Vectors & Matrices

Solutions to Problem Sheet 7

1. (i) To evaluate B^2, we directly apply the definition of matrix multiplication,

$$B^2 = BB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} (0)(0) + (1)(0) & (0)(1) + (1)(0) \\ (0)(0) + (0)(0) & (0)(1) + (0)(0) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = O_{2\times 2} \, .$$

(ii) By the definition of matrix addition, it is clear that

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1+0 & 0+1 \\ 0+0 & 1+0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = I_2 + B \, .$$

We can therefore combine this fact with Theorem 7.1.14 a) in the notes to obtain

$$\begin{aligned} A^2 &= AA \\ &= (I_2 + B)(I_2 + B) \\ &= I_2 I_2 + I_2 B + B I_2 + BB \\ &= I_2^2 + B + B + B^2 \\ &= I_2 + 2B + O_{2\times 2} \\ &= I_2 + 2B \, , \end{aligned}$$

where we have used the fact that $I_2^2 = I_2$ (a result of Theorem 7.1.14 c)) and $B^2 = O_{2\times 2}$ (from part (i)). We can now use the given definition of B to evaluate

$$A^2 = I_2 + 2B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 + 2(0) & 0 + 2(1) \\ 0 + 2(0) & 1 + 2(0) \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \, .$$

(iii) Extending the result demonstrated in part (ii), we get

$$\begin{aligned} A^3 &= A^2 A \\ &= (I_2 + 2B)(I_2 + B) \\ &= I_2 I_2 + I_2 B + 2BI_2 + 2BB \\ &= I_2^2 + B + 2B + 2B^2 \\ &= I_2 + 3B + O_{2\times 2} \\ &= I_2 + 3B\ . \end{aligned}$$

We have shown that $A = I_2 + B$, $A^2 = I_2 + 2B$ and $A^3 = I_2 + 3B$. It therefore makes sense to suggest the formulation $A^n = I_2 + nB$. We now prove this formula using induction. Let $P(n)$ be the claim that $A^n = I_2 + nB$. Our base case is $n = 1$, proven in part (ii). Next, we assume that the statement holds for some $n = k$, that is, $A^k = I_2 + kB$. Finally,

$$\begin{aligned} A^{k+1} &= A^k A \\ &= (I_2 + kB)(I_2 + B) \\ &= I_2 I_2 + I_2 B + kBI_2 + kBB \\ &= I_2^2 + B + kB + kB^2 \\ &= I_2 + (k+1)B + O_{2\times 2} \\ &= I_2 + (k+1)B\ . \end{aligned}$$

Hence, if $A^k = I_2 + kB$, then $A^{k+1} = I_2 + (k+1)B$, demonstrating the inductive step. Combined with the base case, this proves the result. For any $n \in \mathbb{N}$, we have

$$A^n = I_2 + nB = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}\ .$$

(iv) To begin, we note that for any $n, m \in \mathbb{N}$,

$$\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} = A^n A^m = A^{n+m} = \begin{pmatrix} 1 & n+m \\ 0 & 1 \end{pmatrix}\ .$$

Even though this property has only been proved for natural values n and m, we can speculate that it extends to the case $n = 1$ and $m = -1$. This will, however, require a separate proof

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

Similarly, for $n = -1$ and $m = 1$,

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

Hence, taking $A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, we have just shown that $AA^{-1} = A^{-1}A = I_2$, and so this A^{-1} is indeed the inverse of A.

2. Suppose $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{m\times n}$. By the definition of matrix addition, $A+B = (c_{ij})_{m\times n}$, where for each i, j, $c_{ij} = a_{ij} + b_{ij}$. We now apply the definition of the transpose of a matrix to obtain

$$(A + B)^T = (c_{ji})_{n\times m} = (a_{ji} + b_{ji})_{n\times m} .$$

The definition of the transposition operation also gives us $A^T = (a_{ji})_{n\times m}$ and $B^T = (b_{ji})_{n\times m}$. We can sum these matrices to get

$$A^T + B^T = (a_{ji})_{n\times m} + (b_{ji})_{n\times m} = (a_{ji} + b_{ji})_{n\times m} .$$

