Vectors and Matrices

Shing Tak Lam*

April 15, 2021

This document is intended for revision purposes. As a result, it does not contain any exposition. This is based off lectures given by Dr Jonathan Evans in Michaelmas 2020, but the order of content, as well as some of the proofs have been modified after the fact, primarily to provide simpler proofs for theorems. Note that this also contains theorems from examples sheets, as some are useful elsewhere.

Vectors and Matrices is on *Paper 1*.

Contents

*stl45@cam.ac.uk

1 Complex Numbers

Definition of $\mathbb{C}$ and elementary properties shall be assumed. The complex conjugate is dnenoted by $\overline{z}$ in this course.

Theorem (Fundamental Theorem of Algebra). *A polynomial over $\mathbb{C}$ of degree n has n roots in $\mathbb{C}$, counted with multiplicity.*

Theorem (de Moivre's Theorem). *For any $n \in \mathbb{Z}$,*

$$(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta)$$

Proof. The case for $n = 0$ is trivial. For the inductive case, $(\cos\theta + i\sin\theta)^{n+1} = (\cos(n\theta) + i\sin(n\theta))(\cos\theta + i\sin\theta) = (\cos(n\theta)\cos\theta - \sin(n\theta)\sin\theta) + i(\sin(n\theta)\cos\theta + \cos(n\theta)\sin\theta) = \cos((n+1)\theta) + i\sin((n+1)\theta)$.

Note that $(\cos\theta + i\sin\theta)^{-m} = (\cos(m\theta) + i\sin(m\theta))^{-1} = \cos(m\theta) - i\sin(m\theta) = \cos(-m\theta) + i\sin(-m\theta)$. □

Definition (exp, cos, sin). For $z \in \mathbb{C}$, we define

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$
$$\cos(z) = \frac{\exp(iz) + \exp(-iz)}{2}$$
$$\sin(z) = \frac{\exp(-iz) - \exp(-iz)}{2i}$$

Proposition.

$$\cos(z) = 1 - \frac{z^2}{2} + \frac{z^4}{4!} - \dots$$

Proposition.

$$\sin(z) = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots$$

Proposition.

$$\exp(z + w) = \exp(z)\exp(w)$$

Lemma. *(i)* $e^{x+iy} = e^x(\cos y + i\sin y)$

(ii) $\exp(\mathbb{C}) = \mathbb{C}\backslash\{0\}$

(iii) $e^z = 1 \iff z = 2n\pi i$ *for some* $n \in \mathbb{N}$.

Proof. (i). By definitions.

(ii) By definitions and properties of the real exponential.

(iii) Using (i) and standard properties of sin and cos. □

Definition (Roots of Unity)**.** The n-th roots of unity are the solutions to $z^n - 1 = 0$.

Proposition. *The* n*-th roots of unity are of the form* ω^k, *where* $0 \leq k < n$, $\omega = \exp\left(\frac{2\pi i}{n}\right)$.

Proof. Clearly they satisify $(\omega^k)^n - 1 = 0$. By the Fundamental Theorem of Algebra, these must be the only roots. □

Definition (log)**.** For $z \in \mathbb{C}$, we define

$$\log(z) = \log|z| + i \arg(z)$$

Note this is multivalued as $\arg(z)$ is multivalued.

Definition (Complex Exponentiation)**.**

$$z^a = \exp(a \log z)$$

Note this is multivalued as $\log(z)$ is multivalued.

Proposition. *A line in* $\mathbb{C}$ *through* z_0 *and parallel to* w *is given by*

$$\overline{w}z - w\overline{z} = \overline{w}z_0 - w\overline{z_0}$$

Proposition. *A circle in* $\mathbb{C}$ *with centre* a *and radius* r *is given by*

$$|z - a| = r$$

2 Vectors in $\mathbb{R}^3$

Definition (Scalar Triple Product)**.** $[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a})$ etc.

Definition (Vector Triple Product)**.** $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$.

Proposition. *A line in* $\mathbb{R}^3$ *has the form* $\mathbf{r} = \mathbf{a} + \lambda\mathbf{u}$, *or* $\mathbf{r} \times \mathbf{u} = \mathbf{c}$, *where* $\mathbf{u}$ *and* $\mathbf{c}$ *are constant vectors.*

Proposition. *A plane in* $\mathbb{R}^3$ *through* $\mathbf{a}$ *and with normal* $\mathbf{n}$ *is given by* $\mathbf{n} \cdot \mathbf{r} = \mathbf{n} \cdot \mathbf{a} =$ *constant.*

Definition (Kronecker Delta)**.**

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Definition (Levi-Civita Epsilon)**.**

$$\varepsilon_{ijk} = \begin{cases} 1 & \text{if } (i\ j\ k) \text{ is an even permutation} \\ -1 & \text{if } (i\ j\ k) \text{ is an odd permutation} \\ 0 & \text{if } (i\ j\ k) \text{ is not a permutation} \end{cases}$$

Proposition.

$$\begin{aligned} \varepsilon_{ijk}\varepsilon_{pqr} = {} & \delta_{ip}\delta_{jq}\delta_{kr} - \delta_{jp}\delta_{iq}\delta_{kr} \\ & + \delta_{jp}\delta_{kq}\delta_{ir} - \delta_{kp}\delta_{jq}\delta_{ir} \\ & + \delta_{kp}\delta_{iq}\delta_{jr} - \delta_{ip}\delta_{kq}\delta_{jr} \end{aligned}$$

Proof. Note that both sides are totally antisymmetric in i, j, k and in p, q, r, which suggests that both sides are equal up to multiplcation by a constant. By substituting $i = p = 1$, $j = q = 2$ and $k = r = 3$, we get that the left hand side and the right hand side are both 1. Therefore the left hand side and the right hand side must be equal. □

Proposition.

$$\varepsilon_{ijk}\varepsilon_{pqk} = \delta_{ip}\delta_{jq} - \delta_{jp}\delta_{iq}$$

Proof. Let $r = k$ in the general identity above. Also note that $\delta_{kk} = 3$. □

Proposition.

$$\varepsilon_{ipk}\varepsilon_{ipq} = 2\delta_{kq}$$

Proof. By permuting the indices, we have that $\varepsilon_{ipk}\varepsilon_{ipq} = \varepsilon_{kip}\varepsilon_{qip} = \delta_{qk}\delta_{ii} - \delta_{ki}\delta_{qi} = 3\delta_{kq} - \delta_{kq} = \delta_{kq}$. □

3 Vector Spaces

Definition (Vector Space)**.** A vector space over a field $\mathbb{F}$ is a set V, where V is an abelian group under addition, and

- $\lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}$.
- $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$
- $\lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}$
- $1\mathbf{v} = \mathbf{v}$

4 Inner Product Space

In this section, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and for $x \in \mathbb{R}$, $\overline{x} = x$. Note further that in this course, we're using "Physicist's notation", where the variable which is conjugate linear is opposite to what Pure Mathematics uses.

Definition (Inner Product Space)**.** An inner product space is a vector space V over a field $\mathbb{F}$ together with a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ satisfying

- $\langle \mathbf{z}, \lambda\mathbf{u} + \mu\mathbf{w} \rangle = \lambda\langle \mathbf{z}, \mathbf{u} \rangle + \mu\langle \mathbf{z}, \mathbf{w} \rangle$
- $\langle \lambda\mathbf{u} + \mu\mathbf{w}, \mathbf{z} \rangle = \overline{\lambda}\langle \mathbf{u}, \mathbf{z} \rangle + \overline{\mu}\langle \mathbf{w}, \mathbf{z} \rangle$
- $\langle \mathbf{w}, \mathbf{z} \rangle = \overline{\langle \mathbf{w}, \mathbf{z} \rangle}$
- $\langle \mathbf{x}, \mathbf{x} \rangle \in \mathbb{R}$, and ≥ 0. Furthermore, it is 0 if and only if $\mathbf{x} = \mathbf{0}$.

