Linera Algebra 2H

typed up by B. S. H. Mithrandir

- *Last compiled: July 2024*
- Blended from notes of C. Keaton and J. R. Parker, Durham
- This was part of the Durham core second year modules. Involves introduction to vector spaces, linear maps, and matrices as linear maps
- I have personally changed the ordering quite a bit, since the ordering in my original I have from the original notes is obscenely jumpy and makes no logical sense to me...(if I were teaching this I would probably have matrices before linear maps)

1 Vector spaces

Recall that a **field** $\mathbb{F}$ equipped with an addition $+$ and a multiplication $\times$ operation satisfies the following axioms:

- operations are associative
- operations are commutative
- there exists additive and multiplicative identity
- there exists additive and multiplicative inverse (except for the additive identity)
- multiplication is distributive over addition

The usual rationals $\mathbb{Q}$, reals $\mathbb{R}$ and complex numbers $\mathbb{C}$ are common fields.

A set V is a **vector space** over a field $\mathbb{F}$ if, for $\boldsymbol{v}, \boldsymbol{u} \in V$ and $\lambda, \mu \in \mathbb{F}$,

- V form an abelian group under addition $+$
- for all $\lambda \boldsymbol{v} \in V$
- $(\lambda + \mu)\boldsymbol{v} = \lambda \boldsymbol{v} + \mu \boldsymbol{v}$
- $\lambda(\boldsymbol{u} + \boldsymbol{v}) = \lambda \boldsymbol{u} + \lambda \boldsymbol{v}$
- $\lambda(\mu \boldsymbol{v}) = (\lambda \mu)\boldsymbol{v}$
- there exists an identity element 1 where $1\boldsymbol{v} = \boldsymbol{v}$.

Some examples are

- $\mathbb{F} = \mathbb{R}$ and $V = \mathbb{C}$
- $\mathbb{F} = \mathbb{R}$ and $V = \mathbb{R}^n$ is the Eucliean n-space
- $\mathbb{F} = \mathbb{R}$ and $V = C[0, 1]$ the set of continuous functions defined over the interval $[0, 1]$
- $\mathbb{F} = \mathbb{R}$ and V the solutions to the ODE $y'' - y' + y = 0$

- $\mathbb{F} = \mathbb{R}$ and $V = \mathbb{R}[t]$, the set of polynomials in t with coefficients in $\mathbb{R}$

Lemma 1.0.1 *If V is a vector space over $\mathbb{F}$, then for all $\boldsymbol{v} \in V$, we have $0\boldsymbol{v} = \mathbf{0}$.*

Proof We have $0\boldsymbol{v} = (0+0)\boldsymbol{v} = 0\boldsymbol{v} + 0\boldsymbol{v}$, so subtracting accordingly leads to the above result. ■

Lemma 1.0.2 *As above, but for all $\lambda \in \mathbb{F}$, we have $\lambda\mathbf{0} = \mathbf{0}$.* □

Lemma 1.0.3 *For all $\boldsymbol{v} \in V$, $(-1)\boldsymbol{v} = -\boldsymbol{v}$, the additive inverse of $\boldsymbol{v}$.*

Proof $\mathbf{0} = 0\boldsymbol{v} = (1 + (-1))\boldsymbol{v} = \boldsymbol{v} + (-1)\boldsymbol{v}$, so that $(-1)\boldsymbol{v} = -\boldsymbol{v}$. ■

Let $U \subseteq V$, then U is a **subspace** of V over $\mathbb{F}$ if U forms a vector space under the same addition and scalar multiplication as V.

Instead of checking that a given subspace U really satisfies all the vector space axioms, the follow theorem reduces the amount of work required.

