$

We can see that the third rows of A and B are identical, and so these can be removed from our considerations. Next, we observe the second rows of the respective matrices. It is clear that if we multiply the third row of A by 2 and add it to the second row, we retrieve the second row of B. Applying this same operation to the identity matrix I_3 gives us

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix},$$

and so (by Theorem 7.6.5 in the lecture notes), the effect of left-multiplying A by this matrix E_2 is

$$E_2 A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & -1 & 7 \\ -2 & -3 & 5 \\ 1 & 1 & -2 \end{pmatrix} = \begin{pmatrix} 3 & -1 & 7 \\ 0 & -1 & 1 \\ 1 & 1 & -2 \end{pmatrix}.$$

We have found an elementary matrix that maps A to a 3×3 matrix with the same second and third rows as B, only the first row remains. We can obtain the first row of B by taking the first row of A and adding the third row of A to it. Again, applying this same operation to the identity matrix I_3 gives

$$E_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Hence, left-multiplying a matrix by the elementary matrix E_1 replaces its first row by the sum of its first and third rows. In particular,

$$\begin{aligned} E_1E_2A &= E_1(E_2A) \\ &= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & -1 & 7 \\ 0 & -1 & 1 \\ 1 & 1 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 0 & 5 \\ 0 & -1 & 1 \\ 1 & 1 & -2 \end{pmatrix} \\ &= B\ . \end{aligned}$$

(iii) We have seen that a point (x, y, z) is in the intersection of the planes Π_1, Π_2 and Π_3 if and only if it satisfies all three of the Cartesian equations that identify these planes. By part (i), this is equivalent to the values x, y, z being the coordinates of any vector $\mathbf{x}$ that solves the matrix equation $A\mathbf{x} = \mathbf{b}$.

By Lemma 7.6.1, the solution set of the matrix equation $A\mathbf{x} = \mathbf{b}$ is equal to the solution set of an equation $MA\mathbf{x} = M\mathbf{b}$, for any invertible 3×3 matrix M. Theorem 7.6.6 tells us that all elementary matrices are invertible, and this includes our matrices E_1 and E_2, defined in part (ii). By Theorem 7.1.19, their product E_1E_2 is also invertible, and so we can take our invertible M to be the product E_1E_2.

Hence, the system $E_1E_2A\mathbf{x} = E_1E_2\mathbf{b}$ has the same set of solutions as $A\mathbf{x} = \mathbf{b}$. But we have shown in part (ii) that $B = E_1E_2A$, and so $B\mathbf{x} = E_1E_2\mathbf{b}$ has the same solution set as $A\mathbf{x} = \mathbf{b}$. We evaluate

$$E_1E_2\mathbf{b} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 8 \\ 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 6 \\ -1 \\ -2 \end{pmatrix} ,$$

and call this resulting vector $\mathbf{c}$. In summary, we have shown that any point in the intersection of planes Π_1, Π_2 and Π_3 is given by a solution of the equation $A\mathbf{x} = \mathbf{b}$, and that any solution of this equation is a solution of the equation $B\mathbf{x} = \mathbf{c}$ (and vice versa).

We can reverse this argument to find that the planes defined by the Cartesian equations resulting from this equation,

$$4x + 5z = 6 ,$$
$$-y + z = -1 ,$$
$$x + y - 2z = -2 ,$$

have an intersection given by solutions of $B\mathbf{x} = \mathbf{c}$, which is equivalent to the set of points in the intersection of planes Π_1, Π_2 and Π_3.

2. (i) The 4×4 Type I elementary matrices can be enumerated by taking the identity matrix I_4 and swapping different pairs of rows. They are therefore given by

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

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

The first three of these elementary matrices are obtained by swapping the first row of I_4 with the second, third and fourth rows (respectively). Since this exhausts all swaps that use the first row, we move onto the second. The next two elementary matrices swap the second row with the third and fourth (respectively). Note that we do not swap with the first row, since that case was handled within the first three matrices. Finally, we do the only remaining swap; the third row with the fourth.

