Chapter 7

Matrices

Definition. An $m \times n$ *matrix* is an array of numbers set out in m rows and n columns, where $m, n \in \mathbb{N} = \{0, 1, 2, 3, \ldots\}$. (See Section 7.9 for the case $m = 0$ or $n = 0$.)

Examples. 1. $\begin{pmatrix} 1 & -1 & 5 \\ 2 & 0 & 6 \end{pmatrix}$ has 2 rows and 3 columns, and so it is a 2×3 matrix.

2. $\begin{pmatrix} 1 & 0 \\ 7 & -1 \\ \sqrt{2} & 3 \\ 3 & 1 \end{pmatrix}$ is a 4×2 matrix.

3. $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ is a 3×3 matrix.

4. A vector $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ is a 3×1 matrix.

The word *matrix* was introduced into mathematics in 1850 by James Joseph Sylvester (1814–1897), though the idea of writing out coefficients of equations as rectangular arrays of numbers dates back to antiquity (there are Chinese examples from about 200 BC), and Carl Friedrich Gauß (1777–1855) used this notation in his work on simultaneous equations. Sylvester and Arthur Cayley (1821–1895) developed the theory of matrices we use today, and which we shall investigate in this chapter.

We write $A = (a_{ij})_{m \times n}$ to mean that A is an $m \times n$ matrix whose (i, j)-entry is a_{ij}, that is, a_{ij} is in the i-th row and j-th column of A. For an $m \times n$ matrix A, we write A_{ij} or $A_{i,j}$ for the (i, j)-entry of A; thus if $A = (a_{ij})_{m \times n}$ then $A_{ij} = a_{ij}$. If $A = (a_{ij})_{m \times n}$ then

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix},$$

and we say that A has *size* $m \times n$. An $n \times n$ matrix is called a *square* matrix.

Example. Let
$$A = (a_{ij})_{2\times 2} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 4 & 3 \end{pmatrix}.$$
Then A is a square matrix of size 2×2. The $(1,2)$-entry of A is $a_{12} = -1$, and the $(2,2)$-entry of A is $a_{22} = 3$.

Example. We write out in full $A = (a_{ij})_{3\times 2}$ with $a_{ij} = i(i+j)$.
$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix},$$
with
$$\begin{array}{ll} a_{11} = 1(1+1) = 2, & a_{12} = 1(1+2) = 3, \\ a_{21} = 2(2+1) = 6, & a_{22} = 2(2+2) = 8, \\ a_{31} = 3(3+1) = 12, & a_{32} = 3(3+2) = 15, \end{array}$$
and so
$$A = \begin{pmatrix} 2 & 3 \\ 6 & 8 \\ 12 & 15 \end{pmatrix}.$$

Definition. Matrices A and B are *equal* if they have the same size and the same (i,j)-entry for every possible value of i and j. That is, if $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{p\times q}$, then A and B are equal if and only if $p = m$, $q = n$ and $a_{ij} = b_{ij}$ for all i, j with $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$.

Notation. 0_{mn} is the $m \times n$ matrix with every entry 0. The matrix 0_{mn} may also be denoted by $0_{m,n}$ or $0_{m\times n}$, especially in cases of ambiguity. (For example, does 0_{234} mean $0_{23\times 4}$ or $0_{2\times 34}$?) We call $0_{mn} = 0_{m,n} = 0_{m\times n}$ the *zero* $m \times n$ matrix.
For example, $0_{23} = 0_{2,3} = 0_{2\times 3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ is the zero 2×3 matrix.

Notation. I_n is the $n\times n$ matrix with $(1,1)$-entry $=$ $(2,2)$-entry $= \cdots =$ (n,n)-entry $= 1$ and all other entries 0. We call I_n the *identity* $n \times n$ matrix. Note that I_n is always a square matrix. For example, we have:
$$\mathrm{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathrm{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Another way to express this definition is that I_n is the unique $n \times n$ matrix having all its (top-left to bottom-right) diagonal entries equal to 1 and all other entries 0. Note that by the *diagonal* of a square matrix, we always mean its top-left to bottom-right diagonal; we are not generally interested in the bottom-left to top-right diagonal.

