# Vectors and Matrices                 IA Mich

Read hermitian matrix diagonalisable proof, and supo notes

# 1   Complex numbers

- Definitions

    - Construct $\mathbb{C}$ from $\mathbb{R}$ with element $i$ such that $i^2 = -1$
    - Addition, multiplication, Complex conjugate ($\bar{z}$ or $z^*$)
    - Modulus, argument (multi-valued, restrict using principal value), Argand diagram

- Properties

    - $\mathbb{C}$ is a field (with $+ - \times \nabla \cdot$, associative, commutative, distributive)
    - Fundamental Theorem of Algebra
    - Geometric representation of addition, subtraction, complex conjugation
    - Composition property: $|z_1 z_2| = |z_1||z_2|$
    - Triangle inequality: $|z_1 + z_2| \leq |z_1| + |z_2|$, and alternatively $|z_1 - z_2| \geq ||z_1| - |z_2||$
    - Multiplying complex numbers: moduli multiply, arguments add
    - De Moivre's Theorem

- Exponential and Trig functions

    - Define $\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$, $\sin(z), \cos(z)$ in terms of $e^z$, consistent with usual rules, thus can have exponential form of complex numbers
    - Roots of unity: $z^n = 1 \iff z = \exp(\frac{2k\pi i}{n}), k = 0, 1, ..., n-1$

- Logarithm and Complex Powers

    - Define log as inverse of exp
    - $z = |z|e^{i\arg(z)} \implies \log(z) = \log|z| + i\arg(z)$ (multi-valued as $\arg(z)$ is multi-valued, can restrict using principal value)
    - Complex power: $z^\alpha = e^{\alpha \log(z)}$, generally infinitely many values, may be finite in special cases (integer/rational exponent)

- Lines and Circles

  - Line: parametric (point and direction), and non-parametric (eliminate by taking complex conjugate)
  - Circles: $|z - c| = \rho$ (or expanded) or $z = c + \rho e^{i\theta}$

- Geometric transformations described in $\mathbb{C}$: translation, scaling, rotation, reflection, inversion; relation to Möbius transformation

# 2 Vectors in 3D

- Euclidean space: vectors have direction and magnitudes

- Addition and scalar multiplication (Geometric)

  - Parallelogram construction, scalar multiplication by extending
  - Properties: identity, inverses, commutativity, associativity, distributivity
  - Linear combination and span of vectors
  - Parallel: one vector being a scalar multiple of the other, otherwise the two vectors span a plane (allow $\mathbf{0} \parallel \mathbf{a}$ for all $\mathbf{a}$ in this definition)

- Scalar/dot/inner product (Geometric)

  - Definition: $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta$, using magnitude and angle between the vectors
  - Properties: symmetric, positive definite, scalar multiple can 'move around', distributive
  - Orthogonality: $\mathbf{a} \perp \mathbf{b} \iff \mathbf{a} \cdot \mathbf{b} = 0$ (allow $\mathbf{0} \perp \mathbf{a}$ for all $\mathbf{a}$ in this definition)
  - Resolving a vector along another one: component of $\mathbf{b}$ along $\mathbf{a}$ is given by (in magnitude)

$$|\mathbf{b}| \cos \theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|}$$

- Orthonormal Bases

  - A set of orthonormal vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ (or $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$) are each of unit length and mutually perpendicular: $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$
  - Equivalent to choosing Cartesian axes along the 3 directions

- Basis $\{\mathbf{e}_i\}$: linearly independent and span: Any vector can be represented uniquely by a linear combination of the three vectors: $\mathbf{a} = \sum_i a_i \mathbf{e}_i$, components determined by dotting $\mathbf{a}$ with each of the basis vectors

- Notion of components: $(a_1, a_2, a_3)$ or $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$

- Vector/Cross product (Geometric)

  - Definition: $\mathbf{a} \times \mathbf{b} = |\mathbf{a}||\mathbf{b}| \sin\theta \hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ (unit normal vector) and $\theta$ are defined in a right-handed sense

