Part IA - Vectors and Matrices

Definitions

Lectured by N. Peake

Michaelmas 2014

Complex numbers
Review of complex numbers, including complex conjugate, inverse, modulus, argument and Argand diagram. Informal treatment of complex logarithm, n-th roots and complex powers. de Moivre's theorem. [2]

Vectors
Review of elementary algebra of vectors in $\mathbb{R}^3$, including scalar product. Brief discussion of vectors in $\mathbb{R}^n$ and $\mathbb{C}^n$; scalar product and the CauchySchwarz inequality. Concepts of linear span, linear independence, subspaces, basis and dimension.

Suffix notation: including summation convention, δ_{ij} and ϵ_{ijk}. Vector product and triple product: definition and geometrical interpretation. Solution of linear vector equations. Applications of vectors to geometry, including equations of lines, planes and spheres. [5]

Matrices
Elementary algebra of 3×3 matrices, including determinants. Extension to $n \times n$ complex matrices. Trace, determinant, non-singular matrices and inverses. Matrices as linear transformations; examples of geometrical actions including rotations, reflections, dilations, shears; kernel and image. [4]

Simultaneous linear equations: matrix formulation; existence and uniqueness of solutions, geometric interpretation; Gaussian elimination. [3]

Symmetric, anti-symmetric, orthogonal, hermitian and unitary matrices. Decomposition of a general matrix into isotropic, symmetric trace-free and antisymmetric parts. [1]

Eigenvalues and Eigenvectors
Eigenvalues and eigenvectors; geometric significance. [2]

Proof that eigenvalues of hermitian matrix are real, and that distinct eigenvalues give an orthogonal basis of eigenvectors. The effect of a general change of basis (similarity transformations). Diagonalization of general matrices: sufficient conditions; examples of matrices that cannot be diagonalized. Canonical forms for 2×2 matrices. [5]

Discussion of quadratic forms, including change of basis. Classification of conics, cartesian and polar forms. [1]

Rotation matrices and Lorentz transformations as transformation groups. [1]

Contents

1 Complex numbers

1.1 Basic properties

Definition (Complex number)**.** A *complex number* is a number $z \in \mathbb{C}$ of the form $z = a + ib$ with $a, b \in \mathbb{R}$, where $i^2 = -1$. We write $a = \mathrm{Re}(z)$ and $b = \mathrm{Im}(z)$.

Definition (Complex conjugate)**.** The *complex conjugate* of $z = a + ib$ is $\bar{z} = z^* = a - ib$.

Definition (Argand diagram)**.** An *Argand diagram* is a diagram in which a complex number $z = x + iy$ is represented as by a vector $\mathbf{p} = \begin{pmatrix} x \\ y \end{pmatrix}$. Addition of vectors corresponds to vector addition and $\bar{z}$ is the reflection of z in the x-axis.

Definition (Modulus and argument of complex number)**.** The *modulus* of $z = x + iy$ is $r = |z| = \sqrt{x^2 + y^2}$. The *argument* is $\theta = \arg z = \tan^{-1}(y/x)$. The modulus is the length of the vector in the Argand diagram, and the argument is the angle between z and the real axis. We have

$$z = r(\cos\theta + i\sin\theta)$$

Clearly the pair (r, θ) uniquely describes a complex number z, but each complex number $z \in \mathbb{C}$ can be described by many different θ since $\sin(2\pi + \theta) = \sin\theta$ and $\cos(2\pi + \theta) = \cos\theta$. Often we take the *principle value* $-\pi < \theta \leq \pi$.

1.2 Complex exponential function

Definition (Exponential function)**.** The *exponential function* is defined as

$$\exp(z) = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

Definition (Sine and cosine functions)**.** Define, for all $z \in \mathbb{C}$,

$$\begin{aligned} \sin z &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} &&= z - \frac{1}{3!}z^3 + \frac{1}{5!}z^5 + \cdots \\ \cos z &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} &&= 1 - \frac{1}{2!}z^2 + \frac{1}{4!}z^4 + \cdots \end{aligned}$$

1.3 Roots of unity

Definition (Roots of unity)**.** The n-th *roots of unity* are the roots to the solution $z^n = 1$ for $n \in \mathbb{N}$. Since this is a polynomial of order n, there are n roots of unity. The n-th roots of unity are $\exp\left(2\pi i \frac{k}{n}\right)$ for $k = 0, 1, 2, 3 \cdots n - 1$.