Theorem 1.0.4 *U is a subspace of V if and only if $U \neq \emptyset$, and for all $\boldsymbol{u}, \boldsymbol{v} \in U$ and $\lambda \in \mathbb{F}$, we have $\boldsymbol{u} + \boldsymbol{v} \in U$ and $\lambda\boldsymbol{v} \in U$.*

Proof These conditions are obviously necessary for U to be a vector space. Conversely, if $\boldsymbol{u}, \boldsymbol{v} \in U$, then $\boldsymbol{u}, -\boldsymbol{v} = (-1)\boldsymbol{v} \in U$ by above lemma.s Since $\boldsymbol{u} - \boldsymbol{v} \in U$, $U \neq \emptyset$, and U is a group under addition. The condition $\lambda vin U$ ensures that U is closed. Since the relevant axioms hold in V, they also hold in U. ■

Corollary 1.0.5 *Let U and V be subspaces of W, then $U \cap V$ is a subspace of W.*

1.1 Bases

We say the vectors $\{\boldsymbol{v}_i\}$ are **linearly independent** if

$$\lambda_1\boldsymbol{v}_1 + \ldots + \lambda_n\boldsymbol{v}_n = 0 \tag{1.1}$$

only has trivial solutions $\lambda_i = 0$ for all i, if $\boldsymbol{v}_i \neq \boldsymbol{v}_j$ for $i \neq j$.

Example 1. Show

$$S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right\}$$

are linearly independent.

We have $a + 2c = 0$, $b = 0$, and $a + b + c = 0$. This ends up reducing down accordingly to $a = b = c = 0$, so S is linearly independent.

2. Shown

$$S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\}$$

is linearly dependent.

Note that the third vector is the sum of the first and second vector.

3. Show that $S = \{\sin nx \mid n \in \mathbb{N}\}$ is linearly independent.

Note that we have the following orthogonality rule:

Essentially from compound angle formula.

$$\int_{-\pi}^{\pi} \sin px \sin qx \, \mathrm{d}x = \begin{cases} \pi, & p = q, \\ 0, & p \neq q. \end{cases}$$

Suppose there exists some $\{b_i\}$ where $\sum_{k=1}^{N} b_k \sin kx = 0$ then we have, for some $p \in \{1, \dots N\}$,

$$\sum_{k=1}^{N} b_k \int_{-\pi}^{\pi} \sin px \sin kx \, \mathrm{d}x = 0,$$

so by the orthogonality condition, then $b_p \pi = 0$, and hence $b_p = 0$ for arbitrary p, and so we have linear independence.

This is one step towards saying sines and cosines forms the Fourier basis, which is important for the theorem of *Fourier transforms*.

The **span** of $\{\boldsymbol{v}_i\}$ is defines as the set containing all linear combinations of the **spanning set** $\{\boldsymbol{v}_i\}$, i.e.

$$\operatorname{Span}(\{\boldsymbol{v}_i\}) = \left\{ \sum_{i=1}^{n} \lambda_i \boldsymbol{v}_i \mid \lambda_i \in \mathbb{F} \right\}. \tag{1.2}$$

Example 1. $S = \{\boldsymbol{e}_1, \boldsymbol{e}_2, (2, 3, 0)\}$ spans only the xy plane, because the third vector is not linearly independent of the other two.

2. Does $S = \{\sin nx \mid n \in \mathbb{N}\}$ span $C^\infty(\mathbb{R})$?

Suppose a function $f \in \operatorname{Span}(S)$, then if S spans $C^\infty(\mathbb{R})$, we should have, for all $x \in \mathbb{R}$,

$$f(x) = \sum_{k=1}^{n} b_k \sin kx.$$

Here $f(0) = 0$, but note that $\cos(x) \in C^\infty(\mathbb{R})$, but $\cos 0 = 1$, so $g \notin \operatorname{Span}(S)$, so the answer is no by a counter-example.

The set $\{\boldsymbol{v}_i\}$ is a **basis** of some vector space U if the spanning set is linearly independent, and spans U.

2 Linear maps

2.1 Spaces of linear maps

2.2 Dual space

2.3 Matrices

Mostly going to assume we are working with matrices defined over the reals, unless stated otherwise.

2.4 Elementary matrices

2.5 Rank

Let A be a $m \times n$ matrix. Then the **row rank** of A is the dimension of the subspaces spanned by the row vectors of A. The **column rank** is defined similarly.

Theorem 2.5.1 *For every matrix, the row rank is equal to the column rank (so we only need to talk about* rank*).*

Proof Think of A as a linear map $\theta : \mathbb{R}^n \to \mathbb{R}^m$, so there exists non-singular X and Y where

$$\mathsf{B} = \mathsf{XAY} = \begin{pmatrix} \mathsf{I}_r & 0 \\ 0 & 0 \end{pmatrix},$$

where r is the row rank of θ. The transpose of B has the same rank, although that is now the column rank of B^T. ■

We define the **nullity** of A to be the dimension of the kernel, i.e., the dimension of the solution set $\mathsf{A}\boldsymbol{x} = \mathbf{0}$.