  - Signed area of parallelogram, with sign giving orientation

  - Can be seen as rotating and scaling a vector at the same time (p.15)

  - Define a orthonormal right-handed set by cross product, then cross products can be expressed in components

- Triple Products

  - Scalar triple product: signed volume of parallelepiped, test for coplanar vectors

  - Vector triple product: $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$

- Lines, planes and vector equations

  - Line: $\mathbf{r} = \mathbf{a} + \lambda\mathbf{u}$ (parametric form); $\mathbf{u} \times \mathbf{r} = \mathbf{c} =$ constant (non-parametric)

  - Plane: $\mathbf{r} = \mathbf{a} + \lambda\mathbf{u} + \lambda\mathbf{v}$, with $\mathbf{u}$ and $\mathbf{v}$ non-parallel in the plane (parametric form); $\mathbf{n} \cdot \mathbf{r} = \mathbf{n} \cdot \mathbf{a} = \kappa$, with $\mathbf{n}$ normal vector to the plane;

  - Perpendicular distance of plane from origin $= \dfrac{|\kappa|}{|\mathbf{n}|}$

  - Vector equations: eliminate $\mathbf{r}$ by taking products; consider cases; let $\mathbf{r}$ be a linear combination of constant vectors and solve for coefficients, complete square

- Index notation and summation convention (Algebraic)

– Free indices appear exactly once in every term: represent components; Dummy indices appear exactly twice in each term, and are summed up;

– Nothing should appear more than twice

– Kronecker Delta $\delta_{ij}$: represent identity matrix $I$

– Levi-Civita Epsilon $\epsilon_{ijk}$: antisymmetric, commonly used in cross product, determinant

– Identities: $\epsilon_{ijk}\epsilon_{pqk} = \delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp}$, "same minus different"

# 3  Vectors in general, $\mathbb{R}^n$ and $\mathbb{C}^n$ (Algebraic)

- $\mathbb{R}^n$: set of real n-tuples, form linear combination by addition and (real) scalar multiplication

  – Component expression correspond to a standard basis

  – Inner product on $\mathbb{R}^n$: symmetric, bilinear (linear in both arguments), positive definite ($\mathbf{x} \cdot \mathbf{x} \geq 0$, equality holds $\iff \mathbf{x} = \mathbf{0}$)

  – Parallel, orthogonality, and (Euclidean) norm of vectors

  – Cauchy-Schwarz inequality: $|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, with equality iff $\mathbf{x} = \lambda\mathbf{y}$ or $\mathbf{y} = \lambda\mathbf{x}$ for $\lambda \in \mathbb{R}$; thus can define angle between vectors in $\mathbb{R}^n$ by inner product

  – Triangle inequality: $|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$

  – Vector inner product as matrix multiplication: $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T\mathbf{y}$

  – Inner product carries to higher dimensions from $\mathbb{R}^3$, but cross product cannot (due to $\epsilon_{ij...k}$ with too many free indices)

  – Instead, $\epsilon_{ij}x_iy_j$ defines an alternative scalar product in 2D, which generalises to scalar triple product in 3D

- Vector spaces

  – Defining vector spaces with axioms that enable addition, scalar multiplication (e.g. distributive, associative, identity)

  – For a set of vectors $\mathbf{v}_i \in V$, a vector space, define $\text{span}\{\mathbf{v}_i\} = \{\sum_i \lambda_i\mathbf{v}_i\}$, the set of all linear combinations

  – The span of a set of vectors forms a subspace of $V$ (a subset and itself a vector space); a subspace must contain $\mathbf{0}$

- Linearly independent: $\sum_i \lambda_i \mathbf{v}_i = \mathbf{0} \implies \lambda_i = 0$ for all $i$

- Linear dependence: $\sum_i \lambda_i \mathbf{v}_i = \mathbf{0}$ has non-trivial solutions; i.e. some vector in the set can be written as a linear combination of the others

- Inner products obey the three properties: symmetric, bilinear, positive definite