7.1 Addition of matrices

If $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{m\times n}$ then $A + B$ is defined to be the $m \times n$ matrix whose (i, j)-entry is $a_{ij} + b_{ij}$.

Note. We can only add matrices of the same size.

If $A = (a_{ij})_{m\times n}$, the *negative* of A, written $-A$, is defined to be the $m \times n$ matrix whose (i, j)-entry is $-a_{ij}$. Thus $-A = D$, where $D = (d_{ij})_{m\times n}$ with $d_{ij} = -a_{ij}$.

Examples. We have

$$\begin{pmatrix} 1 & 2 & -3 \\ 4 & -5 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 2 \\ -2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & -1 \\ 2 & -4 & 9 \end{pmatrix}$$

and

$$-\begin{pmatrix} 1 & -2 \\ -3 & 4 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 3 & -4 \end{pmatrix}.$$

7.2 Rules for matrix addition

The next theorem states that matrix addition is associative and commutative. Furthermore, the is an identity element for matrix addition, and each matrix has an inverse under the operation of matrix addition.

Theorem 7.1. Let A, B and C be $m \times n$ matrices. Then

(i) $(A + B) + C = A + (B + C)$,

(ii) $A + B = B + A$,

(iii) $A + 0_{mn} = 0_{mn} + A = A$, and

(iv) $A + (-A) = (-A) + A = 0_{mn}[= 0_{m\times n}]$.

Proof. (i): Let $A = (a_{ij})_{m\times n}$, $B = (b_{ij})_{m\times n}$ and $C = (c_{ij})_{m\times n}$. Then $A + B$ is an $m \times n$ matrix whose (i, j)-entry is $a_{ij} + b_{ij}$ and so $(A + B) + C$ is an $m \times n$ matrix whose (i, j)-entry is $(a_{ij} + b_{ij}) + c_{ij}$.

Similarly, $B + C$ is an $m \times n$ matrix whose (i, j)-entry is $b_{ij} + c_{ij}$ and so $A + (B + C)$ is an $m \times n$ matrix whose (i, j)-entry is $a_{ij} + (b_{ij} + c_{ij})$. The result follows, since

$$(a_{ij} + b_{ij}) + c_{ij} \;=\; a_{ij} + (b_{ij} + c_{ij})$$

by the associative law for addition of real numbers.

Proofs of (ii), (iii) and (iv): exercises for you. □

7.3 Scalar multiplication of matrices

Let $A = (a_{ij})_{m\times n}$ and let α be a scalar (that is, a real number). Then αA is defined to be the $m \times n$ matrix whose (i, j)-entry is αa_{ij}. Thus $\alpha A := B$ where $B = (b_{ij})_{m\times n}$ with $b_{ij} = \alpha a_{ij}$ for all i and j. Note that $(-1)A = -A$.

Examples. We have

$$3\begin{pmatrix} -1 & 3 \\ 2 & 5 \\ -7 & 6 \end{pmatrix} = \begin{pmatrix} -3 & 9 \\ 6 & 15 \\ -21 & 18 \end{pmatrix}$$

and

$$(-2)\begin{pmatrix} 1 & -2 \\ 3 & 4 \end{pmatrix} + 4\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -2 & 4 \\ -6 & -8 \end{pmatrix} + \begin{pmatrix} 0 & 4 \\ 4 & 0 \end{pmatrix} = \begin{pmatrix} -2 & 8 \\ -2 & -8 \end{pmatrix}.$$

7.4 Rules for scalar multiplication

Theorem 7.2. Let A and B be $m \times n$ matrices and let α and β be scalars. Then:

(i) $\alpha(A + B) = \alpha A + \alpha B$,

(ii) $(\alpha + \beta)A = \alpha A + \beta A$,

(iii) $\alpha(\beta A) = (\alpha\beta)A$,

(iv) $1A = A$, and

(v) $0A = 0_{mn}[= 0_{m\times n}]$.

