Required for: MATH40003 Linear Algebra and Groups

Based on the lectures of Charlotte Kestner and David Evans, Imperial College London

Part I: Linear Algebra

1 Linear Transformations (Introduction)

In general, a transformation L is linear if

$$L(\alpha u_1 + \beta u_2) = \alpha L u_1 + \beta L u_2.$$

2 Systems of Linear Equations and Matrices

2.0.1 Definition

We can write a system of m linear equations in n unknowns as $A_{m\times n}x = b$ where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \vdots \\ a_{m1} & & \cdots & a_{mn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

or as the augmented matrix

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & & & b_2 \\ \vdots & & \ddots & \vdots & \vdots \\ a_{m1} & & \cdots & a_{mn} & b_m \end{array}\right).$$

2.0.2 Note

A system of linear equations in n unknowns defines a subset of n-space.

2.0.3 Definition

A system of linear equations is homogeneous if $b_1 = b_2 = ... = b_m = 0$.

2.1 Operations on Matrices

2.1.1 Definition

Let $A = [a_{ij}]_{m\times n} \in M_{m\times n}(\mathbb{R})$ and $B = [b_{ij}]_{m\times n} \in M_{m\times n}(\mathbb{R})$ (i.e. $a_{ij}, b_{ij} \in \mathbb{R}\ \forall\, i, j$). Define the matrix sum as $C = [c_{ij}]_{m\times n} \in M_{m\times n}(\mathbb{R})$ such that

$$c_{ij} = a_{ij} + b_{ij}.$$

2.1.2 Definition

Let $A = [a_{ij}]_{m\times n} \in M_{m\times n}(\mathbb{R})$ and $\lambda \in \mathbb{R}$. The scalar multiple of A by λ is

$$\lambda A = [\lambda a_{ij}]_{m\times n}.$$

2.1.3 Definition

Let $A = [a_{ij}]_{m\times n} \in M_{m\times n}(\mathbb{R})$ and $B = [b_{ij}]_{n\times p} \in M_{n\times p}(\mathbb{R})$. Define the matrix product as $C = [c_{ij}]_{m\times p} \in M_{m\times p}(\mathbb{R})$ such that

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

2.1.4 Theorem

$A(\lambda B) = \lambda(AB)\ \forall\, A \in M_{m\times n},\ B \in M_{n\times p}$.

2.1.5 Theorem

$A(B + C) = AB + AC\ \forall\, A \in M_{m\times n},\ B \in M_{n\times p},\ C \in M_{n\times p}$.

2.1.6 Corollary

$A(\alpha v_1 + \beta v_2) = \alpha A v_1 + \beta A v_2\ \forall\, A \in M_{m\times n}(\mathbb{R}),\ v_1, v_2 \in M_{n\times 1}$, so the matrix A is a linear transformation, which can also be thought of as the map

$$\begin{aligned} A : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ v &\longmapsto Av. \end{aligned}$$

2.1.7 Theorem

$(AB)C = A(BC)\ \forall\, A \in M_{m\times n}(\mathbb{R}),\ B \in M_{n\times p}(\mathbb{R}),\ C \in M_{p\times q}(\mathbb{R})$.

2.2 Row Operations

2.2.1 Definition: Elementary Row Operations

Elementary row operations can be performed on matrices and there are three permitted operations:

- Multiply an entire row by $k \in \mathbb{R} \setminus \{0\}$.
- Add a multiple of one entire row to another entire row.
- Swap two rows.

Row operations preserve the system's solutions and they each have an inverse.

2.2.2 Definition

Two systems of equations are equivalent if their augmented matrices are interchangeable via elementary row operations or they are both inconsistent.

2.2.3 Definition

A matrix is in row echelon form $\iff$

i Every zero row is at the bottom.

ii The first non-zero entry of every row is 1.

iii The first non-zero entry in the $(i + 1)^{th}$ row is strictly to the right of that in the i^{th} row.

and is in reduced row echelon form if additionally, the first non-zero entry of every row is the only non-zero entry in its column.

2.2.4 Definition

An elementary matrix is a matrix which can be obtained by performing one elementary row operation on the identity matrix (see 2.2.8).

2.2.5 Theorem

Let $A \in M_{m\times n}$ and E be an elementary $m\times m$ matrix. Left multiplying by E represents the same row operation with which E can be obtained from I_m.

2.2.6 Definition

A matrix is square if it is a member of $M_{n\times n}$ for some $n \in \mathbb{N}$.

2.2.7 Definition

A square matrix is
upper triangular if $a_{ij} = 0$ for $i > j$;
lower triangular if $a_{ij} = 0$ for $i < j$;
diagonal if $a_{ij} = 0$ for $i \neq j$.

2.2.8 Definition

The identity matrix I_n is an $n \times n$ matrix with every diagonal entry equal to 1 and every other entry equal to 0. Note that $AI_n = I_nA = A$.

2.3 Inverses and Transposes of Matrices

2.3.1 Definition

If for $A \in M_{n\times n}$, $\exists\, B \in M_{n\times n}$ such that $AB = BA = I_n$ then A is invertible and B is the inverse of A written A^{-1}. A^{-1} is unique.

2.3.2 Definition

Non-invertible matrices are singular.

2.3.3 Definition

The transpose of $A = [a_{ij}]_{m\times n} \in M_{m\times n}$ is defined $A^T = [a_{ji}]_{n\times m}$.

2.3.4 Theorem

For $A \in M_{m\times n}$ and $B \in M_{n\times p}$, $(AB)^{-1} = B^{-1}A^{-1}$ (clear with some thought) and $(AB)^T = B^TA^T$ (provable with sum expansions).

2.3.5 Theorem

For any elementary matrix, the inverse exists and is also elementary.

2.3.6 Theorem

If $E_kE_{k-1}...E_2E_1A = I_n$ for $A \in M_{n\times n}$ and elementary matrices $E_i \in M_{n\times n}\ \forall\, i$, then $\exists\, A^{-1} = E_kE_{k-1}...E_2E_1$.

2.3.7 Theorem

For $A \in M_{n\times n}$, the following are equivalent.

$$\exists\, A^{-1} = A^T \iff A^TA = AA^T = I_n \iff (Ax)^T(Ay) = x^Ty$$

(i.e A preserves the inner product - see definition 8.3.1).

2.3.8 Definition

$A \in M_{n\times n}$ is orthogonal $\iff A^{-1} = A^T$.

2.3.9 Definition

$A \in M_{n\times n}$ is symmetric $\iff A = A^T$.

2.3.10 Note

The set of $m \times n$ matrices over the field $\mathbb{F}$ is sometimes written $\mathbb{F}^{m\times n}$, but we will continue to use $M_{m\times n}(\mathbb{F})$ since, as in some of the results above, we may wish to keep the field of a matrix general and unspecified (i.e. simply $\in M_{m\times n}$).

3 Vector spaces

3.0.1 Definition

Recall the axioms of a field. A vector space over a field $\mathbb{F}$ (or an $\mathbb{F}$-vector space) is a set $V \neq \emptyset$ which, for the two maps below, satisfies the axioms that follow.

$$\begin{aligned} 1 \qquad \oplus : V \times V &\longrightarrow V \\ (v_1, v_2) &\longmapsto v_1 \oplus v_2 \\ &\text{vector addition} \end{aligned}$$

$$\begin{aligned} 2 \qquad \odot : \mathbb{F} \times V &\longrightarrow V \\ (r, v_1) &\longmapsto r \odot v_1 \\ &\text{scalar multiplication} \end{aligned}$$

where $v_1, v_2 \in V$ and $r \in \mathbb{F}$.

$\forall\, u, v, w \in V,\ r, s \in \mathbb{F}$...

V1 $u \oplus v = v \oplus u$ (commutativity of vector addition).

V2 $(u \oplus v) \oplus w = u \oplus (v \oplus w)$ (associativity of vector addition).

V3 $\exists\, 0_v \in V$ (zero vector) such that $u \oplus 0_v = u$ (additive identity of V).

V4 $\exists -u \in V$ such that $u \oplus -u = 0_v$ (additive inverse in V).

V5 $(rs) \odot u = r \odot (s \odot u)$ (associativity of scalar multiplication).

V6 $1 \odot u = u$ (scalar multiplicative identity).

