

# MATH1002

Mingyuan Ba

February 29, 2024

## 1 Week 1

1. A vector  $\mathbf{v}$  is a directed line segment that corresponds to a displacement from a point  $A$  to a point  $B$ . The vector from  $A$  to  $B$  is denoted  $\overrightarrow{AB}$ , with  $A$  the initial point or tail and  $B$  the terminal point or head.
2. The set of all points in  $n$ -dimensional space corresponds to the set of all vectors with tails at the origin  $O$ . To each point  $A$ , there corresponds the vector  $\overrightarrow{OA}$ , called the position vector of  $A$ .
3. If  $A$  is the point in the plane with coordinates  $(a_1, a_2)$ , then  $\overrightarrow{OA} = [a_1, a_2]$ . The entries  $a_1$  and  $a_2$  are the components of the vector  $\overrightarrow{OA}$ . We often use the column vector  $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$  rather than the row vector  $[a_1, a_2]$ . Similarly if  $A$  is the point in  $n$ -dimensional space with coordinates  $(a_1, a_2, \dots, a_n)$ ,

then the vector  $\overrightarrow{OA}$  is  $[a_1, a_2, \dots, a_n]$  or  $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ , and this vector has  $i$ th component  $a_i$ .

4. Two vectors are equal if and only if their corresponding components are equal. Geometrically, two vectors are equal if and only if they have the same length and the same direction, equivalently, we could translate one on to the other. A vector with its tail at the origin is in standard position.

*assignment2 - 2.3*

5. The set of all vectors with  $n$  components is denoted  $\mathbb{R}^n$ . This is the same as the set of all ordered  $n$ -tuples, written as row or column vectors.
6. If  $\mathbf{u} = [u_1, u_2]$  and  $\mathbf{v} = [v_1, v_2]$  then their sum  $\mathbf{u} + \mathbf{v}$  is the vector  $\mathbf{u} + \mathbf{v} = [u_1 + v_1, u_2 + v_2]$ . Similarly in  $\mathbb{R}^n$ , if  $\mathbf{u} = [u_1, u_2, \dots, u_n]$  and  $\mathbf{v} = [v_1, v_2, \dots, v_n]$ , then

$$\mathbf{u} + \mathbf{v} = [u_1 + v_1, u_2 + v_2, \dots, u_n + v_n]$$

7. Head-to-Tail Rule for Vector Addition: Translate  $\mathbf{v}$  so its tail is at the head of  $\mathbf{u}$ . Then  $\mathbf{u} + \mathbf{v}$  is the vector from the tail of  $\mathbf{u}$  to the head of  $\mathbf{v}$ .
8. Parallelogram Rule: If  $\mathbf{u}$  and  $\mathbf{v}$  are in standard position, then  $\mathbf{u} + \mathbf{v}$  is the vector in standard position along the diagonal of the parallelogram determined by  $\mathbf{u}$  and  $\mathbf{v}$ .
9. The zero vector in  $\mathbb{R}^n$  is  $\mathbf{0} = [0, 0, \dots, 0] = \overrightarrow{OO}$ . This vector has length 0 and its direction is not defined. For all vectors  $\mathbf{u}$  in  $\mathbb{R}^n$ ,  $\mathbf{u} + \mathbf{0} = \mathbf{u}$ .
10. A scalar is a real number. Given a scalar  $c$  and a vector  $\mathbf{v}$ , the scalar multiple  $c\mathbf{v}$  is the vector obtained by multiplying each component of  $\mathbf{v}$  by  $c$ . So if  $\mathbf{v} = [v_1, v_2]$  then  $c\mathbf{v} = [cv_1, cv_2]$ . Similarly in  $\mathbb{R}^n$ , if  $\mathbf{v} = [v_1, v_2, \dots, v_n]$  then  $c\mathbf{v} = [cv_1, cv_2, \dots, cv_n]$ . For all vectors  $\mathbf{v}$ ,  $0\mathbf{v} = \mathbf{0}$  and  $1\mathbf{v} = \mathbf{v}$ , and for all scalars  $c$  and  $d$ ,  $c(d\mathbf{v}) = (cd)\mathbf{v}$ . Geometrically,  $c\mathbf{v}$  points in the same direction as  $\mathbf{v}$  if  $c > 0$  and the opposite direction if  $c < 0$ , and  $c\mathbf{v}$  has length  $|c|$  times the length of  $\mathbf{v}$ . Two vectors are scalar multiples of each other if and only if they are parallel.
11. The negative of a vector  $\mathbf{v}$  is the vector  $(-1)\mathbf{v} = -\mathbf{v}$ . This has the same length as  $\mathbf{v}$  but points in the opposite direction. For all vectors  $\mathbf{v}$ ,  $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$ . If  $A$  and  $B$  are points then  $\overrightarrow{AB} = -\overrightarrow{BA}$ .
12. We define vector subtraction by  $\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$ .
13. If  $P$  and  $Q$  are points in  $\mathbb{R}^n$  then  $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}$ . (xiv) Commutative Law of Addition: For all vectors  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathbb{R}^n$ ,  $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ .
14. Associative Law of Addition: For all vectors  $\mathbf{u}$ ,  $\mathbf{v}$  and  $\mathbf{w}$  in  $\mathbb{R}^n$ ,  $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ .
15. Distributive Laws: For all scalars  $c$  and  $d$  and all vectors  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathbb{R}^n$ ,  $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$  and  $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ .

## 2 Week 2

1. A vector  $\mathbf{v}$  is a linear combination of vectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  if there are scalars  $c_1, c_2, \dots, c_k$  such that

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k$$

The scalars  $c_1, c_2, \dots, c_k$  are the coefficients of this linear combination.

2. The dot product of  $\mathbf{u} = [u_1, u_2, \dots, u_n]$  and  $\mathbf{v} = [v_1, v_2, \dots, v_n]$  is the scalar

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

3. Commutativity of dot product: For all vectors  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathbb{R}^n$ ,  $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ .
4. Distributivity of dot product: For all vectors  $\mathbf{u}, \mathbf{v}$  and  $\mathbf{w}$  in  $\mathbb{R}^n$ ,  $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$ .
5. For all scalars  $c$  and all vectors  $\mathbf{v}$  and  $\mathbf{w}$  in  $\mathbb{R}^n$ ,  $(c\mathbf{v}) \cdot \mathbf{w} = c(\mathbf{v} \cdot \mathbf{w}) = \mathbf{v} \cdot (c\mathbf{w})$ .
6. The length of a vector  $\mathbf{v} = [v_1, v_2, \dots, v_n]$  in  $\mathbb{R}^n$  is the scalar

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}$$

*assignment1 - 1.3*

7. For all scalars  $c$  and all vectors  $\mathbf{v}$  in  $\mathbb{R}^n$ ,  $\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$ .
8. A unit vector is a vector of length one. Given any nonzero vector  $\mathbf{v}$ , the vector  $\left(\frac{1}{\|\mathbf{v}\|}\right)\mathbf{v}$  is the unit vector in the same direction as  $\mathbf{v}$ , called the normalisation of  $\mathbf{v}$ .

*W3-tut - 2*

9. In  $\mathbb{R}^2$  the unit vectors  $\mathbf{e}_1 = [1, 0]$  and  $\mathbf{e}_2 = [0, 1]$  lie along the positive  $x$ - and  $y$ -axes, respectively. In  $\mathbb{R}^3$  the unit vectors  $\mathbf{e}_1 = [1, 0, 0]$ ,  $\mathbf{e}_2 = [0, 1, 0]$  and  $\mathbf{e}_3 = [0, 0, 1]$  lie along the positive  $x$ -,  $y$ - and  $z$ -axes, respectively. In  $\mathbb{R}^n$  we have the standard unit vectors  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ , where  $\mathbf{e}_i$  has a 1 in its  $i$ th component and 0 elsewhere.
10. Cauchy-Schwarz Inequality: For all vectors  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathbb{R}^n$ ,  $|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\|\|\mathbf{v}\|$ .
11. Triangle Inequality: For all vectors  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathbb{R}^n$ ,  $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ .
12. The distance between vectors  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathbb{R}^n$  is  $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$ .
13. The angle  $\theta$  between nonzero vectors  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathbb{R}^n$  is given by

$$\cos \theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}, \quad \text{with } 0 \leq \theta \leq \pi$$

14. Two vectors  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathbb{R}^n$  are orthogonal if  $\mathbf{u} \cdot \mathbf{v} = 0$ , meaning that the angle between them is a right angle.

*assignment1 - 2.1*

15. If  $\mathbf{u}$  and  $\mathbf{v}$  are vectors in  $\mathbb{R}^n$  and  $\mathbf{u} \neq \mathbf{0}$ , the projection of  $\mathbf{v}$  onto  $\mathbf{u}$  is the vector

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|^2}\right)\mathbf{u} = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}}\right)\mathbf{u}$$

*assignment1 - 2.2*

### 3 Week 3

1. In  $\mathbb{R}^3$ , the cross product of  $\mathbf{u} = [u_1, u_2, u_3]$  and  $\mathbf{v} = [v_1, v_2, v_3]$  is the vector

$$\mathbf{u} \times \mathbf{v} = [u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1]$$

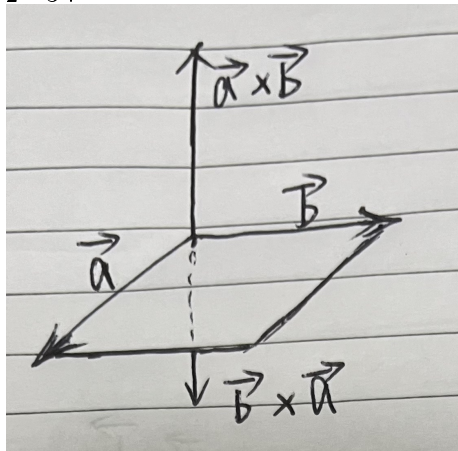
Handwritten derivation of the cross product formula for two vectors  $\mathbf{u} = [u_1, u_2, u_3]$  and  $\mathbf{v} = [v_1, v_2, v_3]$ . The derivation shows the expansion of the determinant:

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \mathbf{i}(u_2v_3 - u_3v_2) - \mathbf{j}(u_1v_3 - u_3v_1) + \mathbf{k}(u_1v_2 - u_2v_1)$$

W3-tut-1

2. The cross product  $\mathbf{u} \times \mathbf{v}$  is orthogonal to both  $\mathbf{u}$  and  $\mathbf{v}$ .
3. The direction of  $\mathbf{u} \times \mathbf{v}$  is given by the right-hand rule.