Proof. (i): Let $A = (a_{ij})_{m\times n}$, $B = (b_{ij})_{m\times n}$. Then $A + B$ is an $m \times n$ matrix with (i, j)-entry $a_{ij} + b_{ij}$ and $\alpha(A+B)$ is an $m\times n$ matrix with (i, j)-entry $\alpha(a_{ij} + b_{ij}) = \alpha a_{ij} + \alpha b_{ij}$. But αA is an $m \times n$ matrix with (i, j)-entry αa_{ij} and αB is an $m \times n$ matrix with (i, j)-entry αb_{ij}, and so $\alpha A + \alpha B$ is an $m \times n$ matrix with (i, j)-entry $\alpha a_{ij} + \alpha b_{ij}$.

(iii): Let $A = (a_{ij})_{m\times n}$. Then βA is an $m \times n$ matrix with (i, j)-entry βa_{ij} and so $\alpha(\beta A)$ is an $m \times n$ matrix with (i, j)-entry $\alpha(\beta a_{ij}) = (\alpha\beta)a_{ij}$. But $(\alpha\beta)A$ is also an $m \times n$ matrix with (i, j)-entry $(\alpha\beta)a_{ij}$.

Proofs of (ii), (iv), (v): exercises for you. □

7.5 Matrix multiplication

This is rather more interesting and more complicated than scalar multiplication. It will become clearer later why we define matrix multiplication in the way that we do.

Let $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{n\times p}$. Then the *product* AB is the $m \times p$ matrix with (i, j)-entry

$$\sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{1n}b_{nj}.$$

Thus $AB = C$, where $C = (c_{ij})_{m\times p}$ with

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{1n}b_{nj}.$$

Observe that to get the (i,j)-entry of AB, we focus on the i-th row of A and the j-th column of B:

$$\left(\boxed{a_{i1} \quad a_{i2} \quad \ldots \quad a_{in}}\right)\left(\quad \boxed{\begin{matrix} b_{ij} \\ b_{2j} \\ \vdots \\ b_{nj} \end{matrix}} \quad\right),$$

and we form their dot product, that is the sum $a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{1n}b_{nj}$.

Note. The matrix product AB of matrices A and B is defined if and only if the number of columns of A is the same as the number of rows of B. In the case when $A = (a_{ij})_{m\times n}$ and $B = (a_{ij})_{n\times p}$ we note that A has n columns and B has n rows.

Example 1. Let

$$A = \begin{pmatrix} 1 & -1 \\ 3 & 7 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & 9 & 5 \\ 6 & 4 & 8 \end{pmatrix}.$$

Now A is a 2×2 matrix and B is a 2×3 matrix. They can be multiplied since the number of columns of A equals the number of rows of B. The product AB will be a 2×3 matrix as it will have the same number of rows as A and the same number of columns as B. Note that BA is *not* defined since the number of columns (3) of B is not equal to the number of rows (2) of A.

To get the $(1,3)$-entry of AB for example, we pick out the 1st row of A and the 3rd column of B:

$$\begin{pmatrix} \boxed{1 \quad -1} \\ * \quad\quad * \end{pmatrix}\begin{pmatrix} * & * & \boxed{\begin{matrix} 5 \\ 8 \end{matrix}} \\ * & * & \end{pmatrix},$$

and we form their dot product, the sum $1\times 5 + (-1)\times 8 = -3$. In full:

$$AB = \begin{pmatrix} 1\times 2 + (-1)\times 6 & 1\times 9 + (-1)\times 4 & 1\times 5 + (-1)\times 8 \\ 3\times 2 + 7\times 6 & 3\times 9 + 7\times 4 & 3\times 5 + 7\times 8 \end{pmatrix} = \begin{pmatrix} -4 & 5 & -3 \\ 48 & 55 & 71 \end{pmatrix}.$$

Example 2. We have

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

Just as in the first example, the product of the matrices in the reverse order is not defined.

Example 3. We have

$$\begin{pmatrix} -6 & 4 \\ -9 & 6 \end{pmatrix} \begin{pmatrix} 6 & -4 \\ 9 & -6 \end{pmatrix} = \begin{pmatrix} 6 & -4 \\ 9 & -6 \end{pmatrix} \begin{pmatrix} -6 & 4 \\ -9 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0_{22} = 0_{2\times 2}.$$

Example 4. Let $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Then

$$AB = \begin{pmatrix} 1 & 2 \\ 4 & 6 \end{pmatrix} \quad \text{and} \quad BA = \begin{pmatrix} 3 & 2 \\ 7 & 4 \end{pmatrix},$$

and so $AB \neq BA$.