The right-hand side of this identity is equal to the right-hand side of the identity above, we can therefore equate the two left-hand sides to obtain

$$(A + B)^T = A^T + B^T .$$

3. The property that $A^T = 48A^{-1}$ is equivalent to the statement that

$$A^T A = 48A^{-1}A = 48I_3 .$$

So to prove this equality, it suffices to show the above. The transpose of A can easily be computed as

$$A^T = \begin{pmatrix} 2\sqrt{3} & 0 & -6 \\ 3 & 6 & \sqrt{3} \\ 3\sqrt{3} & -2\sqrt{3} & 3 \end{pmatrix} ,$$

leaving us with

$$\begin{aligned} A^T A &= \begin{pmatrix} 2\sqrt{3} & 0 & -6 \\ 3 & 6 & \sqrt{3} \\ 3\sqrt{3} & -2\sqrt{3} & 3 \end{pmatrix} \begin{pmatrix} 2\sqrt{3} & 3 & 3\sqrt{3} \\ 0 & 6 & -2\sqrt{3} \\ -6 & \sqrt{3} & 3 \end{pmatrix} \\ &= \begin{pmatrix} (2\sqrt{3})^2 + 0 + (-6)^2 & (2\sqrt{3})(3) + 0 + (-6)(\sqrt{3}) & (2\sqrt{3})(3\sqrt{3}) + 0 + (-6)(3) \\ (3)(2\sqrt{3}) + 0 + (\sqrt{3})(-6) & 3^2 + 6^2 + \sqrt{3}^2 & (3)(3\sqrt{3}) + (6)(-2\sqrt{3}) + (\sqrt{3})(3) \\ (3\sqrt{3})(2\sqrt{3}) + 0 + (3)(-6) & (3\sqrt{3})(3) + (-2\sqrt{3})(6) + (3)(\sqrt{3}) & (3\sqrt{3})^2 + (-2\sqrt{3})^2 + 3^2 \end{pmatrix} \\ &= \begin{pmatrix} 48 & 0 & 0 \\ 0 & 48 & 0 \\ 0 & 0 & 48 \end{pmatrix} \\ &= 48 I_3 , \end{aligned}$$

giving us the result. The inverse of A is therefore given by

$$A^{-1} = \frac{1}{48} A^T = \frac{1}{48} \begin{pmatrix} 2\sqrt{3} & 0 & -6 \\ 3 & 6 & \sqrt{3} \\ 3\sqrt{3} & -2\sqrt{3} & 3 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{24} & 0 & \frac{-1}{8} \\ \frac{1}{16} & \frac{1}{8} & \frac{\sqrt{3}}{48} \\ \frac{\sqrt{3}}{16} & \frac{-\sqrt{3}}{24} & \frac{1}{16} \end{pmatrix} .$$

4. Firstly, note that as A is a $m \times n$ matrix, A^T is a $n \times m$ matrix, and hence the product $A^T A$ is well-defined with size $n \times n$. We can now use the formula given in Theorem 7.2.3 d) for the transpose of a matrix product to obtain

$$(A^T A)^T = A^T (A^T)^T .$$

Theorem 7.2.3 a) tells us that the transpose of a transpose of a matrix is equal to the original matrix, i.e. $(A^T)^T = A$, and so

$$(A^T A)^T = A^T A \ ,$$

meaning that the matrix $A^T A$ is equal to its own transpose. Hence, by the definition of matrix symmetry, $A^T A$ is symmetric.

5. To quantify the effect of the matrix

$$A = \begin{pmatrix} x & 1 & 2 \\ 0 & y & -1 \\ 0 & 0 & z \end{pmatrix}$$

on a linear system, we apply it to a general 3×3 matrix

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

to get

$$\begin{pmatrix} x & 1 & 2 \\ 0 & y & -1 \\ 0 & 0 & z \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} xa_{11} + a_{21} + 2a_{31} & a_{12} + a_{22} + 2a_{32} & a_{13} + a_{23} + 2a_{33} \\ ya_{21} - a_{31} & ya_{22} - a_{32} & ya_{23} - a_{33} \\ za_{31} & za_{32} & za_{33} \end{pmatrix} .$$

It's clear that the effect of left-multiplication by A is that the third row is multiplied by a factor of z, the second row is multiplied by a factor of y and then has the third row subtracted from it, and the first row is multiplied by a factor of x and then has the second row added it to it followed by two times the third row.

The salient point here is that the effect of an upper diagonal matrix is that each row is unaffected by the values in the rows above it, e.g. the second row is modified according to its own values, and the values in the third, but is not affected by the values in the first. We now focus on the equation in the question, evaluating the left-hand side of this equation gives

$$\begin{pmatrix} x & 1 & 2 \\ 0 & y & -1 \\ 0 & 0 & z \end{pmatrix} \begin{pmatrix} -1 & 8 & -5 \\ 2 & -3 & 2 \\ 3 & 0 & -4 \end{pmatrix} = \begin{pmatrix} -x+2+2(3) & 8x-3+2(0) & -5x+2+2(-4) \\ 2y-3 & -3y-0 & 2y-(-4) \\ 3z & 0z & -4z \end{pmatrix}$$

$$= \begin{pmatrix} -x+8 & 8x-3 & -5x-6 \\ 2y-3 & -3y & 2y+4 \\ 3z & 0 & -4z \end{pmatrix}.$$