V7 $r \odot (u \oplus v) = r \odot u \oplus r \odot v$ (distributive law 1).

V8 $(r + s) \odot u = r \odot u \oplus s \odot u$ (distributive law 2).

3.0.2 Definition

Let V be an $\mathbb{F}$-vector space. $W \subseteq V$ is a subspace of V if $W \neq \emptyset$ and, $\forall\, u, v, \in W,\ r \in \mathbb{F}$...

S1 $u \oplus v \in W$ (W is closed under vector addition).

S2 $r \odot u \in W$ (W is closed under scalar multiplication).

We will denote this $W \leqslant V$.

3.0.3 Theorem

Let V be an $\mathbb{F}$-vector space. $0_v \in W \quad \forall\, W \leqslant V$.

3.0.4 Theorem

Let V be an $\mathbb{F}$-vector space. $U \leqslant V \wedge W \leqslant V \implies U \cap W \leqslant V$. Note that this does <u>not</u> imply $U \cup W \leqslant V$.

3.0.5 Theorem

Let V be an $\mathbb{F}$-vector space. Any subspace of V is an $\mathbb{F}$-vector space.

4 Span, Linear Independence and Bases

4.1 Spanning Sets

4.1.1 Definition

Let V be an $\mathbb{F}$-vector space and let $v_1, ..., v_n \in V$. $\alpha_1 v_1 + ... + \alpha_n v_n$ is a linear combination of $v_1, ..., v_n$ where $\alpha_1, ..., \alpha_n \in \mathbb{F}$.

4.1.2 Definition

The span of $\{v_1, ..., v_n\}$ is the set of linear combinations of $v_1, ..., v_n$: $\text{Span}(\{v_1, ..., v_n\}) = \{\alpha_1 v_1 + ... + \alpha_n v_n \mid \alpha_1, ..., \alpha_n \in \mathbb{F}\}$, sometimes written $\langle v_1, ..., v_n \rangle$.

4.1.3 Note

Conventionally, $\text{Span}(\emptyset) := \{0_v\}$.

4.1.4 Lemma

Let V be an $\mathbb{F}$-vector space and let $v_1, ..., v_n \in V$. $\text{Span}(\{v_1, ..., v_n\}) \leqslant V$.

4.1.5 Definition

Let V be an $\mathbb{F}$-vector space. If $S \subseteq V$ with $\text{Span}(S) = V$, then S is a spanning set for V (S spans V).

4.2 Linear Independence

4.2.1 Definition

Let V be an $\mathbb{F}$-vector space. $v_1, ..., v_n \in V$ are linearly independent if $\alpha_1 v_1 + ... + \alpha_n v_n = 0_v \implies \alpha_1 = \alpha_2 = ... = \alpha_n = 0$. Linear dependence is the negation of this statement.

4.2.2 Note

Let $\{v_1, ..., v_n\} = S$ Suppose $\alpha_i \neq 0$ for one or more $i \in [1, ..., n-1]$ and that $\alpha_1 v_1 + ... + \alpha_n v_n = 0_v$. Then the vectors $v_1, ..., v_n$ are linearly dependent by definition and we can write $\frac{-1}{\alpha_n}(\alpha_1 v_1 + ... + \alpha_{n-1} v_{n-1}) = v_n$ $(\alpha_n \neq 0)$ which is equivalent to saying that one of the vectors $v \in S$ is a linear combination of vectors in $S \setminus \{v\}$. This is a common way to show that S is linearly dependent.

4.2.3 Corollary

Let $v_1, ..., v_n$ be linearly independent vectors in an $\mathbb{F}$-vector space V. If $v_{n+1} \notin \text{Span}(\{v_1, ..., v_n\})$, then $v_1, ..., v_{n+1}$ are linearly independent.

4.3 Bases

4.3.1 Definition

Let V be an $\mathbb{F}$-vector space. A set that spans V and is linearly independent is a basis of V. V is finite-dimensional $\iff$ it has a finite basis.

4.3.2 Corollary

For a field $\mathbb{F}$, let e_i be the column vector of zeroes with a 1 in the i^{th} row. $e_1, ..., e_n$ is a basis of $\mathbb{F}^n$.

4.3.3 Theorem

Let V be an $\mathbb{F}$-vector space. $S = \{v_1, ..., v_n\}$ is a basis of $V \iff$ each vector $v \in V$ is expressed uniquely as a linear combination of vectors in S.
Note: the span gives rise to the existence, and the linear independence gives rise to the uniqueness.

4.3.4 Definition

Let V be an n-dimensional $\mathbb{F}$-vector space and $B = \{v_1, ..., v_n\}$ be a basis. For a vector $v = \alpha_1 v_1 + ... + \alpha_n v_n$, $\alpha_1, ..., \alpha_n$ are called the coordinates of v (with respect to B).
Note that by Theorem 4.3.3, the map

$$V \longrightarrow \mathbb{F}^n$$
$$v \longmapsto \alpha_1, ..., \alpha_n$$

is a bijection.

4.4 Dimension

4.4.1 Lemma: Steinitz Exchange Lemma

Let V be an $\mathbb{F}$-vector space. Let $X \leqslant V$ and let u be such that $u \in \text{Span}(X)$ but $u \notin \text{Span}(X \setminus \{v\})$ (where $v \in X$). Let $Y = (X \setminus \{v\}) \cup \{u\}$. Span($X$)=Span($Y$).

4.4.2 Theorem

Let V be a vector space. Let $S, T \subseteq V$ (both finite) where S spans V and T is linearly independent. $|S| \geq |T|$.

4.4.3 Corollary

Let V be a finite-dimensional vector space. Let S, T be bases of V. Then S and T are both finite and $|S| = |T|$.

4.4.4 Definition

Let V be a finite-dimensional $\mathbb{F}$-vector space. The dimension of V over $\mathbb{F}$, $\dim_{\mathbb{F}}(V)$, is the cardinality of any basis of V.

4.4.5 Lemmas

For a vector space V, suppose $\dim(V) = n$.

1. Any spanning set of size n is a basis of V.
2. Any linearly independent set of size n is a basis of V.
3. S is a spanning set $\iff$ it contains a basis of V.
4. T is linearly independent $\iff$ it is contained in a basis of V.
5. Any subset of size $< n$ is not spanning.
6. Any subset of size $> n$ is linearly dependent.

4.4.6 Definition

For two spaces U and W, $U + W = \{u + w \mid u \in U,\ w \in W\}$.

4.4.7 Theorem

$U \subseteq U + W$ and $W \subseteq U + W$.

4.4.8 Theorem

Let V be an $\mathbb{F}$-vector space.
$U \leqslant V \wedge W \leqslant V \implies U \cap W \leqslant V$ (Theorem 3.0.4).
$U \leqslant V \wedge W \leqslant V \implies U + W \leqslant V$.

4.4.9 Theorem

Let $U = \text{Span}(\{u_1, ..., u_r\})$, $W = \text{Span}(\{w_1, ..., w_s\})$, then $U + W = \text{Span}(\{u_1, ..., u_r, w_1, ..., w_s\})$.

4.4.10 Theorem

Let V be an $\mathbb{F}$-vector space and let $U, W \leqslant V$.

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

5 Matrix Rank

5.0.1 Definition

Let $A \in M_{m \times n}$.

- The row space of $A = \text{RSp}(A) = \text{Span}(\{\text{Rows of } A\})$.
- The column space of $A = \text{CSp}(A) = \text{Span}(\{\text{Columns of } A\})$.

5.0.2 Definition

- The row rank of $A = \dim(\mathrm{RSp}(A))$.
- The column rank of $A = \dim(\mathrm{CSp}(A))$.

5.0.3 Note

Reducing to row or column echelon form leaves a basis of the row or column space, the cardinality of which gives the row or column rank.

5.0.4 Theorem

Let $A \in M_{m\times n}$. The row rank and column rank of A are equal.

5.0.5 Definition

We can now redefine simply the rank of $A \in M_{m\times n}$, $\mathrm{rank}(A)$, as the row or column rank of A.

