# Geometry and Linear Algebra

Sayan Das (dassayan0013@gmail.com)

December 22, 2023

**Geometry**
Poles and polars. Diameters and conjugate diameters. Systems of conics. Polar equation of a conic referred to a focus as pole. Equations of tangent, normal, chord of contact. Sphere: General Equation. Great circle, Sphere through the intersection of two spheres. Radical Plane, Tangent, Normal. Cone: Right circular cone. General homogeneous second degree equation. Section of cone by a plane as a conic and as a pair of lines. Condition for three perpendicular generators. Reciprocal cone. Cylinder: Generators parallel to either of the axes, general form of equation. Right-circular cylinder. Ellipsoid, Hyperboloid, Paraboloid: Canonical equations only. Tangent planes, Normal, Enveloping cone. Generating lines of hyperboloid of one sheet and hyperbolic paraboloid.

**Vector Spaces**
Vectors in $\mathbb{R}^n$, notions of linear independence and dependence, linear span of a set of vectors, vector subspaces of $\mathbb{R}^n$, basis of a vector subspace. Systems of linear equations, matrices and Gauss elimination, row space, null space, column space, rank of a matrix. Vector spaces (over $\mathbb{R}$ or $\mathbb{C}$), subspaces, algebra of subspaces, quotient spaces, linear combination of vectors, linear span, linear independence, basis and dimension.

**Linear Transformations**
Linear transformations, null space, range, rank and nullity of a linear transformation, matrix representation of a linear transformation, Different notion of matrices, Eigen values, Eigen vectors and characteristic equation of a matrix. Cayley-Hamilton theorem. Algebra of linear transformations. Isomorphisms. Isomorphism theorems, invertibility and isomorphisms, change of coordinate matrix. [50]

## Contents

# §0 Preliminaries

**Definition 0.1** (Kronecker delta)

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

We have

$$\begin{pmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3.$$

So the Kronecker delta represents an identity matrix.

**Definition 0.2**

The **transpose** of an $m \times n$ matrix $A$ is the $n \times m$ matrix $A^T$ obtained by interchanging the rows and columns of $A$:

$$(A^T)_{ij} = A_{ji}.$$

The **conjugate transpose** of an $m \times n$ matrix $A$ is the $n \times m$ matrix $A^\dagger$ defined as:

$$A^\dagger = \overline{A^T}.$$

**Definition 0.3**

$A \in M_{m \times n}(\mathbb{F})$ is:

1. **upper triangular** iff $i > j \implies A_{ij} = 0$.
2. **lower triangular** iff $i < j \implies A_{ij} = 0$.
3. **orthogonal** iff $m = n$ (square) and $AA^T = A^T A = I_n$.
4. **diagonal** iff $i \neq j \implies A_{ij} = 0$.
5. **symmetric** iff $m = n$ (square) and $A = A^T$.
6. **skew-symmetric** or **anti-symmetric** iff $m = n$ (square) and $A = -A^T$.

7. **self-adjoint** or **Hermitian** iff $m = n$ (square), $\mathbb{F} = \mathbb{C}$ and $A = A^\dagger$ (equal to its complex conjugate transpose); in particular, a $2 \times 2$ matrix over $\mathbb{C}$ is Hermitian iff it has the form

$$\begin{pmatrix} z & x+iy \\ x-iy & w \end{pmatrix}$$

s.t. $w, x, y, z \in \mathbb{R}$.

8. **skew-Hermitian** or **anti-Hermitian** iff $m = n$ (square), $\mathbb{F} = \mathbb{C}$ and $A = -A^\dagger$.

9. **positive semi-definite (positive-definite)** iff $A$ is Hermitian and the real number $z^\dagger Az$ is nonnegative (positive) $\forall\ \mathbf{0} \neq \mathbf{z} \in M_{1\times n}(\mathbb{C})$.

10. **normal** iff $m = n$ (square), $\mathbb{F} = \mathbb{C}$ and $AA^\dagger = A^\dagger A$.

11. **unitary** iff $m = n$ (square), $\mathbb{F} = \mathbb{C}$ and $AA^\dagger = A^\dagger A = I_n$.

12. **similar** to $B \in M_{m\times n(\mathbb{F})}$ iff $m = n$ (square) and $B = P^{-1}AP$ for some invertible matrix $P \in M_{n\times n}(\mathbb{F})$.

13. **congruent** to $B \in M_{m\times n(\mathbb{F})}$ iff $m = n$ (square) and $B = P^T AP$ for some invertible matrix $P \in M_{n\times n}(\mathbb{F})$.

**Remark.** All symmetric matrices are Hermitian. All orthogonal matrices are unitary. All unitary matrices are normal.

**Definition 0.4**

The **trace** of $A \in M_{n\times n}(\mathbb{F})$ is defined as the sum of its diagonal entries:

$$\operatorname{tr}(A) = \sum_{i=1}^{n} A_{ii}.$$

**Fact 0.5**

If $A$ is Hermitian or skew-Hermitian then $\operatorname{tr}(A) \in \mathbb{R}$ always.

