Fact Sheet 7 – Matrices

- A *matrix* is a rectangular array of numbers. The numbers are called the *elements* (or entries) of the matrix. The horizontal lines of elements in the matrix are called its *rows* and the vertical lines of elements are called its *columns*.

- The number of rows and number of columns of a matrix determine its *shape*. An $m \times n$ (m-by-n) matrix $\mathbf{A}$ has m rows and n columns:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \bullet & a_{1n} \\ a_{21} & a_{22} & a_{23} & \bullet & a_{2n} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ a_{m1} & a_{m2} & a_{m3} & \bullet & a_{mn} \end{pmatrix},$$

  where we have applied *double index notation* in which the *matrix elements* are denoted a_{ij}, where i is the row number and j the column number of the element.

- A matrix with only one row , $m = 1$, is called a *row matrix*, while a matrix with only one column, $n = 1$, is called a *column matrix*. The terms *(n-dimensional) row vector* and *(m-dimensional) column vector* are sometimes used, because the elements in the matrix can be regarded as the components of vectors in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively.

- **Definition of sum of matrices $\mathbf{A}+\mathbf{B}$:** Two matrices can be *added* if and only if they have the same shape. If $\mathbf{A}$ is an $m \times n$ matrix with elements a_{ij} and $\mathbf{B}$ is an $m \times n$ matrix with elements b_{ij}, then their *sum* $\mathbf{A}+\mathbf{B}$ is defined as the $m \times n$ matrix with the ij th element $a_{ij} + b_{ij}$, i.e., the sum of the corresponding elements in the two matrices:

$$\mathbf{A}+\mathbf{B} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \bullet & a_{1n} \\ a_{21} & a_{22} & a_{23} & \bullet & a_{2n} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ a_{m1} & a_{m2} & a_{m3} & \bullet & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} & \bullet & b_{1n} \\ b_{21} & b_{22} & b_{23} & \bullet & b_{2n} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ b_{m1} & b_{m2} & b_{m3} & \bullet & b_{mn} \end{pmatrix}$$

$$\equiv \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} & \bullet & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} & \bullet & a_{2n}+b_{2n} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & a_{m3}+b_{m3} & \bullet & a_{mn}+b_{mn} \end{pmatrix}.$$

- **Definition of numerical multiple $r\mathbf{A}$:** If $\mathbf{A}$ is an $m \times n$ matrix with elements a_{ij} and r is a number, then the *numerical multiple* $r\mathbf{A}$ is defined as the $m \times n$ matrix with the ij th element ra_{ij}, that is simply multiply all elements in the matrix with r:

$$r\mathbf{A} = r\begin{pmatrix} a_{11} & a_{12} & a_{13} & \bullet & a_{1n} \\ a_{21} & a_{22} & a_{23} & \bullet & a_{2n} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ a_{m1} & a_{m2} & a_{m3} & \bullet & a_{mn} \end{pmatrix} \equiv \begin{pmatrix} ra_{11} & ra_{12} & ra_{13} & \bullet & ra_{1n} \\ ra_{21} & ra_{22} & ra_{23} & \bullet & ra_{2n} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ ra_{m1} & ra_{m2} & ra_{m3} & \bullet & ra_{mn} \end{pmatrix}.$$

  Notice that we write $-\mathbf{A}$ for the numerical multiple $(-1)\mathbf{A}$ and $\mathbf{A}-\mathbf{B}$ as an abbreviation for $\mathbf{A}+(-1)\mathbf{B}$.

- **Properties of matrix addition and numerical multiple.** Let $\mathbf{A}, \mathbf{B}$, and $\mathbf{C}$ be $m \times n$ matrices and introduce the $m \times n$ <u>*null-matrix*</u> $\mathbf{0}$ where all elements are zero. Then:

  1. $r\mathbf{A} + s\mathbf{A} = (r+s)\mathbf{A}$ — Distributive law for numerical multiple.
  2. $r\mathbf{A} + r\mathbf{B} = r(\mathbf{A}+\mathbf{B})$ — Distributive law for numerical multiple.
  3. $r(s\mathbf{A}) = (rs)\mathbf{A}$ — Associative law for numerical multiple.
  4. $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$ — Commutative law for addition.
  5. $(\mathbf{A}+\mathbf{B})+\mathbf{C} = \mathbf{A}+(\mathbf{B}+\mathbf{C})$ — Associative law for addition.
  6. $\mathbf{A}+\mathbf{0} = \mathbf{A}$ — Neutral element for addition.
  7. $\mathbf{A}+(-\mathbf{A}) = \mathbf{0}$ — Inverse element for addition.

  Hence we see that the $m \times n$ matrices form a vector space.