2 - 3 :



W4-tut-4

4. Anti-commutativity of cross product: For all vectors  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathbb{R}^3$ ,  $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$ .
5. Distributivity of cross product: For all vectors  $\mathbf{u}, \mathbf{v}$  and  $\mathbf{w}$  in  $\mathbb{R}^3$ ,  $(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$

6. If  $\mathbf{v}$  and  $\mathbf{w}$  are vectors in  $\mathbb{R}^3$  and  $c$  is a scalar then :  $(c\mathbf{v}) \times \mathbf{w} = c(\mathbf{v} \times \mathbf{w}) = \mathbf{v} \times (c\mathbf{w})$ .
7. For all vectors  $\mathbf{u}$  in  $\mathbb{R}^3$ ,  $\mathbf{u} \times \mathbf{u} = \mathbf{0}$ .
8. If  $\theta$  is the angle between vectors  $\mathbf{v}$  and  $\mathbf{w}$  in  $\mathbb{R}^3$  then

$$\|\mathbf{v} \times \mathbf{w}\| = \|\mathbf{v}\|\|\mathbf{w}\|\sin \theta$$

*W3-tut-4*

9. The area of the parallelogram inscribed by  $\mathbf{v}$  and  $\mathbf{w}$  in  $\mathbb{R}^3$  is  $\|\mathbf{v} \times \mathbf{w}\|$ , and the area of the triangle inscribed by  $\mathbf{v}$  and  $\mathbf{w}$  is  $\frac{1}{2}\|\mathbf{v} \times \mathbf{w}\|$ .

*assignment1 - 1.1 W3-tut-3*

10. A normal vector for a line  $\ell$  is a nonzero vector  $\mathbf{n}$  which is orthogonal to any vector parallel to  $\ell$ .
- W4-tut-5*
11. A direction vector for a line  $\ell$  is a nonzero vector  $\mathbf{d}$  which is parallel to  $\ell$ . If  $A$  and  $B$  are distinct points on  $\ell$  then  $\overrightarrow{AB}$  is a direction vector for  $\ell$
12. The normal form of the equation for a line  $\ell$  in  $\mathbb{R}^2$  is

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0, \quad \text{or , equivalently,} \quad \mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$$

where  $\mathbf{p}$  is a point on  $\ell$ , and  $\mathbf{n}$  is a normal vector for  $\ell$ .

13. The general form of the equation for a line  $\ell$  in  $\mathbb{R}^2$  is

$$ax + by = c$$

where  $\mathbf{n} = [a, b]$  is a normal vector for  $\ell$ .

14. The vector form of the equation for a line  $\ell$  in  $\mathbb{R}^2$  or  $\mathbb{R}^3$  is

$$\mathbf{x} = \mathbf{p} + t\mathbf{d}$$

where  $\mathbf{p}$  is a point on  $\ell$ ,  $\mathbf{d}$  is a direction vector for  $\ell$  and  $t \in \mathbb{R}$ .

(xv) The parametric equations of  $\ell$  are the equations corresponding to the components:

$$x = p_1 + td_1 \quad \text{and} \quad y = p_2 + td_2 \quad \left( \text{and, for lines in } \mathbb{R}^3, \quad z = p_3 + td_3 \right), \quad \text{for } t \in \mathbb{R}.$$

## 4 Week 4

1. The normal form of the equation of a plane  $\mathcal{P}$  in  $\mathbb{R}^3$  is

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0 \quad \text{or equivalently} \quad \mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$$

where  $\mathbf{p}$  is a point on  $\mathcal{P}$  and  $\mathbf{n}$  is a normal vector for  $\mathcal{P}$ .

2. The general form of the equation for a plane  $\mathcal{P}$  in  $\mathbb{R}^3$  is

$$ax + by + cz = d$$

where  $\mathbf{n} = [a, b, c]$  is a normal vector for  $\mathcal{P}$ .

3. Recall that a vector  $\mathbf{v}$  is a linear combination of vectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  if there are scalars  $c_1, c_2, \dots, c_k$  such that

$$\mathbf{v} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k$$

The scalars  $c_1, c_2, \dots, c_k$  are the coefficients of this linear combination.

4. If  $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$  is a set of vectors in  $\mathbb{R}^n$ , then the span of  $\mathbf{v}_1, \dots, \mathbf{v}_k$ , denoted  $\text{span}(S)$  or  $\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ , is the set of all linear combinations of  $\mathbf{v}_1, \dots, \mathbf{v}_k$ . That is

$$\begin{aligned} \text{span}(S) &= \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_k) \\ &= \{\mathbf{v} \in \mathbb{R}^n : \mathbf{v} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k \text{ for some } c_1, c_2, \dots, c_k \in \mathbb{R}\} \end{aligned}$$

If  $\text{span}(S) = \mathbb{R}^n$ , then  $S$  is called a spanning set for  $\mathbb{R}^n$ .

5. A set of vectors  $\mathbf{v}_1, \dots, \mathbf{v}_k$  is linearly independent if the only scalars  $c_1, \dots, c_k$  that satisfy

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k = \mathbf{0}$$

are  $c_1 = c_2 = \dots = c_k = 0$ . A set of vectors that is not linearly independent is called linearly dependent.

6. Two non-zero vectors in  $\mathbb{R}^n$  are linearly independent if and only if they are not parallel.

*W4-tut-2*

7. In  $\mathbb{R}^3$ , the span of any two non-zero non-parallel vectors is the plane through the origin. In other words, let  $\mathbf{u}$  and  $\mathbf{v}$  be non-zero non-parallel vectors in  $\mathbb{R}^3$ , then the set

$$S = \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = s\mathbf{u} + t\mathbf{v}, \text{ for some } s, t \in \mathbb{R}\}$$

is the plane through the origin and parallel to the vectors  $\mathbf{u}$  and  $\mathbf{v}$ .