# §1 Geometry

## §1.1 Direction ratios and cosines

Let $\{\mathbf{e_k}\}_{k=1}^{n}$ denote the *standard basis* in Euclidean $n$-space. Exempli gratia, in 3-space we have $\{\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}\} = \{(1,0,0), (0,1,0), (0,0,1)\}$.

If $\mathbf{u} = \{u_1, \ldots, u_n\}$ be a unit vector in $n$-space, its components are its **direction cosines**. For, if $\theta_k$ be the small angle betwixt $\mathbf{u}$ and $\mathbf{e_k}$, then

$$\mathbf{u_k} = \mathbf{u} \cdot \mathbf{e_k} = \cos\theta_k$$

and a vector $\mathbf{x}$ is proportional to $\mathbf{u}$ iff $\mathbf{x} = \lambda\mathbf{u}$ for some real scalar $\lambda$. So,

$$\mathbf{x_k} = \lambda\mathbf{u_k}\ \forall k.$$

These components of $\mathbf{x}$ are called **direction ratios**. Equivalently, direction ratios are *homogeneous coordinates in projective* $(n-1)$*-space.*

## §1.2 Lines

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

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

By crossing both sides of the equation with $\mathbf{t}$, we have

**Theorem 1.1**

The equation of a straight line through $\mathbf{a}$ and parallel to $\mathbf{t}$ is

$$(\mathbf{x} - \mathbf{a}) \times \mathbf{t} = \mathbf{0} \text{ or } \mathbf{x} \times \mathbf{t} = \mathbf{a} \times \mathbf{t}.$$

## §1.3 Plane

To define a plane $\Pi$, we need a normal $\mathbf{n}$ to the plane and a fixed point $\mathbf{b}$. For any $\mathbf{x} \in \Pi$, the vector $\mathbf{x} - \mathbf{b}$ is contained in the plane and is thus normal to $\mathbf{n}$, i.e. $(\mathbf{x} - \mathbf{b}) \cdot \mathbf{n} = 0$.

**Theorem 1.2**

The equation of a plane through $\mathbf{b}$ with normal $\mathbf{n}$ is given by

$$\mathbf{x} \cdot \mathbf{n} = \mathbf{b} \cdot \mathbf{n}.$$

If $\mathbf{n} = \hat{\mathbf{n}}$ is a unit normal, then $d = \mathbf{x} \cdot \hat{\mathbf{n}} = \mathbf{b} \cdot \hat{\mathbf{n}}$ is the perpendicular distance from the origin to $\Pi$.

Alternatively, if $\mathbf{a}, \mathbf{b}, \mathbf{c}$ lie in the plane, then the equation of the plane is

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

**Example 1.3**

1. Consider the intersection betwixt a line $\mathbf{x} \times \mathbf{t} = \mathbf{a} \times \mathbf{t}$ with the plane $\mathbf{x} \cdot \mathbf{n} = \mathbf{b} \cdot \mathbf{n}$.
   Cross $\mathbf{n}$ on the right with the line equation to obtain

   $$(\mathbf{x} \cdot \mathbf{n})\mathbf{t} - (\mathbf{t} \cdot \mathbf{n})\mathbf{x} = (\mathbf{a} \times \mathbf{t}) \times \mathbf{n}$$

   Eliminate $\mathbf{x} \cdot \mathbf{n}$ using $\mathbf{x} \cdot \mathbf{n} = \mathbf{b} \cdot \mathbf{n}$

   $$(\mathbf{t} \cdot \mathbf{n})\mathbf{x} = (\mathbf{b} \cdot \mathbf{n})\mathbf{t} - (\mathbf{a} \times \mathbf{t}) \times \mathbf{n}$$

   Provided $\mathbf{t} \cdot \mathbf{n}$ is non-zero, the point of intersection is

   $$\mathbf{x} = \frac{(\mathbf{b} \cdot \mathbf{n})\mathbf{t} - (\mathbf{a} \times \mathbf{t}) \times \mathbf{n}}{\mathbf{t} \cdot \mathbf{n}}.$$

   Exercise: what if $\mathbf{t} \cdot \mathbf{n} = 0$?

2. Shortest distance betwixt two lines. Let $L_1$ be $(\mathbf{x} - \mathbf{a}_1) \times \mathbf{t}_1 = \mathbf{0}$ and $L_2$ be $(\mathbf{x} - \mathbf{a}_2) \times \mathbf{t}_2 = \mathbf{0}$.