Lemma 2.5.2 *The following are equivalent if* A *is a square $n \times n$ matrix:*

1. A *is invertible,*
2. *rank*$(\mathsf{A}) = n$,

3. $null(\mathsf{A}) = 0$.

Proof If A is invertible then the corresponding linear map θ is an isomorphism, and we have the result by a previous corollary. reference back ■

2.6 *Trace and determinant*

2.7 *Adjugate*

2.8 *Eigenvalues and eigenvectors*

Let V be a vector space, and $\theta : V \to V$ be a linear map. Then $\boldsymbol{v} \in V - \{\mathbf{0}\}$ is an **eigenvector** of θ if

$$\theta(\boldsymbol{v}) = \lambda \boldsymbol{v}\,, \tag{2.1}$$

where λ is the corresponding **eigenvalue**. The subset

$$V_\lambda = \{\boldsymbol{v} \in V \mid \theta(\boldsymbol{v}) = \lambda \boldsymbol{v}\} \tag{2.2}$$

is a vector space called the **eigenspace** corresponding to the eigenvalue λ.

Clearly if A is the corresponding matrix to θ then the eigenvalues and eigenvectors would correspond accordingly.

Two $n \times n$ matrices A and B are called **similar** if there exists some non-singular Y where $\mathsf{B} = \mathsf{Y}^{-1}\mathsf{A}\mathsf{Y}$.

Theorem 2.8.1 *If A corresponds to the linear map $\theta : V \to V$, then the roots of the **characteristic polynomial** $|\lambda\mathsf{I} - \mathsf{A}| = 0$ are the eigenvalues of θ. This is independent of the basis of V.*

Proof Let $\boldsymbol{v}$ be an eigenvector, so that $\boldsymbol{v} = \sum_i v_i \boldsymbol{e}_i$. If λ is the corresponding eigenvalue, and A is the representation of θ in the basis $\{\boldsymbol{e}_i\}$, then

$$\mathsf{A}\boldsymbol{v} = \lambda \boldsymbol{v} \quad \Leftrightarrow \quad (\lambda\mathsf{I} - \mathsf{A})\boldsymbol{v} = \mathbf{0}.$$

Since $\boldsymbol{v}$ is non-trivial by definition of it being an eigenvector, then $\lambda\mathsf{I} - \mathsf{A}$ is singular, so $|\lambda\mathsf{I} - \mathsf{A}| = 0$.

For basis independence, we know that

$$\begin{aligned} |\lambda\mathsf{I} - \mathsf{Y}^{-1}\mathsf{A}\mathsf{Y}| &= |\lambda\mathsf{Y}^{-1}\mathsf{I}\mathsf{Y} - \mathsf{Y}^{-1}\mathsf{A}\mathsf{Y}| \\ &= |\mathsf{Y}^{-1}(\lambda\mathsf{I} - \mathsf{A})\mathsf{Y}| \\ &= |\mathsf{Y}^{-1}| \cdot |\lambda\mathsf{I} - \mathsf{A}| \cdot |\mathsf{Y}|, \end{aligned}$$

but by the property that $|\mathsf{Y}^{-1}| \cdot |\mathsf{Y}| = 1$, we have $|\lambda\mathsf{I} - \mathsf{Y}^{-1}\mathsf{A}\mathsf{Y}| = |\lambda\mathsf{I} - \mathsf{A}| = 0$. ■

Example For

$$\mathsf{A} = \begin{pmatrix} 2 & 2 \\ 0 & 4 \end{pmatrix}, \quad \lambda\mathsf{I} - \mathsf{A} = \begin{pmatrix} \lambda - 2 & -2 \\ 0 & \lambda - 4 \end{pmatrix}$$

and the eigenvalues satisfy the quadratic $(\lambda - 2)(\lambda - 4) = 0$, so $\lambda = 2, 4$. For $\lambda = 2$, the eigenvector satisfies

