1 Matrices and matrix algebra

1.1 Examples of matrices

□ Definition: A matrix is a rectangular array of numbers and/or variables. For instance

\[
A = \begin{pmatrix}
  4 & -2 & 0 & -3 & 1 \\
  5 & 1.2 & -0.7 & x & 3 \\
  \pi & -3 & 4 & 6 & 27 \\
\end{pmatrix}
\]

is a matrix with 3 rows and 5 columns (a $3 \times 5$ matrix). The 15 entries of the matrix are referenced by the row and column in which they sit: the (2,3) entry of $A$ is $-0.7$. We may also write $a_{23} = -0.7$, $a_{24} = x$, etc. We indicate the fact that $A$ is $3 \times 5$ (this is read as "three by five") by writing $A_{3 \times 5}$. Matrices can also be enclosed in square brackets as well as large parentheses. That is, both

\[
\begin{pmatrix}
  2 & 4 \\
  1 & -6 \\
\end{pmatrix}
\quad \text{and} \quad
\begin{bmatrix}
  2 & 4 \\
  1 & -6 \\
\end{bmatrix}
\]

are perfectly good ways to write this $2 \times 2$ matrix.

Real numbers are $1 \times 1$ matrices. A vector such as

\[
v = \begin{pmatrix}
  x \\
  y \\
  z \\
\end{pmatrix}
\]

is a $3 \times 1$ matrix. We will generally use upper case Latin letters as symbols for general matrices, boldface lower case letters for the special case of vectors, and ordinary lower case letters for real numbers.

□ Definition: Real numbers, when used in matrix computations, are called scalars.

Matrices are ubiquitous in mathematics and the sciences. Some instances include:

• Systems of linear algebraic equations (the main subject matter of this course) are normally written as simple matrix equations of the form $Ax = y$.

• The derivative of a function $f : R^3 \to R^2$ is a $2 \times 3$ matrix.

• First order systems of linear differential equations are written in matrix form.

• The symmetry groups of mathematics and physics, some of which we’ll look at later, are groups of matrices.

• Quantum mechanics can be formulated using infinite-dimensional matrices.
1.2 Operations with matrices

Matrices of the same size can be added or subtracted by adding or subtracting the corresponding entries:

\[
\begin{pmatrix}
  2 & 1 \\
  -3 & 4 \\
  7 & 0
\end{pmatrix}
+ 
\begin{pmatrix}
  6 & -1.2 \\
  \pi & x \\
  1 & -1
\end{pmatrix}
= 
\begin{pmatrix}
  8 & -0.2 \\
  \pi - 3 & 4 + x \\
  8 & -1
\end{pmatrix}.
\]

□ **Definition:** If the matrices \(A\) and \(B\) have the same size, then their sum is the matrix \(A + B\) defined by

\[(A + B)_{ij} = a_{ij} + b_{ij}.
\]

Their difference is the matrix \(A - B\) defined by

\[(A - B)_{ij} = a_{ij} - b_{ij}.
\]

□ **Definition:** A matrix \(A\) can be multiplied by a scalar \(c\) to obtain the matrix \(cA\), where

\[(cA)_{ij} = ca_{ij}.
\]

This is called scalar multiplication. We just multiply each entry of \(A\) by \(c\). For example

\[-3 \begin{pmatrix}
  1 & 2 \\
  3 & 4
\end{pmatrix} = \begin{pmatrix}
  -3 & -6 \\
  -9 & -12
\end{pmatrix}.
\]

□ **Definition:** The \(m \times n\) matrix whose entries are all 0 is denoted \(0_{mn}\) (or, more often, just by \(0\) if the dimensions are obvious from context). It’s called the zero matrix.

□ **Definition:** Two matrices \(A\) and \(B\) are equal if all their corresponding entries are equal:

\[A = B \iff a_{ij} = b_{ij} \text{ for all } i, j.
\]

□ **Definition:** If the number of columns of \(A\) equals the number of rows of \(B\), then the product \(AB\) is defined by

\[(AB)_{ij} = \sum_{s=1}^{k} a_{is} b_{sj}.
\]

Here \(k\) is the number of columns of \(A\) or rows of \(B\).