W4-tut-6

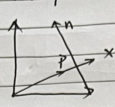
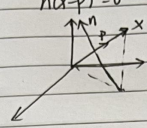
8. The vector form of the equation for a plane  $\mathcal{P}$  in  $\mathbb{R}^3$  is

$$\mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}$$

where  $\mathbf{p}$  is a point on  $\mathcal{P}$ ,  $\mathbf{u}$  and  $\mathbf{v}$  are direction vectors for  $\mathcal{P}$  (meaning  $\mathbf{u}$  and  $\mathbf{v}$  are nonzero and parallel to  $\mathcal{P}$ , but not parallel to each other), and  $s, t \in \mathbb{R}$ . The parametric equations of  $\mathcal{P}$  are the equations corresponding to the components:

$$x = p_1 + su_1 + tv_1, \quad y = p_2 + su_2 + tv_2 \quad \text{and} \quad z = p_3 + su_3 + tv_3.$$

summarize for vector equations :

	$\mathbb{R}^2$	$\mathbb{R}^3$ (plane)
normal form	$n(x-p)=0$ 	$n(x-p)=0$ 
general form	$ax+by=c$ $[a, b] \leftarrow \text{normal vector}$	$ax+by+cz=d$ $[a, b, c] \leftarrow \text{normal vector}$
vector form	$\mathbf{x} = \mathbf{p} + t\mathbf{d}$ $\downarrow$ a point      direction vector	$\mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}$ $\downarrow$ $\downarrow$ a point      direction vector
Parametric equations	$\begin{cases} x = p_1 + td_1 \\ y = p_2 + td_2 \end{cases}$	$\begin{cases} x = p_1 + su_1 + tv_1 \\ y = p_2 + su_2 + tv_2 \\ z = p_3 + su_3 + tv_3 \end{cases}$

W3-tut-5-7 W4-tut-1 W4-tut-3 W4-tut-7-8

9. If  $\mathbf{u}$  and  $\mathbf{v}$  are direction vectors for a plane  $\mathcal{P}$ , then  $\mathbf{u} \times \mathbf{v}$  is a normal vector for  $\mathcal{P}$ .
10. If  $A, B$  and  $C$  are three non-collinear points on a plane  $\mathcal{P}$ , then  $\overrightarrow{AB}$  and  $\overrightarrow{AC}$  are direction vectors for  $\mathcal{P}$ , hence  $\overrightarrow{AB} \times \overrightarrow{AC}$  is a normal vector to  $\mathcal{P}$ .





## 5 W5

1. A linear equation in variables  $x_1, x_2, \dots, x_n$  has the form

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b$$

where  $a_1, a_2, \dots, a_n, b$  are constants.

2. A system of linear equations has the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

where the  $a_{ij}$  and  $b_i$ , for  $1 \leq i \leq m$  and  $1 \leq j \leq n$ , are constants.

The system is homogeneous if  $b_1 = b_2 = \dots = b_m = 0$ .

3. Every system of linear equations has either no solutions, one solution, or infinitely many solutions. If it has no solutions, then the system is called inconsistent. If it has at least one solution, then the system is called consistent.
4. To simplify the solution of a system of linear equations an augmented matrix can be used. It is a matrix

$$\left[ \begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right]$$

One can apply the following elementary row operations to the augmented matrix without changing the solutions of the initial system:

- (a)  $R_i \leftrightarrow R_j$  (row swap)

$$\left[ \begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} & b_i \\ \dots & \dots & \dots & \dots & \dots \\ a_{j1} & a_{j2} & \dots & a_{jn} & b_j \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right] \longrightarrow \left[ \begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ \dots & \dots & \dots & \dots & \dots \\ a_{j1} & a_{j2} & \dots & a_{jn} & b_j \\ \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} & b_i \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right]$$

- (b)  $R_i \mapsto \lambda R_i$  ( $\lambda \neq 0$ , row scale)

$$\left[ \begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{in} & b_i \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right] \longrightarrow \left[ \begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \lambda a_{i1} & \lambda a_{i2} & \cdots & \lambda a_{in} & \lambda b_i \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right]$$

(c)  $R_i \mapsto R_i + \lambda R_j$  (row add)

$$\left[ \begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{in} & b_i \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} & b_j \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right] \longrightarrow \left[ \begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} + \lambda a_{j1} & a_{i2} + \lambda a_{j2} & \cdots & a_{in} + \lambda a_{jn} & b_i + \lambda b_j \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} & b_j \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right]$$

5. An augmented matrix is in row echelon form if
  - (a) rows of zeroes appear at the bottom
  - (b) the first nonzero entries of consecutive rows appear further to the right
  - (b) Each leading entry is equal to 1
6. The process of Gaussian elimination applies elementary row operations (row reduction) to transform the augmented matrix of a system into row echelon form, after which the associated system is solved using back substitution:
  - (a) the leading variables corresponding to leading entries are evaluated one equation at a time from the bottom towards the top,
  - (b) parameters are assigned to each nonleading variable (if any).
7. An augmented matrix is in reduced row echelon form if
  - (a) It is in row echelon form.
  - (b) The leading entry in each nonzero row is a 1 (called a leading 1 ).
  - (c) Each column containing a leading 1 has zeros everywhere else.
8. The process of Gauss-Jordan elimination row reduces the augmented matrix to reduced row echelon form, after which the process of back substitution simplifies. However, Gauss-Jordan elimination is usually less efficient in terms of the overall number of arithmetic operations used than Gaussian elimination.