The distance of closest approach $s$ is along a line perpendicular to both $L_1$ and $L_2$, i.e. 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 Vector Spaces

## §2.1 Vector Spaces and Subspaces

**Definition 2.1** (Vector Space)

An **$\mathbb{F}$-vector space**, **linear space** or **vector space over $\mathbb{F}$** is an *abelian group of vectors* $(V, +)$ with a binary operation (*scalar multiplication*) defined from an *underlying field of scalars* $\mathbb{F}$ to $V$ as $\mathbb{F} \times V \to V : (\lambda, \mathbf{v}) \mapsto \lambda\mathbf{v}$ s.t. the foll. axioms hold:

1. *Vector addition*: $\mathbf{u}, \mathbf{v} \in V \implies \mathbf{u} + \mathbf{v} \in V$.
   a) *Commutativity*: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}\ \forall u, v \in V$.
   b) *Associativity*: $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}\ \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.
   c) *Zero vector*: $\exists \mathbf{0} \in V$ (called the *origin*) s.t. $\mathbf{v} = \mathbf{v} + \mathbf{0} = \mathbf{0} + \mathbf{v}\ \forall \mathbf{v} \in V$.
   d) *Additive Inverse*: $\forall \mathbf{v} \in V\ \exists (-\mathbf{v}) \in V : \mathbf{v} + (-\mathbf{v}) = (-\mathbf{v}) + \mathbf{v} = \mathbf{0}$.
2. *Scalar multiplication*: $\lambda \in \mathbb{F},\ \mathbf{v} \in V \implies \lambda\mathbf{v} \in V$.
   a) *Multiplicative Identity*: $\mathbf{v} = \mathbf{v} \cdot 1 = 1 \cdot \mathbf{v}\ \forall \mathbf{v} \in V$.
   b) *Associativity*: $(\lambda\mu)\mathbf{v} = \lambda(\mu\mathbf{v})\ \forall \mathbf{v} \in V$.
3. *Distributivity*:
   a) *Distributivity in $V$*: $\lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v}\ \forall \lambda \in \mathbb{F},\ \mathbf{u}, \mathbf{v} \in V$.
   b) *Distributivity in $\mathbb{F}$*: $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}\ \forall \lambda, \mu \in \mathbb{F},\ \mathbf{v} \in V$.

**Example 2.2**

The set of $n \times n$ complex Hermitian matrices is

1. *an $\mathbb{R}$-vector space.*
2. *not a $\mathbb{C}$-vector space* as the identity matrix $I_n$ is Hermitian but $iI_n$ is not.

**Theorem 2.3** (Cancellation Law)

Let $V$ be an $\mathbb{F}$-vector space and $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ s.t. $\mathbf{x} + \mathbf{z} = \mathbf{y} + \mathbf{z}$, then $\mathbf{x} = \mathbf{y}$.

*Proof.* From the vector space axioms, $\exists(-z) \in V$ s.t. $z + (-z) = 0$. (additive inverse)

Thus,

$$\mathbf{x}, \mathbf{y}, \mathbf{z} \in V : \mathbf{x} + \mathbf{z} = \mathbf{y} + \mathbf{z} \implies (\mathbf{x} + \mathbf{z}) + (-\mathbf{z}) = (\mathbf{y} + \mathbf{z}) + (-\mathbf{z})$$

$$\implies \mathbf{x} + (\mathbf{z} + (-\mathbf{z})) = \mathbf{y} + (\mathbf{z} + (-\mathbf{z})) \implies \mathbf{x} + \mathbf{0} = \mathbf{y} + \mathbf{0} \implies \mathbf{x} = \mathbf{y}.$$

□

**Definition 2.4** (Subspace)

Let $V$ be an $\mathbb{F}$-vector space. A subset $U \subseteq V$ is an **$\mathbb{F}$-subspace**, or **subspace**, of $V$ iff *$U$ is also an $\mathbb{F}$-vector space.*

**Notation** (Subspace)**.** We use $W \leq V$ to denote that $W$ is a subspace of $V$ (over $\mathbb{F}$ unless stated otherwise).

**Theorem 2.5**

*Let $V$ be an $\mathbb{F}$-vector space and $W \subseteq V$. Then $W$ is an $\mathbb{F}$-subspace of $V$ iff the foll. hold (for the operations defined in $V$):*

1. $\mathbf{0} \in W$.
2. $\mathbf{x}, \mathbf{y} \in W \implies \mathbf{x} + \mathbf{y} \in W$.
3. $\lambda \in \mathbb{F}, \mathbf{x} \in W \implies \lambda\mathbf{x} \in W$.

*Proof.* If $W \leq V$ then $W$ is an $\mathbb{F}$-vector space with the operations defined in $V$ so the conditions 2. and 3. hold trivially. Moreover, there must be some $\mathbf{0}' \in W$ s.t. $\mathbf{x} = \mathbf{x} + \mathbf{0}' = \mathbf{0}' + \mathbf{x}$ $\forall x \in W$; but we also have that $\mathbf{x} \in W \subseteq V$ so $\mathbf{x} = \mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x}$. Thus,

$$\begin{aligned} \mathbf{x} = \mathbf{x} &\iff \mathbf{0}' + \mathbf{x} = \mathbf{0} + \mathbf{x} \\ &\iff \mathbf{0}' + \mathbf{x} + (-\mathbf{x}) = \mathbf{0} + \mathbf{x} + (-\mathbf{x}) \\ &\iff \mathbf{0}' + \mathbf{0} = \mathbf{0} + \mathbf{0} \iff \mathbf{0}' = \mathbf{0}. \end{aligned}$$