1.4 Complex logarithm and power

Definition (Complex logarithm)**.** The *complex logarithm* $\log z$ is a solution to $e^{\omega} = z$. i.e. $\omega = \log z$. Writing $z = re^{i\theta}$, we have $\log z = \log(re^{i\theta}) = \log r + i\theta$. This can multi-valued for different values of θ and, as above, we should select the θ that satisfies $-\pi < \theta \leq \pi$.

Definition (Complex power)**.** The *complex power* z^α for $z, \alpha \in \mathbb{C}$ is defined as $z^\alpha = e^{\alpha \log z}$. This complex power can be multi-valued, as $z^\alpha = e^{\alpha \log |z|} e^{i\alpha\theta} e^{2in\pi\alpha}$ (there are finitely many values if $\alpha \in \mathbb{Q}$, infinitely many otherwise). Nevertheless, make z^α single-valued by insisting $-\pi < \theta \leq \pi$.

1.5 De Moivre's theorem

1.6 Lines and circles in $\mathbb{C}$

2 Vectors

2.1 Definition and basic properties

Definition (Vector)**.** A *vector* $\mathbf{v}$ has a (positive) length and direction. If $|\mathbf{v}| = 0$, then $\mathbf{v} = \mathbf{0}$.

Vectors can be added together or multiplied by a scalar in $\mathbb{R}$ or $\mathbb{C}$. Vector addition satisfies the following axioms:

(i) (Commutativity) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$

(ii) (Associativity) $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$

(iii) (Identity) There is a vector $\mathbf{0}$ such that $\mathbf{a} + \mathbf{0} = \mathbf{a}$.

(iv) (Inverse) For all vectors $\mathbf{a}$, there is a vector $(-\mathbf{a})$ such that $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$.

Scalar multiplication satisfies the following axioms:

(i) $\lambda\mathbf{a}$ is either parallel ($\lambda > 0$) to or anti-parallel ($\lambda < 0$) to $\mathbf{a}$.

(ii) $\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}$.

(iii) $(\lambda + \mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a}$.

(iv) $\lambda(\mu\mathbf{a}) = \mu(\lambda\mathbf{a}) = (\mu\lambda)\mathbf{a}$.

(v) $1\mathbf{a} = \mathbf{a}$.

Definition (Unit vector)**.** A *unit vector* is a vector with length 1. We write a unit vector as $\hat{\mathbf{v}}$.

2.2 Scalar product

2.2.1 Geometric picture ($\mathbb{R}^2$ and $\mathbb{R}^3$ only)

Definition (Scalar/dot product)**.** $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta$. It satisfies the following properties:

(i) $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$

(ii) $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2 \geq 0$

(iii) $\mathbf{a} \cdot \mathbf{a} = 0$ iff $\mathbf{a} = \mathbf{0}$

(iv) If $\mathbf{a} \cdot \mathbf{b} = 0$ and $\mathbf{a}, \mathbf{b} \neq 0$, then $\mathbf{a} \perp \mathbf{b}$.

2.2.2 General algebraic definition

Definition (Inner/scalar product)**.** In a real vector space V, the *inner product* or *scalar product* is a map $V \times V \to \mathbb{R}$ that satisfies the following axioms. It is written as $\mathbf{x} \cdot \mathbf{y}$ or $\langle \mathbf{x} | \mathbf{y} \rangle$ (Dirac bra-ket notation).

(i) (Symmetry) $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$

(ii) (Linearity in 2nd argument) $\mathbf{x} \cdot (\lambda\mathbf{y} + \mu\mathbf{z}) = \lambda\mathbf{x} \cdot \mathbf{y} + \mu\mathbf{x} \cdot \mathbf{z}$

(iii) (Positive definite) $\mathbf{x} \cdot \mathbf{x} \geq 0$ with equality iff $\mathbf{x} = \mathbf{0}$.