Equating this resulting matrix with the right-hand side of the equation in the question, we have

$$\begin{pmatrix} -x+8 & 8x-3 & -5x-6 \\ 2y-3 & -3y & 2y+4 \\ 3z & 0 & -4z \end{pmatrix} = \begin{pmatrix} 11 & -27 & 9 \\ 1 & -6 & 8 \\ 9 & 0 & -12 \end{pmatrix}.$$

To find the values of x, y, z, we equate the entries between both of these matrices. For instance, since the $(1, 1)$ entries must be equal, we have $-x + 8 = 11$, giving $x = -3$. We can check that this value satisfies the equalities given through the remaining entries in the first rows: $8x - 3 = 8(-3) - 3 = -27$ and $-5x - 6 = -5(-3) - 6 = 9$. Similarly, the second and third rows give us $y = 2$ and $z = 3$. Hence,

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

Applying this matrix to the column vector $\mathbf{x}$ is straightforward, since the dependence of each row on the others is directed "downwards" (the resulting vector will have a first element dependent on all elements of this vector, whereas the second element will depend only on the second and third, and the third only on itself). More explicitly,

$$A\mathbf{x} = \begin{pmatrix} -3 & 1 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 5 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} (-3)(5)+(1)(2)+(2)(4) \\ (2)(2)+(-1)(4) \\ (3)(4) \end{pmatrix} = \begin{pmatrix} -5 \\ 0 \\ 12 \end{pmatrix}.$$

6. Again, in order to determine whether or not A is invertible, it would be helpful to observe the effect A has on a matrix. All we require for left-multiplication by A is that the right-hand matrix have n rows. For simplicity, we choose this matrix to have a single column, i.e. the right-hand matrix is a column vector.

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_{11}b_1 + 0b_2 + \ldots + 0b_n \\ 0b_1 + a_{22}b_2 + \ldots + 0b_n \\ \vdots \\ 0b_1 + 0b_2 + \ldots + a_{nn}b_n \end{pmatrix} = \begin{pmatrix} a_{11}b_1 \\ a_{22}b_2 \\ \vdots \\ a_{nn}b_n \end{pmatrix} .$$

It is clear that the effect of left-multiplying the column vector by A is that the i^{th} row is rescaled by a factor of a_{ii} (where a_{ii} is the i^{th} element along the diagonal of A). It can be shown that this property also holds when A is applied to matrices with more than a single column.

Since we now see that the i^{th} row of the resulting column vector is dependent **only** on the data in the i^{th} column of the input vector, the question of invertibility can be reduced to the invertibility of the mapping across each individual row. It is a trivial result that, as long as the factor $\alpha \in \mathbb{R}$ is non-zero, the process of multiplying by α can be inverted by multiplying by $\frac{1}{\alpha}$.

Hence, we can assert that the inverse map of A is the map that multiplies the i^{th} row of the input vector by $\frac{1}{a_{ii}}$ (assuming none of the values a_{ii} is equal to zero). We have seen that this type of linear map is uniquely identified with diagonal matrices. We can therefore take

$$A^{-1} = \begin{pmatrix} \frac{1}{a_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{a_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{a_{nn}} \end{pmatrix} ,$$

and demonstrate that it satisfies the properties to be the inverse of A. Indeed,

$$
\begin{aligned}
AA^1 &= \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \frac{1}{a_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{a_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{a_{nn}} \end{pmatrix} \\
&= \begin{pmatrix} (a_{11})(\frac{1}{a_{11}}) + 0 & (a_{11})(0) + 0 & \dots & (a_{11})(0) + 0 \\ 0 + (a_{22})(0) + 0 & 0 + (a_{22})(\frac{1}{a_{22}}) + 0 & \dots & 0 + (a_{22})(0) + 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 + (a_{nn})(0) & 0 + (a_{nn})(0) & \dots & 0 + (a_{nn})(\frac{1}{a_{nn}}) \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \\
&= I_n \ ,
\end{aligned}
$$

with $A^{-1}A = I_n$ demonstrated in a similar way.

Therefore, if this matrix A^{-1} exists, it is indeed the inverse of A. We finally discuss conditions for existence. As mentioned earlier, all we need is that the values $\frac{1}{a_{ii}}$ are defined for each i. This is equivalent to the condition that each diagonal entry a_{ii} is non-zero. Hence, diagonal matrices are invertible if and only if there is no zero entry along the diagonal.