5.0.6 Theroem

For $A \in M_{n\times n}$, the following are equivalent.

$$\begin{aligned} & \mathrm{rank}(A) = n \\ \iff & \text{The rows of } A \text{ form a basis} \iff \text{The columns of } A \text{ form a basis} \\ \iff & A \text{ is invertible.} \end{aligned}$$

6 Linear Transformations

6.0.1 Definition

For two $\mathbb{F}$-vector spaces V and W, let T be a function

$$T : V \longrightarrow W.$$

T is a linear transformation $\iff$

T1 $\forall\, v_1, v_2 \in V,\ T(v_1 + v_2) = T(v_1) + T(v_2)$ (T preserves addition).

T2 $\forall\, v_1 \in V,\ r \in \mathbb{F},\ T(rv_1) = rT(v_1)$ (T preserves scalar multiplication).

6.0.2 Theorem

Let $A \in M_{m\times n}(\mathbb{F})$, and define $T : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ such that $T(v) = Av\ \forall\, v \in \mathbb{F}^n$. T is a linear transformation.

6.0.3 Theorem

$\forall$ linear transformations $T : \mathbb{F}^n \longrightarrow \mathbb{F}^m\ \exists\, A \in M_{m\times n}(\mathbb{F})$ such that $T(v) = Av\ \forall\, v \in \mathbb{F}^n$.

6.0.4 Theorem

Let V and W be $\mathbb{F}$-vector spaces and let $T : V \longrightarrow W$ be linear.

1. $T(0_v) = 0_w$.
2. Let $v = \lambda_1 v_1 + ... + \lambda_n v_n$. $T(v) = \lambda_1 T(v_1) + ... + \lambda_n T(v_n)$.

6.0.5 Theorem

Let V and W be $\mathbb{F}$-vector spaces. Let $v_1, ..., v_n$ be a basis for V and let $w_1, ..., w_n \in W$. $\exists$ a unique linear $T : V \longrightarrow W$ such that $T(v_i) = w_i\ \forall\, i$.

6.1 Image and Kernel

6.1.1 Definition

Let V and W be $\mathbb{F}$-vector spaces and let $T : V \longrightarrow W$ be a linear transformation.
The image of T is $\operatorname{Im} T = \{T(v) \mid v \in V\} \subseteq W$.
The kernel of T is $\operatorname{Ker} T = \{v \in V \mid T(v) = 0_w\} \subseteq V$.

6.1.2 Theorem

Let $T : V \longrightarrow W$ be a linear transformation.

$$\operatorname{Im} T \leqslant W \text{ and } \operatorname{Ker} T \leqslant V.$$

6.1.3 Lemma

Let $T : V \longrightarrow W$ be a linear transformation, let $v_1, v_2 \in V$.

$$T(v_1) = T(v_2) \iff v_1 - v_2 \in \operatorname{Ker} T.$$

6.1.4 Lemma

Let $T : V \longrightarrow W$ be a linear transformation, let $\{v_1, ..., v_n\}$ be a basis for V.

$$\operatorname{Im} T = \operatorname{Span}(\{T(v_1), ..., T(v_n)\}).$$

6.1.5 Lemma

Let $A \in M_{m\times n}(\mathbb{F})$ and let $T : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ be such that $T(v) = Av \; \forall \, v \in \mathbb{F}^n$.

1. $\operatorname{Ker} T$ is the set of solutions to $Av = 0$.

2a. $\operatorname{Im} T = \operatorname{CSp}(A)$.

2b. $\dim(\operatorname{Im} T) = \operatorname{rank}(A)$.

6.1.6 Theorem: Rank-Nullity Theorem

Let $T : V \longrightarrow W$ be a linear transformation.

$$\dim(V) = \dim(\operatorname{Im} T) + \dim(\operatorname{Ker} T).$$

6.1.7 Corollary

Let $A \in M_{m\times n}(\mathbb{F})$ written also as the linear transformation $A : \mathbb{F}^n \longrightarrow \mathbb{F}^m$. The dimension of the set of solutions x to the homogeneous system

$$A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = 0_{\mathbb{F}^m}$$

is given by

$$\dim(\text{Set of solutions}) = \dim(\operatorname{Ker} A)) \overset{6.1.6}{=} \dim(\mathbb{F}^n) - \dim(\operatorname{Im} A) \overset{6.1.5.2b}{=} n - \operatorname{rank}(A).$$

6.2 Representing Vectors and Transformations With Respect to a Basis

6.2.1 Definition

Let V be an n-dimensional $\mathbb{F}$-vector space. Let $B = \{v_1, ..., v_n\}$ be a basis for V. Let $v = \lambda_1 v_1 + ... + \lambda_n v_n \in V$.
The vector of v with respect to B is

$$[v]_B = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

6.2.2 Theorem

Let V be an n-dimensional $\mathbb{F}$-vector space and $B = \{v_1, ..., v_n\}$ be a basis. The map

$$\begin{aligned} T : V &\longrightarrow \mathbb{F}^n \\ v &\longmapsto [v]_B \end{aligned}$$

is bijective (by 4.3.3) and linear.

6.2.3 Definition

Let V, W be finite-dimensional $\mathbb{F}$-vector spaces with respective bases $B = \{b_1, ..., b_n\}$ and $C = \{c_1, ..., c_m\}$. Let $T : V \longrightarrow W$ be linear. $\exists$ bijections (by 6.2.2) mapping $v \in V,\ w \in W$ to $[v]_B$ and $[w]_C$, so we have a construction as shown:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, leftrightarrow] & W \arrow[d, leftrightarrow] \\ \mathbb{F}^n & \mathbb{F}^m \end{tikzcd}$$

Therefore $\exists$ a map $\mathbb{F}^n \longrightarrow \mathbb{F}^m$ which is linear (by linearity of composites), and hence $\exists$ a matrix A representing it (by 6.0.3) such that $A[v]_B = [T(v)]_C$, the columns of which are the vectors with respect to C of $T(b_i \in B)$. A is called the matrix of T with respect to B and C, written ${}_C[T]_B$, so note that we can now write ${}_C[T]_B[v]_B = [T(v)]_C$.

6.2.4 Theorem

Let V be an $\mathbb{F}$-vector space and let $B = \{b_1, ..., b_n\}$ and $C = \{c_1, ..., c_n\}$ be bases. For $j \in \{1, ..., n\}$, $b_j = \lambda_{1j}c_1 + ... + \lambda_{nj}c_n$. Let

$$P = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & \ddots & \vdots \\ \lambda_{n1} & \cdots & \lambda_{nn} \end{pmatrix}$$

(The matrix with $[b_j]_C$ as the j^{th} column).

1. $P = {}_C[X]_C$ where $X : V \longrightarrow V$ is the unique linear transformation such that $X(c_j) = b_j \ \forall\, j$.
2. $\forall\, v \in V$, $P[v]_B = [v]_C$.
3. $P = {}_C[Id]_B$ where $Id : V \longrightarrow V$ is the identity transformation.

6.2.5 Definition

P, defined as above, is the change of basis matrix from B to C.

6.2.6 Theorem

Let V be an $\mathbb{F}$-vector space, let $B = \{b_1, ..., b_n\}$ and $C = \{c_1, ..., c_n\}$ be bases and let P be the change of basis matrix from B to C.

1. P is invertible and its inverse is the change of basis matrix from C to B (i.e. ${}_C[Id]_B^{-1} = {}_B[Id]_C$).
2. For a linear transformation $T : V \longrightarrow V$, ${}_C[T]_C = P_B[T]_BP^{-1}$ $(= {}_C[Id]_{BB}[T]_{BB}[Id]_C)$.

6.2.7 Note

Change of basis is transitive, i.e., letting D be a third basis, ${}_D[Id]_{CC}[Id]_B = {}_D[Id]_B$, so when finding the change of basis matrix from F to G, it is often easiest to find that from F to the standard basis E and find that from G to E (as these methods involve simply expressing elements of F and G as vectors with respect to the standard basis), invert the latter (see 7.1) and multiply ${}_G[Id]_{EE}[Id]_F = {}_G[Id]_F$.

7 Determinants

7.0.1 Definition

Let $A \in M_{n\times n}(\mathbb{F})$ $(n > 1)$. The ij-minor of A is $A_{ij} \in M_{(n-1)\times(n-1)}(\mathbb{F})$ formed by deleting row i and column j from A.