## 6 W6

1. A matrix is an array of numbers, called entries. The plural of matrix is matrices. If a matrix  $M$  has  $m$  rows and  $n$  columns then we say that  $M$  is  $m \times n$ . We call a matrix  $M$  square if  $M$  is  $n \times n$  for some  $n$ . The entry lying in the  $i$  th row and  $j$  th column is called the  $(i, j)$ -entry.
2. A matrix consisting of one row is called a row matrix. A matrix consisting of one column is called a column matrix.
3. To add or subtract using two matrices of the same size, simply add or subtract the corresponding entries. To form the negative of a matrix, simply take the negatives of its entries.
4. To multiply a matrix by a scalar, simply multiply all of its entries by the scalar.
5. The zero matrix has all of its entries equal to 0 , and is denoted by  $0$  , or by  $0_{m \times n}$  if it is  $m \times n$  and its size needs to be emphasised.
6. The identity matrix is a square matrix with diagonal entries equal to 1 , and all entries off the diagonal equal to 0. The identity matrix is denoted by  $I$ , or by  $I_n$  if it is  $n \times n$  and its size needs to be emphasised.
7. If  $A$  is an  $m \times n$  matrix and  $B$  is an  $n \times p$  matrix, then the product  $AB$  is defined and is an  $m \times p$  matrix. The  $(i, j)$ -entry of  $AB$  is

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

i.e. the dot product of the  $i$  th row of  $A$  and the  $j$  th column of  $B$ , for each  $1 \leq i \leq m$  and  $1 \leq j \leq p$ .

8. The transpose of an  $m \times n$  matrix  $A$  is the  $n \times m$  matrix  $A^T$  obtained by interchanging the rows and columns of  $A$ . So the  $i$  th column of  $A^T$  is the  $i$  th row of  $A$  for all  $i$ .
9. If  $A, B$ , and  $C$  are matrices of the appropriate size (so all of the following sums and products make sense), and  $\lambda$  and  $\mu$  are scalars, then the following properties hold:

$$\begin{aligned} A + B &= B + A, & (A + B) + C &= A + (B + C), & A + 0 &= 0 + A = A, \\ -(-A) &= A, & A + (-A) &= A - A = 0, & \lambda(\mu A) &= (\lambda\mu)A, \\ \lambda(A + B) &= \lambda A + \lambda B, & (\lambda + \mu)A &= \lambda A + \mu A, & IA &= AI = A, \\ (AB)C &= A(BC), & A(B + C) &= AB + AC, & (A + B)C &= AC + BC, \\ \lambda(BC) &= (\lambda B)C = B(\lambda C), & 0A &= 0 = A0. \end{aligned}$$

10. Warning: Matrix multiplication is not in general commutative. In most cases, we have  $AB \neq BA$ .
11. If  $A$  and  $B$  are matrices of the appropriate size (so all of the following sums and products make sense), and  $\lambda$  and  $\mu$  are scalars, then the following properties hold:

$$\begin{aligned} (A^T)^T &= A, & (A+B)^T &= A^T + B^T, & (\lambda A)^T &= \lambda (A^T) \\ (AB)^T &= B^T A^T, & (A^n)^T &= (A^T)^n & \text{for all nonnegative integers } n \end{aligned}$$

## 7 W7

1. The inverse of a matrix  $A$  is a matrix  $A^{-1}$  such that, for some positive integer  $n$ ,

$$AA^{-1} = A^{-1}A = I_n$$

**Only square matrices have inverses.** When it exists, **the inverse  $A^{-1}$  is unique.**

2. Only half of the definition needs to be checked, in the following sense: if  $A$  is a square matrix and  $AB = I$  or  $BA = I$ , then

$$AB = BA = I$$

so that the inverse  $A^{-1}$  exists and equals  $B$ .

3. Let  $A$  be an invertible matrix. Then  $A^T$  is invertible and  $(A^T)^{-1} = (A^{-1})^T$ .
4. A matrix is invertible if its inverse exists. If  $A$  and  $B$  are invertible matrices of the same size, then  $AB$  is invertible and  $(AB)^{-1} = B^{-1}A^{-1}$ .
5. For all integers  $m, n$  and all nonzero scalars  $\lambda$ ,

$$A^m A^n = A^{m+n}, \quad (A^{-1})^{-1} = A, \quad (A^m)^n = A^{mn}, \quad (\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}$$

6. The determinant of a  $1 \times 1$  matrix  $[a]$  is simply the entry  $a$ .

7. The determinant of a  $2 \times 2$  matrix  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  is

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

8. Let  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ . Then  $A$  is invertible if and only if  $\det(A) \neq 0$ , in which case  $A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$
9. A square matrix  $A$  is invertible if and only if the augmented matrix  $[A \mid I]$  can be row reduced to  $[I \mid B]$ , in which case  $A^{-1} = B$
10. If a system of equations can be expressed in the form  $A\mathbf{x} = \mathbf{b}$  where  $A$  is invertible, then  $\mathbf{x} = A^{-1}\mathbf{b}$ .