- **Definition of matrix multiplication $\mathbf{AB}$:** Two matrices $\mathbf{A}$ and $\mathbf{B}$ can be <u>*multiplied*</u> if and only if the number of columns in $\mathbf{A}$ equals the number of rows in $\mathbf{B}$. If $\mathbf{A}$ is an $m \times p$ matrix with elements a_{ij} and $\mathbf{B}$ is an $p \times n$ matrix with elements b_{ij}, then the <u>*matrix product*</u> $\mathbf{AB}$ is defined as the $m \times n$ matrix with the ij th element $\sum_{k=1}^{p} a_{ik} b_{kj}$, that is, the sum of the products of the elements from the i th row in $\mathbf{A}$ with the elements from the j th column in $\mathbf{B}$.

$$\mathbf{AB} = \begin{pmatrix} a_{11} & a_{12} & \bullet & \bullet & a_{1p} \\ a_{21} & a_{22} & \bullet & \bullet & a_{2p} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ a_{m1} & a_{m2} & \bullet & \bullet & a_{mp} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \bullet & b_{1n} \\ b_{21} & b_{22} & \bullet & b_{2n} \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ b_{p1} & b_{p2} & \bullet & b_{pn} \end{pmatrix}$$

$$\equiv \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1p}b_{p1} & a_{11}b_{12} + a_{12}b_{22} + \cdots + a_{1p}b_{p2} & \bullet & a_{11}b_{1n} + a_{12}b_{2n} + \cdots + a_{1p}b_{pn} \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2p}b_{p1} & a_{21}b_{12} + a_{22}b_{22} + \cdots + a_{2p}b_{p2} & \bullet & a_{21}b_{1n} + a_{22}b_{2n} + \cdots + a_{2p}b_{pn} \\ \bullet & \bullet & \bullet & \bullet \\ a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mp}b_{p1} & a_{m1}b_{12} + a_{m2}b_{22} + \cdots + a_{mp}b_{p2} & \bullet & a_{m1}b_{1n} + a_{m2}b_{2n} + \cdots + a_{mp}b_{pn} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=1}^{p} a_{1k}b_{k1} & \sum_{k=1}^{p} a_{1k}b_{k2} & \bullet & \sum_{k=1}^{p} a_{1k}b_{kn} \\ \sum_{k=1}^{p} a_{2k}b_{k1} & \sum_{k=1}^{p} a_{2k}b_{k2} & \bullet & \sum_{k=1}^{p} a_{2k}b_{kn} \\ \bullet & \bullet & \bullet & \bullet \\ \sum_{k=1}^{p} a_{mk}b_{k1} & \sum_{k=1}^{p} a_{mk}b_{k2} & \bullet & \sum_{k=1}^{p} a_{mk}b_{kn} \end{pmatrix}.$$

- Matrix multiplication is, in general, <u>*non-commutative*</u>, i.e. $\mathbf{AB} \neq \mathbf{BA}$. Indeed, it is possible that two matrices can be multiplied in one order but not in the other.

  Note that we write $\mathbf{A}^k = \underbrace{\mathbf{AA}\cdots\mathbf{A}}_{\text{k factors}}$ and can now define the exponential of a matrix by

$$\exp \mathbf{A} \equiv \sum_{k=1}^{\infty} \frac{\mathbf{A}^k}{k!}.$$

The following properties hold for any number r and matrices for which the indicated matrix operations are defined:

1. $(\mathbf{A}+\mathbf{B})\mathbf{C}=\mathbf{AC}+\mathbf{BC}$. Matrix multiplication is distributive over addition.
2. $\mathbf{A}(\mathbf{B}+\mathbf{C})=\mathbf{AB}+\mathbf{AC}$. Matrix multiplication is distributive over addition.
3. $(r\mathbf{A})\mathbf{B}=r(\mathbf{AB})=\mathbf{A}(r\mathbf{B})$
4. $\mathbf{A}(\mathbf{BC})=(\mathbf{AB})\mathbf{C}$. Matrix multiplication is associative.
5. $\mathbf{0A}=\mathbf{A0}=\mathbf{0}$ Zero's element for matrix multiplication,
6. $\mathbf{IA}=\mathbf{AI}=\mathbf{A}$ Neutral element for matrix multiplication.