(ii) Just as with the previous enumeration, for any $n \times n$ matrix, we can swap the first row with any of the remaining $(n-1)$ rows. This gives $(n-1)$ possible swaps, and exhausts all possibilities involving swapping the first row. Similarly, we count $(n-2)$ ways of swapping the second row, etc, until we eventually have only 1 way of swapping the $(n-1)^{th}$ and n^{th} rows.

We can sum over all of these possibilities in reverse to add the 1 possibility of swapping the last two rows with the 2 possibilities from the last three rows, etc, to enumerate a total of

$$\sum_{k=0}^{n-1} k = \frac{n(n-1)}{2}$$

possible swaps. Since each different swap can be identified with a distinct Type I elementary matrix, we have a total of $\frac{1}{2}n(n-1)$ Type I elementary matrices of size $n \times n$. (Alternatively, one could simply say that the out of n rows, we choose a total of 2 to swap for each elementary matrix, and so there are

$$\binom{n}{2} = \frac{n(n-1)}{2}$$

possible $n \times n$ Type I elementary matrices.)

(iii) We know that any $n \times n$ Type I elementary matrix E_1 can be obtained by swapping two rows of I_n. If we introduced a second Type I elementary matrix E_2, then by Theorem 7.6.5, the product E_1E_2 will be equal to a matrix formed by swapping two rows of E_2. However, since E_2 is already formed by swapping two rows of I_n, we can identify E_1E_2 as a matrix obtained by swapping two pairs of rows in the identity matrix I_n.

We extend this argument to obtain the result that for any $m \in \mathbb{N}$, the product of m Type I elementary matrices, $E_1E_2\dots\dots E_m$, will give a matrix equal to I_n with m pairs of its rows swapped around. This matrix could also be identified as I_n with its rows permuted (that is, reordered in some way). Let the tuple $(r_1, r_2, \dots, r_n)$ denote the ordering of the rows of I_n in such a matrix. For instance, the matrix I_n itself would be identified as $(1, 2, \dots, n)$, and swapping the first two rows of I_n would result in the matrix given by $(2, 1, \dots, n)$.

We now show that **any** matrix generated by permuting the rows of I_n can be attained as the product of Type I elementary matrices. That is, for all row permutations $(r_1, r_2, \ldots, r_n)$ (where $r_i \in \{1, \ldots, n\}$, and $r_i \neq r_j$ for $i \neq j$), there exist $n \times n$ Type I elementary row matrices $E_1, E_2, \ldots, E_m$ such that their product $E_1 E_2 \ldots E_m$ results in a matrix identified by $(r_1, r_2, \ldots, r_n)$.

We proceed by induction. For the case $n = 2$, the result is clear. There are only two ways of permuting the rows of I_2, $(1, 2)$ (which is equal to I_2 itself) and $(2, 1)$ (which is I_2 with its rows flipped around). There only exists a single 2×2 Type I elementary matrix,

$$E = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} ,$$

and this matrix is equal to the permutation $(2, 1)$ (hence, this permutation is obtained as the product of a single elementary matrix). Since $E^2 = I_2$, the permutation $(1, 2)$ is obtained by taking the product of two Type I elementary matrices. This shows that the result holds for $n = 2$. We now assume that for some $k \in \mathbb{N}$, all row permutations of the identity matrix, $(r_1, r_2, \ldots, r_k)$, can expressed as the product of some number of Type I elementary matrices. We now use this assumption to prove the same must hold for $n = k + 1$.

