# Math Tripos Part IA: Vectors and Matrices

Michael Li

May 21, 2015

**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 Cauchy-Schwarz inequality. Concepts of linear span, linear independence, subspaces, basis and dimension.

Suffix notation: including summation convention, $\delta_{ij}$ and $\epsilon_{ijk}$. 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 $3 \times 3$ 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 $2 \times 2$ 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.** A *complex number* is a number $z \in \mathbb{C}$ of form $a + bi$ or $re^{i\theta} = r(\cos\theta + i\sin\theta)$, where $a, b \in \mathbb{R}$, $i = \sqrt{-1}$. *Modulus* is defined as $|z| = r = \sqrt{a^2 + b^2}$ and *argument* is $\theta = \arg z = tan^{-1}(\frac{y}{x})$. We take the argument to within the *principle value* of $-\pi < \theta \leq \pi$ to ensure unique representation. The *complex conjugate* of $z$ is $\bar{z} = z^* = a - ib$.

**Definition.** An *Argand diagram* is a 2D diagram with real numbers on the "x-axis" and complex numbers on the "y-axis". A complex number is represented by vector $\mathbf{p} = \begin{pmatrix} x \\ y \end{pmatrix}$.

### 1.1.1 Basic Theorems of Complex Numbers

(i) (Triangle Inequality) $|z_1 + z_2| \leq |z_1| + |z_2|$

(ii) (De Movire's Theorem+Extension) $\cos n\theta + i\sin n\theta = (cos\theta + i\sin\theta)^n$. (Proof by induction) Also, multiply the moduli and add the arguments when multiplying two complex numbers.

## 1.2 Complex exponential function

The exponential function is defined as usual (see D.E.). Using the lemma below (expand to see it is trivially true):

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} a_{mn} = \sum_{r=0}^{\infty}\sum_{m=0}^{r} a_{r-m,m}$$

We can prove that $e^{z_1}e^{z_2} = e^{z_1+z_2}$:

*Proof.*

$$\begin{aligned}\exp(z_1)\exp(z_2) &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{z_1^m}{m!}\frac{z_2^n}{n!} \\ &= \sum_{r=0}^{\infty}\sum_{m=0}^{r}\frac{z_1^{r-m}}{(r-m)!}\frac{z_2^m}{m!} = \sum_{r=0}^{\infty}\frac{1}{r!}\sum_{m=0}^{r}\frac{r!}{(r-m)!m!}z_1^{r-m}z_2^m \\ &= \sum_{r=0}^{\infty}\frac{(z_1+z_2)^r}{r!}\end{aligned}$$

□

**Theorem.** When $\sin z = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1}$ and $\cos z = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n}$ (defined in power series), $e^{iz} = \cos z + i \sin z$

*Proof.* The result is immediate after expansion and comparison of the two power series. □

### 1.3 Roots of Unity

**Definition.** The $n$-th *roots of unity* are roots of $z^n = 1$ for $n \in \mathbb{N}$, or $\exp\left(2\pi i \frac{k}{n}\right)$ for $k = 0, \cdots n - 1$.

**Proposition.** If $\omega = \exp\left(\frac{2\pi i}{n}\right)$, then $1 + \omega + \omega^2 + \cdots + \omega^{n-1} = 1$

*Proof.*

(i) Coefficient of $z^{n-1}$ is 0 so by Vieta's Theorem, the sum of roots is 0.

(ii) $z^n = 1 \Rightarrow z^n - 1 = 0 \Rightarrow (z-1)(z^{n-1} + \cdots + 1) = 0$. Clearly $z \neq 1$, so $z^{n-1} + \cdots + 1 = 0$

□

### 1.4 Complex logarithm and power

**Definition.** Logarithm is defined as $\log z = \log re^{i\theta} = \log r + i\theta$. Power is defined as $z^\alpha = e^{\alpha \log z}$. We use the principle value to make the two single-valued.

### 1.5 Lines and circles in $\mathbb{C}$

**Theorem.** Straight line in $\mathbb{C}$ is $z = z_0 + \lambda\omega$ for $\lambda \in \mathbb{R}$. Since $\frac{z - z_0}{\omega} = \lambda = \bar{\lambda} = \frac{\bar{z} - \bar{z_0}}{\bar{\omega}}$, we have:

$$z\bar{\omega} - \bar{z}\omega = z_0\bar{\omega} - \bar{z}_0\omega$$

For circles, we have:

$$|z - c| = \rho$$

## 2 Vectors

### 2.1 Definition and basic properties

**Definition.** A *vector* $\mathbf{v}$ has length and direction. If $|\mathbf{v}| = 0$, then $\mathbf{v} = \mathbf{0}$. A *unit vector* is one with length 1, denoted $\hat{\mathbf{v}}$. Vectors have addition and scalar multiplication. Vectors with addition forms a group and *Scalar multiplication* satisfies axioms:

(i) $(\lambda + \mu)\mathbf{a} = (\lambda + \mu)\mathbf{a}, \lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}$.

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

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

**Note.** $\mathbb{R}^n$ is a vector space (addition+scalar multiplication). A vector space has to contain $\mathbf{0}$!

## 2.2 Scalar product

### 2.2.1 Geometric Picture

**Definition** (Scalar 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 $\mathbf{a} \perp \mathbf{b}$.

The projection of $\mathbf{b}$ onto $\mathbf{a}$ is given by $|\mathbf{b}| \cos\theta\hat{\mathbf{a}} = (\hat{\mathbf{a}} \cdot \mathbf{b})\hat{\mathbf{a}}$

### 2.2.2 General algebraic definition

**Definition.** In a vector space $V$, the *inner product* maps $V \times V \to \mathbb{R}$, denoted $\mathbf{x} \cdot \mathbf{y}$ or $\langle \mathbf{x} | \mathbf{y} \rangle$ with:

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

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

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

**Note.** linearity in 1st argument is only true for $\mathbb{R}^n$. Also $|\mathbf{a}|$=$\sqrt{\mathbf{a} \cdot \mathbf{a}}$ is the norm of the vector.

### 2.3 Cauchy-Schwarz inequality

**Theorem** (Cauchy-Schwarz inequality)**.** $\mathbf{x}\cdot\mathbf{y} \leq |\mathbf{x}||\mathbf{y}|$ for all $\mathbf{x},\mathbf{y} \in \mathbb{R}$.

*Proof.*
$$|\mathbf{x}-\lambda\mathbf{y}|^2 \geq 0 \Rightarrow \lambda^2|\mathbf{y}|^2 - \lambda(2\mathbf{x}\cdot\mathbf{y}) + |\mathbf{x}|^2 \geq 0$$

Viewing this as a quadratic in $\lambda$, the quadratic is non-negative and thus the determinant $\Delta \leq 0$

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

□

**Example.**

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

### 2.4 Vector product

**Definition.** 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:

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

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

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

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

The area of a triangle $ABC$ is given by $\frac{1}{2}|\overrightarrow{OA}||\overrightarrow{OB}|\sin\theta = \frac{1}{2}|\overrightarrow{OA}\times\overrightarrow{OB}|$.

### 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 $[\mathbf{a},\mathbf{b},\mathbf{c}] = [\mathbf{b},\mathbf{c},\mathbf{a}] = [\mathbf{c},\mathbf{a},\mathbf{b}] = -[\mathbf{b},\mathbf{a},\mathbf{c}] = -[\mathbf{a},\mathbf{c},\mathbf{b}] = -[\mathbf{c},\mathbf{b},\mathbf{a}]$. It represents volume of a parallelopiped with sides $\mathbf{a},\mathbf{b},\mathbf{c}$ in a right-handed system.

## 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 $\mathbf{r}$, there exists some $\lambda, \mu \in \mathbb{R}$ such that $\mathbf{r} = \lambda\mathbf{a} + \mu\mathbf{b}$. The coefficients are unique if $\mathbf{a} \neq \kappa\mathbf{b}$. (Proof below).

*Proof.* Suppose that $\mathbf{r} = \lambda\mathbf{a}+\mu\mathbf{b} = \lambda'\mathbf{a}+\mu'\mathbf{b}$. So $\mathbf{a} = \frac{\mu'-\mu}{\lambda-\lambda'}\mathbf{b}$, a contradiction. □

**Definition.** Two vectors $\mathbf{a}$ and $\mathbf{b}$ are *linearly independent* if for $\alpha, \beta \in \mathbb{R}$, $\alpha\mathbf{a} + \beta\mathbf{b}$ iff $\alpha = \beta = 0$. A *basis* is a set of vectors that spans $\mathbb{R}^2$ and are linearly independent, e.g.$\{\hat{\mathbf{i}}, \hat{\mathbf{j}}\}$.