Using this, we can define norms, as well as what it means for two vectors to be orthogonal. Furthermore, the Cauchy-Schwarz inequality holds in any inner product space.

5 Vectors in $\mathbb{R}^n$

Definition (Inner Product)**.** For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\mathbf{x} \cdot \mathbf{y} = x_i y_i$$

Proposition. *The inner product is symmetric, bilinear and positive definite.*

Definition (Norm)**.** For $\mathbf{x} \in \mathbb{R}^n$, the norm of $\mathbf{x}$ is defined by

$$|\mathbf{x}|^2 = ||\mathbf{x}||^2 = \mathbf{x} \cdot \mathbf{x}$$

Theorem (Cauchy-Schwarz Inequality). *For all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$$

with equality if and only if $\mathbf{x}$ *and* $\mathbf{y}$ *are parallel.*

Proof. If $\mathbf{y} = \mathbf{0}$, the result is trivial. Otherwise, consider $\|\mathbf{x} - \lambda\mathbf{y}\|$.

$$\|\mathbf{x} - \lambda\mathbf{y}\|^2 = (\mathbf{x} - \lambda\mathbf{y}) \cdot (\mathbf{x} - \lambda\mathbf{y}) = \|x\| - 2\lambda\mathbf{x} \cdot \mathbf{y} + \lambda^2\|\mathbf{y}\| \geq 0$$

This is a quadratic in λ, and as it is always non-negative, it has at most one real root. Consider the discriminant, we get that

$$4(\mathbf{x} \cdot \mathbf{y})^2 - 4\|\mathbf{x}\|^2\|\mathbf{y}\|^2 \leq 0$$

Hence

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$$

and equality holds if and only if $\mathbf{x} = \lambda\mathbf{y}$. □

Proposition (Triangle Inequality). *For all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

Proof.

$$\|\mathbf{x} + \mathbf{y}\|^2 = (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\mathbf{x} \cdot \mathbf{y} \leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + \|x\|\|y\| = (\|x\| + \|y\|)^2$$

□

6 Linear Indepdence

In this section, let V be a (real) vector space.

Definition (Linear Indepdence). . Let $\mathbf{v}_1, \ldots, \mathbf{v}_r \in V$. The $\mathbf{v}_i$ are linearly independent if

$$\sum_{i=1}^{r} \lambda_i \mathbf{v}_i = \mathbf{0}$$

if and only if $\lambda_i = 0$ for all i.

Lemma. *In any real inner product space* V*, if* $\mathbf{v}_1, \ldots, \mathbf{v}_r \neq \mathbf{0}$ *and orthogonal, then they are linearly independent.*

Proof. If $\sum_i \lambda_i \mathbf{v}_i = \mathbf{0}$, then $\langle \mathbf{v}_j, \sum_i \lambda_i \mathbf{v}_i \rangle = \sum_i \lambda_i \langle \mathbf{v}_j, \mathbf{v}_i \rangle = \lambda_j \|\mathbf{v}_j\| = 0$. So $\lambda_j = 0$ for all j, and the vectors are linearly independent. □

7 Basis and Dimension

Definition (Basis). For a vector space V, a basis $\mathcal{B}$ is a set such that

- $\text{span}(\mathcal{B}) = V$
- $\mathcal{B}$ is linearly independent.

Theorem. *If* $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ *and* $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ *are bases for* V*, then* $n = m$*.*

Proof. For each a, $\mathbf{f}_a = \sum_i A_{ai}\mathbf{e}_i$. Similarly, $\mathbf{e}_i = \sum_a B_{ia}\mathbf{f}_a$. So $\mathbf{f}_a = \sum_i A_{ai} \sum_b B_{ib}\mathbf{f}_b = \sum_b (\sum_i A_{ai}B_{ib})\mathbf{f}_b$. As the $\mathbf{f}$s are linearly independent, $\sum_i A_{ai}B_{ib} = \delta_{ab}$. Similarly, from $\mathbf{e}$, we get that $\sum_a B_{ia}A_{aj} = \delta_{ij}$. Then $\sum_{i,a} A_{ai}B_{ia} = \sum_a \delta_{aa} = m$, but we also have that $\sum_{i,a} A_{ai}B_{ia} = \sum_i \delta_{ii} = n$. So $m = n$. □

Definition (Dimension)**.** We define the dimension $\dim(V)$ of a vector space V as the size of any basis for V.

Proposition ((A version of the) Steinitz Exchange Lemma)**.** *Let V be a vector space with $\dim(V) = n$, with*

- $Y = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ *with* $\operatorname{span} Y = V$.
- $X = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ *that are linearly independent.*

Then $k \le n \le m$, and

(i) A basis can be found as a subset of Y, by discarding vectors as necessary.

(ii) X can be extended to a basis by adding vectors from Y as necessary.

Proof. If Y is linearly independent, then Y is a basis, and $n = m = \dim V$. If Y is linearly dependent, then without loss of generality (by reordering the $\mathbf{w}_i$), we may write $\mathbf{w}_m = \sum_{i=1}^{m-1} \lambda_i \mathbf{w}_i$. Then $\operatorname{span} Y = \operatorname{span}(Y \backslash \{\mathbf{w}_m\})$. Repeat this until a basis is found. Consequently, we must have that $n \le m$.

If X spans V, then X is a basis, and $k = n$. Otherwise, there exists $\mathbf{u}_{k+1} \in V$ which is not in $\operatorname{span} X$. But as $\mathbf{u}_{k+1} \notin \operatorname{span} X$, $\sum_{i=1}^{k+1} \lambda_i \mathbf{u}_i = \mathbf{0}$ must mean that $\lambda_i = 0$ for all i. Hence $X \cup \{\mathbf{u}_{k+1}\}$ is linearly independent. Furthermore, we may choose $\mathbf{u}_{k+1} \in Y$, as if $Y \subseteq \operatorname{span} X$, then $\operatorname{span} Y \subseteq \operatorname{span} X$ and $\operatorname{span} X = V$. Repeat this until a basis is obtained. □

8 $\mathbb{C}^n$

Again, note that in this course we're using "Physicists' Notation" and the conjugation is opposite in Pure Mathematics.

Definition (Complex Inner Product)**.**

$$\langle \mathbf{z}, \mathbf{w} \rangle = \sum_{i=1}^{n} \overline{z_i} w_i$$

satisfies the axioms of an inner product.

9 Linear Maps

Definition. Let V and W be vector spaces, then $T : V \to W$ is a linear map if

$$T(\lambda \mathbf{v} + \mu \mathbf{w}) = \lambda T(\mathbf{v}) + \mu T(\mathbf{w})$$

Proposition. *A linear map is completely determined by its action on a basis.*

$$T(\mathbf{v}) = T\left(\sum_{i=1}^{n} v_i \mathbf{e}_i\right) = \sum_{i=1}^{n} v_i T(\mathbf{e}_i)$$

Definition (Image)**.**

$$\operatorname{Im}(T) = \{\mathbf{w} \in W : \exists \mathbf{v} \in V, T(\mathbf{v}) = \mathbf{w}\}$$

Definition (Kernel)**.**

$$\ker(T) = \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}\}$$

Lemma. $\ker T$ *is a subspace of V,* $\operatorname{Im} T$ *is a subspace of W.*

Definition (Rank)**.**

$$\operatorname{rank} T = \dim \operatorname{Im} T$$

Definition (Nullity).

$$\text{null } T = \dim \ker T$$

Theorem (Rank-Nullity).

$$\dim V = \text{rank } T + \text{null } T$$

Proof. Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ be a basis for $\ker T$, and extending this to a basis of V by adding on $\{\mathbf{e}_{k+1}, \ldots, \mathbf{e}_n\}$. We claim that $\mathcal{B} = \{T(\mathbf{e}_{k+1}), \ldots, T(\mathbf{e}_n)\}$ is a basis for $\text{Im } T$.