7.0.2 Definition

Let $A = [a_{ij}]_{n\times n} \in M_{n\times n}(\mathbb{F})$. The determinant of A is

$$\det(A) = \begin{cases} a_{11}, & n = 1 \\ \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det(A_{1j}), & n > 1. \end{cases}$$

$\det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ is often written $\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$

7.0.3 Theorem

Let $A \in M_{n\times n}(\mathbb{F})$. Let B be the matrix obtained by multiplying row l of A by α. $\det(B) = \alpha \det(A)$.

7.0.4 Theorem

Let $A, B, C \in M_{n\times n}(\mathbb{F})$ be identical other than that row l of C is equal to the sum of row l of A and B. $\det(C) = \det(A) + \det(B)$.

7.0.5 Corollary

The transformation

$$\begin{aligned} M_{n\times n}(\mathbb{F}) &\longrightarrow \mathbb{F} \\ A &\longmapsto \det(A) \end{aligned}$$

is linear on rows of A.

7.0.6 Theorem

Let $A \in M_{n\times n}(\mathbb{F})$ $(n \geq 2)$. Let rows l and $l+1$ of A be identical. $\det(A) = 0$.

7.0.7 Theorem

$\det(I_n) = 1 \; \forall n$.

7.0.8 Theorem: Effects of the Elementary Operations on Determinant

Let $A, B \in M_{n\times n}(\mathbb{F})$.

1. Suppose multiplying a row of A by α gives B. $\det(B) = \alpha \det(A)$ (theorem 7.0.3).
2. Suppose adding a multiple of one row of A to another gives B. $\det(B) = \det(A)$.
3. Suppose swapping two rows of A gives B. $\det(B) = -\det(A)$.

7.0.9 Corollary

Further to 7.0.6, if any two rows of $A \in M_{n\times n}(\mathbb{F})$ are identical, $\det(A) = 0$.

7.0.10 Definition

If $B \in M_{n\times n}(\mathbb{F})$ can be obtained from $A \in M_{n\times n}(\mathbb{F})$ by a sequence of elementary row operations, A and B are row-equivalent.

7.0.11 Corollary

$A \in M_{n\times n}(\mathbb{F})$ and $B \in M_{n\times n}(\mathbb{F})$ are row-equivalent $\implies \exists\, \lambda \in \mathbb{F},\ \lambda \neq 0$ such that $\det(B) = \lambda \det(A)$.

7.0.12 Definition

$A \in M_{n\times n}(\mathbb{F})$ is singular if $\exists\, v \in \mathbb{F}^n$ $(v \neq 0_{\mathbb{F}^n})$ such that $Av = 0$ (0 is an eigenvalue of A - see 8.0.2), and A is non-singular otherwise.

7.0.13 Theorem

Let $A \in M_{n\times n}(\mathbb{F})$.

$$\begin{aligned} & A \text{ is invertible} \\ \iff & A \text{ is non-singular} \\ \iff & A \text{ has linearly independent rows} \\ \iff & A \text{ is row-equivalent to } I_n \\ \iff & \det(A) \neq 0. \end{aligned}$$

7.0.14 Theorem

Let $A \in M_{n\times n}(\mathbb{F})$ $(n > 1)$. The determinant of A is, in fact, given for any i by

$$\textstyle\sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

(definition 7.0.2 is extended to any row, even numbered rows beginning with a negative term).

7.0.15 Corollary

Let $A \in M_{n\times n}(\mathbb{F})$ be lower or upper triangular. $\det(A) = a_{11}a_{22}...a_{nn} = \prod_{i=1}^{n} a_{ii}$. Note: Gaussian elimination followed by application of this corollary is a quick way to find the determinant.

7.0.16 Lemma

Let $A \in M_{n\times n}(\mathbb{F})$ be non-singular. By row-equivalence to I_n (7.0.13), $\exists$ elementary matrices $E_1, ..., E_n$ such that $A = E_1...E_n$.

7.0.17 Lemma

Let $A, E \in M_{n\times n}(\mathbb{F})$ and let E be elementary. $\det(EA) = \det(E)\det(A)$.

7.0.18 Lemma

AB is singular $\iff$ A is singular $\vee$ B is singular.

7.0.19 Theorem: Product of Determinants

Let $A, B \in M_{n\times n}(\mathbb{F})$. $\det(AB) = \det(A)\det(B)$.

7.0.20 Corollary

Let $A \in M_{n\times n}(\mathbb{F})$ be non-singular. $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

7.0.21 Theorem

Let $A \in M_{n\times n}(\mathbb{F})$. $\det(A^T) = \det(A)$.

7.0.22 Corollary

Elementary column operations have the same effects on determinant as elementary row operations.

7.0.23 Theorem

7.0.21 allows the extension of the definition to expansion about any row <u>or</u> column:

$$\det(A) = \underset{\forall\, i}{\textstyle\sum_{j=1}^{n}(-1)^{i+j}a_{ij}}\det(A_{ij}) = \underset{\forall\, j}{\textstyle\sum_{i=1}^{n}(-1)^{i+j}a_{ij}}\det(A_{ij}).$$

7.0.24 Theorem: Vandermonde Determinant

Let $n \geq 2$ and let $\alpha_1, ..., \alpha_n \in \mathbb{F}$ (the field of the matrix).

$$\begin{vmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \cdots & \alpha_n^{n-1} \end{vmatrix} = \prod_{1\leq i<j\leq n} (\alpha_j - \alpha_i).$$

7.1 Inverting a Matrix

7.1.1 Definition

Let $A \in M_{n\times n}(\mathbb{F})$. The ij-cofactor of A is $(-1)^{i+j}\det(A_{ij})$.

7.1.2 Definition

The adjugate of A is $\mathrm{adj}(A) = C^T$ where $C = [c_{ij}]$, the matrix of ij-cofactors of A.

7.1.3 Theorem

$\mathrm{adj}(A)A = \det(A)I_n$.

7.1.4 Corollary: Formula of the Inverse

$A^{-1} = \frac{1}{\det(A)}\mathrm{adj}(A)$ (which is of course undefined where $\det(A) = 0$).

7.1.5 Corollary

$A \in M_{n\times n}(\mathbb{Z})$ and $\det(A) = \pm 1 \implies A^{-1} \in M_{n\times n}(\mathbb{Z})$.

7.2 The Determinant of a Linear Transformation

7.2.1 Definition

Let V be finite-dimensional $\mathbb{F}$-vector space, let B be a basis of V and let $T : V \longrightarrow V$ be a linear transformation. $\det(T) := \det({}_B[T]_B)$.

7.2.2 Theorem

$\det(T)$, defined as in 7.2.1, does not depend on the choice of B.

8 Eigenvalues and Eigenvectors

8.0.1 Definition

Let V be a vector space over $\mathbb{F}$ and let $T : V \longrightarrow V$ be a linear transformation. $v \in V$ ($v \neq 0_V$) is an eigenvector of $T \iff T(v) = \lambda v$ for some $\lambda \in \mathbb{F}$. λ is the eigenvalue of v.

8.0.2 Definition

Let V be a vector space over $\mathbb{F}$ and let $T : V \longrightarrow V$ be a linear transformation. $\lambda \in \mathbb{F}$ is an eigenvalue of $T \iff T(v) = \lambda v$ for some $v \in V$ ($v \neq 0_V$). v is the eigenvector of λ.

8.0.3 Theorem

$T : \mathbb{F}^n \longrightarrow \mathbb{F}^n$ has n distinct eigenvalues $\in \mathbb{F} \implies \exists$ a basis for $\mathbb{F}^n$ of eigenvectors of T.

8.0.4 Lemmas

Let $T : \mathbb{F}^n \longrightarrow \mathbb{F}^n$ and suppose $\exists$ a basis B (in order) for $\mathbb{F}^n$ of eigenvectors of T.

1. ${}_B[T]_B$ is the diagonal matrix with entries equal to the eigenvalues of vectors in B.
2. ${}_E[Id]_B$ is the matrix with columns equal to the vectors in B (where E is the standard basis in $\mathbb{F}^n$).

8.0.5 Theorem: Spectral Decomposition

Let T and B be defined as above. Note that the matrix representing T ($A \in M_{n\times n}(\mathbb{F})$) can be written as ${}_E[T]_E$ where E is the standard basis in $\mathbb{F}^n$. By 6.2.6.2, ${}_B[T]_B$ is equal to ${}_B[Id]_{EE}[T]_{EE}[Id]_B$ and so ${}_E[T]_E = A = {}_E[Id]_{BB}[T]_{BB}[Id]_E$.