If the summation sign is confusing, this could also be written as

\[(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj}.
\]
Example:

\[
\begin{pmatrix}
1 & 2 & 3 \\
-1 & 0 & 4 \\
1 & 3
\end{pmatrix}
\begin{pmatrix}
-1 & 0 \\
4 & 2 \\
1 & 3
\end{pmatrix}
= \begin{pmatrix}
1 \cdot -1 + 2 \cdot 4 + 3 \cdot 1 & 1 \cdot 0 + 2 \cdot 2 + 3 \cdot 3 \\
-1 \cdot -1 + 0 \cdot 4 + 4 \cdot 1 & -1 \cdot 0 + 0 \cdot 2 + 4 \cdot 3
\end{pmatrix}
= \begin{pmatrix}
10 & 13 \\
5 & 12
\end{pmatrix}
\]

If \( AB \) is defined, then the number of rows of \( AB \) is the same as the number of rows of \( A \), and the number of columns is the same as the number of columns of \( B \):

\[
A_{m \times n}B_{n \times p} = (AB)_{m \times p}.
\]

Why define multiplication like this? The answer is that this is the definition that corresponds to what shows up in practice.

Example: Recall from calculus (Exercise!) that if a point \((x, y)\) in the plane is rotated counterclockwise about the origin through an angle \(\theta\) to obtain a new point \((x', y')\), then

\[
x' = x \cos \theta - y \sin \theta \\
y' = x \sin \theta + y \cos \theta.
\]

In matrix notation, this can be written

\[
\begin{pmatrix}
x' \\
y'
\end{pmatrix}
= \begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}
\begin{pmatrix}
x \\
y
\end{pmatrix}.
\]

If the new point \((x', y')\) is now rotated through an additional angle \(\phi\) to get \((x'', y'')\), then

\[
\begin{pmatrix}
x'' \\
y''
\end{pmatrix}
= \begin{pmatrix}
\cos \phi & -\sin \phi \\
\sin \phi & \cos \phi
\end{pmatrix}
\begin{pmatrix}
x' \\
y'
\end{pmatrix}
= \begin{pmatrix}
\cos \phi -\sin \phi & \cos \theta -\sin \theta \\
\sin \phi & \cos \phi
\end{pmatrix}
\begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}
\begin{pmatrix}
x \\
y
\end{pmatrix}
= \begin{pmatrix}
\cos(\theta + \phi) & -\sin(\theta + \phi) \\
\sin(\theta + \phi) & \cos(\theta + \phi)
\end{pmatrix}
\begin{pmatrix}
x \\
y
\end{pmatrix}.
\]

This is obviously correct, since it shows that the point has been rotated through the total angle of \(\theta + \phi\). So the right answer is given by matrix multiplication as we’ve defined it, and not some other way.

**Matrix multiplication is not commutative:** in English, \( AB \neq BA \), for arbitrary matrices \( A \) and \( B \). For instance, if \( A \) is \(3 \times 5\) and \( B \) is \(5 \times 2\), then \( AB \) is \(3 \times 2\), but \( BA \) is not defined. Even if both matrices are square and of the same size, so that both \( AB \) and \( BA \) are defined and have the same size, the two products are not generally equal.

**Exercise:** Write down two \(2 \times 2\) matrices and compute both products. Unless you’ve been very selective, the two products won’t be equal.
Another example: If
\[ A = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \text{ and } B = \begin{pmatrix} 1 & 2 \end{pmatrix}, \]
then
\[ AB = \begin{pmatrix} 2 & 4 \\ 3 & 6 \end{pmatrix}, \text{ while } BA = (8). \]

Two fundamental properties of matrix multiplication:

1. If \( AB \) and \( AC \) are defined, then \( A(B + C) = AB + AC \).
2. If \( AB \) is defined, and \( c \) is a scalar, then \( A(cB) = c(AB) \).