Note. In general, the case $AB \neq BA$ is more common than the case $AB = BA$. In any case, matrix multiplication is not commutative. Also AB can be defined without BA being defined, and vice-versa. Moreover, AB and BA need not have the same size, for if we let $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{n\times m}$ then AB is an $m \times m$ matrix while BA is an $n \times n$ matrix, so that AB and BA have different sizes when $m \neq n$.

7.6 Rules for matrix multiplication

Theorem 7.3. Let $A = (a_{ij})_{m\times n}$, $B = (b_{ij})_{m\times n}$ $X = (x_{ij})_{n\times p}$ and $Y = (y_{ij})_{n\times p}$. Then:

(i) $(A + B)X = AX + BX$ and $A(X + Y) = AX + AY$; and

(ii) $\alpha(AX) = (\alpha A)X = A(\alpha X)$ for every scalar α.

Now let $A = (a_{ij})_{m\times n}$, $B = (b_{ij})_{n\times p}$ and $C = (c_{ij})_{p\times q}$. Then:

(iii) $(AB)C = A(BC)$; and

(iv) $\mathrm{I}_m A = A\mathrm{I}_n = A$.

Proof. (i): Let $C = A + B$. Then $C = (c_{ij})_{m\times n}$ with $c_{ij} = a_{ij} + b_{ij}$. We have that $(A + B)X = CX$ is an $m \times p$ matrix, with (i, j)-entry

$$\begin{aligned}
\textstyle\sum_{k=1}^{n} c_{ik}x_{kj} &= c_{i1}x_{1j} + c_{i2}x_{2j} + \cdots + c_{in}x_{nj} \\
&= (a_{i1} + b_{i1})x_{1j} + (a_{i2} + b_{i2})x_{2j} + \cdots + (a_{in} + b_{in})x_{nj} \\
&= (a_{i1}x_{1j} + a_{i2}x_{2j} + \cdots + a_{in}x_{nj}) + (b_{i1}x_{1j} + b_{i2}x_{2j} + \cdots b_{in}x_{nj}) \\
&= ((i,j)\text{-entry of } AX) + ((i,j)\text{-entry of } BX)
\end{aligned}$$

Since, in addition, $AX + BX$ has the same size, $m \times p$, as $(A + B)X$, we conclude that $(A + B)X = AX + BX$. The proof that $A(X + Y) = AX + AY$ is similar, and is left as an exercise, as is the proof of (ii).

(iii): Let $AB = X$, where $X = (x_{ij})_{m\times p}$, and let $BC = Y$, where $Y = (y_{ij})_{n\times q}$. Then $(AB)C = XC$ is an $m \times q$ matrix with (i, j)-entry

$$x_{i1}c_{1j} + x_{i2}c_{2j} + \cdots + x_{ip}c_{pj}.$$

But

$$\begin{aligned} x_{i1} &= a_{i1}b_{11} + a_{i2}b_{21} + \cdots + a_{in}b_{n1}, \\ x_{i2} &= a_{i1}b_{12} + a_{i2}b_{22} + \cdots + a_{in}b_{n2}, \\ &\cdots \\ x_{ip} &= a_{i1}b_{1p} + a_{i2}b_{2p} + \cdots + a_{in}b_{np}. \end{aligned}$$

Thus the (i, j)-entry of $(AB)C$ is

$$(a_{i1}b_{11} + \cdots + a_{in}b_{n1})c_{1j} + (a_{i1}b_{12} + \cdots + a_{in}b_{n2})c_{2j} + \cdots + (a_{i1}b_{1p} + \cdots + a_{in}b_{np})c_{pj}.$$