Definition. The *norm* of a vector, written as $|\mathbf{a}|$ or $||\mathbf{a}||$, is defined as

$$|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}}$$

2.3 Cauchy-Schwarz inequality

2.4 Vector product

Definition (Vector/cross product)**.** Consider $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$. Define the *vector product*

$$\mathbf{a} \times \mathbf{b} = |\mathbf{a}||\mathbf{b}| \sin\theta \hat{\mathbf{n}},$$

where $\hat{\mathbf{n}} \perp \mathbf{a}, \mathbf{b}$ in a right-handed sense. The vector product satisfies the following properties:

(i) $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$.

(ii) $\mathbf{a} \times \mathbf{a} = \mathbf{0}$.

(iii) $\mathbf{a} \times \mathbf{b} = \mathbf{0} \Rightarrow \mathbf{a} = \lambda \mathbf{b}$ for some $\lambda \in \mathbb{R}$.

(iv) $\mathbf{a} \times (\lambda \mathbf{b}) = \lambda(\mathbf{a} \times \mathbf{b})$.

(v) $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$.

2.5 Scalar triple product

Definition (Scalar triple product)**.** The *scalar triple product* is defined as

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}).$$

We have $[\mathbf{a}, \mathbf{b}, \mathbf{c}] = [\mathbf{b}, \mathbf{c}, \mathbf{a}] = [\mathbf{c}, \mathbf{a}, \mathbf{b}] = -[\mathbf{b}, \mathbf{a}, \mathbf{c}] = -[\mathbf{a}, \mathbf{c}, \mathbf{b}] = -[\mathbf{c}, \mathbf{b}, \mathbf{a}]$.

2.6 Spanning sets and bases

2.6.1 2D space

Definition (Spanning set)**.** A set of vectors $\{\mathbf{a}, \mathbf{b}\}$ *spans* $\mathbb{R}^2$ if for all vectors $\mathbf{r} \in \mathbb{R}^2$, there exists some $\lambda, \mu \in \mathbb{R}$ such that $\mathbf{r} = \lambda \mathbf{a} + \mu \mathbf{b}$. In two vectors span $\mathbb{R}^2$ if $\mathbf{a} \times \mathbf{b} \neq 0$.

Definition (Linearly independent vectors in $\mathbb{R}^2$)**.** Two vectors $\mathbf{a}$ and $\mathbf{b}$ are *linearly independent* if for $\alpha, \beta \in \mathbb{R}$, $\alpha \mathbf{a} + \beta \mathbf{b}$ iff $\alpha = \beta = 0$. In $\mathbb{R}^2$, $\mathbf{a}$ and $\mathbf{b}$ are linearly independent if $\mathbf{a} \times \mathbf{b} \neq 0$.

Definition (Basis of $\mathbb{R}^2$)**.** A set of vectors is a *basis* of $\mathbb{R}^2$ if it spans $\mathbb{R}^2$ and are linearly independent.

2.6.2 3D space

2.6.3 $\mathbb{R}^n$ space

Definition (Linearly independent vectors)**.** A set of vectors $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3} \cdots \mathbf{v_m}\}$ are *linearly independent* if

$$\sum_{i=1}^{m} \lambda_i \mathbf{v}_i = \mathbf{0} \Rightarrow \forall i(\lambda_i = 0).$$

Definition (Spanning set). A set of vectors $\{\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3} \cdots \mathbf{u_m}\} \subseteq \mathbb{R}^n$ is a *spanning set* of $\mathbb{R}^n$ if

$$\forall \mathbf{x} \in \mathbb{R}^n \left(\exists \lambda_i \left(\sum_{i=1}^{m} \lambda_i \mathbf{u}_i = \mathbf{x} \right) \right)$$

Definition (Basis vectors). A *basis* of $\mathbb{R}^n$ is a linearly independent spanning set. The standard basis of $\mathbb{R}^n$ is $\mathbf{e}_1 = (1, 0, 0, \cdots 0), \mathbf{e}_2 = (0, 1, 0, \cdots 0) \cdots \mathbf{e}_n = (0, 0, 0, \cdots, 1)$.