- Bases and dimension

  - Basis: linearly independent and spans the vector space $V$; thus each point in $V$ has a unique linear combination

  - Any set of bases of $V$ has the same number of elements, defined to be $\dim V$

  - A spanning set with more vectors than $\dim V$ is not linearly dependent (can remove one); a linearly independent set with less vectors than $\dim V$ does not span (can add one)

  - A set of mutually orthogonal non-zero vectors is linearly independent (dot with each vector, everything else vanish)

- Vectors in $\mathbb{C}^n$

  - $\mathbb{C}^n$: Vector space with linear combinations (complex scalar multiple), thus has $n$ basis vectors

  - Inner product: $(\mathbf{z}, \mathbf{w}) = \sum_j \bar{z}_j w_j$

  - Inner product as matrix multiplication: $(\mathbf{z}, \mathbf{w}) = \mathbf{z}^\dagger \mathbf{w}$

  - Properties: hermitian ($(\mathbf{z}, \mathbf{w}) = \overline{(\mathbf{w}, \mathbf{z})}$), anti-linear/linear (in first/second argument), positive definite; scalar products in $\mathbb{R}^2$ recovered in inner product on $\mathbb{C}$

  - Define length/norm, orthogonality, orthonormal basis, linear independence (where $\mathbf{z}^\dagger = \bar{\mathbf{z}}^T$, a row vector with components $\bar{z}_i$)

# 4 Matrices and Linear Map

- Linear transformation/map

  - A function $T : V \rightarrow W$, between vector spaces $V$ (with dim $n$) and $W$ (with dim $m$): satisfying $T(\lambda \mathbf{x} + \mu \mathbf{y}) = \lambda T(\mathbf{x}) + \mu T(\mathbf{y})$, where $\lambda$ and $\mu$ real/complex (a field in general)

- – Image, kernel, rank, nullity, rank-nullity theorem
- – Linear combination and composition of linear maps

- Matrices as linear maps

  - – Matrices define a transformation by $x_i' = M_{ij}x_j$ (summation)
  - – The i-th basis vector is mapped to the i-th column; image of matrix is the span of columns
  - – $x_i' = (\mathbf{R}_i)_j x_j = \mathbf{R}_i \cdot \mathbf{x}$: dotting the i-th row with $\mathbf{x}$ gives the i-th component of $\mathbf{x}'$; kernel of matrix is $\{\mathbf{x} : \mathbf{R}_i \cdot \mathbf{x} = 0 \text{ for all i}\}$

- Geometric examples of linear map: $\mathbb{R}^3 \to \mathbb{R}^3$

  - – Rotation (about a unit vector $\hat{\mathbf{n}}$)
  - – Reflection (in a plane through $\mathbf{0}$ with unit normal $\hat{\mathbf{n}}$)
  - – Dilation (along three axes)
  - – Shear

- Matrices in general

  - – Given vector spaces $V$ (dim $n$) with basis $\{\mathbf{e_1}, \mathbf{e_2}, ..., \mathbf{e_n}\}$, and $W$ (dim $m$) with basis $\{\mathbf{f_1}, \mathbf{f_2}, ..., \mathbf{f_m}\}$, the linear map $T : V \to W$ w.r.t. to the given bases can be represented by an $n \times m$ matrix $M$, where $T(\mathbf{e_i}) = \sum_a M_{ai}\mathbf{f_a}$, i.e. image of $\mathbf{e_i}$ is the $i$-th column of $M$ with respect to basis $\{\mathbf{f_a}\}$
  - – $T(\mathbf{x})_i = \sum_j M_{ij}x_j =$ dot product of $i$-th row of $M$ with $\mathbf{x}$
  - – Given bases of each vector space, vectors and transformation become components and matrices ($\mathbb{F}^n, \mathbb{F}^m$ and $M_{n \times m}(\mathbb{F})$)
  - – Matrix addition, scalar multiplication and matrix multiplication represent adding, scaling and composition of linear maps
  - – Invertible/non-singular matrices: left/right inverse, same inverse for square matrices