For any $\mathbf{x} \in \text{Im } T$,

$$\mathbf{x} = T(\mathbf{v}) = T\left(\sum_{i=1}^{n} v_i \mathbf{e}_i\right) = \sum_{i=1}^{n} v_i T(\mathbf{e}_i) = \sum_{i=1}^{k} v_i T(\mathbf{e}_i) + \sum_{i=k+1}^{n} v_i T(\mathbf{e}_i) = \sum_{i=k+1}^{n} v_i T(\mathbf{e}_i) \in \text{span}\,\mathcal{B}$$

Now suppose if $\sum\limits_{i=k+1}^{n} \lambda_i T(\mathbf{e}_i) = \mathbf{0}$. Then $T\left(\sum\limits_{i=k+1}^{n} \lambda_i \mathbf{e}_i\right) = \mathbf{0}$. So $\sum\limits_{i=k+1}^{n} \lambda_i \mathbf{e}_i \in \ker T$. This means that for some μ_i, we have $\sum\limits_{i=k+1}^{n} \lambda_i \mathbf{e}_i = \sum_{i=1}^{k} \mu_i \mathbf{e}_i$. As the $\mathbf{e}_i$s form a basis, we must have that $\lambda_i = 0$ for all i. So $\mathcal{B}$ is a basis.

Consequently, $\text{rank } T = n - k$, $\text{null } T = k$ and $\text{rank } T + \text{null } T = n = \dim V$. □

9.1 Geometrical Examples

Definition (Rotation). An anticlockwise rotation about an axis given by a unit vector $\mathbf{n}$ is given by

$$T(\mathbf{x}) = (\cos\theta)\mathbf{x} + (1 - \cos\theta)(\mathbf{n} \cdot \mathbf{x})\mathbf{n} + (\sin\theta)\mathbf{n} \times \mathbf{x}$$

Definition (Projection). A projection onto a plane with unit normal $\mathbf{n}$ is defined by

$$T(\mathbf{x}) = \mathbf{x} - (\mathbf{x} \cdot \mathbf{n})\mathbf{n}$$

Definition (Reflection). A reflection across a plane with unit normal $\mathbf{n}$ is defined by

$$T(\mathbf{x}) = \mathbf{x} - 2(\mathbf{x} \cdot \mathbf{n})\mathbf{n}$$

Definition (Dilation). Given scale factors $\alpha, \beta, \gamma > 0$, a dilation is defined by

$$T(\mathbf{e}_1) = \alpha\mathbf{e}_1$$
$$T(\mathbf{e}_2) = \beta\mathbf{e}_2$$
$$T(\mathbf{e}_3) = \gamma\mathbf{e}_3$$

Definition (Shear). Let $\mathbf{a}, \mathbf{b}$ be orthogonal unit vectors in $\mathbb{R}^3$, and $\lambda \in \mathbb{R}$. Define a shear parallel to $\mathbf{a}$ with scale factor λ by

$$T(\mathbf{x}) = \mathbf{x} + \lambda\mathbf{a}(\mathbf{x} \cdot \mathbf{b})$$

10 Matrices

Definition (Matrix-Vector Multiplication). If a matrix M represents the action of a linear map T, then

$$T(\mathbf{x}) = M\mathbf{x}$$

and

$$[T(\mathbf{x})]_a = M_{ai} x_i$$

Definition (Matrix Multiplication)**.** Matrix multiplication is given by

$$[AB]_{ij} = A_{ia}B_{aj}$$

Definition (Transpose)**.** The transpose of a matrix M, denoted by M^T is given by

$$[M^T]_{ij} = M_{ji}$$

Definition (Hermitian Conjugate)**.** The hermitian conjugate of a complex matrix M, denoted by $M^\dagger$ is given by

$$[M^\dagger]_{ij} = \overline{M_{ji}}$$

Definition (Symmetric, Antisymmetric Matrices)**.** A square matrix S is symmetric if $S^T = S$. A square matrix A is antisymmetric if $A^T = -A$.

Definition (Hermitian, Antihermitian Matrices)**.** A square complex matrix M is hermitian if $M^\dagger = M$. It is antihermitian if $M^\dagger = -M$.

Proposition. *The inner product can be written in terms of the Hermitian Conjugate (if we regard* 1×1 *matrices and scalars to be equivalent).*

$$\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{z}^\dagger \mathbf{w}$$

Definition (Trace)**.** For q square matrix M, the trace is defined to be

$$\text{tr}(M) = M_{ii}$$

Definition (Orthogonal Matrix)**.** A real square matrix U is orthogonal if $U^T U = I$. Equivalently, $U^T = U^{-1}$.

Proposition. *U is orthogonal if and only if its columns are orthonormal vectors.*

Proof.

$$[U^T U]_{ij} = [U^T]_{ia} U_{aj} = U_{ai} U_{aj} = \delta_{ij}$$

□

Proposition. *U is orthonormal if and only if its rows are orthonormal vectors.*

Proposition. *U is orthogonal if and only if it preserves the real inner product.*

Proof.

$$(U\mathbf{x}) \cdot (U\mathbf{y}) = (U\mathbf{x})^T U\mathbf{y} = \mathbf{x}^T U^T U\mathbf{y} = \mathbf{x}^T \mathbf{y} = \mathbf{x} \cdot \mathbf{y}$$

The reverse implication can be checked by assuming $\mathbf{x}^T U^T U\mathbf{y} = \mathbf{x}^T \mathbf{y}$ and calculating the entries of $U^T U$. □

Definition (Unitary Matrix)**.** A complex square matrix U is unitary if $U^\dagger U = I$, or equivalently $U^\dagger = U^{-1}$.

Proposition. *U is unitary if and only if it preserves the complex inner product.*

11 Determinants and Inverses

Definition (Levi-Civita Symbol).

$$\underbrace{\varepsilon_{ij\ldots l}}_{n \text{ indices}} = \begin{cases} 1 & \text{if } (i\ j\ \ldots\ l) \text{ is an even permutation} \\ -1 & \text{if } (i\ j\ \ldots\ l) \text{ is an odd permutation} \\ 0 & \text{if } (i\ j\ \ldots\ l) \text{ is not a permutation} \end{cases}$$

Proposition. *If σ is a permutation of $\{1, \ldots, n\}$, then*

$$\varepsilon_{\sigma(1)\ldots\sigma(n)} = \operatorname{sign}(\sigma)$$

Definition (Alternating Form). Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in $\mathbb{R}^n$ or $\mathbb{C}^n$, the alternating form is defined to be

$$[\mathbf{v}_1, \ldots, \mathbf{v}_n] = \varepsilon_{ij\ldots l}(\mathbf{v}_1)_i(\mathbf{v}_2)_j \ldots (\mathbf{v}_n)_l$$

Proposition. *The alternating form is multilinear.*

Proposition. *The alternating form is totally antisymmetric. For any permutation σ,*

$$[\mathbf{v}_{\sigma(1)}, \ldots, \mathbf{v}_{\sigma(n)}] = \operatorname{sign}(\sigma)[\mathbf{v}_1, \ldots, \mathbf{v}_n]$$