Definition (Orthonormal basis). A basis $\{\mathbf{e}_i\}$ is *orthonormal* if $\mathbf{e}_i \cdot \mathbf{e_j} = 0$ if $i \neq j$ and $\mathbf{e}_i \cdot \mathbf{e_i} = 1$ for all i, j. (alternatively, $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$, c.f. Kronecker Delta)

Definition (Dimension of vector space). The *dimension* of a vector space is the number of vectors in its basis.

Definition (Scalar product). The *scalar product* of $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ is defined as $\mathbf{x} \cdot \mathbf{y} = \sum \mathbf{x}_i \mathbf{y}_i$.

2.6.4 $\mathbb{C}^n$ space

Definition ($\mathbb{C}^n$). $\mathbb{C}^n = \{(z_1, z_2, \cdots z_n) : z_i \in \mathbb{C}\}$. It has the same standard basis as $\mathbb{R}^n$ but the scalar product is defined differently. For $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$, $\mathbf{u}, \mathbf{v} = \sum \bar{\mathbf{u}}_i \mathbf{v}_i$. The scalar product has the following properties:

(i) $\mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{u}}$

(ii) $\mathbf{u} \cdot (\lambda \mathbf{v} + \mu \mathbf{w}) = \lambda(\mathbf{u} \cdot \mathbf{v}) + \mu(\mathbf{u} \cdot \mathbf{w})$

(iii) $\mathbf{u} \cdot \mathbf{u} \geq 0$ and $\mathbf{u} \cdot \mathbf{u} = 0$ iff $\mathbf{u} = \mathbf{0}$

2.7 Vector subspaces

Definition (Vector subspace). A *vector subspace* of a vector space V is a subset of V that is also a vector space under the same operations. Both V and $\{0\}$ are subspaces of V. All others are proper subspaces.

A subset $U \subseteq V$ is a subspace iff

(i) $\mathbf{x}, \mathbf{y} \in U \Rightarrow (\mathbf{x} + \mathbf{y}) \in U$.

(ii) $\mathbf{x} \in U \Rightarrow \lambda \mathbf{x} \in U$ for all scalars λ.

(iii) $\mathbf{0} \in U$.

This can be simply written as U is non-empty and for all $\mathbf{x}, \mathbf{y} \in U$, $(\lambda \mathbf{x} + \mu \mathbf{y}) \in U$.

2.8 Suffix notation

Definition (Kronecker delta). δ_{ij} (2 free suffices i and j, i.e. 2nd rank tensor).

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

We have

$$\begin{pmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{I}$$

Definition (Alternating symbol ϵ_{ijk})**.** Consider rearrangements of $1, 2, 3$. We can divide them into even and odd permutations. Even permutations include $(1, 2, 3)$, $(2, 3, 1)$ and $(3, 1, 2)$. These are permutations obtained by performing two (or no) swaps of the elements of $(1, 2, 3)$. (Alternatively, it is any "rotation" of $(1, 2, 3)$)

The odd permutations are $(2, 1, 3)$, $(1, 3, 2)$ and $(3, 2, 1)$. They are the permutations obtained by one swap only.

Define

$$\epsilon_{ijk} = \begin{cases} +1 & ijk \text{ is even permutation} \\ -1 & ijk \text{ is odd permutation} \\ 0 & \text{otherwise (i.e. repeated suffices)} \end{cases}$$

ϵ_{ijk} has 3 free suffices We have $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = +1$ and $\epsilon_{213} = \epsilon_{132} = \epsilon_{321} = -1$. $\epsilon_{112} = \epsilon_{111} = \cdots = 0$.

2.8.1 Spherical trigonometry

2.9 Geometry

2.9.1 Lines

2.9.2 Plane

2.10 Vector equations

3 Linear maps

3.1 Examples

3.1.1 Rotation in $\mathbb{R}^3$

3.1.2 Reflection in $\mathbb{R}^3$

3.2 Linear Maps

Definition (Domain, codomain and image of map)**.** Consider sets A and B and mapping $T : A \to B$ such that $x \in A$ is mapped into a unique $x' = T(x) \in B$. A is the *domain* of T and B is the *co-domain* of T. Typically, we have $T : \mathbb{R}^n \to \mathbb{R}^m$ or $T : \mathbb{C}^n \to \mathbb{C}^m$.