Multiplying out all brackets, we get that the (i, j)-entry of $(AB)C$ is

$$\sum_{s=1}^{p}\sum_{r=1}^{n}(a_{ir}b_{rs})c_{sj},$$

which is to say the sum of all terms $(a_{ir}b_{rs})c_{sj}$ as r varies over $1, 2, \ldots, n$ and s varies over $1, 2, \ldots, p$. (There are no issues with how the terms are ordered or grouped since addition of real numbers is both associative and commutative.)
A similar calculation shows that the (i, j)-entry of the $m \times q$ matrix $AY = A(BC)$ is

$$\sum_{r=1}^{n}\sum_{s=1}^{p}a_{ir}(b_{rs}c_{sj}).$$

But by the associative law for multiplying real numbers we have

$$(a_{ir}b_{rs})c_{sj} = a_{ir}(b_{rs}c_{sj})$$

for all i, j, r and s. Thus $(AB)C$ and $A(BC)$ are both $m \times q$ matrices and their (i, j)-entries are the same for all possible i and j. Hence $(AB)C = A(BC)$.

(iv): We prove that $\mathrm{I}_m A = A$, and leave the proof that $A\mathrm{I}_n = A$ as an exercise. Now I_m is an $m \times m$ matrix and so $\mathrm{I}_m A$ is an $m \times n$ matrix. The i-th row of I_m has zero entries everywhere except for the (i, i)-entry, which is 1. Thus $\mathrm{I}_m A$ has (i, j)-entry

$$0a_{1j} + \cdots + 0a_{i-1,j} + 1a_{ij} + 0a_{i+1,j} + \cdots + 0a_{mj} = a_{ij},$$

which is also the (i, j)-entry of A. □

Note. Since $(AB)C = A(BC)$ we can just write ABC for this product. In fact, matrix multiplication is associative in the strong sense that if one of $(AB)C$ and $A(BC)$ exists, then so does the other, and they are equal. But matrix multiplication is not commutative: we have seen examples where AB exists but BA does not, and also examples where both AB and BA exist and $AB = BA$ (Example 3) and $AB \neq BA$ (Example 4). There is also no anti-commutativity property for matrix multiplication.

Example. Let

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} -1 & -1 \\ 2 & 3 \end{pmatrix}.$$

We compute $(AB)C$ and $A(BC)$ and check that they are the same:

$$(AB)C = \left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} -1 & -1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix};$$

and

$$A(BC) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \left(\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 2 & 3 \end{pmatrix} \right) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -3 & -4 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}.$$

Note. There is a different matrix product, known as the *Hadamard product* or *pointwise product*, which takes two $m \times n$ matrices A and B and produces the $m \times n$ matrix $A \circ B$, where $(A \circ B)_{ij} = A_{ij}B_{ij}$ for all i and j. While this alternative matrix product has some mathematical utility, it is nowhere near as useful as the standard (and more complicated) matrix product. I leave it you to discern what properties hold for the operation $\circ$.

There also yet another type of matrix product, known as the *Kronecker product* or *tensor product*, which can be applied to any two matrices. I am not going to tell you how to define this product, but the interested reader can easily find this information on the World Wide Web.

7.7 Some useful notation

For an $n \times n$ matrix A we define $A^0 = \mathrm{I}_n$, $A^1 = A$, $A^2 = AA$, $A^3 = AAA$, and so on. As a consequence of the (generalised) associative law for matrix multiplication we have $A^pA^q = A^{p+q}$ and $(A^p)^q = A^{pq}$ for all $p, q \in \mathbb{N} = \{0, 1, 2, \ldots\}$. Note that A^pA^q means $(A^p)(A^q)$. For $m \times n$ matrices B and C we write $B - C$ for $B + (-C)$.

7.8 Inverses of matrices

Every nonzero real number α has a multiplicative inverse, that is, if $0 \neq \alpha \in \mathbb{R}$ there exists $\beta \in \mathbb{R}$ such that $\alpha\beta = \beta\alpha = 1$. What about matrices?

Definition 7.4. An $n \times n$ matrix A is said to be *invertible* if there is some $n \times n$ matrix B such that $AB = BA = \mathrm{I}_n$.

Example 1. We have

$$\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{I}_2$$

and

$$\begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{I}_2.$$

Thus $A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$ is invertible (and so is $B = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$ by the same calculation).