## 8 W8

1. If a system of equations can be expressed in the form  $A\mathbf{x} = \mathbf{b}$  where  $A$  is invertible, then  $\mathbf{x} = A^{-1}\mathbf{b}$ .
2. An elementary matrix is any matrix that can be obtained by performing an elementary row operation on an **identity matrix**.
3. Suppose  $E$  is an elementary matrix obtained by performing an elementary row operation on the identity matrix  $I_n$ . If  $A$  is an  $n \times p$  matrix and we perform the same elementary row operation on  $A$ , then the resulting matrix is the same as the product  $EA$ .
4. The determinant of a  $1 \times 1$  matrix  $[a]$  is simply the entry  $a$ .
5. The determinant of a  $2 \times 2$  matrix  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  is  $\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$ .
6. Expanding along any row or down any column of a square matrix  $A$  produces the same real number, called the determinant of  $A$ , denoted by  $\det(A)$  (or  $\det A$ ) or  $|A|$  :  
If  $A = [a_{i,j}]$  is an  $n \times n$  matrix, then

$$\det(A) = \sum_{k=1}^n a_{i,k} C_{i,k} = \sum_{k=1}^n a_{k,j} C_{k,j}$$

where  $C_{ij} = (-1)^{i+j} \det A_{ij}$  is the  $(i, j)$ -cofactor of  $A$ , and  $A_{ij}$  is the  $(i, j)$ -minor of  $A$ , defined as the  $(n-1) \times (n-1)$  matrix obtained from  $A$  by removing row  $i$  and column  $j$ .

7. The determinant of a  $3 \times 3$  matrix  $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}$  is

$$\det(A) = |A| = a \begin{vmatrix} e & f \\ h & k \end{vmatrix} - b \begin{vmatrix} d & f \\ g & k \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

Note that this is the expansion along the first row (where the smaller determinant arises by ignoring the row and column of the entry being used as a coefficient), but expanding along any row or column is valid, as long as we remember to multiply the determinant of the  $(i, j)$ -minor of  $A$  by  $(-1)^{i+j}$ . So, for a  $3 \times 3$  matrix, we get the following pattern of positive and negative signs:

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$$

8. If  $A$  is triangular, in the sense that all entries above or below the diagonal are zero, then  $\det(A)$  is the product of the diagonal elements.
9. Determinant method for cross products: If  $\mathbf{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3$  and  $\mathbf{w} = w_1\mathbf{e}_1 + w_2\mathbf{e}_2 + w_3\mathbf{e}_3$  then

$$\begin{aligned}\mathbf{u} \times \mathbf{w} &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ u_1 & u_2 & u_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \\ &= (u_2w_3 - u_3w_2)\mathbf{e}_1 - (u_1w_3 - u_3w_1)\mathbf{e}_2 + (u_1w_2 - u_2w_1)\mathbf{e}_3 \\ &= [u_2w_3 - u_3w_2, u_3w_1 - u_1w_3, u_1w_2 - u_2w_1]\end{aligned}$$



## 9 W9

1. The determinant of a  $1 \times 1$  matrix  $[a]$  is simply the entry  $a$ .
2. The determinant of a  $2 \times 2$  matrix  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  is  $\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$ .
3. Expanding along any row or down any column of a square matrix  $A$  produces the same real number, called the determinant of  $A$ , denoted by  $\det(A)$  (or  $\det A$ ) or  $|A|$  :  
If  $A = [a_{i,j}]$  is an  $n \times n$  matrix, then

$$\det(A) = \sum_{k=1}^n a_{i,k} C_{i,k} = \sum_{k=1}^n a_{k,j} C_{k,j}$$

where  $C_{ij} = (-1)^{i+j} \det A_{ij}$  is the  $(i, j)$ -cofactor of  $A$ , and  $A_{ij}$  is the  $(i, j)$ -minor of  $A$ , defined as the  $(n-1) \times (n-1)$  matrix obtained from  $A$  by removing row  $i$  and column  $j$ .

4. A quick way to remember the signs is the following "checkerboard" pattern

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

5. If  $A$  is triangular, in the sense that all entries above or below the diagonal are zero, then  $\det(A)$  is the product of the diagonal elements.
6.  $A$  is invertible if and only if  $\det(A) \neq 0$ .
7. Multiplicative property of the determinant: For any  $n \times n$  matrices  $A$  and  $B$ , we have

$$\det(AB) = \det(A) \det(B)$$

8. If  $A$  is invertible then  $\det(A^{-1}) = \frac{1}{\det(A)}$ .
9. Recall that the transpose  $A^T$  of a matrix  $A$  is the matrix obtained by interchanging rows and columns of  $A$ . Since the determinant of a matrix can be computed by expanding along any row or column of the matrix, taking the transpose of a matrix does not change its determinant. That is,

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

10. If  $B$  is obtained from a matrix  $A$  by swapping two rows of  $A$  then

$$\det(B) = -\det(A)$$

11. If  $B$  is obtained from a matrix  $A$  by adding a scalar multiple of one row of  $A$  to another row of  $A$ , then

$$\det(B) = \det(A)$$

12. If  $B$  is obtained from a matrix  $A$  by multiplying a row of  $A$  by a scalar  $\mu \in \mathbb{R}$ , then

$$\det(B) = \mu \det(A)$$

13. If  $A$  is an  $n \times n$  matrix, then  $\det(kA) = k^n \det(A)$ .

14. Let  $A$  be an  $n \times n$  matrix. The system  $A\mathbf{x} = \mathbf{b}$  has a unique solution if and only if  $\det(A) \neq 0$ .

15. Let  $A$  be a square matrix,  $\mathbf{x}$  a nonzero column vector, and  $\lambda$  a scalar such that

$$A\mathbf{x} = \lambda\mathbf{x}$$

Then  $\lambda$  is called an eigenvalue of  $A$ , and  $\mathbf{x}$  is called an eigenvector of  $A$  corresponding to the eigenvalue  $\lambda$ . An  $n \times n$  matrix has at most  $n$  distinct eigenvalues.