Definition (Linear map)**.** Let V, W be real (or complex) vector spaces, and $T : V \to W$. Then T is a *linear map* if

(i) $T(\mathbf{a} + \mathbf{b}) = T(\mathbf{a}) + T(\mathbf{b})$ for all $\mathbf{a}, \mathbf{b} \in V$.

(ii) $T(\lambda\mathbf{a}) = \lambda T(\mathbf{A})$ for all $\lambda \in \mathbb{R}$ (or $\mathbb{C}$).

Equivalently, we have $T(\lambda\mathbf{a} + \mu\mathbf{b}) = \lambda T(\mathbf{a}) + \mu T(\mathbf{b})$.

Definition (Image and kernel of map)**.** The *image* of a map $f : U \to V$ is the subset of V $\{f(\mathbf{u}) : \mathbf{u} \in U\}$. The *kernel* is the subset of U $\{\mathbf{u} \in U : f(\mathbf{u}) = \mathbf{0}\}$.

3.3 Rank and nullity

Definition (Rank of linear map)**.** The *rank* of a linear map $f : U \to V$, denoted as $r(f)$, is the dimension of the image of f.

Definition (Nullity of linear map)**.** The *nullity* of f, denoted $n(f)$ is the dimension of the kernel of f.

3.4 Matrices

Definition (Matrix)**.** Consider a general linear map $\alpha : \mathbb{R}^n \to \mathbb{R}^m$, writing $\mathbf{x}' = \alpha(\mathbf{x})$ with $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{x}' \in \mathbb{R}^m$. Write $\mathbf{x} = x_j\mathbf{e}_j$, where $\mathbf{e}_i$ a is basis of $\mathbb{R}^n$. Then $\mathbf{x}' = \alpha(x_j\mathbf{e}_j) = x_j\alpha(\mathbf{e}_j)$. So $x_i' = [\alpha(\mathbf{e}_j)]_i x_j$. Write $x_i' = A_{ij}x_j$, with $A_{ij} = [\alpha(\mathbf{e}_j)]_i$. A is the *matrix* for α. We write

$$A = \{A_{ij}\} = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & A_{ij} & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix}$$

We have A_{ij} is in the ith row and jth column. A is an $m \times n$ matrix. We can write $\mathbf{x}' = A\mathbf{x}$.

3.4.1 Examples

3.4.2 Matrix Algebra

Definition (Addition of matrices)**.** Consider two linear maps $\alpha, \beta : \mathbb{R}^n \to \mathbb{R}^m$. The sum of α and β is defined by

$$\begin{aligned}(\alpha + \beta)(\mathbf{x}) &= \alpha(\mathbf{x}) + \beta(\mathbf{x})\\ (A+B)_{ij}x_j &= A_{ij}x_j + B_{ij}x_j\\ (A+B)_{ij} &= A_{ij}B_{ij}\end{aligned}$$

Definition (Scalar multiplication of matrices)**.** Define $(\lambda\alpha)\mathbf{x} = \lambda[\alpha(\mathbf{x})]$. So $(\lambda A)_{ij} = \lambda A_{ij}$.

Definition (Matrix multiplication)**.** Consider maps $\alpha : \mathbb{R}^\ell \to \mathbb{R}^n$ and $\beta : \mathbb{R}^n \to \mathbb{R}^m$. The composition is $\beta\alpha : \mathbb{R}^\ell \to \mathbb{R}^m$. Take $\mathbf{x} \in \mathbb{R}^\ell \mapsto \mathbf{x}'' \in \mathbb{R}^m$. Then $\mathbf{x}'' = (BA)\mathbf{x} = B\mathbf{x}'$, where $\mathbf{x}' = A\mathbf{x}$. Using suffix notation, we have $x_i'' = (B\mathbf{x}')_i = b_{ik}x_k' = B_{ik}A_{kj}x_j$. But $x_i'' = (BA)_{ij}x_j$. So

$$(BA)_{ij} = B_{ik}A_{kj}.$$

Generally, an $m \times n$ matrix multiplied by an $n \times \ell$ matrix gives an $m \times \ell$ matrix. $(BA)_{ij}$ is the ith row of $\mathbf{B}$ dotted with the jth column of $\mathbf{A}$.