Proof. Suffices to check that it holds for a transposition, as every permutation can be written as a product of transpositions. Let $\tau = (p\ q)$, where $p < q$. Then

$$\begin{aligned}
&[\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}, \mathbf{v}_q, \mathbf{v}_{p+1}, \ldots, \mathbf{v}_{q-1}, \mathbf{v}_p, \mathbf{v}_{q+1}, \ldots, \mathbf{v}_n] \\
&\quad = \sum_\sigma \operatorname{sign}(\sigma)(\mathbf{v}_1)_{\sigma(1)} \ldots (\mathbf{v}_p - 1)_{\sigma(p-1)}(\mathbf{v}_q)_{\sigma(q)}(\mathbf{v}_{p+1})_{\sigma(p+1)} \ldots (\mathbf{v}_{q-1})_{\sigma(q-1)}(\mathbf{v}_p)_{\sigma(p)}(\mathbf{v}_{q+1})_{\sigma(q+1)} \ldots (\mathbf{v}_n)_{\sigma(n)} \\
&\quad = \sum_{\sigma'} \operatorname{sign}(\sigma)(\mathbf{v}_1)_{\sigma'(1)} \ldots (\mathbf{v}_p - 1)_{\sigma'(p-1)}(\mathbf{v}_q)_{\sigma'(q)}(\mathbf{v}_{p+1})_{\sigma'(p+1)} \ldots (\mathbf{v}_{q-1})_{\sigma'(q-1)}(\mathbf{v}_p)_{\sigma'(p)}(\mathbf{v}_{q+1})_{\sigma'(q+1)} \ldots (\mathbf{v}_n)_{\sigma'(n)}
\end{aligned}$$

Where $\sigma' = \sigma\tau$, and summing over all σ is the same as summing over all σ'. As $\operatorname{sign}(\sigma\tau) = -\operatorname{sign}(\sigma)$, we get that

$$[\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}, \mathbf{v}_q, \mathbf{v}_{p+1}, \ldots, \mathbf{v}_{q-1}, \mathbf{v}_p, \mathbf{v}_{q+1}, \ldots, \mathbf{v}_n] = -[\mathbf{v}_1, \ldots, \mathbf{v}_n]$$

as expected. □

Proposition.

$$[\mathbf{e}_1, \ldots \mathbf{e}_n] = 1$$

where $\{\mathbf{e}_i\}$ is the standard basis for $\mathbb{R}^n$, that is, $(\mathbf{e}_i)_j = \delta_{ij}$.

Proposition. *If two of the vectors are the same, then the alternating form is zero.*

Proposition. $[\mathbf{v}_1, \ldots, \mathbf{v}_n] \neq 0$ *if and only if* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *are linearly independent.*

s

Proof. First suppose if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ were linearly dependent. Without loss of generality, say $\mathbf{v}_n = \sum_{i=1}^{n-1} \lambda_i \mathbf{v}_i$. Using the multilinearity of the alternating form, and the fact that if two of the vectors were the same then the alternating form is zero, we get that $[\mathbf{v}_1, \ldots, \mathbf{v}_n] = 0$.

Now suppose if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent. Then they span $\mathbb{R}^n$. Hence the standard basis vectors can be written as linear combinations of the $\mathbf{v}_i$s. We can write $\mathbf{e}_i = U_{ai}\mathbf{v}_a$. Thus

$$[\mathbf{e}_1, \ldots, \mathbf{e}_n] = [U_{a1}\mathbf{v}_a, \ldots, U_{bn}\mathbf{v}_b] = U_{a1} \ldots U_{bn}[\mathbf{v}_a, \ldots, \mathbf{v}_b] = U_{a1} \ldots U_{bn}\varepsilon_{a\ldots b}[\mathbf{v}_1, \ldots, \mathbf{v}_n]$$

But $[\mathbf{e}_1, \ldots, \mathbf{e}_n] = 1$, so we must have that $[\mathbf{v}_1, \ldots, \mathbf{v}_n] \neq 0$. □

11.1 Determinants

Definition (Determinant). If M is a $n \times n$ matrix, with columns $\mathbf{C}_a = M\mathbf{e}_a$, $\det M$ is defined to be

$$\det M = [\mathbf{C}_1, \dots, \mathbf{C}_n] = [M\mathbf{e}_1, \dots, M\mathbf{e}_n] = \varepsilon_{i\dots j} M_{i1} \dots M_{jn} = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) M_{\sigma(1)1} \dots M_{\sigma(n)n}$$

Proposition. *If M has rows $\mathbf{R}_a$, then*

$$\det M = [\mathbf{R}_1, \dots, \mathbf{R}_n]$$

Consequently, $\det M = \det\left(M^T\right)$

Proof. Note that $(\mathbf{C}_a)i = M_{ia} = (\mathbf{R}_i)_a$. Then $M_{\sigma(1)1} \dots M_{\sigma(n)n} = M_{1\sigma^{-1}(1)} \dots M_{n\sigma^{-1}(n)}$, and $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1})$, so we are done. □

Definition (Minor). For a matrix M, we define the minor M^{ia} as the determinant of the matrix obtained by removing the i-th row and a-th column of M.

Proposition. *For fixed a,*

$$\det M = \sum_i (-1)^{i+a} M_{ia} M^{ia}$$

Proposition. *For fixed i,*

$$\det M = \sum_a (-1)^{i+1} M_{ia} M^{ia}$$

Proposition (Column/Row Scaling). *If $\mathbf{R}_i \mapsto \lambda \mathbf{R}_i$, or $\mathbf{C}_i \mapsto \lambda \mathbf{C}_i$, then* $\det M \mapsto \lambda \det M$.

Proposition (Column/Row operations). *If $\mathbf{R}_i \mapsto \mathbf{R}_i + \lambda \mathbf{R}_j$ $(i \neq j)$, or $\mathbf{C}_i \mapsto \mathbf{C}_i + \lambda \mathbf{C}_j$ $(i \neq j)$, then* $\det M$ *is fixed.*

Proposition (Column/Row swaps). *If two columns/rows are swapped, then* $\det M \mapsto -\det M$.

Lemma.

$$\varepsilon_{i_1 \dots i_n} M_{i_1 a_1} \dots M_{i_n a_n} = (\det M) \varepsilon_{a_1 \dots a_n}$$

Proof. Note both sides are antisymmetric in $a_1, \dots, a_n$, so they are equal up to a constant. Set $a_i = 1$ to find the constant is 1. □

Alternative Proof. Let $\sigma(i) = a_i$, then $\sigma \in S_n$ and

$$\begin{aligned}
\varepsilon_{i_1 \dots i_n} M_{i_1 a_i} \dots M_{i_n a_n} &= \varepsilon_{i_1 \dots i_n} M_{i_1 \sigma(1)} \dots M_{i_n \sigma(n)} \\
&= \varepsilon_{i_1 \dots i_n} M_{i_1 1} \dots M_{i_n n} \operatorname{sign}(\sigma) \\
&= (\det M) \operatorname{sign}(\sigma) \\
&= (\det M) \varepsilon_{a_1 \dots a_n}
\end{aligned}$$

□

Theorem.

$$\det(MN) = \det M \det N$$

Proof.

$$\begin{aligned}
\det(MN) &= \varepsilon_{i_1 \dots i_n} (MN)_{i_1 1} \dots (MN)_{i_n n} \\
&= \varepsilon_{i_1 \dots i_n} (M_{i_1 k_1} N_{k_1 1}) \dots (M_{i_n k_n} N_{k_n n}) \\
&= \varepsilon_{i_1 \dots i_n} M_{i_1 k_1} \dots M_{i_n k_n} N_{k_1 1} \dots N_{k_n n} \\
&= (\det M) \varepsilon_{k_1 \dots k_n} N_{k_1 1} \dots N_{k_n n} \\
&= \det M \det N
\end{aligned}$$

□

11.2 Cofactors

Definition. If a matrix M has columns $\mathbf{C}_a$, then the cofactor Δ_{ia} is

$$\Delta_{ia} = [\mathbf{C}_1, \dots, \mathbf{C}_{a-1}, \mathbf{e}_i, \mathbf{C}_{a+1}, \dots, \mathbf{C}_n]$$

Proposition.

$$\Delta_{ia} = (-1)^{i+a} M_{ia}$$

Proposition. *For a fixed,*

$$\det M = \sum_i M_{ia} \Delta_{ia}$$

Proposition.

$$[\mathbf{C}_1, \dots, \mathbf{C}_{a-1}, \mathbf{C}_b, \mathbf{C}_{a+1}, \dots, \mathbf{C}_n] = \begin{cases} \det M & \text{if } a = b \\ 0 & \text{otherwise} \end{cases} = \delta_{ab} \det M$$

Definition (Cofactor Matrix)**.** For a matrix M, define the cofactor matrix Δ with entries Δ_{ia}.