- Operation and properties of matrices

  - – Transpose and hermitian conjugate: (need not be square matrix) properties under adding, scaling and product of matrices
  - – Symmetric ($A^T = A$), anti-symmetric ($A^T = -A$), hermitian ($A^\dagger = A$), anti-hermitian ($A^\dagger = -A$)

- – Trace, decomposition of square matrices into antisymmetric + symmetric traceless + isotropic (multiple of identity)

- Orthogonal and Unitary matrices

  - – Orthogonal matrix: $A^T = A^{-1}$, inner product preserved, rows orthonormal, columns orthonormal
  - – General 2×2 orthogonal matrices: rotation or reflection (distinguished by determinant of matrix)
  - – Unitary matrix: $A^\dagger = A^{-1}$, complex inner product preserved

# 5  Determinants and Inverses

- Determinant

  - – Factor by which the signed (n-dim) volume is scaled by the transformation
  - – $\det A \neq 0 \iff$ linear dependence
  - – Sum over indices: $\det A = \sum\limits_{i_1, i_2, \ldots, i_n} \varepsilon_{i_1 i_2 \ldots i_n} A_{i_1 1} A_{i_2 2} \ldots A_{i_n n}$, extended from 3D case
  - – Sum over permutations in $S_n$: $\det A = \sum\limits_{\sigma \in S_n} \varepsilon(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \ldots A_{\sigma(n)n}$, sign of permutation $\varepsilon(\sigma)$
  - – Scaling a row/column, scaling the whole matrix, row and column operations/exchanges, transpose property, multiplicative property
  - – Expand along column or row

- Minors, cofactors and inverses

  - – Minor: $M^{ij} = $ det of matrix $M$ with $i$-th row and $j$-th col deleted
  - – Cofactor of $\Delta_{ij} = (-1)^{i+j} M^{ij}$ (signed minor)
  - – Adjugate: transpose of the cofactor matrix, so $M^{-1} = \frac{1}{\det M} \Delta^T$

- System of linear equations $A\mathbf{x} = \mathbf{b}$, where $\mathbf{x}, \boldsymbol{b} \in \mathbb{R}^n$, and $A \in M_{n \times n}(\mathbb{R})$

  - – Existence and nature of solutions:
    - ∗ $\det A \neq 0 \iff \mathrm{Im}A = \mathbb{R}^n \iff \mathrm{Ker}A = \{\mathbf{0}\} \iff$ unique solution

* $\det A = 0$ and $\mathbf{b} \in \mathrm{Im} A$ $\iff$ infinitely many solution
* $\det A = 0$ and $\mathbf{b} \notin \mathrm{Im} A$ $\iff$ no solutions
* General solution (if there exists particular solution $\mathbf{x_0}$) is $\mathbf{x} = \mathbf{x_0} + \mathrm{Ker} A$

- Can interpret each equation as a (hyper-)plane in $\mathbb{R}^n$, intersection is the set of solutions

* Homogeneous equation: each plane passes through origin, so must have at least one solution: $\mathbf{0}$
* Inhomogeneous: may not pass through origin, solution not guaranteed

- Gaussian Elimination: reduce to echelon form, (variables that are not pivot columns have 'degree of freedom')

# 6 Eigenvalues and eigenvectors

- Introduction

  - Given a linear map $T : V \to V$, an eigenvector $\mathbf{v} \in V \backslash \{\mathbf{0}\}$ with eigenvalue $\lambda \in \mathbb{F}$ (where $V$ is defined over) satisfies $T(\mathbf{v}) = \lambda \mathbf{v}$
  - Given a basis, can express the map with a matrix $A$, then have $A\mathbf{v} = \lambda\mathbf{v}$ $\iff$ $\chi_A(\lambda) = \det(A - \lambda I) = 0$ (characteristic polynomial)
  - $\chi_A(\lambda) = 0$ has $n$ roots (counted with multiplicity) in $\mathbb{C}$, by FTA; deduce $\mathrm{tr} A$, $\det A$ from the equation
  - Eigenspace $E_\lambda$, geometric multiplicity $m_\lambda$, algebraic multiplicity $M_\lambda$, defect $\Delta_\lambda$
  - Eigenvectors with distinct eigenvalues are linearly independent