Definition (Transpose of matrix)**.** If A is an $m \times n$ matrix, the *transpose* A^T is an $n \times m$ matrix defined by $(A^T)_{ij} = A_{ji}$.

Definition (Hermitian conjugate)**.** Define $A^\dagger = (A^T)^*$. Similarly, $(AB)^\dagger = B^\dagger A^\dagger$.

Definition (Symmetric matrix)**.** A matrix is *symmetric* if $A^T = A$.

Definition (Hermitian matrix)**.** A matrix is *Hermitian* if $A^\dagger = A$. (The diagonal of a Hermitian matrix must be real).

Definition (Anti/skew symmetric matrix)**.** A matrix is *anti-symmetric* or *skew symmetric* if $A^T = -A$. The diagonals are all zero.

Definition (Skew-Hermitian matrix)**.** A matrix is *skew-Hermitian* if $A^\dagger = -A$. The diagonals are pure imaginary.

Definition (Trace of matrix)**.** The *trace* of an $n \times n$ matrix A is the sum of the diagonal. $\mathrm{tr}(A) = A_{ii}$

Definition (Identity matrix)**.** $I = \delta_{ij}$.

3.4.3 Decomposition of an $n \times n$ matrix

3.4.4 Matrix inverse

Definition (Inverse of matrix)**.** Consider an $m \times n$ matrix A and $n \times m$ matrices B and C. If $BA = I$, then B is the *left inverse* of A. If $AC = I$, then C is the *right inverse* of A. If A is square ($n \times n$), then $B = B(AC) = (BA)C = C$, i.e. the left and right inverses coincide. Both are denoted by A^{-1}, the *inverse* of A. Therefore we have

$$AA^{-1} = A^{-1}A = I.$$

Definition (Invertible matrix)**.** If A has an inverse, then A is *invertible.*

Definition (Orthogonal and unitary matrices)**.** A real $n\times n$ matrix is *orthogonal* if $A^TA = AA^T = I$, i.e. $A^T = A^{-1}$. A complex $n \times n$ matrix is *unitary* if $U^\dagger U = UU^\dagger = I$, i.e. $U^\dagger = U^{-1}$.

3.5 Determinants

3.5.1 Permutations

Definition (Fixed point)**.** A *fixed point* of ρ is a k such that $\rho(k) = k$. e.g. in $\begin{pmatrix}1 & 2 & 3 & 4\\ 4 & 1 & 3 & 2\end{pmatrix}$, 3 is the fixed point. By convention, we can omit the fixed point and write as $\begin{pmatrix}1 & 2 & 4\\ 4 & 1 & 2\end{pmatrix}$.

Definition (Disjoint permutation)**.** Two permutations are *disjoint* if numbers moved by one are fixed by the other, and vice versa. e.g. $\begin{pmatrix}1 & 2 & 4 & 5 & 6\\ 5 & 6 & 1 & 4 & 2\end{pmatrix} = \begin{pmatrix}2 & 6\\ 6 & 2\end{pmatrix}\begin{pmatrix}1 & 4 & 5\\ 5 & 1 & 4\end{pmatrix}$, and the two cycles on the right hand side are disjoint. Disjoint permutations commute, but in general non-disjoint permutations do not.

Definition (Transposition and k-cycle)**.** $\begin{pmatrix}2 & 6\\ 6 & 2\end{pmatrix}$ is a *2-cycle* or a *transposition*, and we can simply write $(2\ 6)$. $\begin{pmatrix}1 & 4 & 5\\ 5 & 1 & 4\end{pmatrix}$ is a 3-cycle, and we can simply write $(1\ 5\ 4)$. (1 is mapped to 5; 5 is mapped to 4; 4 is mapped to 1)

Definition (Sign of permutation)**.** The *sign* of a permutation $\epsilon(\rho)$ is $(-1)^r$, where r is the number of 2-cycles when ρ is written as a product of 2-cycles. If $\epsilon(\rho) = +1$, it is an even permutation. Otherwise, it is an odd permutation. Note that $\epsilon(\rho\sigma) = \epsilon(\rho)\epsilon(\sigma)$ and $\epsilon(\rho^{-1}) = \epsilon(\rho)$.