11.3 Adjugate and Inverses

Definition (Adjugate Matrix)**.**

$$\mathrm{Adj}(M) = \tilde{M} = \Delta^T$$

Proposition.

$$\tilde{M}M = (\det M)I$$

Proof.

$$[\tilde{M}M]_{ab} = \tilde{M}_{ai} M_{ib} = \Delta_{ia} M_{ib} = \delta_{ab} \det M = (\det M) I_{ab}$$

□

Definition (Inverse Matrix)**.** If $\det M \neq 0$, then $M^{-1} = \dfrac{1}{\det M}\tilde{M}$.

12 Systems of Linear Equations

Proposition. *For a linear system $A\mathbf{x} = \mathbf{b}$, there are three possibilities*

- $\det A \neq 0$ *and* $\mathbf{x} = A^{-1}\mathbf{b}$ *is the unique solution.*
- $\det A = 0$ *and* $\mathbf{b} \notin \mathrm{Im}\, A$*, and there is no solution.*
- $\det A = 0$ *and* $\mathbf{b} \in \mathrm{Im}\, A$*, then* $\mathbf{x} = \mathbf{x}_0 + \mathbf{u}$ *is a solution, where* $A\mathbf{x}_0 = \mathbf{b}$*,* $A\mathbf{u} = \mathbf{0}$.

Proof. Note that $\det A \neq 0 \iff \mathrm{Im}\, A = \mathbb{R}^n \iff \ker A = \{\mathbf{0}\}$, and there is a solution if and only if $\mathbf{b} \in \mathrm{Im}\, A$.

If $\det A \neq 0$, then $\mathbf{x} = A^{-1}\mathbf{b}$ is a solution. Suppose if $\mathbf{x}'$ was also a solution. Then $A(\mathbf{x} - \mathbf{x}') = \mathbf{0}$, so $\mathbf{x} - \mathbf{x}' \in \ker A = \{\mathbf{0}\}$, and $\mathbf{x} = \mathbf{x}'$.

If $\det A = 0$ and $b \in \mathrm{Im}\, A$, then $\mathrm{rank}\, A < n$ and $\mathrm{null}\, A > 0$. Then choosing any $u \in \ker A$ and adding it to a solution will yield another solution. □

12.1 Gaussian Elimination

Definition (Gaussian Elimination)**.** For a general system $A\mathbf{x} = \mathbf{b}$, reorder the equations (1 to m) and variables (rename/reorder $x_1, \ldots, x_n$ to $y_1, \ldots, y_n$) to get the coefficient of y_1 in equation 1 to be

$$B_{11}^{(1)} \neq 0$$

Eliminate y_1 from all of the other equations by subtracting off multiples of equation 1. Then reorder equations/variables such that the coefficient of y_2 in equation 2 is

$$B_{22}^{(2)} \neq 0$$

and so on.

This process stops when all coefficients in equations $r+1, \ldots, n$ are 0.

Proposition. *$r \leq m$, $r \leq n$.*

Proposition. *If $r < m$, then a solution exists only if $c_{r+1}^{(r)} = \cdots = c_n^{(r)} = 0$*

Proposition. *If $r < n$, and a solution exists, then $y_{r+1}, \ldots, y_n$ are undetermined.*

Using Gaussian Elimination, the values for $x_r, \ldots, x_1$ can be determined as $B_{jj}^{(j)} \neq 0$ for $j = 1, \ldots, r$.

Proposition. *If $r = m < n$, then there is no constraint coming from the b_is, so there are infinitely many solutions.*

Proposition. *If $r = n < m$, then if the constraint coming from the b_is is satisfied, there is a unique solution. Otherwise there is no solution.*

In matrices, the new system is $M\mathbf{y} = \mathbf{c}$, where M is a $m \times n$ matrix in echelon form, that is

$$M = \left(\begin{array}{c|c} \hat{M} & \\ \hline 0 & 0 \end{array}\right)$$

where $\hat{M}$ is a $r \times r$ upper-triangular matrix, that is

$$M = \begin{pmatrix} M_1 1 & \ldots & M_{1r} \\ & \ddots & \vdots \\ 0 & & M_{rr} \end{pmatrix}$$

In the process of Gaussian Elimination, reordering variables corresponds to column swaps, reordering equations corresponds to row swaps and subtractions corresponds to row operations.

Then $r = \operatorname{rank} M = \operatorname{rank} A = \#$ linearly independent columns $= \#$ linearly independent rows.

If $n = m$, then $\det A = \pm \det M$, and in addition, if $r = n = m$, then $\det A = \pm M_{11} M_{22} \ldots M_{nn} \neq 0$.

13 Eigenvalues and Eigenvectors

Definition (Eigenvector, Eigenvalue)**.** For a linear map $T : V \to W$, $\mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}$ is an eigenvector if there exists $\lambda \in \mathbb{R}$ such that

$$T(\mathbf{v}) = \lambda \mathbf{v}$$

λ is known as the eigenvalue.

Definition (Characteristic Polynomial)**.** For a matrix A, the characteristic polynomial $\chi_A(t)$ is given by

$$\chi_A(t) = \det(A - tI)$$

Proposition. *λ is an eigenvalue of A if and only if $\chi_A(\lambda) = 0$.*

Proposition. *$\chi_A(t)$ has n roots over $\mathbb{C}$, counted with multiplicity.*

Proposition. $\text{tr}(A) = -$*coefficient of t^{n-1} in $\chi_A(t)$*

Proposition. $\det A =$ *product of eigenvalues* $= \chi_A(0)$

13.1 Multiplicities

Definition (Eigenspace)**.** For an eigenvalue λ of a matrix A, define the eigenspace E_λ by

$$E_\lambda = \{\mathbf{v} : A\mathbf{v} = \lambda\mathbf{v}\}$$

Definition (Geometric Multiplicity)**.** For an eigenvalue λ of a matrix A, define the geometric multiplicty by

$$m_\lambda = \dim E_\lambda = \text{null}(A - \lambda I)$$

Definition (Algebraic Multiplicity)**.** For an eigenvalue λ of a matrix A, define the algebraic multiplicity by

$$M_\lambda = \text{Multiplicity of } (t - \lambda) \text{ in } \chi_A(t)$$

Proposition.

$$M_\lambda \geq m_\lambda$$

Proof.

$$\mathbf{v}_1, \ldots, \mathbf{v}_r$$

where $r = \dim E_\lambda = m_\lambda$. Extend this to a basis for $\mathbb{R}^n$ (or $\mathbb{C}^n$) by adding vectors

$$\mathbf{w}_{r+1}, \ldots, \mathbf{w}_n$$

From now on, we will take the indices with ranges $i, j = 1, \ldots, r$ and $a, b = r+1, \ldots, n$. Then define a matrix P with columns

$$\begin{aligned} \mathbf{C}_i(P) &= \mathbf{v}_i && \text{for } i = 1, \ldots, r \\ \mathbf{C}_a(P) &= \mathbf{w}_a && \text{for } a = r+1, \ldots, n \end{aligned}$$

Then $A\mathbf{C}_i = A\mathbf{v}_i = \lambda\mathbf{C}_i$ and

$$\begin{aligned} A\mathbf{C}_a = A\mathbf{w}_a &= \sum_i B_{ia}\mathbf{v}_i + \sum_b B_{ba}\mathbf{w}_a \\ &= \sum_i B_{ia}\mathbf{C}_i + \sum_b B_{ba}\mathbf{C}_b \end{aligned}$$

where B_{ia}, B_{ba} are arbitrary constants, $\sum_i = \sum_{i=1}^r$, $\sum_b = \sum_{b=r+1}^n$. This means that $AP = PB$, where B is a matrix of the form

- $B_{ij} = \lambda\delta_{ij}$
- $B_{ai} = 0$
- B_{ia}, B_{ab} are unknowns.