8.0.6 Corollary

Let $A, D, P \in M_{n\times n}(\mathbb{F})$ be the matrices in the previous theorem such that $A = PDP^{-1}$ (with the eigenvalues in D in the correct order corresponding to the eigenvectors in P).

$$A^k = \underbrace{PDP^{-1}PDP^{-1}...PDP^{-1}}_{k} = PD^kP^{-1} \quad (k \in \mathbb{N}).$$

8.0.7 Corollary

$A = PDP^{-1} \implies A^{-1} = (PDP^{-1})^{-1} = PD^{-1}P^{-1}$.

8.0.8 Corollary: Determinant From Eigenvalues

$A = PDP^{-1} \implies \det(A) = \det(P)\det(D)\det(P^{-1}) = \det(D) = \prod$ eigenvalues of A.

8.0.9 Theorem

Let V be a finite-dimensional $\mathbb{F}$-vector space, let B be a basis of V and let $T : V \longrightarrow V$.

1. λ is an eigenvalue of $T \iff \lambda$ is an eigenvalue of ${}_B[T]_B$.
2. v is an eigenvector of $T \iff [v]_B$ is an eigenvector of ${}_B[T]_B$.

8.0.10 Note

Every transformation that has eigenvectors has infinitely many, since $T(v) = \lambda v \implies T(\mu v) = \lambda\mu v \;\forall\, \mu$. 'The eigenvectors of T' refers to any set of linearly independent eigenvectors of T. The set of eigenvectors for a particular λ (an eigenspace) is often written $E_\lambda = \{v \in V \mid T(v) = \lambda v\} \leqslant V$ (or $\{v \in \mathbb{F}^n \mid Av = \lambda v\}$ for $A \in M_{n\times n}(\mathbb{F})$). Under this notation, λ is not eigenvalue $\iff E_\lambda = \{0\}$.

8.1 The Characteristic Polynomial

8.1.1 Definition

The characteristic polynomial of $A \in M_{n\times n}(\mathbb{F})$ is $\chi_A(x) = \det(xI_n - A)$, where x is a variable taking values in $\mathbb{F}$.

8.1.2 Definition

The characteristic polynomial of a linear transformation $T : V \longrightarrow V$ is $\chi_T(x) = \chi_{({}_B[T]_B)}(x)$ where B is a basis for V. Note that definition 8.1.1 is essentially the case where the basis B is standard.

8.1.3 Theorem

The characteristic polynomial of a transformation T, defined as in 8.1.2, does not depend on the choice of B (and so may be found using a representative A).

8.1.4 Theorem

1. Let $T : V \longrightarrow V$ be linear. $\lambda \in \mathbb{F}$ is an eigenvalue of $T \iff \chi_T(\lambda) = 0$.
2. Let $A \in M_{n\times n}(\mathbb{F})$. $\lambda \in \mathbb{F}$ is an eigenvalue of $A \iff \chi_A(\lambda) = 0$.

8.1.5 Corollary

$A \in M_{n\times n}(\mathbb{F})$ has at most n distinct eigenvalues.

8.2 Diagonalisation

8.2.1 Definition

Let V be a vector space over $\mathbb{F}$ and let $T : V \longrightarrow V$ be a linear transformation. T is diagonalisable $\iff \exists$ a basis of V consisting of eigenvectors of T.

8.2.2 Theorem

Let V be a vector space over $\mathbb{F}$ and let $T : V \longrightarrow V$ be a linear transformation. T is diagonalisable $\iff \exists$ a basis $B = v_1, ..., v_n$ of V such that ${}_B[T]_B$ is diagonal. Or in terms of matrices, $A \in M_{n\times n}(\mathbb{F})$ is diagonalisable over $\mathbb{F} \iff \exists\, P \in M_{n\times n}(\mathbb{F})$ such that $P^{-1}AP$ is diagonal. Both of these results follow from the ideas required to prove 8.0.4-6.

8.2.3 Note: Diagonalisable Matrices

A matrix may be diagonalisable over one field but not another. Suppose $A \in M_{n\times n}(\mathbb{R})$ has $< n$ linearly independent eigenvectors $\in \mathbb{R}^n \implies A$ is not diagonalisable over $\mathbb{R}$, but if A has n linearly independent eigenvectors $\in \mathbb{C}^n$, then it $\underline{\text{is}}$ diagonalisable over $\mathbb{C}$.

8.2.4 Theorem

Let V be a vector space over $\mathbb{F}$ and let $T : V \longrightarrow V$ be a linear transformation. Suppose $v_1, ..., v_n$ are any eigenvectors of T where $T(v_i) = \lambda_i v_i$, and that the $\{\lambda_i\}$ are distinct, then the $\{v_i\}$ are linearly independent. 8.0.3 is a corollary of this theorem and implies also that T is diagonalisable by definition.

8.2.5 Note

The conditions in the theorem above (and in 8.0.3) are sufficient but not necessary for diagonalisability. For example E_{λ^*} may have a basis of multiple linearly independent eigenvectors (i.e. $\dim(E_{\lambda^*}) > 1$).

8.2.6 Theorem

Let r be the number of distinct eigenvalues of $T : V \longrightarrow V$. $\sum_{i=1}^{r} \dim(E_{\lambda_i}) = \dim(V)$ is a necessary and sufficient condition for diagonalisability (and in fact $\bigcup_{i=1}^{r} B_i$ is a basis for V where B_i is a basis for E_{λ_i}). This theorem is equivalent to saying that the algebraic and geometric multiplicities of each eigenvalue are the same (see next definition). 8.2.4 is the special case where $r = \dim(V)$ (and the algebraic and geometric multiplicities of each eigenvalue are all 1).

8.2.7 Definition

The algebraic multiplicity of λ is the number of times it appears as a root of $\chi_T(x) = 0$. The geometric multiplicity of λ is $\dim(E_\lambda)$.

8.2.8 Theorem

$A \in M_{n\times n}(\mathbb{R})$ is symmetric $\implies A$ is diagonalisable, and $\exists$ a diagonalising matrix (P such that $P^{-1}AP$ is diagonal) which is orthogonal ($P^T = P^{-1}$ - definition 2.3.8). The proof relies on results in the following sections - see 8.4.5.

8.3 Orthogonality in $\mathbb{R}^n$ and the Gram-Schmidt Process

8.3.1 Definition

The inner product of $u, v \in \mathbb{R}^n$ is $u \cdot v := u^T v = \sum_{i=1}^{n} u_i v_i$.

8.3.2 Theorem

1. $(\alpha u + \beta v) \cdot w = \alpha u \cdot w + \beta v \cdot w$ (linearity in the first argument).

2. $u \cdot (\alpha v + \beta w) = \alpha u \cdot v + \beta u \cdot w$ (linearity in the second argument).

8.3.3 Definition

$u, v \in \mathbb{R}^n$ are orthogonal $\iff u \cdot v = 0$.

8.3.4 Definition

The norm of $u \in \mathbb{R}^n$ is $||u|| := \sqrt{u \cdot u} = \sqrt{\sum_{i=1}^n u_i^2}$. $||u - v||$ is the 'distance' of u from v.

8.3.5 Definition

$u \in \mathbb{R}^n$ is normal $\iff ||u|| = 1$.

8.3.6 Theorem

2. $||u|| = 0 \iff u = \mathbf{0}$.
3. $||\alpha u|| = |\alpha| \, ||u||$.
3. $\left|\left|\frac{u}{||u||}\right|\right| = 1$.

8.3.7 Theorem: Cauchy-Schwarz Inequality

Let $u, v \in \mathbb{R}^n$. $||u|| \, ||v|| \geq |u \cdot v|$.

8.3.8 Corollary

1. Triangle inequality: $||u + v|| \leq ||u|| + ||v||$.
2. Metric triangle inequality: $||u - v|| \leq ||u - w|| + ||v - w||$.

8.3.9 Definition

$\{u_1, ..., u_n\}$ is an orthonormal set $\iff ||u_i|| = 1 \; \forall \, i$ and $u_i \cdot u_j = 0 \; \forall \, i \neq j$.