So 1. holds.
Now assume that 1., 2., 3. hold. As the vector space axioms $1.(a), (b), 2., 3.$ hold for all vectors in $V$ over $\mathbb{F}$, they automatically hold for any subset $W \subseteq V$. Also condition 1. satisfies axiom $1.(c)$, so we only need to show that axiom $1.(d)$ holds (existence of additive inverses). Sps $\mathbf{x} \in W$, then $(-1)\mathbf{x} = -\mathbf{x} \in W$ by condition 3.; therefore $W \leq V$. □

**Example 2.6** 1. For a vector space $V$, $V \leq V$ itself and $\{\mathbf{0}\} \leq V$ (the **zero subspace** of $V$).

2. The space of polynomial functions over the field $\mathbb{F}$ is a subspace of the space of all functions from $\mathbb{F}$ to $\mathbb{F}$: $P_n(\mathbb{F}) \leq \mathcal{F}(\mathbb{F}, \mathbb{F})$.
3. The set of $n \times n$ symmetric matrices over $\mathbb{F}$ form a subspace of $M_{n \times n}(\mathbb{F})$.
4. The set of $n \times n$ Hermitian matrices *do not* form a subspace of $M_{n \times n}(\mathbb{C})$. For, if $A$ is Hermitian then $\operatorname{tr}(A) \in \mathbb{R}$ always but in general $\operatorname{tr}(iA)$ need not be real.

**Theorem 2.7**

*Any intersection of subspaces of an $\mathbb{F}$-vector space $V$ is a subspace of $V$.*

*Proof.* Let $\mathcal{C}$ be an arbitrary collection of subspaces of $V$, and let $W = \bigcap \mathcal{C}$.

From Thm. 2.8 every subspace in the collection $\mathcal{C}$ contains $\mathbf{0}$, so $\mathbf{0} \in \bigcap \mathcal{C} = W$.

Thus, $W \neq \varnothing$; thus let $\mathbf{x}, \mathbf{y} \in W$ and let $\lambda \in \mathbb{F}$.

By Thm. 2.8 if $\mathbf{x}, \mathbf{y}$ are in every subspace in $\mathcal{C}$ then $\lambda\mathbf{x}$ and $\mathbf{x}+\mathbf{y}$ as well, thus $\lambda\mathbf{x}, \mathbf{x}+\mathbf{y} \in W$. Thus $W \leq V$. □

## §2.2 Linear Combinations and Systems of Linear Equations

**Definition 2.8**

Let $V$ be an $\mathbb{F}$-vector space and $\varnothing \neq S \subseteq V$. A vector $\mathbf{v} \in V$ is called a **linear combination** of vectors of $S$ if there exist a finite number of vectors $\mathbf{u_1}, \dots, \mathbf{u_n} \in \mathbf{S}$ and scalars $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ s.t.

$$\mathbf{v} = \sum_{k=1}^{n} \lambda_k \mathbf{u_k}$$

and we say that $\mathbf{v}$ is a linear combination of $\mathbf{u_1}, \dots, \mathbf{u_n}$ with coefficients $\lambda_1, \dots, \lambda_n$.

**Remark.** The origin $\mathbf{0}$ in any vector space $V$ is trivially a linear combination of any $\varnothing \neq S \subseteq V$.

## §2.3 Linear Dependence and Linear Independence

## §2.4 Bases and Dimension

**Definition 2.9** (Basis)

Let $V$ be an $\mathbb{F}$-vector space. A subset $S \subseteq V$ is called a **basis of**, or **basis for**, $V$ iff

1. S is linearly independent.
2. $\operatorname{span}(S) = V$.

**Example 2.10**

1. Since $\operatorname{span}(\varnothing) = \{\mathbf{0}\}$ and $\varnothing$ is linearly independent, $\varnothing$ is a basis for the zero vector space $\{\mathbf{0}\}$.
2. In $\mathbb{F}^n$ let $\mathbf{e_1} = (1, 0, \dots, 0), \mathbf{e_2} = (0, 1, \dots, 0), \dots, \mathbf{e_n} = (0, 0, \dots, 1)$ and define $S = \{\mathbf{e_1}, \dots, \mathbf{e_n}\}$. Then $\operatorname{span}(S) = \mathbb{F}^n$ and $S$ is linearly independent, so $S$ is a basis of $\mathbb{F}^n$ (in fact the **standard basis**).
3. In $M_{m\times n}(\mathbb{F})$ let $E^{ij} = \mathbf{e_i}^T\mathbf{e_j}$, i.e. the $mxn$ matrix whose only nonzero entry is $E_{ij} = 1$. Then $\{E^{ij} : 1 \leq i \leq m, 1 \leq j \leq n\}$ is a basis for $M_{m\times n}(\mathbb{F})$.
4. In $P_n(\mathbb{F})$ the set $\{1, x, x^2, \dots, x^n\}$ is a basis. We call this the *standard basis* for $P_n(\mathbb{F})$.
5. In $P(\mathbb{F})$ the set $\{1, x, x^2, \dots\}$ is a basis. Observe that a basis need not be finite.