Therefore,

$$P^{-1}AP = B = \left(\begin{array}{c|c} \lambda I & \\ \hline 0 & \hat{B} \end{array}\right)$$

where I is the $r \times r$ identity matrix, and $\hat{B}$ is the $(n-r) \times (n-r)$ matrix with entries B_{ab}. Then, as A and B are similar, we have that $\chi_A(t) = \chi_B(t) = \det(B - tI)$, and

$$\det(B - tI) = \left|\begin{array}{c|c} (\lambda - t)I & \\ \hline 0 & \hat{B} - tI \end{array}\right|$$

If we expand the determinant, then as the bottom left part of the matrix is zero, we get that $\det(B - tI) = \det((\lambda - t)I) \det\left(\hat{B} - tI\right) = (\lambda - t)^r \det\left(\hat{B} - tI\right)$. Thus, the algebraic multiplicity is at least $r = m_\lambda = \dim E_\lambda$. □

13.2 Linear Independence

Proposition. *Let $\mathbf{v}_1, \ldots, \mathbf{v}_r$ be eigenvectors of a matrix A, with eigenvalues $\lambda_1, \ldots, \lambda_r$. If the eigenvalues are distinct, say $\lambda_i \neq \lambda_j$ fo $i \neq j$, then the eigenvectors are linearly independent.*

Proof. Let $\mathbf{w} = \sum_{i=1}^{r} \alpha_i v_i$. Then $(A - \lambda I)\mathbf{w} = \sum_{i=1}^{r} (\lambda_i - \lambda)\mathbf{v}_i$. Now suppose if the eigenvectors were linearly dependent. Then there exists a linear combination $\mathbf{w} = \mathbf{0}$. Furthermore, let p represent the number of non-zero α_is. Clearly $p \geq 2$.

Now choose the linear combination $\mathbf{w}$ such that p is least. Without loss of generality, let $\alpha_1 \neq 0$. Then $(A - \lambda_1 I)\mathbf{w} = \sum_{j=2}^{r} \alpha_j (\lambda_j - \lambda_1)\mathbf{v}_j = \mathbf{0}$. But this is a linear combination with $p - 1$ non-zero coefficients. Contradicting the minimality of p. □

Alternative Proof. Fix k, then

$$\left(\prod_{j \neq k} (A - \lambda_j I)\right) \mathbf{w} = \alpha_k \left(\prod_{j \neq k} (\lambda_j - \lambda_k)\right) \mathbf{v}_k = \mathbf{0}$$

as the other $\mathbf{v}_i$s are multiplied by $\lambda_i - \lambda_i = 0$. As the $\mathbf{v}_k$s are non-zero, and $\lambda_j - \lambda_k \neq 0$, we must have that $\alpha_k = 0$, and this holds for all k. □

Proposition. *Let $\mathcal{B}_{\lambda_i}$ be a basis for E_{λ_i}, then*

$$\bigcup_{i=1}^{r} \mathcal{B}_{\lambda_i}$$

is linearly independent.

Proof. Consider $\mathbf{w} = \mathbf{w}_1 + \cdots + \mathbf{w}_r$, where $\mathbf{w}_i \in E_{\lambda_i}$. By a similar argument to the previous proposition, $\mathbf{w}_i = 0$. Then we must have that the coefficients for each of the basis vectors is zero. □

14 Diagonalisation

Proposition. *For an $n \times n$ matrix A, the following are equivalent.*

- *There exists a basis of eigenvectors* $\mathbf{v}_1, \dots, \mathbf{v}_n$
- *There exists a matrix P with $P^{-1}AP = D$, where* $D = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$

Proof. Let the columns of P be $\mathbf{C}_i$. Letting $\mathbf{C}_i = \mathbf{v}_i$, we get that

$$P^{-1}AP = D \iff AP = PD \iff \forall i, \lambda_i \mathbf{v}_i = \lambda_i \mathbf{v}_i$$

□

Definition (Diagonalisable)**.** A matrix M is diagonalisable if either of the above conditions holds.

Proposition. *For an $n \times n$ matrix M, M is diagonalisable if it has n distinct eigenvalues.*

Proof. If we have n distinct eigenvalues, we must have n distinct linearly independent eigenvectors, which forms a basis. □

Proposition. *A is diagonalisable if and only if for each eigenvalue λ_i, $M_{\lambda_i} = m_{\lambda_i}$.*

Proof. Clearly $\sum_i M_{\lambda_i} = n$. Now consider $\bigcup_i B_{\lambda_i}$. We have shown that this is linearly independent, and the size is $\sum_i m_{\lambda_i} = n$. So it is a basis of eigenvectors. □

15 Similar Matrices

Definition (Similar)**.** Matrices A and B are similar if $B = P^{-1}AP$ for some matrix P>

Proposition. *If A and B are similar, then*

- $\operatorname{tr} A = \operatorname{tr} B$
- $\det A = \det B$
- $\chi_A(t) = \chi_B(t)$

Proposition. *Similarity is an equivalence relation.*

Remark. Similarity is also called conjugate, especially in Groups. Note that two matrices are similar if they represent the same linear map but with respect to different bases.

16 Hermitian and Symmetric Matrices

Theorem. *If a matrix A is hermitian, then every eigenvalue is real, and eigenvectors with distinct eigenvalues are orthogonal.*

Proof. As A is hermitian,

$$\mathbf{v}^\dagger(A\mathbf{v}) = (A\mathbf{v})^\dagger \mathbf{v} \implies \mathbf{v}^\dagger(\lambda \mathbf{v}) = (\lambda \mathbf{v}^\dagger)\mathbf{v} \implies \bar{\lambda}\mathbf{v}^\dagger\mathbf{v} = \lambda \mathbf{v}^\dagger \mathbf{v} \implies \bar{\lambda} = \lambda \implies \lambda \in \mathbb{R}$$

Furthermore,

$$\mathbf{v}^\dagger(A\mathbf{w}) = (A\mathbf{v})^\dagger \mathbf{w} \implies \mathbf{v}^\dagger(\mu \mathbf{w}) = (\lambda \mathbf{v})^\dagger \mathbf{w} \implies \mu \mathbf{v}^\dagger \mathbf{w} = \lambda \mathbf{v}^\dagger \mathbf{w} \implies \mathbf{v}^\dagger \mathbf{w} = 0$$

□

Proposition. *For a real symmetric matrix A acting on $\mathbb{C}^n$, for each eigenvalue λ there is a real eigenvector* $\mathbf{v}$.