### 2.6.2 $\mathbb{R}^n$ Space

**Definition.** 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.** 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

$$\forall \mathbf{x} \in \mathbb{R}^n \ \exists \lambda_i \text{ such that } \sum_{i=1}^{m} \lambda_i \mathbf{u}_i = \mathbf{x}$$

**Definition** (Basis vectors)**.** A *basis* of $\mathbb{R}^n$ is a linearly independent spanning set. The standard basis of $\mathbb{R}^n$ is $\mathbf{e}_1 = (1, 0, 0, \cdots 0), \cdots \mathbf{e}_n = (0, 0, 0, \cdots, 1)$. The number of vectors in basis of vector space is called the *dimension* of vector space. It is $\{\mathbf{e}_i\}$ is *orthonormal* if $\mathbf{e}_i \cdot \mathbf{e_j} = 0$ for $i \neq j$ and $\mathbf{e}_i \cdot \mathbf{e_i} = 1$ for all $i$.

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

### 2.6.3 $\mathbb{C}^n$ space

**Definition.** $\mathbb{C}^n = \{(z_1, z_2, \cdots z_n) : z_i \in \mathbb{C}\}$. $\mathbf{u} \cdot \mathbf{v} = \sum \bar{\mathbf{u}}_i \mathbf{v}_i$. Scalar product has same properties except $\mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{u}}$ and $(\lambda\mathbf{u} + \mu\mathbf{v}) \cdot \mathbf{w} = \bar{\lambda}\mathbf{u} \cdot \mathbf{w} + \bar{\mu}\mathbf{v} \cdot \mathbf{w}$.

### 2.7 Vector subspaces

**Definition** (Vector subspace)**.** A non-empty subset $U$ of a vector space $V$ is $U$ is a *vector subspace* with the same operations as $V$. Both $V$ and $\{0\}$ are subspaces of $V$. All others are proper subspaces.

A subset $U$ of $V$ is a subspace if $U$ is non-empty and for all $\mathbf{x}, \mathbf{y} \in U$, $(\lambda \mathbf{x} + \mu \mathbf{y}) \in U$.

**Example.** $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ is a basis of $\mathbb{R}^3$, then $\{\mathbf{a}+\mathbf{c}, \mathbf{b}+\mathbf{c}\}$ is a basis of the 2D subspace

Suppose $x, y \in span\{\mathbf{a}+\mathbf{c}, \mathbf{b}+\mathbf{c}\}$. Then

$$\mathbf{x} = \alpha_1(\mathbf{a}+\mathbf{c}) + \beta_1(\mathbf{b}+\mathbf{c}), \mathbf{y} = \alpha_2(\mathbf{a}+\mathbf{c}) + \beta_2(\mathbf{b}+\mathbf{c})$$
$$\lambda\mathbf{x} + \mu\mathbf{y} = (\lambda\alpha_1 + \mu\alpha_2)(\mathbf{a}+\mathbf{c}) + (\lambda\beta_1 + \mu\beta_2)(\mathbf{b}+\mathbf{c})$$

So it is a subspace. To check the two vectors is a basis, . $\alpha(\mathbf{a}+\mathbf{c}) + \beta(\mathbf{b}+\mathbf{c}) = \alpha\mathbf{a} + \beta\mathbf{b} + (\alpha+\beta)\mathbf{c} = 0$ Since $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is a basis $\Rightarrow \alpha = \beta = 0 \Rightarrow \mathbf{a}+\mathbf{c}, \mathbf{b}+\mathbf{c}$ L.I. Thus they are a basis (clearly spans).

### 2.8 Suffix Notation

We can write vector equations in form: $\mathbf{a} = \mathbf{b} \to a_i = b_i$ $\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b} \to c_i = \alpha a_i + \beta b_i$

A vector has one free suffix$i$. A scalar has no free suffix. consider scalar $\mathbf{x} \cdot \mathbf{y} = \sum x_i y_i$

Here $i$ is repeated once - a dummy suffix. (the dummy suffix doesn't matter)

Einstein's summation convention: "drop $\sum$ if we have repeated suffix, summation understood."

#### Rules

(i) Suffice appears once in a term: free suffix

(ii) suffix appears twice in a term: dummy suffix, summed over

(iii) If it appears $\geq 3$ times then it is WRONG!

**Example.** $[(\mathbf{a} \cdot \mathbf{b})\mathbf{c} - (\mathbf{c} \cdot \mathbf{a})\mathbf{b}] = a_j b_j c_i - a_j c_j b_i$ summing over $j$ understood

**Definition.** The *Kronecker Delta* is $\delta_{ij}$

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

Properties: $a_i \delta_{ij} = a_j$, $\delta_{ij}\delta_{jk} = \delta_{ik}$, $\delta_{ii} = 3$ (or $n$ in $\mathbf{r}^n$), $a_p \delta_{pq} b_q = a_p b_p$

**Definition.** The *Alternating symbol* $\epsilon_{ijk}$
$\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$ (Even permutations)
$\epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1$ (Odd permutations)
It is 0 if any of the numbers repeat.

**Properties**

(i) $\epsilon_{ijk}\delta_{jk} = \epsilon_{ijj} = 0$

(ii) If $a_{jk} = a_{kj}$ $\epsilon_{ijk}a_{jk} = \epsilon_{ijk}a_{kj} = -\epsilon_{ikj}a_{kj} = -\epsilon_{ijk}a_{jk} = 0$

(iii) $(\mathbf{a} \times \mathbf{b})_i = \epsilon_{ijk}a_jb_k$, in 3D, $\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$

**Lemma.** $\epsilon_{ijk}\epsilon_{ipq} = \delta_{jp}\delta_{kq} - \delta_{jq}\delta_{kp}$ (Equally $\epsilon_{ijk}\epsilon_{pqk} = \delta_{ip}\delta_{jq} - \delta_{jp}\delta_{iq}$, $\epsilon_{ijk}\epsilon_{pjq} = \delta_{ip}\delta_{kq} - \delta_{iq}\delta_{kp}$)

*Proof.* Proof by exhaustion.

$$\begin{aligned} RHS &= +1 \text{ if } j = p \text{ and } k = q \\ &= -1 \text{ if } j = q \text{ and } k = p \\ &= 0 \text{ otherwise} \end{aligned}$$

LHS: summing over $i$, only nonzero terms are when $j, k \neq i$ and $p, q \neq$ i If $j = p$ and $k = q$, LHS $(-1)^2 or 1^2$=1. If $j = q$ and $k = p$, LHS $1 * (-1)$ or $-1 * (1) = -1$. All other possibilities $\to$ 0. ☐

### 2.8.1 Vector Triple Product

$$\begin{aligned} \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) &= \epsilon_{ijk}\epsilon_{kpq}a_jb_pc_q = \epsilon_{ijk}\epsilon_{pqk}a_jb_pc_q \\ &= (\delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp})a_jb_kc_q = a_jb_ic_j - a_jc_ib_j \\ &= (\mathbf{a} \cdot \mathbf{c})b_i - (\mathbf{a} \cdot \mathbf{b})c_i = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c} \end{aligned}$$

but $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{b} \cdot \mathbf{c})\mathbf{a}$. (vector products are not associative)

### 2.8.2 Spherical Trig

**Lemma.** $(\mathbf{a} \times \mathbf{b})(\mathbf{a} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{c}) - (\mathbf{a} \cdot \mathbf{b})(\mathbf{a} \cdot \mathbf{c})$

*Proof.* $(\mathbf{a} \times \mathbf{b})(\mathbf{a} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b})_i(\mathbf{a} \times \mathbf{c})_i = \epsilon_{ijk}a_jb_k\epsilon_{ipq}a_pc_q = (\delta_{jp}\delta_{kq} - \delta_{jq}\delta_{kp})a_jb_ka_pc_q = a_jb_ka_jc_k - a_jb_ka_kc_j = (\mathbf{a} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{c}) - (\mathbf{a} \cdot \mathbf{b})(\mathbf{a} \cdot \mathbf{c})$ ☐

Consider unit sphere, centre O, points a,b,c on surface such that $|a| = |b| = |c| = 1$

Consider circle centre O containing A and B (angle is $\delta(A, B)$) arc length AB=$\delta(A, B)$

$\alpha$ is the angle between planes OAB and OAC=angle between normals to planes $OAB$ and $OAC$=angle between $\mathbf{a} \times \mathbf{b}$ and $\mathbf{a} \times \mathbf{c}$

$$\cos\alpha = \frac{(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{a} \times \mathbf{c})}{|\mathbf{a} \times \mathbf{b}||\mathbf{a} \times \mathbf{c}|} = \frac{\mathbf{b} \cdot \mathbf{c} - (\mathbf{a} \cdot \mathbf{b})(\mathbf{a} \cdot \mathbf{c})}{|\mathbf{a} \times \mathbf{b}||\mathbf{a} \times \mathbf{c}|}$$

From this we can have the cosine rule for spherical triangles:

$$\cos\alpha \sin\delta(A, B) \sin\delta(A, C) = \cos\delta(B, C) - \cos\delta(A, B) cos\delta(A, C)$$

We can deduce from this that $\alpha > 60$ for any nontrivial equilateral spherical triangle.