8.3.10 Theorem

An orthonormal set of vectors is linearly independent.

8.3.11 Lemma

A matrix is orthogonal $\iff$ its columns form an orthonormal set.

8.3.12 Theorem: The Gram-Schmidt Process

Let $v_1, ..., v_r$ be linearly independent vectors $\in \mathbb{R}^n$. $\exists$ an orthonormal set $\{u_1, ..., u_r\} \subseteq \mathbb{R}^n$ such that $\forall \, i \leq r$, $\text{Span}(\{u_1, ..., u_i\}) = \text{Span}(\{v_1, ..., v_i\})$. The 'process' refers to the inductive construction involved in the proof, which can of course be used in application to find such an orthonormal set.

8.3.13 Corollary

If $v \in \mathbb{R}^n$ is normal, $\exists$ an orthogonal matrix with v as its first column.

8.3.14 Corollary

Any subspace of $\mathbb{R}^n$ has an orthonormal basis.

8.4 Real Symmetric Matrices

8.4.1 Theorem: The Fundamental Theorem of Algebra

Any non-constant polynomial with coefficients in $\mathbb{C}$ has a root in $\mathbb{C}$. The proof of this theorem is not examinable in this course.

8.4.2 Lemmas

(1. A is orthogonal $\iff (Au)\cdot(Av) = u\cdot v \;\forall\, u, v$ (theorem 2.3.7).)

2. A is symmetric $\iff (Au)\cdot v = u\cdot(Av) \;\forall\, u, v$.

8.4.3 Lemma

$A \in M_{n\times n}(\mathbb{R})$ is symmetric $\implies$ any root of $\chi_A(x)$ is real.

8.4.4 Corollary

$A \in M_{n\times n}(\mathbb{R})$ is symmetric $\implies$ A has a real eigenvalue.

8.4.5 Note

Theorem 8.2.8 can now be proved using some of the above properties and Gram-Schmidt, and we can see from lemma 8.3.11 and the definition of diagonalisability that it is equivalent to saying: A is symmetric $\implies \exists$ an orthonormal eigenvector basis of $\mathbb{R}^n$, since a matrix with columns equal to such a basis will diagonalise A and be orthogonal.

8.4.6 Lemma

$A \in M_{n\times n}(\mathbb{R})$ is symmetric and u, v are eigenvectors with eigenvalues $\lambda \neq \mu \implies u\cdot v = 0$.

8.4.7 Note

In actually finding an orthogonal matrix which diagonalises a given symmetric matrix $A \in M_{n\times n}(\mathbb{R})$, one can find the eigenspaces of A, find a basis for each, and find an orthonormal basis for each using Gram-Schmidt. Combining all these bases gives a basis for $\mathbb{R}^n$ (by 8.2.6), and the combined basis is also orthonormal by the lemma above.

Part II: Groups

1 Groups and Subgroups

1.1 Binary Operations

1.1.1 Definition

A binary operation $*$ on a set S is a function $S \times S \longrightarrow S^*$. It assigns an element $a * b \in S^*$ to every ordered pair $(a, b) \in S^2$.

1.2 Groups

1.2.1 Definition

G is a group with respect to the binary operation $* \iff \forall\, g, h, i \in G...$

G1 $g * h \in G$ (closure axiom).

G2 $(g * h) * i = g * (h * i)$ (associativity axiom).

G3 $\exists\, e \in G,\ \forall\, g \in G,\ g * e = g = e * g$ (identity axiom).

G4 $\forall\, g \in G,\ \exists\, f \in G,\ g * f = e = f * g$ (existence of inverses).

We may explicitly write the group together with the binary operation: '$(G, *)$ is a group'.

1.2.2 Definition

A group $(G, *)$ is abelian $\iff \forall\, g, h \in G,\ g * h = h * g$ (commutativity of $*$).

1.2.3 Theorem

Let G be a group. The identity in G is unique and the inverse associated to each element in G is unique.

1.2.4 Note

$*$ is used above for the group operation to be completely abstract and avoid confusion with arithmetic operators, however $\cdot$ or $+$ (sometimes for abelian groups) are often used or may be ommited altogether. By 1.2.3, we may write the unique inverse of g as g^{-1} from now on.

1.2.5 Lemma: Equations in Groups

Let G be a group and let $g, h \in G$. $\forall\, x, y \in G...$

1. $gx = h \iff x = g^{-1}h$.
2. $yg = h \iff y = hg^{-1}$.

1.2.6 Lemma: Inverse of a Product

Let G be a group and let $g_1, ..., g_n \in G$. $(g_1 ... g_n)^{-1} = g_n^{-1} ... g_1^{-1}$.

1.3 Symmetric Groups

1.3.1 Definition

Let $X \neq \emptyset$. A permutation of X is a bijection $\sigma : X \longrightarrow X$.

1.3.2 Note

We often write the permutation σ of $X = \{x_1, ..., x_n\}$ in Cauchy two-line form (below) or, if the elements of X have an intuitive and unambiguous order (e.g. $X = \{1, ..., n\}$), we may simply write the lower line in parentheses (not to be confused with disjoint cycle form - section 3.1).

$$\sigma = \begin{pmatrix} x_1 & \cdots & x_n \\ \sigma(x_1) & \cdots & \sigma(x_n) \end{pmatrix}.$$

1.3.3 Definition

The set of all permutations of X is denoted by $\mathrm{Sym}(X)$, and if $X = \{1, ..., n\}$, $\mathrm{Sym}(X)$ may be written $\mathrm{Sym}(n)$ or S_n.

1.3.4 Theorem

$(\mathrm{Sym}(X), \circ)$ is a group, called the symmetric group on X.

1.4 Powers on Group Elements

1.4.1 Definition

For a group G and an element $g \in G$, $g^0 := e$ (the identity in G), $g^{n+1} := g^n g \; \forall\, n \in \mathbb{N}$ and $g^{-n} := (g^{-1})^n \; \forall\, n \in \mathbb{N}$. The notation used for the group operation may affect how we write this definition. If $+$ is used for the group operation, we may write $g + ... + g = ng$ (rather than $g \cdot ... \cdot g = g^n$).

1.4.2 Lemma

1. $g^m g^n = g^{m+n}$ $(\forall\, m, n \in \mathbb{Z})$.
2. $(g^m)^n = g^{mn}$ $(\forall\, m, n \in \mathbb{Z})$.

1.5 Subgroups

1.5.1 Definition

Let $(G, *)$ be a group. A subset $H \subseteq G$ is a subgroup of $G \iff (H, *)$ is also a group (satisfies **G1-4** in 1.2.1). We will denote this $H \leqslant G$ (similarly to the notation for a subspace).

1.5.2 Theorem: Test for a Subgroup

Let $(G, *)$ be a group and let $H \subseteq G$. $H \leqslant G \iff$

i $H \neq \emptyset$.

ii $h_1, h_2 \in H \implies h_1 * h_2 \in H$.

iii $h \in H \implies h^{-1} \in H$.

(The group axioms on H together with $H \subseteq G$ are equivalent to the above conditions).

1.5.3 Corollary

Suppose $H \leqslant G$.

1. $e_H = e_G$ (the identities in G, H are the same; I will assume this notation as standard from here on).
2. Inverses are the same in H as in G.

1.5.4 Definition

Let G be a group. The cyclic subgroup generated by $g \in G$ is $\langle g \rangle := \{g^n \mid n \in \mathbb{Z}\}$. G is cyclic $\iff \exists\, g \in G,\ \langle g \rangle = G$ (and such a g is a generator of G).

1.5.5 Definition

Let G be a group and let $g \in G$. g has finite order $\iff \exists\, n \in \mathbb{N} \setminus \{0\},\ g^n = e$ (the identity in G). The smallest such n is the order of g, $\mathrm{ord}(g)$, and if there is no such n, g has inifite order.

1.5.6 Theorem

Let G be a group and let $g \in G$ have finite order n.

1. $\forall\, l, m \in \mathbb{Z},\ g^l = g^m \iff n | (l - m) \iff l \equiv m \mod n$.
2. Corollary of 1: $\forall\, m \in \mathbb{Z},\ g^m = e \iff n | m \iff m \equiv 0 \mod n$.
3. $\mathrm{ord}(g) = |\langle g \rangle|$.