Take any row permutation of I_{k+1}, $(r_1, r_2, \ldots, r_{k+1})$. There must contain a row in this matrix that contains the $(k+1)^{th}$ row of I_{k+1}, and we label this as row i. There exists a $(k+1) \times (k+1)$ Type I elementary matrix E that swaps rows i and $(k+1)$. Multiplying our matrix $(r_1, r_2, \ldots, r_{k+1})$ by this E therefore swaps rows r_i and r_{k+1}, giving the matrix $(r_1, r_2, \ldots, k+1)$ (since the i^{th} row of this matrix was taken to be the $(k+1)^{th}$ row of the identity matrix, we have $r_i = k + 1$).

This resulting permutation $(r_1, r_2, \ldots, r_k, k+1)$ has represents a matrix equal to the identity matrix I_{k+1} with the first k of its rows permuted, and its $(k+1)^{th}$ row in place. Our inductive assumption shows that any permutation of k rows can be obtained as the product of Type I elementary matrices $E_1, E_2, \ldots, E_m$. Hence, if we take the product of $E_1 E_2 \ldots E_m$ (the matrix product that permutes the first k rows of I_{k+1} into place) and E (the matrix required to swap the $(k+1)^t h$ row into place), we get a matrix $E_1 E_2 \ldots E_m E$ that takes the identity matrix I_{k+1} and permutes its rows into the order $(r_1, r_2, \ldots, r_{k+1})$, as desired.

Therefore, **all** row permutations of I_n can be expressed as the product of $n \times n$ Type I elementary matrices. Since applying a Type I elementary matrix to such a row permutation would only give another row permutation of I_n, the set of row permutations of I_n gives **all** products of $n \times n$ Type I elementary matrices. We know that there are $n!$ permutations of n objects, and so there are $n!$ matrices that can be expressed as the product of $n \times n$ Type I elementary matrices.

(iv) Per the argument given in part (iii), the 3×3 matrices that can be expressed as the product of Type I elementary matrices are exactly the matrices generated by permuting the rows of I_3. We therefore enumerate them as:

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

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

There are $3! = 6$ of these permutations, as expected.

3. (i) If we have the quadratic defined by $q(x) = ax^2 + bx + c$, then

$$\begin{aligned} q(-3x) &= a(-3x)^2 + b(-3)x + c \\ &= 9ax^2 - 3bx + c \\ &= (9a)x^2 + (-3b)x + c\,, \end{aligned}$$

represented by the vector

$$\begin{pmatrix} 9a \\ -3b \\ c \end{pmatrix}.$$

To get from $\mathbf{q}$ to this new vector, we must multiply the first row by a factor of 9, and the second by a factor of -3. These operations can be encoded as the following Type II elementary matrices,

$$E_1 = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} , \qquad \text{and} \qquad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

We have

$$\begin{aligned} E_1 E_2 \mathbf{q} &= \begin{pmatrix} 9 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\ &= \begin{pmatrix} 9 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\ &= \begin{pmatrix} 9a \\ -3b \\ c \end{pmatrix} , \end{aligned}$$

showing that $E_1 E_2$ is indeed the correct mapping.

(ii) If we take linear functions to be of the form $f(x) = ax + b$ (for some real values $a, b \in \mathbb{R}$), then we can represent each function as vectors

$$\mathbf{f} = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 .$$

We have

$$(2x - 3)f(x) = (2x - 3)(ax + b) = (2a)x^2 + (-3a + 2b)x + (-3b) ,$$

which is represented by the vector

$$\begin{pmatrix} 2a \\ -3a + 2b \\ -3b \end{pmatrix} \in \mathbb{R}^3 . \tag{1}$$

Given that we're mapping from elements of $\mathbb{R}^2$ to elements of $\mathbb{R}^3$, we require M to be a 3×2 matrix. Taking

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{pmatrix} ,$$

the vector $M\mathbf{f}$ must be of the form

$$M\mathbf{f} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} m_{11}a + m_{12}b \\ m_{21}a + m_{22}b \\ m_{31}a + m_{32}b \end{pmatrix} .$$