**Theorem 2.11** (Steinitz Exchange Lemma or Replacement Theorem)

*Let $V$ be an $\mathbb{F}$-vector space generated by a set $G$ containing exactly $n$ vectors, and let $L$ be a linearly independent subset of $V$ containing exactly $m$ vectors. Then $m \leq n$ and there exists a subset $H \subseteq G$ containing exactly $n-m$ vectors s.t. $L \cup H$ generates $V$.*

*Proof.* We proceed by induction on $m$. □

## §2.5 Maximal Linearly Independent Subsets

# §3 Linear Transformations

## §3.1 Linear Transformations, Null Spaces and Ranges

## §3.2 Matrix Representations of Linear Transformations

## §3.3 Composition of Linear Transformations and Matrix Multiplication

## §3.4 Invertibility and Isomorphisms

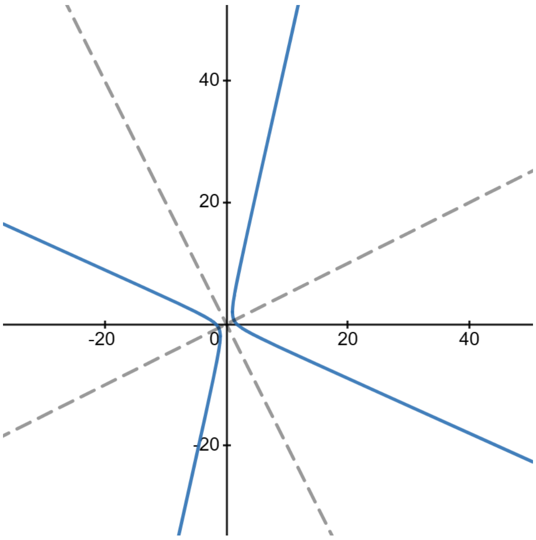
## §3.5 Dual Spaces

# §4 Cayley-Hamilton Theorem

## §4.1 Motivation

**Problem.** *Sps we have the hyperbola:*

$$Q : 2x^2 + 4xy - y^2 = 6.$$




*We want to transform it into the standard form* $\frac{y^2}{b^2} - \frac{x^2}{a^2} = c$*, or* $b^2y^2 - a^2x^2 = c$*.*

**Answer.** Every real symmetric matrix is orthogonally diagonalisable by the spectral theorem (we'll prove this in the next subsection), so take a $2 \times 2$ real symmetric matrix

$$A = \begin{pmatrix} a & h \\ h & b \end{pmatrix}$$

and let

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Then

$$\mathbf{x}^T A\mathbf{x} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & h \\ h & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + hy & hx + by \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$= ax^2 + 2hxy + by^2 = 2x^2 + 4xy - y^2 \ : Q$$

$$\implies a = 2, h = 2, b = -1.$$

So, the matrix representation for the hyperbola $Q$ is

$$A = \begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix}.$$

To diagonalise $A$ we want a non-singular matrix $P$ (i.e. $\det P \neq 0$) s.t.

$$P^{-1}AP = D, P^{-1} = P^T \ (\because PP^T = P^TP = I \text{ as } A \text{ is orthogonally diagonalisable})$$

where $D$ is our required diagonal matrix. Now, we define the *characteristic polynomial* of $A$ as:

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

and equate it to 0 to obtain the *characteristic equation* of $A$:

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

$$\implies \det \begin{pmatrix} 2-\lambda & 2 \\ 2 & -1-\lambda \end{pmatrix} = 0$$

$$(2-\lambda)(-1)(1+\lambda) - 4 = 0 \implies \boxed{\lambda^2 - \lambda - 6 = 0}$$

and the roots of this characteristic equation of $A$ are called the *eigenvalues* of $A$:

$$\lambda^2 - \lambda - 6 = (\lambda + 2)(\lambda - 3) = 0 \implies \boxed{\lambda = -2, 3}.$$

Now, if $\lambda$ is an eigenvalue of $A$, then a non-null vector $\mathbf{v} \in V (= \mathbb{R}^2$ here) is called an *eigenvector* iff

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

and as a matrix can be seen as a representation of a linear transformation, we can write in equivalent linear transformation form:

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

Now, $\lambda = \lambda_1, \lambda_2 = -2, 3$. So,

$$A\mathbf{x} = \lambda_1 \mathbf{x} = -2\mathbf{x} \iff$$

$$\begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -2 \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\iff \begin{cases} 2x + 2y = -2x \\ 2x - y = -2y \end{cases}$$

$$\iff 2x = -y \iff \frac{x}{-1} = \frac{y}{2} = \alpha$$

$$\iff (x, y) = (-\alpha, 2\alpha) = \alpha(-1, 2).$$

Let

$$S_{-2} = \{(x, y) \in \mathbb{R}^2 : -2x = y\} = \{\alpha(-1, 2) : \alpha \in \mathbb{R}\} \subseteq \mathbb{R}^2.$$