Could have done this with the trace and determinant also for a 2×2 matrix.

$$\begin{pmatrix} 0 & -2 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \mathbf{0},$$

which implies $v_2 = 0$ but v_1 is undetermined; we could choose for simplicity for the eigenvector to be

$$\boldsymbol{v}_{\lambda=2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

For $\lambda = 4$, we result in $v_1 = v_2$, so we could choose

$$\boldsymbol{v}_{\lambda=4} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Notice for the above that if we take Y to be the matrix whose column vectors are the eigenvectors, i.e.

$$\mathsf{Y} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \Rightarrow \quad \mathsf{Y}^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix},$$

then notice that

$$\mathsf{Y}^{-1}\mathsf{A}\mathsf{Y} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix},$$

which is just the **diagonal** matrix with the eigenvalues on the main diagonal. This is called **diagonalisation**.

Theorem 2.8.2 *Let* A *be a* $n \times n$ *matrix. The matrix* A *is* ***diagonalisable*** *if there exists a non-singular* Y *such that* A *is similar to a diagonal matrix* D*, i.e.* $\mathsf{D} = \mathsf{Y}^{-1}\mathsf{A}\mathsf{Y}$ *or* $\mathsf{A} = \mathsf{Y}\mathsf{D}\mathsf{Y}^{-1}$*. This is equivalent to the condition that the eigenvectors of* A *has full span.*

Notice we do not require A to be non-singular, nor that the eigenvalues have to be distinct. For example,

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

are respectively singular and have repeated eigenvalues, but are both trivially diagonal (so diagonalisable by I). On the other hand,

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

has eigenvalues $\lambda = \pm i$, so it is definitely not diagonalisable *over the reals*.

Example For

$$A = \begin{pmatrix} 3 & 0 \\ 1 & 3 \end{pmatrix},$$

we have $\lambda = 3$ of multiplicity 2, but additionally only one eigenvector $(1, 0)^T$, so the span of the eigenspace is only dimension 1, and matrix is not diagonalisable.

Example For

$$A = \begin{pmatrix} 0 & -2 & 2 \\ 0 & 2 & 0 \\ -1 & -1 & 3 \end{pmatrix},$$

can show that the eigenvalues are $\lambda = 1$ and $\lambda = 2$ twice. For eigenvector corresponding to $\lambda = 1$ can be chosen to be

$$\boldsymbol{v}_1 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}.$$

For $\lambda = 2$, we reduce down to $v_1 + v_2 - v_3 = 0$, so we could take

$$\boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

which are linearly independent of each other, so the eigenspace has full span, and the transformation matrix could be

$$Y = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Notice we could have made a different choice for $\boldsymbol{v}_{2,3}$, and that

$$Y = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 0 & 1 \\ 2 & -1 & 1 \end{pmatrix}$$

would have also worked.

There are at least two notable uses for this:

Example 1. If we have the system of ODEs

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x} = A\boldsymbol{x}$$

where A is constant and diagonalisable with $\mathsf{A} = \mathsf{Y}\mathsf{D}\mathsf{Y}^{-1}$, then the solutions are

$$\boldsymbol{x} = \mathsf{Y} \begin{pmatrix} \mathrm{e}^{\lambda_1 t} \\ \vdots \\ \mathrm{e}^{\lambda_n t} \end{pmatrix}$$

up to multiplicative constants (which are specified by the initial conditions). This is because

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x} = \mathsf{Y}\mathsf{D}\mathsf{Y}^{-1}\boldsymbol{x} \quad \Rightarrow \frac{\mathrm{d}}{\mathrm{d}t}\mathsf{Y}^{-1}\boldsymbol{x} = \mathsf{D}\mathsf{Y}^{-1}\boldsymbol{x},$$

because Y is constant in time so commutes with the derivative. Then calling $\boldsymbol{y} = \mathsf{Y}^{-1}\boldsymbol{x}$, we just have a set of trivial constant ODEs to solve, then we invert to get $\boldsymbol{x} = \mathsf{Y}\boldsymbol{y}$ which is known.