Definition (Levi-Civita symbol)**.** The *Levi-Civita* symbol by

$$\epsilon_{j_1j_2\cdots j_n} = \begin{cases} +1 & \text{if } j_1j_2j_3\cdots j_n \text{ is an even permutation of } 1, 2, \cdots n\\ -1 & \text{if it is an odd permutation}\\ 0 & \text{if any 2 of them are equal}\end{cases}$$

Clearly, $\epsilon_{\rho(1)\rho(2)\cdots\rho(n)} = \epsilon(\rho)$.

Definition (Determinant)**.** The determinant of an $n \times n$ matrix A is defined as:

$$\det(A) = \sum_{\sigma\in S_n} \epsilon(\sigma)A_{\sigma(1)1}A_{\sigma(2)2}\cdots A_{\sigma(n)n},$$

or equivalently,

$$\det(A) = \epsilon_{j_1j_2\cdots j_n}A_{j_11}A_{j_22}\cdots A_{j_nn}.$$

3.5.2 Properties of determinants

3.5.3 Minors and Cofactors

Definition (Minor and cofactor)**.** For an $n \times n$ matrix A, define A^{ij} to be the $(n-1) \times (n-1)$ matrix in which row i and column j of A have been removed.

The *minor* of the ijth element of A is $M_{ij} = \det A^{ij}$

The *cofactor* of the ijth element of A is $\Delta_{ij} = (-1)^{i+j} M_{ij}$.

4 Matrices and linear equations

4.1 Simple example, 2×2

4.2 Inverse of an $n \times n$ matrix

4.3 Homogeneous and inhomogeneous equations

Definition (Homogeneous equation)**.** If $\mathbf{b} = \mathbf{0}$, then the system is *homogeneous*. Otherwise, it's *inhomogeneous*.

4.3.1 Gaussian elimination

4.4 Matrix rank

Definition (Column and row rank of linear map)**.** The *column rank* of a matrix is the maximum number of linearly independent columns.

The *row rank* of a matrix is the maximum number of linearly independent rows.

4.5 Homogeneous problem $A\mathbf{x} = \mathbf{0}$

4.5.1 Geometrical interpretation

4.5.2 Linear mapping view of $A\mathbf{x} = \mathbf{0}$

4.6 General solution of $A\mathbf{x} = \mathbf{d}$

5 Eigenvalues and eigenvectors

5.1 Preliminaries and definitions

Definition (Multiplicity of root)**.** The root $z = \omega$ has *multiplicity* k if $(z - \omega)^k$ is a factor of $p(z)$ but $(z - \omega)^{k+1}$ is not.

Definition (Eigenvector and eigenvalue)**.** Consider a linear map $\alpha : \mathbb{C}^n \to \mathbb{C}^n$ with associated matrix A. Then $\mathbf{x} \neq \mathbf{0}$ is an eigenvector of A if

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

for some λ. λ is the associated *eigenvalue*. This means that the direction of the eigenvector is preserved by the mapping, but is scaled up by λ.

Definition (Characteristic equation of matrix)**.** The *characteristic equation* of A is

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

Definition (Characteristic polynomial of matrix)**.** The *characteristic polynomial* of A is

$$p_A(\lambda) = \det(A - \lambda I).$$

Definition (Eigenspace)**.** The *eigenspace* denoted as E_λ is the kernel of the matrix $A - \lambda I$, i.e. the set of eigenvectors with eigenvalue λ.

Definition (Algebraic multiplicity of eigenvalue)**.** The *algebraic multiplicity* $M(\lambda)$ or M_λ of an eigenvalue λ is the multiplicity of λ in $p_A(\lambda) = 0$. Clearly (by the fundamental theorem of algebra),

$$\sum_\lambda M(\lambda) = n.$$

If $M(\lambda) > 1$, then the eigenvalue is *degenerate*.

Definition (Geometric multiplicity of eigenvalue)**.** The *geometric multiplicity* $m(\lambda)$ or m_λ of an eigenvalue λ is the dimension of the eigenspace, i.e. the maximum number of linearly independent eigenvectors with eigenvalue λ.

Definition (Defect of eigenvalue)**.** The *defect* Δ_λ of eigenvalue λ is

$$\Delta_\lambda = M(\lambda) - m(\lambda).$$

It can be proven that $\Delta_\lambda > 0$.