- Let $\mathbf{A}$ denote an $m\times n$ matrix with elements a_{ij}. The <u>*transpose*</u> $\mathbf{A}^t$ of the matrix $\mathbf{A}$ is the $n\times m$ matrix obtained by exchanging rows and columns, that is, $a_{ij}^t=a_{ji}, i=1,2,\dots n; j=1,2,\dots m$:

$$\mathbf{A}=\begin{pmatrix} a_{11} & a_{12} & a_{13} & \bullet & a_{1n} \\ a_{21} & a_{22} & a_{23} & \bullet & a_{2n} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ a_{m1} & a_{m2} & a_{m3} & \bullet & a_{mn} \end{pmatrix} \Leftrightarrow \mathbf{A}^t=\begin{pmatrix} a_{11} & a_{21} & \bullet & a_{m1} \\ a_{12} & a_{22} & \bullet & a_{m2} \\ a_{13} & a_{23} & \bullet & a_{m3} \\ \bullet & \bullet & \bullet & \bullet \\ a_{1n} & a_{2n} & \bullet & a_{mn} \end{pmatrix}.\ \text{Notice that } \left(\mathbf{A}^t\right)^t=\mathbf{A}.$$

The following only apply to square matrices, that is, $n\times n$ matrices.

- A <u>*square matrix*</u> has the same number of rows as columns, that is, $m=n$:

$$\mathbf{A}=\begin{pmatrix} a_{11} & a_{12} & \bullet & \bullet & a_{1n} \\ a_{21} & a_{22} & \bullet & \bullet & a_{2n} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ a_{n1} & a_{n2} & \bullet & \bullet & a_{nn} \end{pmatrix}$$

- The (<u>*leading) diagonal*</u> of a square matrix $\mathbf{A}$ runs from top left to bottom right and includes the elements with equal indices, that is, $a_{ii}, i=1,2,\dots n$.

- The <u>*trace*</u> of a $n\times n$ matrix is the sum of the elements in the diagonal $\mathrm{Tr}\mathbf{A}=\sum_{i=1}^{n}a_{ii}$.

- A <u>*symmetric matrix*</u> is a square matrix that is unchanged if the elements are reflected in the leading diagonal. Hence, a square matrix symmetric if and only if it is equal to its transpose, that is, $\mathbf{A}$ is symmetric $\Leftrightarrow$ $\mathbf{A}=\mathbf{A}^t$.

- A <u>*diagonal matrix*</u> is a square matrix with non-zero elements only on the leading diagonal:

$$\mathbf{A}=\begin{pmatrix} a_{11} & 0 & 0 & \bullet & 0 \\ 0 & a_{12} & 0 & \bullet & 0 \\ 0 & 0 & a_{13} & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & 0 \\ 0 & 0 & \bullet & 0 & a_{nn} \end{pmatrix}$$

- The *identity matrix* **I** (sometimes called a *unit matrix*) is a diagonal matrix with all elements on the leading diagonal equal to unity and all others zero.

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 & \bullet & 0 \\ 0 & 1 & 0 & \bullet & 0 \\ 0 & 0 & 1 & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & 0 \\ 0 & 0 & \bullet & 0 & 1 \end{pmatrix}$$

- The *ij*th *minor* of an $n \times n$ matrix $\mathbf{A}$ is the $(n-1)\times(n-1)$ matrix $\mathbf{A}_{ij}$ obtained from $\mathbf{A}$ by removing the *i*th row and the *j*th column. (NB: In some books, the *ij*th minor of an $n \times n$ matrix $\mathbf{A}$ is defined as the *determinant* of the $(n-1)\times(n-1)$ matrix obtained from $\mathbf{A}$ by removing the *i*th row and the *j*th column but we use the former definition.)
- A *determinant* of an $n \times n$ matrix $\mathbf{A}$ is denoted $\det \mathbf{A}$ or $|\mathbf{A}|$ and is defined by $\det \mathbf{A} = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det \mathbf{A}_{1j}$ (expanded by first row), see Fact Sheet 6 for properties.
- The *cofactor of the element* a_{ij} is defined as $C_{ij} = (-1)^{i+j} \det \mathbf{A}_{ij}$ such that the determinant $\det \mathbf{A} = \sum_{j=1}^{n} a_{1j} C_{1j}$.
- The *adjoint* matrix $\text{adj}\,\mathbf{A}$ is obtained by writing down the matrix of the cofactors, and taking the transpose:

$$adj\mathbf{A} \equiv \begin{pmatrix} C_{11} & C_{12} & C_{13} & \bullet & C_{1n} \\ C_{21} & C_{22} & C_{23} & \bullet & C_{2n} \\ C_{31} & C_{32} & C_{33} & \bullet & C_{3n} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ C_{n1} & C_{n2} & C_{n3} & \bullet & C_{nn} \end{pmatrix}^t = \begin{pmatrix} C_{11} & C_{21} & C_{31} & \bullet & C_{n1} \\ C_{12} & C_{22} & C_{32} & \bullet & C_{n2} \\ C_{13} & C_{23} & C_{33} & \bullet & C_{n3} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ C_{1n} & C_{2n} & C_{3n} & \bullet & C_{nn} \end{pmatrix}$$