Clearly, $\mathbf{0} \in S_{-2}$. If $\mathbf{v_1}, \mathbf{v_2} \in S_{-2} = S_{\lambda_1}$, and $\mu \in \mathbb{F}(= \mathbb{R}$ here), then

$$\begin{cases} A\mathbf{v_1} = \lambda_1 \mathbf{v_1} \implies \mu A\mathbf{v_1} = \mu\lambda_1\mathbf{v_1} \\ A\mathbf{v_2} = \lambda_1 \mathbf{v_2} \end{cases}$$

$$\implies \begin{cases} A(\mathbf{v_1} + \mathbf{v_2}) = \lambda_1(\mathbf{v_1} + \mathbf{v_2}) \\ A(\mu\mathbf{v_1}) = \lambda_1(\mu\mathbf{v_1}) \end{cases}$$

Thus, $S_{\lambda_1}$ is a subspace of $V$, called the *eigenspace* of $A$ i.e. the space of all eigenvectors corresponding to the eigenvalue $\lambda_1$. Similarly, let $S_{\lambda_2} = S_3$ be the eigenspace of $A$ corresponding to $\lambda_2 = 3$.

$$\begin{aligned} A\mathbf{x} = \lambda_2\mathbf{x} = 3\mathbf{x} &\iff \\ \begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= 3\begin{pmatrix} x \\ y \end{pmatrix} \\ \iff &\begin{cases} 2x + 2y = 3x \\ 2x - y = 3y \end{cases} \\ \iff x = 2y &\iff \frac{x}{2} = \frac{y}{1} = \alpha \\ \iff (x, y) &= (2\alpha, \alpha) = \alpha(2, 1). \end{aligned}$$

Then

$$S_3 = \{\alpha(2, 1) : \alpha \in \mathbb{R}\}.$$

$S_{-2}$ is nothing but the vector space of the line $y = -2x$ and $S_3$ is nothing but the vector space of the line $2y = x$, which are the minor and major axes of $Q$ respectively. If $\hat{v_1}, \hat{v_2}$ are unit vectors in $S_{\lambda_1}$ and $S_{\lambda_2}$ respectively, then setting $\alpha = 1$:

$$\begin{cases} \mathbf{v_1} = (-1, 2) \in S_{-2} \\ \mathbf{v_2} = (2, 1) \in S_3 \end{cases}$$

$$\implies \begin{cases} \mathbf{v_1} = (-1, 2) \implies \hat{v_1} = \frac{\mathbf{v_1}}{\|\mathbf{v_1}\|} = \frac{1}{\sqrt{5}}(-1, 2) \\ \mathbf{v_2} = (2, 1) \implies \hat{v_2} = \frac{\mathbf{v_2}}{\|\mathbf{v_2}\|} = \frac{1}{\sqrt{5}}(2, 1). \end{cases}$$

Then our required matrix $P$ for diagonalising $A$ is:

$$P = \begin{pmatrix} \hat{v_1}^T & \hat{v_2}^T \end{pmatrix} = \frac{1}{\sqrt{5}}\begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix}$$

and we can verify that $P$ is indeed orthogonal:

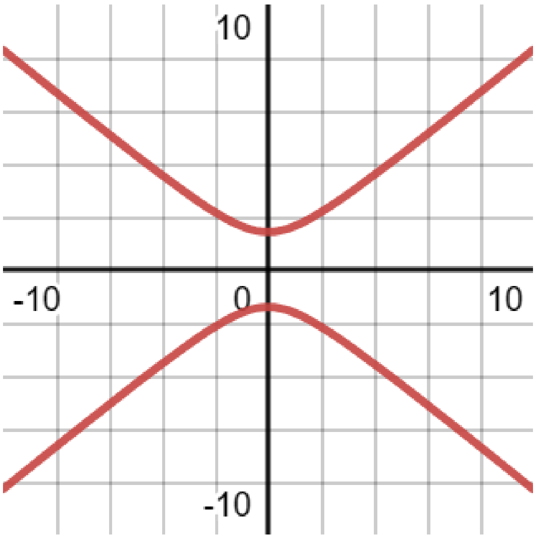
$$\begin{aligned} PP^T &= \frac{1}{\sqrt{5}}\begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix}\frac{1}{\sqrt{5}}\begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2. \end{aligned}$$

Now $P^{-1}AP = P^TAP =$

$$\begin{aligned} &\frac{1}{\sqrt{5}}\begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix}\frac{1}{\sqrt{5}}\begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 2 & -4 \\ 6 & 3 \end{pmatrix}\begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} -10 & 0 \\ 0 & 15 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix} = D. \end{aligned}$$

Thus,

$$\mathbf{x}^T D\mathbf{x} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$
$$= 3y^2 - 2x^2 = 6.$$




What we did here is the same as a change of basis:
from the basis $\{\hat{v}_1, \hat{v}_2\} = \left\{ \frac{1}{\sqrt{5}}(-1, 2), \frac{1}{\sqrt{5}}(2, 1) \right\}$ to the standard basis of $\mathbb{R}^2$.