## 2.9 Geometry

### 2.9.1 Lines

Any line through $\mathbf{a}$ and parallel to $\mathbf{t}$ can be written as

$$\mathbf{x} = \mathbf{a} + \lambda\mathbf{t}.$$

Taking $\times\mathbf{t}$, we have $(\mathbf{x} - \mathbf{a}) \times \mathbf{t} = \mathbf{0}$.

### 2.9.2 Plane

The normal $\mathbf{n}$ of a plane is perpendicular to a vector $x - b$ contained in the plane ($\mathbf{b}$ is a fixed point). Alternatively, we know $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ lie in the plane, so the equation of a plane is:

$$\mathbf{x} \cdot \mathbf{n} = \mathbf{b} \cdot \mathbf{n}. \text{ or } (\mathbf{x} - \mathbf{a}) \cdot [(\mathbf{b} - \mathbf{a}) \times (\mathbf{c} - \mathbf{a})] = \mathbf{0}.$$

**Example.** (Distance between two lines) Let $L_1$ be $(\mathbf{x} - \mathbf{a}_1) \times \mathbf{t}_1 = 0$ and $L_2$ be $(\mathbf{x} - \mathbf{a}_2) \times \mathbf{t}_2 = 0$.

The line of closest approach is perpendicular to both lines and thus parallel to $\mathbf{t}_1 \times \mathbf{t}_2$. The distance $s$ can then be found by projecting $\mathbf{a}_1 - \mathbf{a}_2$ onto $\mathbf{t}_1 \times \mathbf{t}_2$. Thus $s = \left|(\mathbf{a}_1 - \mathbf{a}_2) \cdot \frac{\mathbf{t}_1 \times \mathbf{t}_2}{|\mathbf{t}_1 \times \mathbf{t}_2|}\right|$.

## 2.10 Vector equations

**Example.** $\mathbf{x} - (\mathbf{x} \times \mathbf{a}) \times \mathbf{b} = \mathbf{c}$. Strategy: take the dot or cross of the equation with suitable vectors.

$$(\mathbf{x} - (\mathbf{x} \times \mathbf{a}) \times \mathbf{b}) \cdot \mathbf{b} = \mathbf{c} \cdot \mathbf{b} \Rightarrow \mathbf{x} \cdot \mathbf{b} = \mathbf{c} \cdot \mathbf{b}$$

Substituting this into the original equation, we have

$$\mathbf{x}(1 + \mathbf{a} \cdot \mathbf{b}) = \mathbf{c} + (\mathbf{c} \cdot \mathbf{b})\mathbf{a}$$

If $(1 + \mathbf{a} \cdot \mathbf{b})$ is non-zero, then

$$\mathbf{x} = \frac{\mathbf{c} + (\mathbf{c} \cdot \mathbf{b})\mathbf{a}}{1 + \mathbf{a} \cdot \mathbf{b}}$$

If $(1 + \mathbf{a} \cdot \mathbf{b}) = 0$ & $\mathbf{c} + (\mathbf{c} \cdot \mathbf{b})\mathbf{a} \neq \mathbf{0}$, then a contradiction is reached. Otherwise, $\mathbf{x} \cdot \mathbf{b} = \mathbf{c} \cdot \mathbf{b}$, which is a plane of solutions.

# 3 Matrices+Linear maps

## 3.1 Examples

### 3.1.1 Rotation in 3D

Map $r : \mathbb{R}^3 \to \mathbb{R}^3 (\mathbf{x}' = R(\mathbf{x}))$

First, rotate by $\theta$ about z axis. The matrix is: $R_{ij} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1. \end{pmatrix}$

General case: $\mathbf{x}' = \mathbf{x} cos\theta + (1 - \cos\theta)(\hat{\mathbf{n}} \cdot \mathbf{x})\hat{\mathbf{n}} + \hat{\mathbf{x}} \times \mathbf{x} \sin\theta \qquad (x_i' = R_{ij}x_j)$

$$\Rightarrow R_{ij} = \delta_{ij}\cos\theta + (1 - \cos\theta)n_i n_j - \epsilon_{ijk} n_k \sin\theta$$

### 3.1.2 Reflection in $\mathbf{R}^3$

Reflect $\mathbf{x}$ in plane with normal $\hat{\mathbf{n}}$ through $\mathbf{0}$:

$$\mathbf{x}' = \mathbf{x} - 2(\mathbf{x} \cdot \hat{\mathbf{n}})\hat{\mathbf{n}} \Rightarrow x_i' = x_i - 2x_j\hat{n_j}\hat{n_i} \Rightarrow R_{ij} = \delta_{ij} - 2\hat{n_i}\hat{n_j}$$

(The projection of $\mathbf{x}$ onto plane is $\mathbf{x}' = \mathbf{x} - (\mathbf{x} \cdot \hat{\mathbf{n}})\hat{\mathbf{n}}$ )

### 3.2 Linear maps

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

**Definition.** Let $V, W$ be real (complex) vector spaces, and $T : V \to W$. Then T is a linear map if $T(\lambda\mathbf{a} + \mu\mathbf{b}) = \lambda T(\mathbf{a}) + \mu T(\mathbf{b})$. Translation, rotation, projection and reflection are linear maps.

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

**Example.** Rotation in $\mathbb{R}^3$: Kernel is $\{0\}$, image is $\{\mathbb{R}^3\}$

**Theorem.** Consider a linear map $f : U \to V$, where $U, V$ are vector spaces. Then $\mathrm{Im}(f)$ is a subspace of $V$, and $\ker(f)$ is a subspace of $U$.

*Proof.* If $\mathbf{x}, \mathbf{y} \in \mathrm{Im}(f)$, then $\exists \mathbf{a}, \mathbf{b} \in U$ such that $\mathbf{x} = f(\mathbf{a}), \mathbf{y} = f(\mathbf{b})$. Then $\lambda\mathbf{x}+\mu\mathbf{y} = \lambda f(\mathbf{a})+\mu f(\mathbf{b}) = f(\lambda\mathbf{a}+\mu\mathbf{b})$. Now $\lambda\mathbf{a}+\mu\mathbf{b} \in U$, so $\lambda\mathbf{x}+\lambda\mathbf{y} \in \mathrm{Im}(f)$ and $\mathrm{Im}(f)$ is a subspace of $V$.

If $\mathbf{x}, \mathbf{y} \in \ker(f)$, then $f(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda f(\mathbf{x}) + \mu f(\mathbf{y}) = \lambda\mathbf{0} + \mu\mathbf{0} = \mathbf{0}$. So, $\lambda\mathbf{x} + \mu\mathbf{y} \in \ker(f)$. □

### 3.3 Rank and Nullity

**Definition.** The rank of $f$, denoted $r(f)$, is the dimension of the image of $f$. The nullity of $f$, denoted $n(f)$, is the dimension of the kernel of $f$.

**Theorem** (The Rank Nullity Theorem)**.** For linear map $f : U \to V$, $U, V$ vector spaces

$$r(f) + n(f) = dim(U) \quad 3.9$$

*Proof not in exam.* It extends basis of kernel to basis of $U$ and prove extension is a basis of image. □

**Example.** Calculate kernel and image given by $f(x, y, z) = (x+y+z, 2x-y+5z, x+2z)$.

Step 1: find $\ker f$ through solving equations $\ker f$=any vector of the form$(-2z, z, z)$=span(-2,1,1) $n(f) = 1$

Step 2: Extend basis of $\ker f$ to a basis of whole of $\mathbb{R}^3$: (-2,1,1),(0,1,0),(0,0,1)

Step 3: Apply $f$ to this basis: $f(-2, 1, 1) = \mathbf{0}$, $f(0, 1, 0) = (1, -1, 0)$, $f(0, 0, 1) = (1, 5, 2)$

Step 4: From proof of rank-nullity theorem, (1,-1,0) and (1,5,2) is a basis of the image. $r(f) = 2$

$\operatorname{Im} f = span\{(1, -1, 0), (1, 5, 2)\}$. Normal to plane is (1,1,3)

## 3.4 Matrices

### 3.4.1 Basic properties of Matrices

**Definition** (Matrix)**.** *Matrix* is defined as the form of linear map $\alpha$ in $x'_i = A_{ij} x_j$, with $A_{ij} = [\alpha(\mathbf{e}_j)]_i$. We have $A_{ij}$ is in $i$th row and $j$th column. $A$ is an $m \times n$ matrix.