Example 2. The matrix $\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$ is not invertible, because no matter what entries we take for the matrix $B = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$ we get

$$\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} B = \begin{pmatrix} p+2r & q+2s \\ 0 & 0 \end{pmatrix} \neq \mathrm{I}_2 \quad \text{and} \quad B \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} p & 2p \\ r & 2r \end{pmatrix} \neq \mathrm{I}_2.$$

Theorem 7.5. Let A be an $n \times n$ matrix and let B and C be $n \times n$ matrices such that $AB = BA = \mathrm{I}_n$ and $AC = CA = \mathrm{I}_n$. Then $B = C$.

Proof. We have $B = B\mathrm{I}_n = B(AC) = (BA)C = \mathrm{I}_n C = C$. □

Definition. If A is an invertible $n \times n$ matrix then the unique matrix B such that $AB = BA = \mathrm{I}_n$ is called the *inverse* of A. We denote the inverse of A by A^{-1}. For $n \in \mathbb{N}^+$ we define $A^{-n} := (A^{-1})^n$, the n-th power of the inverse of A. (This definition is consistent for $n = 1$.)

Note. If A is invertible, then $AA^{-1} = A^{-1}A = \mathrm{I}_n$, and so by the definitions of invertible and inverse, we see that A^{-1} is invertible and $(A^{-1})^{-1} = A$. If A is invertible then $A^p A^q = A^{p+q}$ and $(A^p)^q = A^{pq}$ for all $p, q \in \mathbb{Z}$. In particular, $(A^m)^{-1} = A^{-m}$ for all $m \in \mathbb{Z}$. Also, I_n is always invertible (we have ${\mathrm{I}_n}^{-1} = \mathrm{I}_n$), as is any nonzero scalar multiple of I_n. We have ${\mathrm{I}_n}^m = \mathrm{I}_n$ for all $n \in \mathbb{N}$ and $m \in \mathbb{Z}$.

Fact. If A and B are $n \times n$ matrices with $AB = \mathrm{I}_n$, then $BA = \mathrm{I}_n$. Thus if a square matrix has a 1-sided "inverse" it is actually invertible. Now let $m > n$ (the case $m < n$ is similar), and let A be an $m \times n$ matrix and B an $n \times m$ matrix. Then AB is an $m \times m$ matrix, and can never be I_m, and BA is an $n \times n$ matrix, and sometimes we can have $BA = \mathrm{I}_n$.

Theorem 7.6. If A, B are invertible $n \times n$ matrices then AB is invertible, and $(AB)^{-1} = B^{-1}A^{-1}$.

Proof. Using the associative law of matrix multiplication implicitly throughout, we have:

$$AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = A\mathrm{I}_n A^{-1} = AA^{-1} = \mathrm{I}_n$$

and

$$(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = B^{-1}\mathrm{I}_n B = B^{-1}B = \mathrm{I}_n.$$ □

7.9 Matrices with no rows or no columns

Somewhat surprisingly, it makes sense mathematically to consider matrices having no rows or no columns (or both). Since such matrices have no entries then (vacuously) all entries in such matrices are equal to 0. Thus such matrices are $0_{0\times m}$ or $0_{m\times 0}$ for some $m \in \mathbb{N} = \{0, 1, 2, \ldots, \}$. We also have $0_{0\times 0} = \mathrm{I}_0$. Among these matrices, only I_0 is invertible, with ${\mathrm{I}_0}^{-1} = \mathrm{I}_0$. The following relations hold (which are the only possible products involving these matrices):

1. $0_{m\times 0}0_{0\times n} = 0_{m\times n}$, with mn entries;
2. $0_{0\times m}0_{m\times 0} = 0_{0\times 0} = \mathrm{I}_0$;
3. $0_{0\times m}A = 0_{0\times n}$, where $A = (a_{ij})_{m\times n}$; and
4. $B0_{n\times 0} = 0_{m\times 0}$, where $B = (b_{ij})_{m\times n}$.

Note that in the case $0_{m\times 0}0_{0\times n} = 0_{m\times n}$, with $m, n \geqslant 1$, we have taken the product of two matrices having no entries to produce a matrix with a nonzero (mn) number of entries (all of which are 0).