## §4.2 Eigenvalues and eigenvectors

**Definition 4.1** (Diagonalisable)

A linear operator $T$ on a finite-dimensional $\mathbb{F}$-vector space $V$ is said to be **diagonalisable** iff there is an order basis $\beta$ for $V$ s.t. $[T]_\beta$ is a diagonalisable matrix.
A square matrix $A$ is **diagonalisable** iff $L_A$ is diagonalisable.

**Definition 4.2** (Eigenvector, eigenvalue)

Let $T$ be a linear operator $T$ on a finite-dimensional $\mathbb{F}$-vector space $V$. A nonzero vector $\mathbf{v} \in V$ is called an **eigenvector** of $T$ iff there is a scalar $\lambda \in \mathbb{F}$ s.t.

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

The scalar $\lambda$ is called the ***eigenvalue*** *corresponding to the eigenvector* $\mathbf{v}$.
Equivalently, let $A \in M_{n\times n}(\mathbb{F})$. A nonzero vector $\mathbf{v} \in \mathbb{F}^n$ is called an **eigenvector** of $A$ iff

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

for some scalar (**eigenvalue**) $\lambda \in \mathbb{F}$.

**Theorem 4.3** (Characteristic equation)

Let $A \in M_{n\times n}(\mathbb{F})$. Then a scalar $\lambda \in \mathbb{F}$ is an eigenvalue of $A$ iff it satisfies the

**characteristic equation**

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

.

**Definition 4.4** (Characteristic polynomial)

The **characteristic polynomial** $f(\lambda)$ of a matrix $A \in M_{n\times n}(\mathbb{F})$ is defined as:

$$f(\lambda) = det(A - \lambda I_n).$$

**Lemma 4.5**

*Eigenvalues of symmetric matrices are always real.*

*Proof.* We extend the usual inner product to a $\mathbb{C}$-vector space $V$ by:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_j \overline{u}_j v_j$$

which has the foll. properties:

1. $\langle A^\dagger \mathbf{v}, \mathbf{u} \rangle = \langle \mathbf{v}, A\mathbf{u} \rangle$.
2. $\langle \lambda \mathbf{v}, \mathbf{u} \rangle = \overline{\lambda} \langle \mathbf{v}, \mathbf{u} \rangle$.
3. $\langle \mathbf{v}, \lambda \mathbf{u} \rangle = \lambda \langle \mathbf{v}, \mathbf{u} \rangle$.

Now, for nonzero vectors $\mathbf{v} \in V$:

$$\langle \mathbf{v}, \mathbf{v} \rangle = \sum_j \overline{v}_j v_j = \sum_j |v_j|^2 > 0.$$

Thus,

$$\overline{\lambda}\langle \mathbf{v}, \mathbf{v} \rangle = \langle \lambda \mathbf{v}, \mathbf{v} \rangle = \langle A\mathbf{v}, \mathbf{v} \rangle = \langle A^\dagger \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{v}, A\mathbf{v} \rangle = \langle \mathbf{v}, \lambda \mathbf{v} \rangle = \lambda \langle \mathbf{v}, \mathbf{v} \rangle$$

$$\implies \overline{\lambda} = \lambda \iff \lambda \in \mathbb{R}.$$

□

**Lemma 4.6**

1. *Eigenvectors corresponding to distinct eigenvalues are distinct.*
2. *If $A$ is symmetric then eigenvectors corresponding to distinct eigenvalues are orthogonal.*

*Proof.* Let $A\mathbf{x} = \lambda \mathbf{x}$ and $A\mathbf{x} = \mu \mathbf{x}$ s.t. $\lambda \neq \mu$. Thus,

$$\lambda \mathbf{x} = \mu \mathbf{x} \iff \mathbf{x} = \mathbf{0}.$$

This proves the first part of the lemma. For the second part, sps $A\mathbf{u} = \lambda \mathbf{u}$ and

$A\mathbf{v} = \mu\mathbf{v}$

$$\lambda\langle\mathbf{u},\mathbf{v}\rangle = \langle\lambda\mathbf{u},\mathbf{v}\rangle = \langle A\mathbf{u},\mathbf{v}\rangle = \langle\mathbf{u},A^T\mathbf{v}\rangle = \langle\mathbf{u},A\mathbf{v}\rangle = \langle\mathbf{u},\mu\mathbf{v}\rangle = \mu\langle\mathbf{u},\mathbf{v}\rangle$$

$$\implies \langle\mathbf{u},\mathbf{v}\rangle = 0.$$

□

**Theorem 4.7** (Spectral Theorem)

*Every real symmetric matrix is orthogonally diagonalisable.*

*Proof.* Let $A$ be a real $n \times n$ symmetric matrix. By Lemma 4.4, all eigenvalues of $A$ are real and by Lemma 4.5 the eigenvectors corresponding to distinct eigenvalues of $A$ are are orthogonal. If $n = 1$, the theorem holds trivially (set $P = \begin{pmatrix}1\end{pmatrix}$ and $D = A$). If $n \geq 2$ assume the theorem holds $\forall A \in M_{(n-1)\times(n-1)}(\mathbb{R})$, $A$ being symmetric.