In other words, we choose a convenient basis where our problem is diagonal, solve the problem there (which is easy), and then transform it back into the original basis. This kind of problem occurs occasionally in quantum mechanics, where you might want to solve the **time-independent Schrödinger equation**

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = \mathsf{H}|\psi(t)\rangle$$

where H would be matrix representation of the relevant Hamiltonian of the system, and $|\psi(t)\rangle$ would be the vector representation of the **state**. The eigenvalues correspond usually to an energy, and the eigenvectors the energy states.

The Hamiltonian normally has additional properties that make the eigenvalues and eigenvectors have certain properties; more later when we talk about *Hermitian/self-adjoint* matrices.

2. Suppose we have some system that is governed by

$$\boldsymbol{x}_{n+1} = \mathsf{A}\boldsymbol{x}_n,$$

such as a **Markov chain** with discrete states, and so A is a matrix detailing the transition probabilities between states (with all positive entries that are strictly less than 1, since these are supposed to describe probabilities). If A is diagonalisable, then the numerics becomes 'easy', in that to find $\boldsymbol{x}_{n+1}$ from initial data $\boldsymbol{x}_0$, we do

$$\begin{aligned} \boldsymbol{x}_n &= \mathsf{A}^n\boldsymbol{x}_0 \\ &= (\mathsf{Y}\mathsf{D}\mathsf{Y}^{-1})^n\boldsymbol{x}_0 \\ &= (\mathsf{Y}\mathsf{D}\mathsf{Y}^{-1})(\mathsf{Y}\mathsf{D}\mathsf{Y}^{-1})\cdots(\mathsf{Y}\mathsf{D}\mathsf{Y}^{-1})\boldsymbol{x}_0 \\ &= \mathsf{Y}\mathsf{D}(\mathsf{Y}^{-1}\mathsf{Y})\mathsf{D}\cdots(\mathsf{Y}^{-1}\mathsf{Y})\mathsf{D}\mathsf{Y}^{-1}\boldsymbol{x}_0 \\ &= \mathsf{Y}\mathsf{D}^n\mathsf{Y}^{-1}\boldsymbol{x}_0, \end{aligned}$$

by associativity, and noting that D^n is essentially trivial to compute.

In the context of a Markov chain, the transition matrix have eigenvalues with magnitude less than or equal to 1 by the *Perron–Frobenius theorem*, so we have no blow up. This is not necessarily true for population models such as the *Leslie model*, since the entries are positive but do not have to be strictly less than 1.

2.9 Cayley–Hamilton theorem

2.10 Orthogonal matrices

Recall that two vectors are orthogonal to each other if $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^T \boldsymbol{v} = 0$, the latter interpreted in terms of a matrix multiplication of a row vector with a column vector.

Analogously, a matrix is called **orthogonal** if the columns are orthogonal to each other. However, a more useful form of orthogonality might be the following:

Theorem 2.10.1 *A matrix* A *is orthogonal if and only if* $\mathsf{A}^{-1} = \mathsf{A}^T$.

So orthogonal matrices are great, because a transpose is much easier to compute than the inverse.

Proof Let $\boldsymbol{e}_i$ be the i^{th} column of P, then $\mathsf{A}\mathsf{A}^T = (\boldsymbol{e}_i^T \boldsymbol{e}_j) = (\delta_{ij}) = \mathsf{I}$, as required. ■

2.11 Symmetric matrices

A matrix is A is **symmetric** if $\mathsf{A} = \mathsf{A}^T$.

Theorem 2.11.1 *Let* A *be a real symmetric matrix. Then the eigenvalues are all real, and the eigenvectors are mutual orthogonal to each other.* □

Theorem 2.11.2 *Let* A *be a real symmetric matrix. If it is diagonalisable, then the similarity matrix is orthogonal, i.e.* $\mathsf{A} = \mathsf{P}\mathsf{D}\mathsf{P}^{-1} = \mathsf{P}\mathsf{D}\mathsf{P}^T$. □

The theorems are special cases of those given in §2.13, so we defer the proof for later.

Example Diagonalise

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

Can show the eigenvalues are 1, -2 and -2, and that

$$\mathsf{P} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ -1 & 1 & 1 \end{pmatrix}$$

is one possible choice, leading to $\mathsf{D} = \operatorname{diag}(1, -2, -2)$.