- If $\mathbf{A}$ is an $n \times n$ matrix and there exists an $n \times n$ matrix $\mathbf{A}^{-1}$ such that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ then we say that $\mathbf{A}$ is *invertible* and that $\mathbf{A}^{-1}$ is the *inverse* of $\mathbf{A}$.
- $\mathbf{A}$ is invertible $\Leftrightarrow$ $\det \mathbf{A} \neq 0$. There are several ways of obtaining the inverse if it exists, for example by using the formula $\mathbf{A}^{-1} = \dfrac{\text{adj}\,\mathbf{A}}{\det \mathbf{A}}$.
- Note that matrices that are **not** square have no inverse, and nor do square matrices with $\det \mathbf{A} = 0$. A *singular matrix* is a square matrix whose determinant is zero and which therefore has no inverse. (The term is sometimes applied to *any* matrix that has no inverse, which includes all non-square matrices!)
- An *orthogonal matrix* **O** is a square matrix whose inverse is equal to its transpose, that is, $\mathbf{O}^{-1} = \mathbf{O}^t$.

NOTATION

There is no agreed convention for identifying matrices. Some books use ordinary capital letters, while others use bold type, which makes matrices hard to distinguish from vectors. In one sense, this doesn't matter too much, given that vectors are frequently represented by row or column matrices! On the black-board we use capital letters with double underlining.

Supplement – Orthogonal Matrices

This sheet summarises the information in Classwork 5 on orthogonal matrices. The symbol **O** is used here to designate an orthogonal matrix.

Definition

An orthogonal matrix is a square matrix whose inverse is equal to its transpose, that is, $\mathbf{OO}^t = \mathbf{O}^t\mathbf{O} = \mathbf{I}$ such that $\mathbf{O}^{-1} = \mathbf{O}^t$.

It follows from this that the *columns* of **O** represent *orthonormal vectors* (i.e., unit vectors that are orthogonal (i.e. have zero dot product)) because only if this is the case can the product of **O** with its transpose yield the identity matrix. This statement applies equally to the rows of **O**.

Properties

- The determinant of an orthogonal matrix $\det\mathbf{O} = \pm 1$.

  This follows from the observation that $1 = \det\mathbf{I} = \det\mathbf{OO}^t = \det\mathbf{O}\det\mathbf{O}^t = (\det\mathbf{O})^2$. The third step follows from a general property of determinants that the determinant of a product is the product of the determinants, that is, $\det\mathbf{AB} = \det\mathbf{A}\det\mathbf{B}$.

- The dot product of two vectors is invariant (i.e., unchanged) under an orthogonal transformation.
  Consider two vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ that are transformed by the matrix **O** into vectors $\mathbf{b}_1$ and $\mathbf{b}_2$, that is, $\mathbf{b}_1 = \mathbf{Oa}_1$ and $\mathbf{b}_2 = \mathbf{Oa}_2$. Then we find that
  $\mathbf{b}_1 \bullet \mathbf{b}_2 = \mathbf{b}_1^t\mathbf{b}_2 = (\mathbf{Oa}_1)^t\,\mathbf{Oa}_2 = \left(\mathbf{a}_1^t\mathbf{O}^t\right)\mathbf{Oa}_2 = \mathbf{a}_1^t\left(\mathbf{O}^t\mathbf{O}\right)\mathbf{a}_2 = \mathbf{a}_1^t\mathbf{I}\mathbf{a}_2 = \mathbf{a}_1^t\mathbf{a}_2 = \mathbf{a}_1 \bullet \mathbf{a}_2$.

  The conclusion is that $\mathbf{Oa}_1 \bullet \mathbf{Oa}_2 = \mathbf{a}_1 \bullet \mathbf{a}_2 =$.

- The magnitude of a vector is not changed by an orthogonal transformation.

  This follows by considering a special case of the previous property. When $\mathbf{a}_1 = \mathbf{a}_2 \equiv \mathbf{a}$,

  then $\mathbf{b}_1 = \mathbf{b}_2 \equiv \mathbf{b}$, the previous result yields $|\mathbf{b}|^2 = \mathbf{b} \bullet \mathbf{b} = \mathbf{a} \bullet \mathbf{a} = |\mathbf{a}|^2$. The magnitude of the vector is therefore not changed by the transformation.

- Rotation matrices are examples of orthogonal matrices; this makes sense because a rotation does not, of course change the dot product between two vectors not the length of a vector. For pure rotations, $\det\mathbf{R} = +1$; if the operation also involves a reflection, the determinant is –1.