Since we know that we're aiming to receive the vector (1) as output, we can determine the values m_{ij} and get

$$M = \begin{pmatrix} 2 & 0 \\ -3 & 2 \\ 0 & -3 \end{pmatrix} .$$

(iii) For any quadratic $q(x) = ax^2 + bx + c$, we have

$$\begin{aligned} q'(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\left(ax^2 + bx + c\right) \\ &= 2ax + b \\ &= (2a)x + b \, , \end{aligned}$$

which results in a linear function, something that we can represent in $\mathbb{R}^2$ as

$$\begin{pmatrix} 2a \\ b \end{pmatrix} .$$

Given that we're mapping from $\mathbb{R}^3$ to $\mathbb{R}^2$, we require D to be a 2×3 matrix such that

$$D\mathbf{q} = \begin{pmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 2a \\ b \end{pmatrix} .$$

We therefore define determine the values d_{ij} as

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

(iv) If we again take linear functions to be of the form $f(x) = ax+b$ (for some real values $a, b \in \mathbb{R}$), then the anti-derivatives are of the form

$$\begin{aligned} F(x) &= \int f(x)\,\mathrm{d}x \\ &= \int ax + b\,\mathrm{d}x \\ &= \left(\frac{a}{2}\right) x^2 + bx + c\,, \end{aligned}$$

for some constant $c \in \mathbb{R}$. We require the anti-derivative satisfying $F(-1) = 0$, or equivalently

$$\left(\frac{a}{2}\right)(-1)^2 + b(-1) + c = 0\,,$$

determining the constant c as

$$c = -\frac{a}{2} + b\,.$$

Our function F is therefore given by

$$F(x) = \left(\frac{a}{2}\right) x^2 + bx - \frac{a}{2} + b\,,$$

represented by the vector

$$\begin{pmatrix} \frac{a}{2} \\ b \\ -\frac{a}{2} + b \end{pmatrix} \in \mathbb{R}^3\,.$$

We need matrix N to map from vectors in $\mathbb{R}^2$ to vectors in $\mathbb{R}^3$, and it must therefore have size 3×2. We find that

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

gives

$$N\mathbf{f} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \\ -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \frac{a}{2} \\ b \\ -\frac{a}{2} + b \end{pmatrix}.$$

(v) We have

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

$$= \begin{pmatrix} (2)(\frac{1}{2}) + (0)(0) + (0)(-\frac{1}{2}) & (2)(0) + (0)(1) + (0)(1) \\ (0)(\frac{1}{2}) + (1)(0) + (0)(-\frac{1}{2}) & (0)(0) + (1)(1) + (0)(1) \end{pmatrix}$$

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

$$= I_2 \ .$$

This demonstrates that finding a specific anti-derivative of a linear function and then differentiating it results in the original function. Note however, that

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

$$= \begin{pmatrix} (\frac{1}{2})(2) + (0)(0) & (\frac{1}{2})(0) + (0)(1) & (\frac{1}{2})(0) + (0)(0) \\ (0)(2) + (1)(0) & (0)(0) + (1)(1) & (0)(0) + (1)(0) \\ (-\frac{1}{2})(2) + (1)(0) & (-\frac{1}{2})(0) + (1)(1) & (-\frac{1}{2})(0) + (1)(0) \end{pmatrix}$$

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

This shows that differentiating a quadratic function and then finding the anti-derivative equal to zero at $x = -1$ does not,in general, result in the original function. In fact, the matrix product shows that the effect is to change the constant term to become equal to the difference of the coefficients of the linear and quadratic terms.

For example, differentiating $x^2 + 3x + 1$ gives $2x + 3$. Anti-derivatives of this function are of the form $x^2 + 3x + c$, where $c \in \mathbb{R}$. The unique anti-derivative with the property that it equals zero at $x = -1$ is given by $x^2 + 3x + 2$. This is consistent with the matrix ND above, as the constant term is equal to the difference of the coefficient of the linear term, 3, and the coefficient of the quadratic term, 1.