2.12 Rotations

Denote the set of $n \times n$ orthogonal matrices over $\mathbb{R}$ by $\mathrm{O}(n)$, and $\mathrm{SO}(n)$ the subset of $\mathrm{O}(n)$ where the corresponding matrices have determinant $+1$. These turn out to form a group under matrix multiplication, and are known as the **orthogonal** and **special orthogonal** groups of dimension n respectively.

These are actually **Lie groups**, which play an important role in mathematical physics.

Theorem 2.12.1 *Every element of $SO(3)$ represent a rotation about an arbitrary axis in $\mathbb{R}^3$.* □

Corollary 2.12.2 *The characteristic polynomial*

$$\Delta(\lambda) = \lambda^3 - (1 + 2\cos\theta)\lambda^2 + (1 + 2\cos\theta)\lambda - 1$$

tells us the angle of rotation θ (via $\cos\theta$, up to a sign) of a rotation about any axis.

Example Find the matrix which represents a rotation in $\mathbb{R}^3$ about $(1,1,0)^T$ through an angle of $\pi/2$.

Note that we have

$$\mathsf{R}_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \quad \mathsf{R}_y(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix},$$
$$\mathsf{R}_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

are the rotation matrices about the x, y and z axis respectively by an angle θ. Then idea is that the axis is represented by a vector that we can normlalise, and we rotate that onto a standard axis (essentially a co-ordinate transformation), do the rotation by the desired angle (which we know the matrix of), and then undo rotation (or the co-ordinate transformation).

We can take the z axis for definiteness, and we want $\mathsf{P} : (1,1,0)^T/\sqrt{2} \mapsto (0,0,1)^T$. This is supposed to be a rotation, so P^{-1} exists and $\mathsf{P}^{-1} : (0,0,1)^T \mapsto (1,1,0)^T/\sqrt{2}$, or

$$\mathsf{P}^{-1} = \begin{pmatrix} \cdot & \cdot & 1/\sqrt{2} \\ \cdot & \cdot & 1/\sqrt{2} \\ \cdot & \cdot & 0 \end{pmatrix}.$$

The column vectors should be mutually orthogonal to each other, and we can choose them to be orthonormal, and we could consider the simple choice

$$\mathsf{P}^{-1} \overset{!?}{=} \begin{pmatrix} 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 1 & 0 & 0 \end{pmatrix}.$$

Notice however that $|\mathsf{P}^{-1}| = -1$, so to make sure $\mathsf{P} \in \mathrm{SO}(3)$, we can simply multiply the bottom row by -1. The resulting P is then

$$\mathsf{P} = \begin{pmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 1 & 0 & 0 \end{pmatrix},$$

and so

$$\mathsf{A} = \mathsf{P}^{-1}\mathsf{R}_z(\pi/2)\mathsf{P} = \begin{pmatrix} 1/2 & 1/2 & 1/\sqrt{2} \\ 1/2 & 1/2 & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix}$$

will do the job. Can brute force check that that $|\mathsf{A}| = 1$, $\mathsf{A}^T = \mathsf{A}$, and the only solution to $\mathsf{A}\boldsymbol{x} = \boldsymbol{x}$ is $\boldsymbol{x} = a(1,1,0)^T$.

The rotation axis is kept fixed by the rotation about that axis.

2.13 *Hermitian matrices*

remember to introduce somewhere star being the Hermitian transpose

A matrix over $\mathbb{C}$ is **Hermitian** if $\mathsf{A} = \mathsf{A}^* = \overline{\mathsf{A}}^T$. Note that if A is actually real, then this reduces to the definition of a symmetric matrix.

A matrix over $\mathbb{C}$ is **unitary** if $\mathsf{A}^* = \mathsf{A}^{-1}$. Again, if A is actually real, then this reduces to the definition of an orthogonal matrix.

Unitary matrixes are used a lot in quantum mechanics.

Theorem 2.13.1 *Let A be a Hermitian matrix. Then the eigenvalues are all real, and the eigenvectors are mutual orthogonal to each other.*

Compare this with Theorem 2.11.1. In quantum mechanics the Hamiltonians are usually (but not always) chosen to be Hermitian, so the eigenvalues are all real and observable.