- Diagonalisability

  - Diagonalisable: eigenvectors form a basis $\iff$ there exists invertible matrix $P$ such that $P^{-1}AP = D$, with entries being the eigenvalues
  - Criteria:

    * Sufficient but not necessary: an $n \times n$ matrix has $n$ distinct eigenvalues $\implies$ form a basis
    * Sufficient and necessary: $M_\lambda = m_\lambda$ for each $\lambda$ ($n$ linearly independent eigenvectors, form a basis)

- Similar matrix

  - $A$ and $B$ are similar $\iff$ $\exists$ invertible $P$ s.t. $B = P^{-1}AP$ (all n-dimensional), equivalence class
  - Similar matrices have same trace, determinant, characteristic polynomial

- Hermitian and symmetric matrices

  - Hermitian matrices have real eigenvalues, eigenvectors with distinct eigenvalues are orthogonal (complex inner product)
  - Real symmetric: can choose real eigenvectors for each real eigenvalue, the eigenvectors are orthogonal (real inner product)
  - Gram-Schmidt Orthogonalisation: transform any basis into orthonormal basis (e.g. the union of bases of eigenspaces with distinct eigenvalues can be orthogonalised)
  - Any hermitian (real symmetric) matrix is unitarily (orthogonally) diagonalisable: can choose orthonormal set of eigenvectors , i.e. $P$ can be chosen to be unitary (orthogonal)

- Cayley-Hamilton Theorem: any matrix satisfies its own characteristic equation

# 7 Change of Bases, Canonical form and Symmetries

- Change of bases:

  - Transformation matrix: the $i$-th column of P is the $i$-th vector in the new basis w.r.t. the old basis: $\mathbf{e'_i} = \sum_j P_{ji}\mathbf{e_j}$
  - $A' = P^{-1}AP$: similar matrices represent same linear map w.r.t. different bases

- Jordan Canonical Form:

  - Classify square matrices by similarity: 2D case classify by eigenvalue and multiplicities

- Quadratic Forms

  - Given real-symmetric matrix $A$: the quadratic form is $\mathcal{F}(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = x_i A_{ij} x_j$

- – Real-symmetric matrix can be orthogonally diagonalised (by above) with $D = P^{-1}AP$, i.e. has an orthonormal basis of eigenvectors, the principal axes of $\mathcal{F}$

  – Using coordinates w.r.t. the principal axes (i.e. letting $\mathbf{x} = P\mathbf{x}'$), then $\mathcal{F}(\mathbf{x}) = \mathbf{x}'^T D \mathbf{x}' = \sum_i \lambda_i x_i'^2$

  – Signs of eigenvalues and value of $\mathcal{F}(\mathbf{x})$ determine shape

- Quadrics and conics

  – General quadric: $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} + \mathbf{b}^T \mathbf{x} + c = A_{ij} x_i x_j + b_i x_i + c = 0$, defines a surface in $\mathbb{R}^n$

  – Completing the square gives quadratic form

  – Conics: classify up to translation and orthogonal transformation (diagonalising)

- Symmetries and Transformation Groups

  – Orthogonal and special orthogonal group $O(n), SO(n)$: length, volume (and orientation) preserving

  – Can view as transforming vectors or changing orthonormal bases

- 2D Minkowski Space and Lorentz Transformations

  – Define new inner product: Minkowski metric, $\mathbb{R}^n$ becomes Minkowski space

  – Minkowski metric preserved by linear maps, subgroup of matrices with $\det M > 0$ and $M_{00} > 0$ form the Lorentz group

  – General form of matrices in terms of hyperbolic trig, matrix map corresponds to Lorentz transformation, speed of light being maximum possible speed