### 3.4.2 Matrix Algebra

**Definition.** Matrix operations are defined as followed:

(i) *Addition* of Matrices is defined as "element" addition in each Matrix.

(ii) *Scalar multiplication* multiplies the scalar to each element in the matrix.

(iii) *Matrix multiplication* is: $(AB)_{ij} = A_{ik} B_{kj}$, i.e. $i$th row of A dotted with $j$th column of B.

**Definition.** The *transpose* of a matrix is defined by $(A^T)_{ij} = A_{ji}$. It satisfies $(AB)^T = B^T A^T$. A matrix is *symmetric* if $A^T = A$, *anti-symmetric* if $A^T = -A$.

The *Hermitian conjugate* is defined as $A^\dagger = (A^T)^*$. Similar to transpose, $(AB)^\dagger = B^\dagger A^\dagger$. A matrix is *Hermitian* if $A^\dagger = A$, *skew-Hermitian* if $A^\dagger = -A$.

**Definition.** The *trace* of an $n \times n$ matrix A is the sum of terms on the diagonal. $\operatorname{tr}(BC) = \operatorname{tr}(CB)$.

### 3.4.3 Decomposition of $n \times n$ matrix

Any $n \times n$ matrix B can be split as a sum of symmetric and anti-symmetric parts. Write:

$$B_{ij} = \frac{1}{2}(B_{ij} + B_{ji}) + \frac{1}{2}(B_{ij} - B_{ji}) = S_{ij} + A_{ij}$$

Furthermore, we can decompose $S$ into an isotropic part (a scalar multiple of the identity) plus a trace-less part (i.e. a matrix with zero trace): Write

$$S_{ij} = \frac{1}{n} tr(S)\delta_{ij} + (S_{ij} - \frac{1}{n} tr(S)\delta_{ij}) = \frac{1}{n} tr(S) I_{ij} + T_{ij}$$

$$Trace(T_{ij}) = T_{ii} = S_{ii} - \frac{1}{n} tr(S)\delta_{ii} = 0$$

Putting all this together:

$$B = \frac{1}{n}tr(B)I + (\frac{1}{2}(B + B^T) - \frac{1}{n}tr(B)I) + A$$

In 3D, $A$ an be written as: $A = \begin{pmatrix} 0 & a & -b \\ -a & 0 & -c \\ b & c & 0 \end{pmatrix}$. Now we have $\epsilon_{ijk}w_k$ = $\begin{pmatrix} 0 & w_3 & -w_2 \\ -w_3 & 0 & w_1 \\ w_2 & -w_1 & 0 \end{pmatrix}$

#### 3.4.4 Matrix Inverse

Consider $m \times n$ matrix $A$, $n \times m$ matrix $B$, $n \times n$ matrix $C$:

If $BA = I$, then B is the *left inverse* of $A$. If $AC = I$, $C$ is the *right inverse* of $A$. If A is a square, matrix, then $B = C$ and is denoted $A^{-1}$, the *inverse* of A. If $A^{-1}$ exists then a is said to be *invertible.* $(AB)^{-1} = B^{-1}A^{-1}$.

**Definition.** A real $n \times n$ matrix is *orthogonal* if $A^TA = AA^T = I, i.e. A^T = A^{-1}$. e.g. Reflection and rotations. A complex $n \times n$ matrix is *unitary* if $U^+U = UU^+ = I, i.e. U^+ = U^{-1}$

**Note.** Rows/columns of an orthogonal or unitary matrix form an orthonormal set.

### 3.5 Determinants

Consider a linear map $\alpha : \mathbb{R}^3 \to \mathbb{R}^3$, Standard basis $\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3} \to \mathbf{e_1'}, \mathbf{e_2'}, \mathbf{e_3'}$ with $e_i' = Ae_i$ Unit cube formed by the original unit vectors has volume:

$$= \epsilon_{ijk}(e_1')_i(e_2')_j(e_3')_k = \epsilon_{ijk}A_{il}(e_1)_lA_{jm}(e_2)_mA_{kn}(e_3)_n$$

$$= \epsilon_{ijk}A_{i1}A_{j2}A_{k3}$$

This is called the determinant, $\det(A)$. $\det(A)$ is

$$\equiv \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} \quad (3.21)$$

#### 3.5.1 Permutations

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

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

**Definition.** The determinant is defined as:

$$\det(A) = \sum_{\sigma \in S_n} \epsilon(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \cdots A_{\sigma n n} = \det(A) = \epsilon_{j_1 j_2 \cdots j_n} A_{j_1,1} A_{j_2,2} \cdots A_{j_n,n} \text{ 3.25a}$$

### 3.5.2 Properties of determinants

**Proposition.** (i) $det(A) = det(A^T)$

(ii) If B only differs from $A$ in a single row by a scalar $\lambda$, we have $\det(B) = \lambda \det(A)$

(iii) Adding a linear multiple of a row/column to another row/column does not change the determinant.

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

*Proof.* (i) Let $\rho$ be another permutation in $S_n$

$$A_{\sigma(1)1} A_{\sigma 22} \cdots A_{\sigma n n} = A_{\sigma(\rho(1))\rho(1)} \cdots A_{\sigma(\rho(n))\rho(n)}$$

Because RHS is just reordering LHS. Choose $\rho = \sigma^{-1}$, and note that $\epsilon(\sigma) = \epsilon(\rho)$, so:

$$\det(A) = \sum_{\rho \in S_n} A_{1\rho(1)} \cdots A_{n\rho n} = det(A^T)$$

(ii) Every item in sum is multiplied by $\lambda$, so result follows.

(iii) Let B be the matrix after addition, then $\det(B) = \det(A) + \lambda \det(\text{matrix with two identical columns}) = \det(A) \Rightarrow \det(B) = \det(A)$.

(iv)

$$\det(AB) = \sum_{r} \epsilon(\sigma)(AB)_{\sigma(1)1} ... (AB)_{\sigma(n)n}$$

$$= \sum_{\sigma} \epsilon(\sigma) \sum_{k_1, k_2 \cdots, k_n}^{n} A_{\sigma(1)k_1} B_{k_1 1} \cdots A_{\sigma(1)k_1} B_{k_n n}$$

$$= \sum_{k_1, \cdots k_n} B_{k_1 1} \cdots B_{k_n n} \underbrace{\sum_{\sigma \in S_n} A_{\sigma(1)k_1} \cdots A_{\sigma(n)k_n} \epsilon(\sigma)}_{S}$$

If in S two of $k_1, \cdots k_n$ are equal then S is the determinant of a matrix with 2 equal columns. So only need to consider distinct $k_1 \cdots k_n$, thus a permutation, say $\rho(i)$. Hence we continue:

$$= \det A \sum_{\rho} \epsilon(\rho) B_{\rho(1)1} \cdots B_{\rho(n)n} = \det A \det B$$

□

Now from these properties we have $\det A = \pm 1$ if A is orthogonal. If A is unitary, $|\det A| = 1$.