Proof Let λ be an eigenvalue of A, so $\mathsf{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ with $\boldsymbol{v} \neq \boldsymbol{0}$. Now,

$$\lambda\|\boldsymbol{v}\|^2 = \lambda\boldsymbol{v}^*\boldsymbol{v} = \boldsymbol{v}^*(\lambda\boldsymbol{v}) = \boldsymbol{v}^*\mathsf{A}\boldsymbol{v} = \boldsymbol{v}^*\mathsf{A}^*\boldsymbol{v} = (\mathsf{A}\boldsymbol{v})^*\boldsymbol{v} = \overline{\lambda}\boldsymbol{v}^*\boldsymbol{v} = \overline{\lambda}\|\boldsymbol{v}\|^2,$$

so λ is real for non-zero $\boldsymbol{v}$.

Let λ and μ be distinct eigenvalues of A, so there exists associated eigenvectors $\boldsymbol{v}$ and $\boldsymbol{w}$ respectively. Then doing similar manipulations to above,

$$\lambda\boldsymbol{w}^*\boldsymbol{v} = \boldsymbol{w}^*\mathsf{A}\boldsymbol{v} = (\mathsf{A}\boldsymbol{w})^*\boldsymbol{v} = \mu\boldsymbol{w}^*\boldsymbol{v}.$$

Since $\lambda \neq \mu$, then $\boldsymbol{w}^*\boldsymbol{v} = \langle \boldsymbol{v}, \boldsymbol{w}\rangle = 0$, and so the eigenvectors are orthogonal. ∎

Theorem 2.13.2 *Let A be a Hermitian matrix. If it is diagonalisable, then the similarity matrix is unitary, i.e. $\mathsf{A} = \mathsf{PDP}^{-1} = \mathsf{PDP}^*$.*

Proof Let λ_1 be an eigenvalue of A with eigenvector $\boldsymbol{v}$. Let $U = \text{span}\{\boldsymbol{v}\}$. For $\boldsymbol{w} \in U^\perp$, then

$$\langle \boldsymbol{v}, \mathsf{A}\boldsymbol{w}\rangle = \langle \mathsf{A}\boldsymbol{v}, \boldsymbol{w}\rangle = \langle \lambda_1\boldsymbol{v}, \boldsymbol{w}\rangle = \lambda_1\langle \boldsymbol{v}, \boldsymbol{w}\rangle = 0,$$

hence $\mathsf{A}\boldsymbol{w} \in U^\perp$. Supposing $\|\boldsymbol{v}\| = 1$, we choose an orthonormal basis $\boldsymbol{u}_2, \ldots \boldsymbol{u}_n$ for $U^\perp$, which implies that $\{\boldsymbol{v}, \boldsymbol{u}_2, \ldots \boldsymbol{u}_n\}$ is an orthonormal basis for $\mathbb{C}^n$.

Let P_1 be the matrix whose columns are those vectors, which by construction means $\mathsf{P}^*\mathsf{P} = \mathsf{I}$ by orthogonality. Let $\{\boldsymbol{e}_i\}$ be the standard basis of $\mathbb{C}^n$, then

$$\mathsf{P}_1^*\mathsf{A}\mathsf{P}_1\boldsymbol{e}_1 = \mathsf{P}_1^*\mathsf{A}\boldsymbol{v} = \mathsf{P}_1^*\lambda_1\boldsymbol{v} = \lambda_1\mathsf{P}_1^*\boldsymbol{v} = \lambda_1\boldsymbol{e}_1 \quad \Rightarrow \quad \mathsf{P}_1^*\mathsf{A}\mathsf{P}_1 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \mathsf{A}_1 \end{pmatrix}.$$

The procedure can be continued for the next one, so that

$$\mathsf{P}_2^*\mathsf{P}_1^*\mathsf{A}\mathsf{P}_1\mathsf{P}_2 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \mathsf{A}_2 \end{pmatrix},$$

and so we can define $\mathsf{P} = \prod_{i=1}^{n} \mathsf{P}_i$ as the transformation matrix. Since each P_i is unitary, its product is unitary. ■

2.14 *Diagonalising quadratic forms*

2.15 *Symmetric and self-adjoint maps*

2.16 *Reflections*

2.17 *The operator* $\mathrm{d}/\mathrm{d}x^2$

2.18 *Jordan normal form*