1.5.7 Theorem

Let $(G,*)$ be a finite group.

1. Lemma for 2: Every $g \in G$ has finite order.
2. Suppose $H \subseteq G$. $H \leqslant G \iff$

   i $H \neq \emptyset$.

   ii $h_1, h_2 \in H \implies h_1 * h_2 \in H$.

(i.e. the test for subgroups 1.5.2 can be reduced to two conditions for finite groups).

1.5.8 Lemma: Bézout's Identity

Let $a, b \in \mathbb{Z}, d = \gcd(a,b)$. $\exists\, x, y \in \mathbb{Z},\ ax + by = d$ and in fact the integers of the form $ax + by$ are the multiples of d (proved in introductory module).

1.5.9 Theorem

Let G be a cyclic group and let $\langle g \rangle = G$.

1. $H \leqslant G \implies H$ is cyclic.
2. Let $\operatorname{ord}(g) = |G| = n$, and $m \in \mathbb{Z}$. Let $d = \gcd(m,n)$. $\langle g^m \rangle = \langle g^d \rangle$ and $|\langle g^d \rangle| = \frac{n}{d}$ (and so $\langle g^m \rangle = \langle g \rangle = G \iff d = 1 \iff m, n$ are coprime).
3. G has a cyclic subgroup of order $k \leq n \iff k | n$, and such a subgroup is $\langle g^{\frac{n}{k}} \rangle$.

1.5.10 Definition

The Euler totient function is

$$\phi(n) = \left| \left\{ k \in \mathbb{N} \mid 1 \leq k \leq n \wedge \gcd(k,n) = 1 \right\} \right|$$

(the number of natural numbers up to n which are coprime with n).

1.5.11 Corollary

$$\sum_{d|n} \phi(d) = n.$$

(This corollary can be proved using the properties of a cyclic group of order n).

1.5.12 Definition

Let G be a group, let $S \subseteq G$ and $S \neq \emptyset$, and let $S^{-1} = \{g^{-1} \mid g \in S\}$. The subgroup generated by $S \subseteq G$ is $\langle S \rangle := \{g_1 ... g_k \mid k \in \mathbb{N},\ g_1, ..., g_k \in S \cup S^{-1}\}$.

1.5.13 Lemma

1. $\langle S \rangle \leqslant G$.
2. $H \leqslant G$ and $S \subseteq H \implies \langle S \rangle \leqslant H$.

1.5.14 Theorem

Let G be a finite group, $|G| = n$ and $g \in G$. $\operatorname{ord}(g) | n$ and $g^n = e$ (the identity in G).

1.5.15 Theorem: Fermat's Little Theorem

Let p be prime and $x \in \mathbb{Z}$. $x^p \equiv x \mod p$.

1.5.16 Theorem

G has prime order $\implies G$ is cyclic, and $\langle g \rangle = G\ \forall\, g \in G$.

2 Lagrange's Theorem and Cosets

2.1 Cosets

2.1.1 Definition

Let G be a group and $H \leqslant G$. Let $g \in G$. $gH := \{gh \mid h \in H\}$ is a left coset of H in G.

2.1.2 Lemma

Let G be a group and $H \leqslant G$. Let $g_1, g_2 \in G$.

1. $g_2 \in g_1 H \implies g_2 H = g_1 H$.
2. $g_1 H \cap g_2 H \neq \emptyset \implies g_1 H = g_2 H$.

2.1.3 Lemma

Let G be a group and $H \leqslant G$. Let $g \in G$. The map

$$\begin{aligned} H &\longrightarrow gH \\ h &\longmapsto gh \end{aligned}$$

is a bijection.

2.1.4 Corollary

Any two cosets of the same subgroup have the same cardinality.

2.1.5 Theorem: Lagrange's Theorem

Let G be a finite group and $H \leqslant G$.

$$|H| \text{ divides } |G|.$$

2.1.6 Corollary

It follows immediately from the contrapositive that $n \nmid |G| \implies \nexists$ a subgroup of order n.

2.1.7 Definition

The number of left cosets of H in G is the index of H in G.

3 Homomorphisms

3.0.1 Definition

Let G, H be groups. $\phi : G \longrightarrow H$ is a homomorphism $\iff \forall\, g_1, g_2 \in G,\ \phi(g_1 g_2) = \phi(g_1)\phi(g_2)$ (where the operation on the left side takes place within G, and on the right side within H).

3.0.2 Definition

Let G, H be groups and let $\phi : G \longrightarrow H$ be a homomorphism.
The image of ϕ is $\operatorname{Im}\phi = \{\phi(g) \mid g \in G\} \subseteq H$.
The kernel of ϕ is $\operatorname{Ker}\phi = \{g \in G \mid \phi(g) = e_H\} \subseteq G$.

3.0.3 Theorem

Let G, H be groups and let $\phi : G \longrightarrow H$ be a homomorphism. $\operatorname{Im}\phi \leqslant H$ and $\operatorname{Ker}\phi \leqslant G$.

3.0.4 Lemma

Let G, H be groups and let $\phi : G \longrightarrow H$ be a homomorphism.

1. $\phi(e_G) = e_H$.
2. $\phi(g^{-1}) = \big(\phi(g)\big)^{-1}\ \forall\, g \in G$.

3.0.5 Theorem

Let G, H be groups. A homomorphism $\phi : G \longrightarrow H$ is injective $\iff$ $\operatorname{Ker}\phi = \{e_G\}$.

3.0.6 Definition

A homomorphism $\phi : G \longrightarrow H$ which is also a bijection is an isomorphism; this may be denoted by $\phi : G \overset{\sim}{\longrightarrow} H$. G and H are isomorphic $\iff$ $\exists\, \phi : G \overset{\sim}{\longrightarrow} H$ $\iff$ $G \cong H$.

3.0.7 Lemma

Let G, H, K be groups.

1. $\phi : G \longrightarrow H$ is an isomorphism $\implies$ $\phi^{-1} : H \longrightarrow G$ is an isomorphism.
2. $\phi : G \longrightarrow H$ and $\psi : H \longrightarrow K$ are homomorphisms $\implies$ $\psi \circ \phi : G \longrightarrow K$ is a homomorphism (and if ϕ, ψ are isomorphisms, $\psi \circ \phi$ is an isomorphism).

3.0.8 Theorem

Isomorphism ($\cong$) is an equivalence relation among groups.

3.0.9 Theorem

Any two cyclic groups of the same order are isomorphic.

3.0.10 Theorem

Any two non-cyclic groups of order 4 are isomorphic.

3.1 Disjoint Cycle Form

As in definition 1.3.3, throughout this section S_n represents the set of all permutations on $\{1, ..., n\}$ and for convenience, I may write $[n] := \{1, ..., n\}$. I may also use ι to denote the identity permutation.

3.1.1 Definition

Let $f \in S_n$ and $x \in [n]$. f fixes x $\iff$ $f(x) = x$ and f moves x $\iff$ $f(x) \neq x$.

3.1.2 Definition

The support of $f \in S_n$ is $\operatorname{supp}(f) := \{x \in [n] \mid f(x) \neq x\}$ (the set of elements of $[n]$ moved by f).

3.1.3 Corollary

1. $x \in \operatorname{supp}(f) \implies f(x) \in \operatorname{supp}(f)$.
2. $\operatorname{supp}(f) = \operatorname{supp}(f^{-1})$.

3.1.4 Definition

$f, g \in S_n$ are disjoint $\iff$ their supports are disjoint $\iff$ $\operatorname{supp}(f) \cap \operatorname{supp}(g) = \emptyset$.

3.1.5 Lemma

Let $f, g \in S_n$ be disjoint.

1. $fg = gf$.
2. $\forall\, m \in \mathbb{Z},\ (fg)^m = f^m g^m$.

3.1.6 Definition

Let $f \in S_n$ and $r \leq n$.

$$f(i_1) = i_2, ..., f(i_{r-1}) = i_r,\ f(i_r) = i_1 \text{ for distinct } i_1, ..., i_r \in [n] \text{ and } f \text{ fixes all other elements of } [n] \\ \iff f \text{ is an } r\text{-cycle, which me may write } (i_1 i_2 ... i_{r-1} i_r).$$

3.1.7 Note

$(i_1 i_2 ... i_{r-1} i_r) = (i_2 i_3 ... i_r i_1) = ... = (i_r i_1 ... i_{r-2} i_{r-1})$. Note also that cycles may be written with commas so as to eliminate ambiguity, for example when two-digit numbers are present.