7.10 Transposes of matrices

Let A be an $m \times n$ matrix. The *tranpose* of A is the $n \times m$ matrix denoted A^{T}, where $A^{\mathsf{T}} = (a^{\mathsf{T}}_{ij})_{n\times m}$ with $a^{\mathsf{T}}_{ij} = a_{ji}$ for all i and j where $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant m$. Thus we see that transposition interchanges the rôles of rows and columns in a matrix. To transpose a matrix A, we read off the rows of A and write them down as the columns of A^{T}.

Examples. We have ${\mathrm{I}_n}^{\mathsf{T}} = \mathrm{I}_n$ for all n. Also, we have

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^{\mathsf{T}} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2 & 3 & 5 \\ 7 & 11 & 13 \\ 17 & 19 & 23 \end{pmatrix}^{\mathsf{T}} = \begin{pmatrix} 2 & 7 & 17 \\ 3 & 11 & 19 \\ 5 & 13 & 23 \end{pmatrix}.$$

The properties of the transposition operator are summarised below.

Theorem 7.7. Let $A = (a_{ij})_{m\times n}$, $B = (a_{ij})_{m\times n}$ and $C = (c_{ij})_{n\times p}$. Then:

(i) $(A^{\mathsf{T}})^{\mathsf{T}} = A$;

(ii) $(A + B)^{\mathsf{T}} = A^{\mathsf{T}} + B^{\mathsf{T}}$;

(iii) $(-A)^{\mathsf{T}} = -(A^{\mathsf{T}})$;

(iv) $(\lambda A)^{\mathsf{T}} = \lambda(A^{\mathsf{T}})$ for all scalars λ;

(v) $(AC)^\mathsf{T} = C^\mathsf{T} A^\mathsf{T}$; and

(vi) if A is an invertible $n \times n$ matrix, then so is A^T, and $(A^\mathsf{T})^{-1} = (A^{-1})^\mathsf{T}$.

Proof. (i), (ii), (iii) and (iv): Exercises for the reader.

(v): Firstly, we note that A, C, AC, $(AC)^\mathsf{T}$, C^T, A^T and $C^\mathsf{T} A^\mathsf{T}$ have sizes $m \times n$, $n \times p$, $m \times p$, $p \times m$, $p \times n$, $n \times m$ and $p \times m$ respectively, so that $(AC)^\mathsf{T}$ and $C^\mathsf{T} A^\mathsf{T}$ both have the same size, namely $p \times m$. Now we calculate, for $1 \leqslant i \leqslant p$ and $1 \leqslant j \leqslant m$, that:

$$(i, j)\text{-entry of } (AC)^\mathsf{T} = (j, i)\text{-entry of } AC = \sum_{k=1}^{n} a_{jk} c_{ki}.$$

Moreover, we have

$$(i, j)\text{-entry of } C^\mathsf{T} A^\mathsf{T} = \sum_{k=1}^{n} c^\mathsf{T}_{ik} a^\mathsf{T}_{kj} = \sum_{k=1}^{n} c_{ki} a_{jk} = \sum_{k=1}^{n} a_{jk} c_{ki}.$$

Since $(AC)^\mathsf{T}$ and $C^\mathsf{T} A^\mathsf{T}$ have the same size and identical (i, j)-entries for all i and j they are equal.

(vi): Using Part (v), we see that $(A^{-1})^\mathsf{T} A^\mathsf{T} = (AA^{-1})^\mathsf{T} = \mathrm{I}_n{}^\mathsf{T} = \mathrm{I}_n$ and $A^\mathsf{T}(A^{-1})^\mathsf{T} = (A^{-1}A)^\mathsf{T} = \mathrm{I}_n{}^\mathsf{T} = \mathrm{I}_n$. Thus A^T is invertible, with inverse $(A^{-1})^\mathsf{T}$. □

Note. In view of the above theorem, we write $-A^\mathsf{T}$ instead of $(-A)^\mathsf{T}$ or $-(A^\mathsf{T})$; λA^T instead of $(\lambda A)^\mathsf{T}$ or $\lambda(A^\mathsf{T})$; and $A^{-\mathsf{T}}$ instead of $(A^\mathsf{T})^{-1}$ or $(A^{-1})^\mathsf{T}$. Also, if A and B are invertible $n \times n$ matrices then $(AB)^{-\mathsf{T}} = A^{-\mathsf{T}} B^{-\mathsf{T}}$.