16. A scalar  $\lambda$  is an eigenvalue of a square matrix  $A$  if and only if

$$\det(A - \lambda I) = 0$$

17. The expression  $\det(A - \lambda I)$  is a polynomial in  $\lambda$  and is called the characteristic polynomial of  $A$ . The equation  $\det(A - \lambda I) = 0$  is called the characteristic equation of  $A$ . The eigenvalues of a matrix are precisely the solutions to its characteristic equation.

18. The characteristic polynomial of an  $n \times n$  matrix is a degree-  $n$  polynomial, and so it has at most  $n$  distinct roots. Therefore, an  $n \times n$  matrix has **at most**  $n$  distinct eigenvalues.

## 10 W10

1. Let  $A$  be a square matrix,  $\mathbf{x}$  a nonzero column vector, and  $\lambda$  a scalar such that

$$A\mathbf{x} = \lambda\mathbf{x}$$

Then  $\lambda$  is called an eigenvalue of  $A$ , and  $\mathbf{x}$  is called an eigenvector of  $A$  corresponding to the eigenvalue  $\lambda$ . An  $n \times n$  matrix has at most  $n$  distinct eigenvalues.

2. A scalar  $\lambda$  is an eigenvalue of a square matrix  $A$  if and only if

$$\det(A - \lambda I) = 0$$

3. The expression  $\det(A - \lambda I)$  is a polynomial in  $\lambda$  and is called the characteristic polynomial of  $A$ . The equation  $\det(A - \lambda I) = 0$  is called the characteristic equation of  $A$ . The eigenvalues of a matrix are precisely the solutions to its characteristic equation.
4. The characteristic polynomial of an  $n \times n$  matrix is a degree-  $n$  polynomial, and so it has at most  $n$  distinct roots. Therefore, an  $n \times n$  matrix has at most  $n$  distinct eigenvalues.
5. Finding the eigenspace corresponding to the eigenvalue  $\lambda$  of a matrix  $A$  is equivalent to solving the homogeneous system with coefficient matrix  $A - \lambda I$ . After the eigenspace has been found, substituting in particular nonzero values of the parameters yields particular eigenvectors.
6. If  $A$  is a triangular matrix, in the sense that all of the entries above or below the diagonal are zero, then  $\det(A)$  is equal to the product of all of the diagonal entries of  $A$ . **This implies that the eigenvalues of a triangular matrix are simply the diagonal entries of the matrix.**
7. The algebraic multiplicity of an eigenvalue of a matrix is its multiplicity as a root of the characteristic equation. So, for instance, if the characteristic equation is  $(\lambda - 2)(\lambda - 5)^4 = 0$ , then the algebraic multiplicity of  $\lambda = 2$  is 1, and the algebraic multiplicity of  $\lambda = 5$  is 4.
8. The geometric multiplicity of an eigenvalue  $\lambda$  of the matrix  $A$  is the number of free variables in the set of solutions of  $(A - \lambda I)\mathbf{x} = \mathbf{0}$  (or equivalently the number of columns without leading entries in the row echelon form of the augmented matrix  $[A - \lambda I \mid \mathbf{0}]$ ).
9. The geometric multiplicity of each eigenvalue of a square matrix is less than or equal to its algebraic multiplicity.

10. Let  $A = [a_{ij}]$  be an  $n \times n$  matrix. The trace of  $A$  is defined by

$$\operatorname{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^n a_{ii}.$$

In other words, the trace of  $A$  is the sum of all elements on the main diagonal.

11. Let  $\lambda_1, \lambda_2, \dots, \lambda_n$  be a complete set of eigenvalues (repetition included) of the  $n \times n$  matrix  $A$ , then

$$\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n, \quad \text{and} \quad \det(A) = \lambda_1 \lambda_2 \cdots \lambda_n.$$

## 11 W11

1. A square matrix  $D$  is diagonal if all entries off the diagonal are zero.
2. If  $D$  and  $E$  are diagonal matrices, then  $DE$  is also diagonal, and its diagonal entries are simply the products of corresponding diagonal entries of  $D$  and  $E$ .
3. For all integers  $n \geq 0$ , the diagonal entries of  $D^n$  are just the  $n$ th powers of the diagonal entries of  $D$ . If  $D$  is invertible, then the diagonal entries of  $D^{-n}$  are the  $(-n)$ th powers of the diagonal entries of  $D$ .
4. A matrix  $A$  is diagonalisable if there is an invertible matrix  $P$  and a diagonal matrix  $D$  such that

$$A = PDP^{-1}$$

In other words, a matrix  $A$  is diagonalisable if there is an invertible matrix  $P$  such that the matrix  $P^{-1}AP$  is diagonal.

5. Let  $A$  be an  $n \times n$  matrix with eigenvalues  $\lambda_1, \dots, \lambda_n$ , and corresponding eigenvectors  $\mathbf{v}_1, \dots, \mathbf{v}_n$ . Then

$$AP = PD$$

where  $D$  is the diagonal matrix with eigenvalues of  $A$  down the diagonal, and  $P$  is the matrix with corresponding eigenvectors as columns (i.e. in the columns corresponding to the positions of the corresponding eigenvalues in  $D$ ). If  $P$  is invertible, then

$$A = PDP^{-1} \text{ and } D = P^{-1}AP$$

for all integers  $n \geq 0$

6. In fact (v), if the eigenvalues of  $A$  are all distinct, then  $P$  is invertible, and  $A$  is diagonalisable. (Note that it is possible for an  $n \times n$  matrix to be diagonalisable even if it doesn't have  $n$  distinct eigenvalues.)
7. **The Diagonalisation Theorem:** An  $n \times n$  matrix  $A$  is diagonalisable if and only if it has  $n$  linearly independent eigenvectors. **Also**, an  $n \times n$  matrix  $A$  is diagonalisable if and only if the algebraic multiplicity of each eigenvalue is equal to its geometric multiplicity.
8. If  $A$  is diagonalisable, then powers of  $A$  can be found efficiently using the formula