3.1.8 Theorem

The number of distinct r-cycles on a set of size n is $\frac{n!}{r(n-r)!}$.

3.1.9 Lemma

$f = (i_1 i_2 ... i_{r-1} i_r) \iff f^{-1} = (i_r i_{r-1} ... i_2 i_1)$.

3.1.10 Note

Since cycles are permutations, they may be composed like functions (and I will do this right to left, as with functions). It is often useful to simplify products of cycles into products of disjoint cycles, so that lemma 3.1.5 may be applied. Notice that composing two cycles does not necessarily produce a cycle.

3.1.11 Theorem

$f \in S_n$ (and $f \neq \iota$) $\implies \exists$ disjoint cycles $f_1, ..., f_k \in S_n$ such that $f = f_1 ... f_k$, called the disjoint cycle form.

3.1.12 Note

By the closure of S_n and theorem 3.1.11, note that it is always possible to simplify as mentioned in 3.1.10. In practice there is a quick way to do this, but as a more foolproof method, the disjoint cycle form may be found by composing the permutations using Cauchy two-line notation, and then extracting the disjoint cycles, as in the example below. The 'matrix' is not standard notation; I am using it to show composition†.

$$(89)(2468)(349) \overset{\dagger}{=} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 2 & 4 & 9 & 5 & 6 & 7 & 8 & 3 \\ 1 & 4 & 6 & 9 & 5 & 8 & 7 & 2 & 3 \\ 1 & 4 & 6 & 8 & 5 & 9 & 7 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 4 & 6 & 8 & 5 & 9 & 7 & 2 & 3 \end{pmatrix} = (248)(369).$$

3.1.13 Theorem

Let $f \in S_n = f_1 ... f_k$ (in its disjoint cycle form, so $f_1, ..., f_k$ are disjoint cycles). Let $m \in \mathbb{N}$.

1. $f^m = f_1^m ... f_k^m$ and $f^{-1} = f_1^{-1} ... f_k^{-1}$ (corollaries of 3.1.5).
2. $f^m = \iota \iff f_i^m = \iota \;\forall\, 1 \leq i \leq k$.
3. $\text{ord}(f) = \text{lcm}(r_1, ..., r_k)$ where (r_i is the cycle length f_i).

3.1.14 Definition

The cycle shape of $f \in S_n$ is the sequence of its cycle lengths in descending order, including fixed points to be thought of as 1-cycles. Several of the same shape consecutively may be abbreviated using an exponent (so for example $\iota \in S_n$ has cycle shape 1^n).

3.2 The Signature of a Permutation

3.2.1 Definition

For a polynomial $P(x_1, ..., x_n)$ and a permutation $f \in S_n$, we extend the definition of the permutation to the polynomial by defining $f(P) = P(x_{f(1)}, ..., x_{f(n)})$ (the polynomial obtained permuting the variables of P by f).

3.2.2 Definition

For the results ahead, let $\Delta = \prod_{1 \leq i < j \leq n} (x_i - x_j)$ where $n \in \mathbb{N}$ and $x_1, ..., x_n$ are variables, similarly to the Vandermonde determinant; note that Δ is a polynomial in $x_1, ..., x_n$ with coefficients in $\mathbb{R}$.

3.2.3 Lemma

$\forall\, n \geq 2,\; f \in S_n$, either $f(\Delta) = \Delta$ or $f(\Delta) = -\Delta$.

3.2.4 Definition

The signature (or signum, but more commonly sign) of $f \in S_n$ is defined

$$\begin{aligned} \mathrm{sgn} : S_n &\longrightarrow (\{1,-1\},\cdot) \\ f &\longmapsto \begin{cases} 1, & f(\Delta) = \Delta \\ -1, & f(\Delta) = -\Delta. \end{cases} \end{aligned}$$

3.2.5 Lemma

$\forall\, f, g \in S_n$ and polynomials $P(x_1, ..., x_n)$,

1. $f(\alpha P) = \alpha f(P)\ \forall\, \alpha \in \mathbb{R}$.
2. $g(f(P)) = g \circ f(P)$.

3.2.6 Theorem

For $n \geq 2$ (otherwise the result is trivial), sgn as defined in 3.2.4 is a homomorphism.

3.2.7 Theorem

1. f is a 2-cycle $\implies \mathrm{sgn}(f) = -1$.
2. f is an r-cycle $\implies \mathrm{sgn}(f) = (-1)^{r-1}$.

3.2.8 Definition

$f \in S_n$ is even if $\mathrm{sgn}(f) = 1$ and odd if $\mathrm{sgn}(f) = -1$. Beware that this makes a cycle of even length an odd permutation and vice versa.

3.2.9 Theorem

$\{f \in S_n \mid f \text{ is even}\} \leqslant S_n$. The subgroup is called A_n, the alternating group, and its two cosets in S_n are A_n and the set of odd permutations.

3.2.10 Theorem

Further to the proof of 3.2.7.2, every permutation in S_n can be written as a product of 2-cycles. Noting that the parity of the number of such 2-cycles must be the same as the parity of the permutation provides some clarification for this terminology (for example 1 is odd, and any 2-cycle is odd as in 3.2.7).

3.2.11 Theorem

Let $A = [a_{ij}]_{n\times n} \in M_{n\times n}(\mathbb{F})$.

$$\det(A) = \sum_{f \in S_n} \mathrm{sgn}(f) a_{1f(1)} ... a_{nf(n)}.$$

3.3 Dihedral Groups

3.3.1 Definition

For $n \in \mathbb{N}$, $n \geq 3$, the dihedral group D_{2n} is the group of symmetries of a regular n-gon. This is not a rigorous abstract definition (see 3.3.5), but geometrically, a symmetry is either of the two transformations described below. Let the n-gon be defined by its n numbered, evenly spaced vertices on a circle with its centre at $0 \in \mathbb{R}^2$.

i Rotate about the origin through a multiple of $\frac{2\pi}{n}$.

ii Reflect across an axis through the origin and at least one vertex or midpoint of an edge.

3.3.2 Lemma

Let $X \neq \emptyset$, $x \in X$ and $G \leqslant \mathrm{Sym}(X)$, where G is finite. Let $H = \{g \in G \mid g(x) = x\}$ and $Y = \{g(x) \mid g \in G\}$.

1. $H \leqslant G$ and for $g_1, g_2 \in G$, $g_1 H = g_2 H \iff g_1(x) = g_2(x)$.
2. $\frac{|G|}{|H|} = |Y|$.

3.3.3 Theorem

$D_{2n} \leqslant S_n$ and $|D_{2n}| = 2n$ (provided S_n is thought of as the set of permutations of the vertices of the n-gon).

3.3.4 Note

From the result above, it's clear why D_{2n} is denoted this way; be aware that in geometry, the dihedral group may be written D_n.

3.3.5 Definition: Precise Abstract Definition of the Dihedral Group

Let $n \in \mathbb{N},\ n \geq 3$ and let $s, t \in S_n$ be defined by their disjoint cycle forms as follows: if n is even,

$$\begin{aligned} s &= (1)(2,n)(3,n-1)...(\tfrac{n}{2}+1) \\ t &= (1,n)(2,n-1)(3,n-2)...(\tfrac{n}{2},\tfrac{n}{2}+1) \end{aligned}$$

and if n is odd,

$$\begin{aligned} s &= (1)(2,n)(3,n-1)...(\tfrac{n+1}{2},\tfrac{n+1}{2}+1) \\ t &= (1,n)(2,n-1)(3,n-2)...(\tfrac{n+1}{2}) \end{aligned}$$

(for which the geometric interpretation is that s is the reflection which fixes vertex 1 and t is the reflection which swaps vertices 1 and n). The dihedral group is defined $D_{2n} = \langle s, t \rangle$.

3.3.6 Theorem

By the definition above, it is given that $D_{2n} \leqslant S_n$. With some justification, st can be shown to be the cycle $(123...n)$, and thus it can be proved that $|D_{2n}| = 2n$ from this definition also.