### 3.5.3 Minors and Cofactors

**Definition.** For an $n \times n$ matrix A, define $A^{ij}$ to be the $(n-1) \times (n-1)$ matrix with row $i$ and column $j$ removed. The *minor* of the $ij^{th}$ element of A, denoted $M_{ij} = \det A^{ij}$. The *cofactor* of the $ij^{th}$ element of A, denoted $\Delta_{ij} = (-1)^{i-j} M_{ij}$. Thus:

$$\det(A) = \sum_{j_I=1}^{n} A_{j_I 1} \sum_{j_1, \cdots, \overline{j_I}, \cdots J_n}^{n} \epsilon_{j_1, j_2, \cdots j_n} A_{j_1 1} A_{j_2 2} \cdots \overline{A_{j_I I}} \cdots A_{j_n n} \quad (3.37)$$

($\overline{A}$ denotes a symbol missed out of natural sequence.) Let $\sigma \in S_n$ be the permutation which moves $J_I$ to the $I$th position, and leave everything else in its natural order. In all 3 cases, see that $|I - j_I|$ transpositions are made:

$$\epsilon(\sigma) = (-1)^{I - j_I}$$

Now consider the permutation $\rho = \begin{pmatrix} 1 & \cdots & \cdots & \bar{j}_i & \cdots & n \\ j_1 & \cdots & \bar{j}_i & \cdots & \cdots & j_n \end{pmatrix}$

$\rho\sigma$ reorders $(1, \cdots n)$ to $(J_1, J_2 \cdots J_n)$. So $\epsilon(\sigma\rho) = \epsilon_{j_1 \cdots j_n} = \epsilon(\rho)\epsilon(\sigma) = (-1)^{I - j_I} \epsilon_{j_i \cdots \overline{j_I} \cdots j_n}$

$$\det A = \sum_{j_I=1}^{b} A_{j_I I} \sum_{j_1 \cdots \overline{j_I} \cdots j_n} (-1)^{I - j_I} \epsilon_{j_1 \cdots \overline{j_I} \cdots j_n} A_{j_1 1} \cdots \overline{A_{j_I I}} \cdots A_{j_n n}$$

$$= \sum_{j_I=1}^{n} A_{j_I I} (-1)^{I - j_I} M_{j_I I} = \sum_{j_I=1}^{n} A_{j_I I} \Delta_{j_I I} = \sum_{j_I=1}^{n} A_{I j_I} \Delta_{I j_I}$$

This is the *Laplace expansion formula.* So, we can expand on the column/row that is easiest.

# 4 Matrices+ Linear equations

## 4.1 Simple Example, $2 \times 2$

Consider

$$A_{11}x_1 + A_{12}x_2 = d_1$$

$$A_{21}x_1 + A_{22}x_2 = d_2$$

i.e. Solve $Ax = d$. Solving it, we have:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$$

So we have

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}$$

## 4.2 Inverses of a square matrix

Consider the matrix B which replaces the $i$th row of A with $I$th row of A. Then the cofactors expanding on row i in B is same as those expanding in A. so

$$\det B = B_{ik}\Delta_{ik} = A_{Ik}\Delta_{ik} = 0 \text{ for } i \neq I$$

Seeing this is same for rows and combining this with Laplace expansion formula:

$$A_{ik}\Delta_{jk} = \delta_{ij} \det A \Rightarrow A_{ik}(\frac{\Delta_{jk}}{\det A}) = \delta_{ij}$$

**Theorem.** If $\det A \neq 0$ then $A^{-1}$ exists and is given by

$$(A^{-1}_{ij}) = \frac{\Delta_{ji}}{\det A}$$

*Proof.* This immediately follows from previous calculations. □

**Note.** This calculation method needs $O(n \cdot n!)$ calculations, so it is highly inefficient.

### 4.3 Homogeneous and Inhomogeneous Equations

Consider

$$A\mathbf{x} = \mathbf{b} \ (4.8)$$

**Definition.** If $\mathbf{b} = 0$, the system is *homogeneous*, if $\mathbf{b} \neq 0$, the system is *inhomogeneous.*

suppose $\det A \neq 0$, then have unique solution

$$\mathbf{x} = A^{-1}\mathbf{b}$$

**Note.** The columns of $A$ $(\det A \neq 0)$ are images of the standard basis. They form a basis of $\mathbb{R}^n$.

#### 4.3.1 Gaussian Elimination

Consider The matrix system of $n$ variables with $m$ by $n$ matrix. $(m \geq n)$ Assume $A_{11}$ is not zero, (otherwise reorder), we use first equation to eliminate $A_{i1}$ other than $A_{11}$ and use the remaining second equation to clear out the second column leaving first and second term, and then repeat.

**Possibilities**

(i) If some equations on the bottom are $0 =$ sth, as in $r(A) < n$, and if at least one of the remaining constants on the right is non-zero, then it is *inconsistent* (no solutions). The system is *overdetermined.*

(ii) If $r(A) = n \leq m$, then there is an unique solution by back substitution from solving $ax_n = b$ and working from bottom to top. This system is *determined.*

(iii) If $r(A) < n$, but every constant in equation $0 =$ sth is 0, then this system is *consistent* but is *underdetermined.* It has infinitely many solutions and the solutions has dimension $n - r(A)$.

**Relation to** $\det A$ When $m = n$ A is square and since the reduced form is *upper triangular*, as in everything lower than the diagonal is 0. Similarly, a *lower triangular matrix* has nothing above the diagonal.

$\det A$ is just the product of terms on the diagonal for triangular matrices. Thus: If $r(A) < n$, you have $\det A = 0$, and if $r(A) = n$, then $\det A \neq 0$.

### 4.4 Rank of Matrix

Consider linear map $\alpha : \mathbb{R}^n \to \mathbb{R}^m$. Recall rank is $\dim(\mathrm{Im})$. The matrix $A$ of $\alpha$ has the same rank. $A$ acts on the standard basis in $\mathbf{R}^n$. They span the image and are the columns for $A$..

**Definition.** *Column rank* and *row rank* are defined as maximum number of linearly independent columns and rows respectively.

**Theorem.** Row rank equals to column rank for all matrices.

*Proof.* Let $r$ be the row rank of $A$, and write the biggest set of LI rows as $\mathbf{v}_1^T, \mathbf{v}_2^T, \ldots, \mathbf{v}_r^T$ Now denote $i$th row of $A$ as $\mathbf{r}_i^T$. Every row of $A$ can be written as a linear combination of the $\mathbf{v}'s$. Then looking at each element in $\mathbf{v}$, the vectors formed from $(v_i^T)_j$ for each $j$ and elements $i = 1$ to $r$ is a linear combination of the $r$ column vectors. Argument works with columns so result follows. □

### 4.5 Homogeneous problem

Looking at $A\mathbf{x} = 0$ with square A, If $\det A \neq 0$, $\mathbf{x} = 0$ only. So if $A\mathbf{x} = 0$ and $\mathbf{x} \neq 0$, then $\det A = 0$.

#### 4.5.1 Geometrical Interpretation

Consider a $3 \times 3$ system with $r_i^T$ the rows. $A\mathbf{x} = \mathbf{0}$ means:

$$\mathbf{r}_i^T \cdot \mathbf{x} = \mathbf{0}$$

Thus we have 3 planes intersecting, and thus we have three different possibilities:

$$\det A \neq 0 \Rightarrow \text{ Unique solution } \mathbf{x} = \mathbf{0}$$

$$\det A = 0, r(A) = 2/1/0 \Rightarrow \text{ Solutions are on a line/plane/whole of } \mathbb{R}^3$$

#### 4.5.2 Linear mapping view of $Ax = 0$

Linear map $\alpha : \mathbb{R}^n \to \mathbb{R}^n$ , $\mathbf{x} \to \mathbf{x}' = A\mathbf{x}$ The kernel $k(A)$ has dimension $n(A)$.

$n(A) = 0 \Rightarrow$ $A(e_1) \ldots A(e_n)$ is a linearly independent set, so Unique solution $\mathbf{x} = \mathbf{0}$

$n(A) > 0, r(A) < 3, \Rightarrow \dim(\mathrm{Im}) = n - n(A)$ by extending basis of kernel to basis of $\mathbb{R}^3$

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

$$\det A \neq 0 \Rightarrow x = A^{-1}\mathbf{b}$$

$$\det A = 0 \Rightarrow n(A) > 0 \; r(A) < n$$

If $\mathbf{b} \notin \text{Im}(A)$, there is no solution. If $\mathbf{b} \in \text{Im}(A)$: at least one solution. The general solution is:

$$x = x_0 + y$$

Where $y$ is the solution of $A\mathbf{x} = \mathbf{0}$ and $x_0$ is a particular solution of $A\mathbf{x} = \mathbf{y}$ For $n(A) > 0$, then solution has dimension $n(A)$(see dimension of$y$), which is infinitely many.

## 5 Eigenvalues and Eigenvectors

### 5.1 Preliminaries and definitions

**Definition.** 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.** Consider a linear map $\alpha : \mathbb{C}^n \to \mathbb{C}^n$ with associated matrix $A$. Then $\mathbf{x} \neq \mathbf{0}$ is an e*eigenvector* of $A$ if $A\mathbf{x} = \lambda\mathbf{x}$ for some $\lambda$. $\lambda$ is the associated *eigenvalue*. This means that the direction of the eigenvector is preserved by the mapping, but is scaled up by $\lambda$.

**Theorem.** $\lambda$ is an eigenvalue of $\lambda$ iff

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

*Proof.* Clear from definition ( $A\mathbf{x} = \lambda\mathbf{x}$) above. □

**Definition.** The *characteristic equation* of $A$ is $\det(A - \lambda I) = 0$ The *characteristic polynomial* of $A$ is $p_A(\lambda) = \det(A - \lambda I)$.