Proof. Given an eigenvector $\mathbf{v} \in \mathbb{C}^n$, let $\mathbf{v} = \mathbf{u} + i\mathbf{w}$, where $\mathbf{u}, \mathbf{w} \in \mathbb{R}^n$. Then if $A\mathbf{v} = \lambda\mathbf{v}$, we must have that $A\mathbf{u} = \lambda\mathbf{u}$ and $A\mathbf{w} = \lambda\mathbf{w}$, as A and λ are real. In addition, as $\mathbf{v} \neq \mathbf{0}$, at least one of $\mathbf{u}$ and $\mathbf{w}$ must be non-zero. □

17 Unitary and Orthogonal Diagonalisation

Theorem. *Any Hermitian matrix can be diagonalised. In addition, the eigenvectors $\mathbf{u}_i$ can be chosen to be orthonormal, so $\mathbf{u}_i^\dagger \mathbf{u}_j = \delta_{ij}$. Consequently P can be chosen to be unitary.*

Proof. Consider $A : \mathbb{C}^n \to \mathbb{C}^n$, and a subspace V such that $A(V) \subseteq V$, and we can consider $A : V \to V$. We will show that for $m = \dim V \leq n$, V has an orthonormal basis of eigenvectors. The case for $n = 1$ is immediate.

Now given V, v be an eigenvector with eigenvalue λ. Let W be the $m-1$ dimensional subspace $W = \{\mathbf{w} \in V : \mathbf{v}^\dagger \mathbf{w} = 0\}$. $v^\dagger(A\mathbf{w}) = (A\mathbf{v})^\dagger \mathbf{w} = \lambda \mathbf{v}^\dagger \mathbf{w} = 0$, so $A\mathbf{w} \in W$. Hence by our inductive hypothesis A has an orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_{m-1}\}$ formed by eigenvectors of A. Letting $\mathbf{u}_m = \frac{1}{||\mathbf{v}||}\mathbf{v}$, $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ is an orthonormal basis of eigenvectors of A. For $m = n$, we get a basis of eigenvectors, and result follows. □

Proposition. *Any real symmetric matrix can be diagonalied, and the matrix P can be chosen to be orthogonal.*

Proof. Immediate from Theorem above. □

18 Quadratic Forms

Definition (Quadratic Form)**.** A quadratic form $\mathcal{F} : \mathbb{R}^n \to \mathbb{R}$, is a function given by

$$\mathcal{F}(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$$

where A is a real symmetric matrix.

Remark. Strictly speaking A does not have to be symmetric, but any asymmetric part will not contribute.

Definition (Diagonalisation)**.** The matrix A can be diagonalised, where $P^T A P = D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$

Then if we let $\mathbf{x}' = P^T \mathbf{x}$, or equivalently $\mathbf{x} = P\mathbf{x}'$, then

$$\mathcal{F}(\mathbf{x}) = \sum_i \lambda_i (x_i')^2$$

Definition (Principal Axes)**.** The axes given by the columns of P are known as the principal axes of the quadratic form.

19 Cayley-Hamilton

Theorem (Cayley-Hamilton)**.** *If we regard the characteristic polynomial as a formal polynomial, then $\chi_A(A) = 0$.*

Proof for 2×2 matrices. By calculation. □

Proof for Diagonalisable Matrices. Clearly $\chi_A(D) = \begin{pmatrix} \chi_A(\lambda_1) & & \\ & \ddots & \\ & & \chi_A(\lambda_n) \end{pmatrix}$ □

Proof for General Matrices. Let $M = A - tI$. Then $\det M = \det(A - tI) = \chi_A(t)$. The adjugate, $\text{Adj}(M) = \tilde{M}$ can be written as a sum $\tilde{M} = \sum_{i=0}^{n-1} B_i t^i$, where B_i are matrix coefficients.

From previously, we know that $\tilde{M}M = (\det M)I$. As M is also a $n \times n$ matrix, $\det M$ is a degree n polynomial. Say $\det M = \sum_{i=0}^{n} c_i t^i$.

$$\tilde{M}M = \left(\sum_{i=0}^{n-1} B_i t^i\right)(A - tI) = B_0 + (B_1 A - B_0)t + \cdots + (B_{n-1}A + B_{n-2})t^{n-1} - B_{n-1}t^n$$

Now if we compare the coefficients of t^r in $\det M$ and $\tilde{M}M$, we get that

$$\begin{aligned} c_0 I &= B_0 A \\ c_1 I &= B_1 A - B_0 \\ &\vdots \\ c_{n-1} I &= B_{n-1} A - B_{n-2} \\ c_n I &= -B_{n-1} \end{aligned}$$

Then $\chi_A(A) = \det M$ with $t = A$, and this is

$$\begin{aligned} c_0 I + c_1 A + & c_2 A^2 + \cdots + c_{n-1} A^{n-1} + c_n A^n \\ &= B_0 A + (B_1 A^2 - B_0 A) + (B_2 A^3 - B_1 A^2) + \cdots + (B_{n-1} A^n - B_{n-2} A^{n-1}) - B_{n-1} A^n \\ &= 0 \end{aligned}$$

□

20 Changing Bases

Definition (Change of Base Matrix). If $\{\mathbf{e}_i\}$ and $\{\mathbf{e}'_i\}$ are bases, then the change of base matrix P is defined to satisfy

$$\mathbf{e}'_i = P_{ji}\mathbf{e}_j$$

ie the i-th column is $\mathbf{e}'_i$ with respect to the $\{\mathbf{e}_j\}$.

Proposition. *Suppose if $T : V \to W$ is a linear map, $\{\mathbf{e}_i\}$ and $\{\mathbf{e}'_i\}$ are bases for V, $\{\mathbf{f}_a\}$ and $\{\mathbf{f}'_a\}$ are bases for W, and A is the matrix representing T with respect to $\{\mathbf{e}_i\}$ and $\{\mathbf{f}_a\}$. Letting the change of base matrices be P and Q, we get that*

$$A' = Q^{-1}AP$$

Proof. $T(\mathbf{e}'_i) = T(\sum_j P_{ji}\mathbf{e}_j) = \sum_j P_{ji}T(\mathbf{e}_j) = \sum_j \sum_a A_{aj}P_{ji}\mathbf{f}_a$.

But also, $T(\mathbf{e}'_i) = \sum_b A'_{bi}\mathbf{f}'_b = \sum_a \sum_b Q_{ab}A'_{bi}\mathbf{f}_a$.

If we equate the coefficients on both sides, we get that

$$\sum_j A_{aj}P_{ji} = \sum_b Q_{ab}A'_{bi}$$

which means $AP = QA'$, consequently $A' = Q^{-1}AP$. □

Proposition (Change in Components). *If the vector* X *can be written as $x_i\mathbf{e}_i$ and $x'_i\mathbf{e}'_i$, then*

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

21 Jordan Normal Form

Proposition. *Any 2×2 complex matrix A is similar to one of the following*

(i) $A' = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, *where* $\lambda_1 \neq \lambda_2$

(ii) $A' = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$

(iii) $A' = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$

Proof. $\chi_A(t)$ has 2 roots over $\mathbb{C}$. If the two roots are distinct, then we have distinct eigenvalues λ_1, λ_2, and we will have $M_{\lambda_1} = m_{\lambda_1} = M_{\lambda_2} = m_{\lambda_2} = 1$. Then, the two eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ forms a basis. Hence $A' = P^{-1}AP$ with the eigenvectors being the columns of P, and A' will be (i).

For a repeated root λ, where $M_\lambda = m_\lambda = 2$, the above argument applies, and A' will be in the form in (ii).

For a repeated eigenvalue λ, with $M_\lambda = 2$ and $m_\lambda = 1$, let v be an eigenvector, and thet $\mathbf{w}$ be any vector such that $\{\mathbf{v}, \mathbf{w}\}$ is linearly independent.