Sps $\lambda$ is an eigenvalue and $\mathbf{v}$ an eigenvector corresponding to $\lambda$. Set

$$\mathbf{q_1} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$$

and extend $\{\mathbf{q_1}\}$ to an orthonormal (i.e. orthogonal and of norm 1) basis

$$\{\mathbf{q_1},\mathbf{q_2},\ldots,\mathbf{q_n}\}$$

of $\mathbb{R}^n$. Let

$$Q = \begin{pmatrix}\mathbf{q_1} & \mathbf{q_2} & \ldots & \mathbf{q_n}\end{pmatrix}.$$

Then

$$AQ = \begin{pmatrix}\lambda\mathbf{q_1} & A\mathbf{q_2} & \ldots & A\mathbf{q_n}\end{pmatrix} = Q\begin{pmatrix}\lambda & \mathbf{d} \\ 0 & A'\end{pmatrix}$$

for some $\mathbf{d} \in M_{1\times(n-1)}(\mathbb{R})$, $A' \in M_{(n-1)\times(n-1)}(\mathbb{R})$. $Q$ being orthonormal by construction imples that

$$Q^TAQ = \begin{pmatrix}\lambda & \mathbf{d} \\ 0 & A'\end{pmatrix}$$

$$\implies Q^TA^TQ = \begin{pmatrix}\lambda & \mathbf{d} \\ 0 & A'\end{pmatrix}^T = \begin{pmatrix}\lambda & 0 \\ \mathbf{d}^T & A'^T\end{pmatrix}$$

but $A = A^T$ so

$$\begin{pmatrix}\lambda & \mathbf{d} \\ 0 & A'\end{pmatrix} = \begin{pmatrix}\lambda & 0 \\ \mathbf{d}^T & A'^T\end{pmatrix}$$

so $d = 0$ and $A' = A'^T$, which being an $(n-1)\times(n-1)$ real symmetric matrix has (by our induction hypothesis) some $Q', D' \in M_{(n-1)\times(n-1)}(\mathbb{R})$ s.t. $Q'$ is orthogonal, $D'$ is diagonal and $A' = Q'D'Q'^T$.

Hence,

$$Q^TAQ = \begin{pmatrix}\lambda & 0 \\ 0 & Q'D'Q'^T\end{pmatrix} = \begin{pmatrix}1 & 0 \\ 0 & Q'\end{pmatrix}\begin{pmatrix}\lambda & 0 \\ 0 & D'\end{pmatrix}\begin{pmatrix}1 & 0 \\ 0 & Q'^T\end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & D' \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix}^T .$$

It follows that

$$A = Q \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & D' \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix}^T Q^T.$$

Setting $P = Q \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix}$ and $D = \begin{pmatrix} \lambda & 0 \\ 0 & D' \end{pmatrix}$, we obtain $A = PDP^T$. It remains to show that $P^T P = I$. Indeed,

$$\begin{aligned} P^T P &= \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix}^T Q^T Q \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix}^T \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & Q'^T \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & Q'^T Q' \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & I_{n-1} \end{pmatrix} \\ &= I_n \end{aligned}$$

which completes the proof. □

**Lemma 4.8**

*A square matrix $A$ of order $n$ is diagonalisable iff $\mathbb{R}^n$ is the direct sum of its eigenspaces:*

$$\mathbb{R}^n = \bigoplus_{j=1}^{k} S_{\lambda_j}$$

where $\{\lambda_j\}_{j=1}^{k}$ is the set of distinct eigenvalues of $\mathbb{R}^n$.

## §4.3 Cayley-Hamilton Theorem

**Theorem 4.9** (Cayley-Hamilton Theorem)

*Let $T$ be a linear operator on a finite-dimensional $\mathbb{F}$-vector space $V$ and let $f(\lambda)$ be its characteristic polynomial. Then $f(T) = T_0$, the zero transformation; i.e., $T$ satisfies its own characteristic equation.*

**Corollary 4.10** (Cayley-Hamilton Theorem for matrices)

*Let $A \in M_{n\times n}(\mathbb{F})$ and let $f(\lambda)$ be the characteristic polynomial of $A$. Then $f(\lambda) =$ O, the $n \times n$ zero matrix; i.e., every square matrix satisfies its own characteristic equation.*