Note the following:

(i) $p_A(\lambda)$ has degree $n$. So an $n \times n$ matrix has $n$ eigenvalues (accounting for multiplicity).

(ii) If $A$ is real, then all eigenvalues are either real or come in complex conjugate pairs.

(iii) Observe the coefficient of the $\lambda^{n-1}$ of $p_A(\lambda)$. It is $(-1)^{n-1}(A_{11} + \cdots + A_{nn}) = (-1)^{n-1}\operatorname{tr}(A)$ from definition. But $c_{n-1}$ is the sum of roots, i.e. $c_{n-1} = (-1)^{n-1}(\lambda_1 + \lambda_2 + \cdots \lambda_n)$, so

$$\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

Finally, $p_A(0) = \det(A)$. Also it is the product of all roots, i.e. $c_0 = \lambda_1\lambda_2\cdots\lambda_n$. So

$$\det A = \lambda_1\lambda_2\cdots\lambda_n.$$

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

The *algebraic multiplicity* $M(\lambda)$ of an eigenvalue $\lambda$ is the multiplicity of $\lambda$ in $p_A(\lambda) = 0$. Clearly, $\sum_\lambda M(\lambda) = n$.If $M(\lambda) > 1$, then the eigenvalue is *degenerate*.

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

The *defect* $\Delta_\lambda$ of eigenvalue $\lambda$ is

$$\Delta_\lambda = M(\lambda) - m(\lambda) \geq 0$$

## 5.2 Linearly independent eigenvectors

**Theorem.** Suppose $n \times n$ matrix $A$ has *distinct* eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$. Then the corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ are linearly independent.

*Proof.* Suppose $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ are linearly dependent. Then we have nontrivial solution to:

$$d_1\mathbf{x}_1 + d_2\mathbf{x}_2 + \cdots + d_r\mathbf{x}_r = \mathbf{0}.$$

Suppose this is the shortest non-trivial linear combination that gives **0** (may need to re-order $\mathbf{x}_i$). Now apply $(A - \lambda_1 I)$ to the whole equation to obtain

$$d_1(\lambda_1 - \lambda_1)\mathbf{x}_1 + d_2(\lambda_2 - \lambda_1)\mathbf{x}_2 + \cdots + d_r(\lambda_r - \lambda_1)\mathbf{x}_r = 0$$

We know that the first term is **0**, while the others are not (since we assumed $\lambda_i \neq \lambda_j$ for $i \neq j$). So

$$d_2(\lambda_2 - \lambda_1)\mathbf{x}_2 + \cdots + d_r(\lambda_r - \lambda_1)\mathbf{x}_r = 0,$$

and we have found a shorter linear combination that gives **0**. Contradiction. □

**Example.** (i) Consider the reflection $R$ in the plane with normal $\mathbf{n}$. Clearly $R\mathbf{n} = -\mathbf{n}$. the eigenvalue is $-1$ and the eigenvector is $\mathbf{n}$. Then $E_1 = \text{span}\{\mathbf{n}\}$. So $M(-1) = m(-1) = 1$.

If $\mathbf{p}$ is any vector in the plane, $R\mathbf{p} = \mathbf{p}$. So this has an eigenvalue of 1 and eigenvectors being any vector in the plane. So $M(1) = m(1) = 2$. So the eigenvectors form a basis of $\mathbb{R}^3$.

(ii) Consider a rotation $R$ by $\theta$ about $\mathbf{n}$. Since $R\mathbf{n} = \mathbf{n}$, we have an eigenvalue of 1 and eigenspace $E_1 = \text{span}\{\mathbf{n}\}$.

We know there are no other real eigenvalues since rotation changes direction of any other vector. The other eigenvalues turn out to be $e^{\pm i\theta}$. if $\theta \neq 0$, there are 3 distinct eigenvalues and the eigenvectors form a basis of $\mathbb{C}^3$.

**Note.** If $n \times n$ matrix $A$ has $n$ distinct eigenvalues, and hence $n$ linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_n$, then *with respect to this eigenvector basis*, $A$ is diagonal with entries the eigenvalues.

## 5.3 Transformation matrices

How do the components of a vector or a matrix change when we change the basis?

Let $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ and $\{\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \cdots, \tilde{\mathbf{e}}_n\}$ be 2 different bases of $\mathbb{R}^n$ or $\mathbb{C}^n$. Then we can write

$$\tilde{\mathbf{e}_j} = \sum_{i=1}^n P_{ij}\mathbf{e}_i$$

i.e. $P_{ij}$ is the $i$th component of $\tilde{\mathbf{e}_j}$ with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$.

Matrix $P$ has as its columns the vectors $\tilde{\mathbf{e}_j}$ relative to $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$. So $P = (\tilde{\mathbf{e}_1}\ \tilde{\mathbf{e}_2}\ \cdots\ \tilde{\mathbf{e}_n})$ and

$$P(\mathbf{e}_i) = \tilde{\mathbf{e}}_i$$

Similarly, we can write

$$\mathbf{e}_i = \sum_{k=1}^n Q_{ki}\tilde{\mathbf{e}_k}$$

Substituting this into the equation for $\tilde{\mathbf{e}_j}$, we have

$$\tilde{\mathbf{e}_j} = \sum_{i=1}^n \left(\sum_{k=1}^n Q_{ki}\tilde{\mathbf{e}_k}\right) P_{ij} = \sum_{k=1}^n \tilde{\mathbf{e}_k}\left(\sum_{i=1}^n Q_{ki}P_{ij}\right)$$

But $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \cdots, \tilde{\mathbf{e}}_n$ are linearly independent, so this is only possible if

$$\sum_{i=1}^n Q_{ki}P_{ij} = \delta_{kj},$$

which is just a fancy way of saying $QP = I$, or $Q = P^{-1}$.

### 5.3.1 Transformation law for vectors

**Theorem.** Denote vector as $\mathbf{u}$ with respect to $\{\mathbf{e}_i\}$ and $\tilde{\mathbf{u}}$ with respect to $\{\tilde{\mathbf{e}_i}\}$. Then

$$\mathbf{u} = P\tilde{\mathbf{u}} \text{ and } \tilde{\mathbf{u}} = P^{-1}\mathbf{u}$$

*Proof.* This is clear from proof above. □

### 5.3.2 Transformation law for matrix

Consider a linear map $\alpha : \mathbb{C}^n \to \mathbb{C}^n$ with associated $n \times n$ matrix $A$. We have

$$\mathbf{u}' = \alpha(\mathbf{u}) = A\mathbf{u}.$$

Denote $\mathbf{u}$ and $\mathbf{u}'$ w.r.t basis $\{\mathbf{e}_i\}$ , and $\tilde{\mathbf{u}}, \tilde{\mathbf{u}}'$ w.r.t basis$\{\tilde{\mathbf{e}_i}\}$. Now we have:

$$\mathbf{u}' = A\mathbf{u}P\tilde{\mathbf{u}}' = AP\tilde{\mathbf{u}}$$

$$\tilde{\mathbf{u}}' = P^{-1}AP\tilde{\mathbf{u}} = \tilde{A}\tilde{\mathbf{u}}$$

**Theorem.**

$$\tilde{A} = P^{-1}AP.$$

More generally, given $\alpha : \mathbb{C}^m \to \mathbb{C}^n$, given $\mathbf{x} \in \mathbb{C}^m$, $\mathbf{x}' \in \mathbb{C}^n$ with $\mathbf{x}' = A\mathbf{x}$. We know that $A$ is an $n \times m$ matrix. Suppose $\mathbb{C}^m$ has a basis $\{\mathbf{e}_i\}$ and $\mathbb{C}^n$ has a basis $\{\mathbf{f}_i\}$.Now change bases to $\{\tilde{\mathbf{e}_i}\}$ and $\{\tilde{\mathbf{f}_i}\}$.

We know that $\mathbf{x} = P\tilde{\mathbf{x}}$ with $P$ being an $m \times m$ matrix, with $\mathbf{x}' = R\tilde{\mathbf{x}}'$ with $R$ being an $n \times n$ matrix. Combining both of these, we have

$$\tilde{A} = R^{-1}AP$$

.

## 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

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

**Proposition.** Similar matrices have the same trace, determinant, and characteristic polynomial.

*Proof.* Determinant and trace are the coefficients of the characteristic polynomial, so only need to prove last one.

$$\begin{aligned} p_B(\lambda) &= \det(B - \lambda I) \\ &= \det(P^{-1}AP - \lambda P^{-1}IP) = \det(P^{-1}AP - \lambda I) \\ &= \det(P^{-1}(A - \lambda I)P) = \det(A - \lambda I) = p_A(\lambda) \end{aligned}$$

□

## 5.5 Diagonalizable matrices

**Definition** (Diagonalizable matrices)**.** An $n \times n$ matrix $A$ is *diagonalizable* if it is similar to a diagonal matrix.