$$A^n = PD^nP^{-1}$$

for  $n \in \mathbb{Z}$  and  $n \geq 0$

9. In general, if  $A$  and  $B$  are  $n \times n$  matrices, and there exists an invertible  $n \times n$  matrix  $P$  such that  $A = PBP^{-1}$ , then we say that  $A$  is similar to  $B$ , and we write  $A \sim B$ . (Note that we don't need  $A$  or  $B$  to be diagonal in order for  $A$  to be similar to  $B$ .)
10. Leslie model of population growth.
- The population is divided into  $k$  age groups of equal duration.
  - For each age group  $0 \leq i \leq k$ , we are given the birth parameter  $b_i$  equal to the average number of females produced by a female in the age group  $i$  (so  $b_i \geq 0$ ).
  - For each age group  $0 \leq i \leq k-1$ , we are given the survival probability  $s_i$  equal to the probability that a female in age group  $i$  survives into age group  $i+1$  (so  $0 \leq s_i \leq 1$ ). - The Leslie matrix for the model is

$$L = \begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_{k-1} & b_k \\ s_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{k-1} & 0 \end{bmatrix}$$

- If  $\mathbf{x}_n$  is the distribution of the female population among the age groups after  $n$  time periods (so  $\mathbf{x}_0$  is the initial distribution of female population), then

$$\mathbf{x}_n = L^n \mathbf{x}_0$$

(**Note:** This is only an approximation, since the right-hand side might be a vector with noninteger entries, while  $\mathbf{x}_n$  always only has integer entries.)

## 12 W12

1. A probability vector is a vector  $\mathbf{y} = [y_1, \dots, y_k]$  such that  $0 \leq y_i \leq 1$  for all  $1 \leq i \leq k$ , and  $\sum_{i=1}^k y_i = 1$ .
2. We call a matrix  $P$  a stochastic matrix if each column of  $P$  is a probability vector. (So all entries of a stochastic matrix are nonnegative, and the entries in each column sum to 1).
3. We say that a stochastic matrix  $P$  is regular if there exists an integer  $m > 0$  such that all entries of the matrix  $P^m$  are positive.
4. A stochastic process is a collection of random variables representing the random evolution of some system over time.
5. A stochastic process with  $k$  states is called a Markov chain if the transition probabilities from the state  $j$  to the state  $i$  are time-independent. We denote the probability of transitioning from state  $j$  to state  $i$  by  $p_{ij}$ 
  - For every  $1 \leq i, j \leq k$ , we have  $0 \leq p_{ij} \leq 1$
  - For every  $1 \leq j \leq k$ , we have  $\sum_{i=1}^k p_{ij} = 1$ .
  - The transition matrix of the Markov chain is the  $k \times k$  matrix  $P = (p_{ij})$ . (Note that  $P$  is a stochastic matrix.)
  - We denote by  $\mathbf{x}_n$  the population distribution at time  $n$ , and call it a state vector. The initial population distribution (at time zero) is  $\mathbf{x}_0$ . We have

$$\mathbf{x}_n = P^n \mathbf{x}_0, \text{ for all integers } n \geq 0$$

- The probability of transitioning from state  $j$  to state  $i$  in  $n$  steps is equal to the  $(i, j)$ -entry of the matrix  $P^n$
- We say that a Markov chain is regular if its transition matrix  $P$  is regular (that is, if there exists a time  $m$  such that the probability of transitioning from any initial state  $j$  to any state  $i$  in  $m$  steps is positive).
- If the transition matrix  $P$  is regular, then for every initial probability distribution  $\mathbf{x}_0$ , we have  $\mathbf{x}_n = P^n \mathbf{x}_0 \rightarrow \mathbf{x}$ , where  $\mathbf{x}$  is the steady-state vector satisfying  $P\mathbf{x} = \mathbf{x}$ , with nonnegative entries that sum to the total population (which is the sum of the entries of  $\mathbf{x}_0$ ). This is referred to as the asymptotic behaviour (or long-term behaviour) of the Markov chain.
- The steady-state probability vector is the unique vector  $\mathbf{x}$  satisfying  $P\mathbf{x} = \mathbf{x}$ , with nonnegative entries that sum to 1.

- Let  $P$  be a regular transition matrix of a Markov chain. Then as  $k \rightarrow \infty$ , the matrix  $P^k$  approaches the matrix  $L = \begin{bmatrix} \mathbf{x}\mathbf{x} & \cdots & \mathbf{x} \end{bmatrix}$ , where  $\mathbf{x}$  is the **steady-state probability vector** (i.e.  $P\mathbf{x} = \mathbf{x}$ , the entries of  $\mathbf{x}$  are all nonnegative, and the sum of the entries of  $\mathbf{x}$  is 1). We call  $L$  the long-range transition matrix of the Markov chain.

6. If  $P$  is the transition matrix of a Markov chain, then  $\lambda = 1$  is an eigenvalue of  $P$ .
7. Let  $P$  be a transition matrix of a Markov chain with eigenvalue  $\lambda$ . Then  $|\lambda| \leq 1$ . If  $P$  is regular and  $\lambda \neq 1$ , then  $|\lambda| < 1$ . (That is, if  $P$  is regular, then -1 is not an eigenvalue of  $P$ .)