Exercise*: Prove the two properties listed above. (Both these properties can be proven by showing that, in each equation, the \((i, j)\) entry on the right hand side of the equation is equal to the \((i, j)\) entry on the left.)

□ Definition: The transpose of the matrix \( A \), denoted \( A^t \), is obtained from \( A \) by making the first row of \( A \) into the first column of \( A^t \), the second row of \( A \) into the second column of \( A^t \), and so on. Formally,
\[ a_{ij}^t = a_{ji}. \]
So
\[ \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}^t = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}. \]

Here’s one consequence of the non-commutatitivity of matrix multiplication: If \( AB \) is defined, then \( (AB)^t = B^t A^t \) (and not \( A^t B^t \) as you might expect).

Example: If
\[ A = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}, \text{ and } B = \begin{pmatrix} -1 & 2 \\ 4 & 3 \end{pmatrix}, \]
then
\[ AB = \begin{pmatrix} 2 & 7 \\ -3 & 6 \end{pmatrix}, \text{ so } (AB)^t = \begin{pmatrix} 2 & -3 \\ 7 & 6 \end{pmatrix}. \]
And
\[ B^t A^t = \begin{pmatrix} -1 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ 7 & 6 \end{pmatrix} \]
as advertised.
Exercise**: Can you show that \((AB)^t = B^tA^t\)? You need to write out the \((i, j)^{th}\) entry of both sides and then “observe” that they’re equal.

\[\square\]

**Definition**: \(A\) is **square** if it has the same number of rows and columns. An important instance is the **identity matrix** \(I_n\), which has ones on the main diagonal and zeros elsewhere:

**Example**:

\[
I_3 = \begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{pmatrix}.
\]

Often, we’ll just write \(I\) without the subscript for an identity matrix, when the dimension is clear from the context. The identity matrices behave, in some sense, like the number 1. If \(A\) is \(n \times m\), then \(I_nA = A\), and \(AI_m = A\).

\[\square\]

**Definition**: Suppose \(A\) and \(B\) are square matrices of the same dimension, and suppose that \(AB = I = BA\). Then \(B\) is said to be the **inverse** of \(A\), and we write this as \(B = A^{-1}\). Similarly, \(B^{-1} = A\). For instance, you can easily check that

\[
\begin{pmatrix}
2 & 1 \\
1 & 1 \\
\end{pmatrix} \begin{pmatrix}
1 & -1 \\
-1 & 2 \\
\end{pmatrix} = \begin{pmatrix}
1 & 0 \\
0 & 1 \\
\end{pmatrix},
\]

and so these two matrices are inverses of one another:

\[
\begin{pmatrix}
2 & 1 \\
1 & 1 \\
\end{pmatrix}^{-1} = \begin{pmatrix}
1 & -1 \\
-1 & 2 \\
\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}
1 & -1 \\
-1 & 2 \\
\end{pmatrix}^{-1} = \begin{pmatrix}
2 & 1 \\
1 & 1 \\
\end{pmatrix}.
\]

**Example**: Not every square matrix has an inverse. For instance

\[
A = \begin{pmatrix}
3 & 1 \\
3 & 1 \\
\end{pmatrix}
\]

has no inverse.

**Exercise***: Show that the matrix \(A\) in the above example has no inverse. Hint: Suppose that

\[
B = \begin{pmatrix}
a & b \\
c & d \\
\end{pmatrix}
\]

is the inverse of \(A\). Then we must have \(BA = I\). Write this out and show that the equations for the entries of \(B\) are inconsistent.

**Exercise**: Which \(1 \times 1\) matrices are invertible, and what are their inverses?
Exercise: Show that if

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \text{ and } ad - bc \neq 0, \text{ then } A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$ 

Hint: Multiply $A$ by the given expression for $A^{-1}$ and show that it equals $I$. If $ad - bc = 0$, then the matrix is not invertible. You should probably memorize this formula.

Exercise*: Show that if $A$ has an inverse that it’s unique; that is, if $B$ and $C$ are both inverses of $A$, then $B = C$. (Hint: Consider the product $BAC = (BA)C = B(AC)$.)