The requirement that matrix $A$ has $n$ distinct eigenvalues is a *sufficient* condition for diagonalizability as shown above. However, it is *not* necessary.

**Theorem.** Let $\lambda_1, \lambda_2, \cdots, \lambda_r$, with $r \leq n$ be distinct eigenvalues of $A$. Let $B_1, B_2, \cdots B_r$ be bases of the eigenspaces $E_{\lambda_1}, E_{\lambda_2}, \cdots, E_{\lambda_r}$ respectively. Then the set $B = \bigcup_{i=1}^{r} B_i$ is linearly independent.

*Proof.* Write $B_1 = \{\mathbf{x}_1^{(1)}, \mathbf{x}_2^{(1)}, \cdots \mathbf{x}_{m(\lambda_1)}^{(1)}\}$. Then $m(\lambda_1) = \dim(E_{\lambda_1})$, and similarly for all $B_i$.

Consider the following general linear combination of all elements in $B$. Consider the equation

$$\sum_{i=1}^{r} \sum_{j=1}^{m(\lambda_i)} \alpha_{ij} \mathbf{x}_j^{(i)} = 0.$$

The first sum is summing over all eigenspaces, and the second sum sums over the basis vectors in $B_i$. Now apply the matrix

$$\prod_{k=1,2,\cdots,\bar{K},\cdots,r} (A - \lambda_k I)$$

to the above sum, for some arbitrary $K$. We obtain

$$\sum_{j=1}^{m(\lambda_K)} \alpha_{Kj} \left[ \prod_{k=1,2,\cdots,\bar{K},\cdots,r} (\lambda_K - \lambda_k) \right] \mathbf{x}_j^{(K)} = 0.$$

Since $\mathbf{x}^{(j)}$ are linearly independent, since $B_k$ is a basis, $\alpha_{kj} = 0$. So $B$ is linearly independent. □

**Proposition.** $A$ is diagonalizable iff all its eigenvalues have zero defect.

### 5.6 Canonical (Jordan normal) form

**Theorem.** Any $2 \times 2$ complex matrix $A$ is similar to exactly one of

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

*Proof.* For each case:

(i) If $A$ has two distinct eigenvalues, then eigenvectors are linearly independent. Then we can use $P$ formed from eigenvectors as its columns

(ii) If $\lambda_1 = \lambda_2 = \lambda$ and $\dim E_\lambda = 2$, then $A = \lambda I$ and $\tilde{A} = P^{-1}AP = \lambda I$. Since $\tilde{A} = A$, A is *isotropic.*

(iii) If $\lambda_1 = \lambda_2 = \lambda$ , then $A = \lambda I$and $\dim(E_\lambda) = 1$, then $E_\lambda = \text{span}\{\mathbf{v}\}$. Now choose basis of $\mathbb{C}^2$ as $\{\mathbf{v}, \mathbf{w}\}$. So $A\mathbf{w} = \alpha\mathbf{v} + \beta\mathbf{w}$. Hence, if we change basis to $\{\mathbf{v}, \mathbf{w}\}$, then $\tilde{A} = P^{-1}AP = \begin{pmatrix} \lambda & \alpha \\ 0 & \beta \end{pmatrix}$.

However, $\tilde{A}$ has eigenvalue $\lambda$ with algebraic multiplicity 2. From its determinant, we have $\beta = \lambda$. To make $\alpha = 1$, let $\mathbf{u} = (\tilde{A} - \lambda I)\mathbf{w}$. We know $\mathbf{u} \neq \mathbf{0}$ (Why?). Then

$$(\tilde{A} - \lambda I)\mathbf{u} = (\tilde{A} - \lambda I)^2\mathbf{w} = \begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix}\mathbf{w} = \mathbf{0}.$$

So $\mathbf{u}$ is an eigenvector of $\tilde{A}$ with eigenvalue $\lambda$.

We have $\mathbf{u} = \tilde{A}\mathbf{w} - \lambda\mathbf{w}$. Su $\tilde{A}\mathbf{w} = \mathbf{u} + \lambda\mathbf{w}$.

Change basis to $\{\mathbf{u}, \mathbf{w}\}$. Then $A$ with respect to this basis is $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

□

**Proposition.** There exists a similarity transform for every $n \times n$ matrix such that $A$ is similar to a matrix $\tilde{A}$ that satisfies one of the following:

(i) $\tilde{A}_{\alpha\alpha} = \lambda_\alpha$, i.e. the diagonal composes of the eigenvalues.

(ii) $\tilde{A}_{\alpha,\alpha+1} = 0$ or 1.

(iii) $\tilde{A}_{ij} = 0$ otherwise.

**Theorem** (Cayley-Hamilton Theorem)**.** If $\det A \neq 0$, the matrix satisfies its own characteristic equation (constant term interpreted as constant multiples of identity).

*Proof.* We will only prove for diagonalizable matrices here, i.e. $\exists P$ such that $D = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) = P^{-1}AP$. Note that

$$D^i = (P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP) = P^{-1}A^iP.$$

Using result above and the fact that similar matrices have the same characteristic polynomial:

$$p_A(D) = p_A(P^{-1}AP) = P^{-1}[p_D(A)]P = P^{-1}[p_A(A)]P$$

However, we also know that $D^i = \text{diag}(\lambda_1^i, \lambda_2^i, \cdots \lambda_n^i)$. So

$$p_A(D) = \text{diag}(p_A(\lambda_1), p_A(\lambda_2), \cdots, p_A(\lambda_n)) = \text{diag}(0, 0, \cdots, 0)$$

So $0 = p_A(D) = P^{-1}p_A(A)P$ and $p_A(A) = 0$. □

**Remark.** (i) If $A^{-1}$ exists, then we can use this to calculate $A^{-1}$ from positive powers of A.

(ii)

$$e^A = I + A + \frac{A^2}{2!} + \ldots + \frac{A^n}{n!} + \ldots$$

$$P^{-1}e^AP = P^{-1}IP + P^{-1}AP + \frac{P^{-1}A^2P}{2!} + \ldots$$

$$= I + D + \frac{D^2}{2!} + \ldots = diag[e^{\lambda_1}, e^{\lambda_2}, \ldots]$$

So the exponential can be easily calculated if the matrix is diagonalizable.

## 5.7 Eigenvalues +Eigenvectors of a Hermitian matrix

**Theorem.** The eigenvalues of a Hermitian matrix H are real.

*Proof.*

$$H\mathbf{v} = \lambda\mathbf{v}$$

$$\mathbf{v}^\dagger H\mathbf{v} = \lambda\mathbf{v}^\dagger\mathbf{v}$$

Take hermitian conjugate and we have:

$$\mathbf{v}^\dagger H\mathbf{v} = \lambda^*\mathbf{v}^\dagger\mathbf{v}$$

So we have $0 = (\lambda - \lambda^*)\mathbf{v}^\dagger\mathbf{v}$ Since $\mathbf{v}^\dagger\mathbf{v} \neq 0$,:

$$\lambda = \lambda^*$$

So $\lambda$ is real.

□

**Theorem.** The eigenvectors of Hermitian matrix corresponding to distinct eigenvalues are orthogonal.

*Proof.*

$$H\mathbf{v_1} = \lambda_1 v_1$$

$$H\mathbf{v_2} = \lambda_2 v_2$$

$$\mathbf{v_2}^\dagger H\mathbf{v_1} = \lambda_1 v_2^\dagger v_1$$

Conjugating that, we have:

$$\mathbf{v_1}^\dagger H\mathbf{v_2} = \lambda_1 v_1^\dagger v_2$$

But we also have:

$$\mathbf{v_1}^\dagger H\mathbf{v_2} = \lambda_2 v_1^\dagger v_2$$

Subtracting these and noting that $\lambda_1 \neq \lambda_2$, we have:

$$v_1^\dagger v_2 = 0$$

Same argument with other half of proof to show $v_2^\dagger v_1 = 0$ □

This means that if there are $n$ distinct eigenvalues then the eigenvectors form an orthonormal basis. The case of degenerate eigenvalues is more difficult and requires the Gram-Schmidt process.

### 5.8 Gram-Schmidt Orthogonalization

Suppose we have s set of $B = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_2\}$ of linearly independent vectors, want to find an orthogonal set.