Then $A\mathbf{v} = \lambda\mathbf{v}$, and $A\mathbf{w} = \alpha\mathbf{v} + \beta\mathbf{w}$. Therefore the matrix of the map with respect to the basis $\{\mathbf{v}, \mathbf{w}\}$ would $\begin{pmatrix} \lambda & \alpha \\ 0 & \beta \end{pmatrix}$, and then by looking at $\chi_A(t)$, we must have that $\lambda = \beta$, otherwise $\chi_A(t)$ has two distinct roots. We must also have that $\alpha \neq 0$, otherwise $M_\lambda = m_\lambda = 2$.

Now set $\mathbf{u} = \alpha\mathbf{v}$, and note that $A(\alpha\mathbf{v}) = \lambda\alpha\mathbf{v}$, and $A\mathbf{w} = \alpha\mathbf{v} + \lambda\mathbf{w}$. Therefore, with respect to the basis $\{\mathbf{v}, \mathbf{w}\}$, we get that $A' = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, and the matrix is similar to (iii). □

Alternative approach for (iii). If A has characteristic polynomial $\chi_A(t) = (t-\lambda)^2$, but $A \neq \lambda I$, then there exists $\mathbf{w}$ such that $\mathbf{u} = (A-\lambda I)\mathbf{w} \neq \mathbf{0}$. However, $(A-\lambda I)\mathbf{u} = (A-\lambda I)^2\mathbf{w} = 0\mathbf{w} = \mathbf{0}$, by the Cayley-Hamilton theorem. Thus, we have that $A\mathbf{u} = \lambda u$ and $A\mathbf{w} = u + \lambda\mathbf{w}$, and with basis $\{\mathbf{u}, \mathbf{w}\}$, we get the matrix $A' = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$. □

Definition (Jordan Block)**.** For a Jordan Block of size n and parameter λ, let N be a $n \times n$ matrix such that

$$N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}$$

where $N_{ij} = \begin{cases} 1 & \text{if } j = i+1 \\ 0 & \text{otherwise} \end{cases}$

Then $J_n(\lambda) = \lambda I + N = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix}$

Proposition. *The standard basis of $\mathbb{C}^n$, $\{\mathbf{e}_i\}$, under the transformation of N is mapped*

$$\mathbf{e}_n \mapsto \mathbf{e}_{n-1} \mapsto \dots \mapsto \mathbf{e}_1 \mapsto \mathbf{0}$$

Consequently, $N^n = 0$, and N is nilpotent.

Proposition. *$\chi_{J_n(\lambda)}(t) = (\lambda - t)^n$, and $M_\lambda = n$, $m_\lambda = 1$ as $\ker(N) = \text{span}\{\mathbf{e}_1\}$.*

Theorem (Jordan Normal Form)**.** *Any $n \times n$ complex matrix A is similar to a matrix of the following form*

$$A' = \begin{pmatrix} J_{n_1}(\lambda_1) & & & \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{n_r}(\lambda_r) \end{pmatrix}$$

where each block is a Jordan Block, $\lambda_1, \dots, \lambda_r$ are eigenvalues of A and A', and the same eigenvalue may appear in different blocks.

Proof. See IB Linear Algebra and IB Groups, Rings and Modules. □

Proposition. *Clearly $n_1 + \dots + n_r = n$*

Proposition. *A is diagonalisable if and only if A' consists of only 1×1 blocks.*

22 Quadrics

Definition (Quadric)**.** A quadric in $\mathbb{R}^n$ is a hypersurface defined by

$$Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} + \mathbf{b}^T\mathbf{x} + c = 0$$

For A non-zero, real and symmetric, $\mathbf{b} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

Definition (Completing the square)**.** Let $\mathbf{y} = \mathbf{x} + \frac{1}{2}A^{-1}\mathbf{b}$, we get that

$$\mathbf{y}^T A\mathbf{y} = \mathbf{x}^T A\mathbf{x} + b^T\mathbf{x} + \frac{1}{4}\mathbf{b}^T A^{-1}\mathbf{b}$$

Hence $Q(\mathbf{x}) = 0$ if and only if $\mathcal{F}(\mathbf{y}) = k$, where $\mathcal{F}(\mathbf{y}) = \mathbf{y}^T A\mathbf{y}$, $k = \frac{1}{4}\mathbf{b}^T A^{-1}\mathbf{b} - c$.

Hence quadrics are (up to affine isomatries) equivalent to a quadratic form.

22.1 Conics

Definition (Conic)**.** A quadric in $\mathbb{R}^2$ is also known as a conic.

Proposition. *By completing the square and diagonalising A, we get that $\lambda_1(x_1')^2 + \lambda_2(x_2')^2 = k$. The solutions are as follows*

- *If $\lambda_1, \lambda_2 > 0$, then*
  - *If $k > 0$ we have an ellipse*
  - *If $k = 0$ we have a point*
  - *If $k < 0$ there is no solution.*
- *If $\lambda_1 > 0, \lambda_2 < 0$ then*
  - *If $k > 0$ or $k < 0$ we have a hyperbola*
  - *If $k = 0$ we have a pair of lines*

Proposition (Cartesian Forms of Conics)**.**

- *Ellipse* - $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$
- *Parabola* - $y^2 = 4ax$
- *Hyperbola* - $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$

Definition (Eccentricity of Conics).

- Ellipse - $b^2 = a^2(1 - e^2)$ and $0 < e < 1$
- Parabola - $e = 1$
- Hyperbola $b^2 = a^2(e^2 - 1)$, $1 < e$

Definition (Foci of Conics).

- Ellipse - $(\pm ae, 0)$
- Parabola - $(a, 0)$
- Hyperbola - $(\pm ae, 0)$

Definition (Polar Form of Conics). Let $r = \frac{l}{1 + e\cos\theta}$, then

- Ellipse - $l = a(1 - e^2)$
- Parabola - $l = 2a$
- Hyperbola - $l = a(e^2 - 1)$

Definition (Double Cone). A double cone in $\mathbb{R}^3$ is given by

$$((\mathbf{x} - \mathbf{c}) \cdot \mathbf{n})^2 = ||\mathbf{x} - \mathbf{c}||^2 \cos^2 \alpha$$

where $\mathbf{c}$ is the point where the cones meet, $\mathbf{n}$ is a unit vector representing the axis of the cone, and α is the angle of the cone (angle from one side to $\mathbf{n}$).

Proposition. *Conics are the intersection of double cones with planes.*

23 Minkowski Space

For this section, also see Dynamics and Relativity.

Definition (Minkowski Metric). For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T J \mathbf{y}$$

where $J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

Proposition.

$$(\forall \mathbf{x}, \mathbf{y}, \langle M\mathbf{x}, M\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle) \iff M^T J M = J$$

Definition (Lorentz group). The Lorentz group contains the matrices which preserve the Minkowski metric and have determinant 1.

Proposition. *General form of a matrix in the Lorentz group is* $M = \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix}$

Proposition.

$$M(\theta_1 + \theta_2) = M(\theta_1)M(\theta_2)$$

23.1 Special Relativity

In this subsection, we will use units where the speed of light is $c = 1$.

Definition (Lorentz Factor)**.** The Lorentz factor for a velocity v is $\gamma(v) = (1 - v^2)^{1/2}$

Proposition.

$$M(\theta) = \gamma(v) \begin{pmatrix} 1 & v \\ v & 1 \end{pmatrix}$$

where $v = \tanh\theta$

Proposition. *If S' is moving with velocity v_1 with respect to S, and S'' is moving with velocity v_2 with respect to S', then S'' is moving with velocity*

$$\frac{v_1 + v_2}{1 + v_1 v_2}$$

as observed in S.

Proof. Using $v = \tanh\theta$ and $M(\theta_1 + \theta_2) = M(\theta_1)M(\theta_2)$. □