5.2 Linearly independent eigenvectors

5.3 Transformation matrices

5.3.1 Transformation law for vectors

5.3.2 Transformation law for matrix

5.4 Similar matrices

Definition (Similar matrices)**.** Two $n \times n$ matrices A and B are *similar* if there exists an invertible matrix P such that

$$B = P^{-1}AP,$$

i.e. they represent the same map under different bases (or: they are in the same conjugacy classes)

5.5 Diagonalizable matrices

Definition (Diagonalizable matrices)**.** An $n \times n$ matrix A is *diagonalizable* if it is similar to a diagonal matrix. We showed above that this is equivalent to saying the eigenvectors form a basis of $\mathbb{C}^n$.

5.6 Canonical (Jordan normal) form

5.7 Cayley-Hamilton Theorem

5.8 Eigenvalues and eigenvectors of a Hermitian matrix

5.8.1 Gram-Schmidt orthogonalization (non-examinable)

5.8.2 Unitary transformation

5.8.3 Diagonalization of $n \times n$ Hermitian matrices

5.8.4 Normal matrices

Definition (Normal matrix)**.** A *normal matrix* as a matrix that commutes with its own Hermitian conjugate, i.e.

$$NN^\dagger = N^\dagger N$$

Hermitian, real symmetric, skew-Hermitian, real anti-symmetric, orthogonal, unitary matrices are all special cases of normal matrices.

6 Quadratic forms and conics

Definition (Sesquilinear, Hermitian and quadratic forms)**.** A *sesquilinear form* is a quantity $F = \mathbf{x}^\dagger A\mathbf{x} = x_i^* A_{ij} x_j$. If A is Hermitian, then F is a *Hermitian form.* If A is real symmetric, then F is a *quadratic form*

6.1 Quadrics and conics

Definition (Quadric)**.** A *quadric* is an n-dimensional surface defined by the zero of a real quadratic polynomial, i.e.

$$\mathbf{x}^T A\mathbf{x} + \mathbf{b}^T\mathbf{x} + c = 0,$$

where A is a real $n \times n$ matrix, $\mathbf{x}, \mathbf{b}$ are n-dimensional column vectors and c is a constant scalar.

6.1.1 Conic sections $(n = 2)$

6.2 Focus-directrix property

Definition (Conic sections)**.** The *eccentricity* and *scale* are properties of a conic section that satisfy the following:

Let the *foci* of a conic section be $(\pm ae, 0)$ and the *directrices* be $x = \pm a/e$.

The a *conic section* is the set of points with properties that the distance from focus $= e\times$ distance from directrix which is closer to that of focus (unless $e = 1$, where we take the distance to the other directrix).

7 Transformation groups

7.1 Groups of orthogonal matrices

Definition (Orthogonal group)**.** The *orthogonal group* $O(n)$ is the group of orthogonal matrices.

Definition (Special orthogonal group)**.** The *special orthogonal group* is the subgroup of $O(n)$ that consists of all orthogonal matrices with determinant 1.

7.2 Length preserving matrices

7.3 Lorentz transformations

Definition (Minkowski inner product)**.** The *Minkowsky* inner product of 2 vectors $\mathbf{x}$ and $\mathbf{y}$ is

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

where

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

Then $\langle \mathbf{x} | \mathbf{y} \rangle = x_1 y_1 - x_2 y_2$.

Definition (Preservation of inner product)**.** A transformation matrix M preserves the Minkowsky inner product if

$$\langle \mathbf{x} | \mathbf{y} \rangle = \langle M\mathbf{x} | M\mathbf{y} \rangle$$

for all $\mathbf{x}, \mathbf{y}$.

Definition (Lorentz matrix)**.** A *Lorentz matrix* or a *Lorentz boost* is a matrix in the form

$$B_v = \frac{1}{\sqrt{1 - v^2}} \begin{pmatrix} 1 & v \\ v & 1 \end{pmatrix}$$

Here $|v| < 1$, where we have chosen units in which the speed of light is equal to 1. We have $B_v = H_{\tanh^{-1} v}$

Definition (Lorentz group)**.** The *Lorentz group* is a group of all Lorentz matrices under matrix multiplication.