Define the projection of $\mathbf{w}$ onto $\mathbf{v}$ by

$$P_{\mathbf{v}(\mathbf{w})} = \frac{\langle \mathbf{v}|\mathbf{w}\rangle}{\langle \mathbf{v}|\mathbf{v}\rangle}\mathbf{v}$$

We let $\mathbf{v_1} = \mathbf{w_1}$ Then $\mathbf{v_2} = \mathbf{w_2} - P_{\mathbf{v_1}}(\mathbf{w_2})$ Repeat this process to eliminate the "projections on each fixed basis vectors to find a new one". Then, we would get orthogonal set of vectors $\{v_1, ...v_k\}$.

### 5.9 Unitary Transformation and Hermitian Matrices

**Theorem.** Suppose $U$ is the transformation between one orthonormal basis and a new orthonormal basis, then $U$ is Unitary.

**Theorem.** An $n \times n$ Hermitian matrix has precisely $n$ orthogonal eigenvectors.

### 5.10 Normal Matrices

**Definition.** 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.

**Proposition.** It can be shown that:

(i) If $\lambda$ is an eigenvalue of $N$, then so is $\lambda^*$.

(ii) The eigenvectors of distinct eigenvalues are orthogonal.

(iii) A normal matrix can always be diagonalized with an orthonormal basis of eigenvectors.

# 6 Quadratic forms and conics

**Definition.** 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*

**Theorem.** Hermitian forms are real.

*Proof.* $(\mathbf{x}^\dagger H\mathbf{x})^* = (\mathbf{x}^\dagger H\mathbf{x})^\dagger = \mathbf{x}^\dagger H^\dagger\mathbf{x} = \mathbf{x}^\dagger H\mathbf{x}$. So $(\mathbf{x}^\dagger H\mathbf{x})^* = \mathbf{x}^\dagger H\mathbf{x}$.

□

After diagonalization, we have So $F(\mathbf{x}) = \mathbf{x}^\dagger H\mathbf{x} = \mathbf{x}^\dagger UDU^\dagger\mathbf{x} = (\mathbf{x}')^\dagger D\mathbf{x}' = \sum_{i=1}^n \lambda_i |x_i'|^2$. where $\mathbf{x}' = U^\dagger\mathbf{x}$. The eigenvectors are known as the *principal axes.*

### 6.1 Quadrics and conics

**Definition.** A *quadric* is an $n$-dimensional surface defined by the zero of a real quadratic polynomial, as in

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

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

Anti-symmetric matrix has $\mathbf{x}^T A\mathbf{x} = 0$, so we retain the symmetric part of a matrix:

$$S = \frac{1}{2}(A + A^T)$$

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

S is real and symmetric so diagonalizable using orthogonal matrices:

$$S = QDQ^T$$

Take $\mathbf{x}' = Q^T\mathbf{x}$, $\mathbf{b}' = Q^T\mathbf{b}$. Then we have:

$$\mathbf{x}'^T D\mathbf{x}' + \mathbf{b}'^T\mathbf{x}' + c = 0$$

If S is invertible, then we write

$$\mathbf{x}'' = \mathbf{x}' + \frac{1}{2}D^{-1}\mathbf{b}'$$

This shift eliminates the linear term in $\mathbf{x}'$. So we finally have (dropping the superfix, $k$ is a constant):

$$\mathbf{x}^T D\mathbf{x} = k$$

### 6.1.1 Conic sections($n = 2$)

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 = k$$

$\lambda_1\lambda_2 > 0$ Then the solution is an ellipse coinciding with eigenvectors of S. (obviously $k$ needs to have the right sign)

$\lambda_1\lambda_2 < 0$ Solution is hyperbole.

$\lambda_1$ or $\lambda_2 = 0$ We have to re-shift the origin (return to original $\lambda_1 x_1'^2 + b_1' x_1' + b_2' x_2' + c = 0$) and shift origin according to:

$$x_1'' = x_1' + \frac{b_1'}{2\lambda_1}$$

$$x_2'' = x_2' + \frac{c}{b_2'} - \frac{b_1'^2}{4\lambda_2 b_2'} \; b_2' \neq 0$$

$x_1'$ term is removed so we are left with a parabola.

$\lambda_1$ or $\lambda_2 = 0, b_2' = 0$ We have either two straight lines or no solution as other linear term is also eliminated

### 6.2 Focus-directrix property

Consider 2 control parameters: $e$ the *eccentricity* and $a$ the *scale*.

**Definition.** Foci$\equiv (\pm ae, 0)$ two points.

Directrices $\equiv x \equiv \pm\frac{a}{e}$ two lines.

Conic section is set of points with property that the distance from focus = e$\times$distance from directrix which is closer to that focus (unless e=1, then take distance to other directrix)

$e = 0$ It is a circle. (See ellipse below)

$e < 1$ By definition:

$$((x - ae)^2 + y^2)^{\frac{1}{2}} = e(\frac{a}{e} - x)$$

We have an eclipse with semi-major axis $a$ and semi-minor axis $a\sqrt{1 - e^2}$.

$e = 1$ We have $y^2 = 4ax$, which is an equation for the parabola.

$e > 1$ We have an hyperbola. (see ellipse, $1 - e^2 < 0$)

This also works in the polar coordinates. Introduce new parameter l, such that $\frac{l}{e}$ =distance from focus to directrix. Using the equation, we have $l = a|1 - e^2|$.

Use polar coordinates centred on the focus. Using focus-directrix property, we have:

$$r = e(\frac{l}{e} - r\cos\theta)$$

$$r = \frac{l}{1 + e\cos\theta}$$

This is valid for all e.

# 7 Transformation Group

### 7.1 Groups of orthogonal matrices

The set of all $n \times n$ orthogonal matrices $P$ forms a group $O(n)$ under matrix multiplication. (See groups) It has a subgroup $SO(n)$, which are matrices of $n \times n$ of determinant 1.

### 7.2 Length Preserving Matrices

Let $P \in O(n)$, then the following are equivalent:

(i) P orthogonal

(ii) $|P\mathbf{x}| = |x|$

(iii) It preserves the scalar product.

(iv) If $(v_1, \cdots, c_n)$ are orthonormal so are $(Pv_1, \cdots, Pv_n)$.

(v) Columns of P are orthonormal.

*Proof.*

$(i) \Rightarrow (ii)$

$$|P\mathbf{x}|^2 = (Px)^T Px = x^T P^T Px = x^T x = |x|^2$$

$(ii) \Rightarrow (iii)$

$$|x|^2+|y|^2+2(Px)^T Py = |Px|^2+|Py|^2+2(Px)^T Py = |P(x+y)|^2 = |x+y|^2 = |x|^2+|y|^2+2x^T y$$

$(iii) \Rightarrow (iv)$ If $v_1^T v_2 = I$, then clearly from (iii), $(Pv_1)^T (Pv_2) = I$

$(iv) \Rightarrow (v)$ From (iv) and take the set of $v_n$ as the standard basis, then the columns of P are $Pv_1$, which are orthonormal.

$(v) \Rightarrow (i)$ Columns of $P$ orthonormal means $PP^T = I$ so P is orthogonal.

□

### 7.3 Lorentz Transformation

Consider Minkowski (1+1) dimensional spacetime. The Minkowski inner product of 2 vectors are:

$$\langle x|y\rangle = x^T Jy$$

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

This follows that:

$$\langle x|y\rangle = x_1 y_1 - x_2 y_2$$

Now consider transformation matrix that preserves the Minkowski inner product:

$$x^T J y = (Mx)^T JMy = x^T M^T JMy$$
$$\Rightarrow J = M^T JM$$

We can show M is either:

$$\begin{pmatrix} \cosh\alpha & \sinh\alpha \\ \sinh\alpha & \cosh\alpha \end{pmatrix} = H_\alpha \text{ (hyperbolic rotation) or } \begin{pmatrix} \cosh\alpha & -\sinh\alpha \\ \sinh\alpha & -\cosh\alpha \end{pmatrix} = K_{\frac{\alpha}{2}} \text{ (hyperbolic reflection)}$$

Note that we insist $M_{11} > 0$, since if not, we would map future into past (physicists don't like it). Define Lorentz matrix

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

Here $|v| < 1$, units are chosen such that speed of light is 1. Then we have:

$$B_v = H_{\tanh^{-1}(v)}$$

$B_v$ is called a *Lorentz boost*. The set of all $Bv$ forms a group (The Lorentz group), easily proven. For $v_1$ and $v_2$, their "addition" is:

$$v_3 = \tanh(\tanh^{-1} v_1 + \tanh^{-1} v_2) = \frac{v_1 + v_2}{1 + v_1 v_2}$$

$Bv$ is a group of transformations that preserve the Minkowski inner product.