Honours Algebra - Week 7 - Eigenvalues, Eigenvectors and Triangularisations

Antonio León Villares

March 2022

Contents

Based on the notes by Iain Gordon, Sections 4.5 - 4.7

1 Eigenvalues and Eigenvectors

1.1 Definition of Eigenstuffs

- **What is an eigenvalue?**
  - let V be an F-vector space, and define an endomorphism:

$$f : V \to V$$

  - $\lambda \in F$ is an **eigenvalue** of f, if:

$$\exists \underline{v} \in V \;\;:\;\; f(\underline{v}) = \lambda \underline{v}$$

- **What is an eigenvector?**
  - any vector $\underline{v} \in V$ such that $f(\underline{v}) = \lambda \underline{v}$
  - $\underline{v}$ is the **eigenvector of f with eigenvalue λ**
  - **eigenvectors** are not unique; for example, if $F = \mathbb{R}$, and $\underline{v}$ is an eigenvector, then $2\underline{v}$ is also an eigenvector of the same eigenvalue
- **What is an eigenspace?**
  - the set of all **eigenvectors** of endomorphism f with **eigenvalue** $\lambda \in F$:

$$E(\lambda, f) = \{\underline{v} \mid \underline{v} \in V, f(\underline{v}) = \lambda \underline{v}\}$$

1.1.1 Examples

- the set of all **fixed points** of f (i.e all $\underline{x}$ with $f(\underline{x}) = \underline{x}$) is the **eigenspace** corresponding to $\lambda = 1$
- the set of non-zero elements in $ker(f)$ is precisely the eigenspace of non-zero eigenvectors with $\lambda = 0$
- if f is the endomorphism rotating the plane by 90^{o}, this only has eigenvalues in $\mathbb{C}$
- if f is the endomorphism representing a reflection on a line, the associated eigenvalue is 1, and the eigenvectors are all the eigenvectors which lie on the line (so the line is the eigenspace)
- if f is the endomorphism representing a 180^{o} rotation, this will have eigenvalue -1 (since this rotation maps a point to a point diametrically opposite)
- real polynomial differentiation only has $\lambda = 0$, with eigenvectors as the non-zero constant polynomials:

$$p \in \mathbb{R}[x], deg(P) = 0 \implies D(p) = 0 = 0p$$

  For the rest of the chapter, to avoid confusion with matrices, polynomials will be denoted using x (instead of X)

1.1.2 Exercises (TODO)

1. **Let $f : V \to V$ be an endomorphism of an F-vector space V. Show that $E(\lambda, f)$ is a vector subspace of V, for any $\lambda \in F$**

1.2 Theorem: Eigenvalues and Characteristic Polynomials

- **What is the characteristic polynomial?**
  - consider a **commutative ring** R, and let $A \in Mat(n; R)$
  - the **characteristic polynomial of the matrix** A **is the polynomial:**

$$\mathcal{X}_A(x) := det(xI_n - A)$$

  - for example,

$$A = \begin{pmatrix} 2 & 3 \\ -6 & 1 \end{pmatrix} \implies xI_n - A = \begin{pmatrix} 2-x & 3 \\ -6 & 1-x \end{pmatrix} \implies \mathcal{X}_A(x) = (2-x)(1-x)+18 = x^2-3x+20$$

> *Let F be a **field**, and let $A \in Mat(n; F)$.*
> *The **eigenvalues** of the **linear mapping**:*
>
> $$A : F^n \to F^n$$
>
> *are precisely the **roots** of $\mathcal{X}_A$. [Theorem 4.5.8]*

Proof. For any $\lambda \in F$, λ is an eigenvalue of A if and only if:

$$\begin{aligned} &\exists \underline{v} \neq 0 \; : \; A\underline{v} = \lambda \underline{v} \\ \iff &\exists \underline{v} \neq 0 \; : \; (\lambda I_n - A)\underline{v} = 0 \\ \iff &ker(\lambda I_n - A) \neq 0 \end{aligned}$$

If the kernel is non-zero, then notice that we have a non-zero $\underline{v}$:

$$(\lambda I_n - A)\underline{v} = \underline{0}$$

If $(\lambda I_n - A)^{-1}$ existed, then:

$$\underline{v} = (\lambda I_n - A)^{-1}\underline{0} = \underline{0}$$

which is a contradiction. Thus, $(\lambda I_n - A)$ can't be invertible. In other words, λ is an eigenvalue **if and only if**:

$$det(\lambda I_n - A) = 0 \implies \mathcal{X}_A(\lambda) = 0$$

□

1.2.1 Exercises (TODO)

1. **Let F be a field, and $A \in Mat(n; F)$. Show that:**

$$\mathcal{X}_A(x) = x^n - tr(A)x^{n-1} + (-1)^n det(A) + \sum_{i=2}^{n-1} a_i x^{n-i}$$

1.3 Remark: Defining the Characteristic Polynomial of Endomorphisms

Consider an endomorphism $f : V \to V$. If we have ordered bases:

$$\mathcal{A} = (\underline{v}_1, \ldots, \underline{v}_n) \qquad \mathcal{B} = (\underline{w}_1, \ldots, \underline{w}_n)$$

then we could define matrices $A, B \in Mat(n; R)$:

$$A = (a_{ij}) = {}_{\mathcal{A}}[f]_{\mathcal{A}} \qquad B = (b_{ij}) = {}_{\mathcal{B}}[f]_{\mathcal{B}}$$

where the jth columns of the matrices satisfied:

$$f(\underline{v}_j) = \sum_{i=1}^{n} a_{ij}\underline{v}_i \qquad f(\underline{w}_j) = \sum_{i=1}^{n} b_{ij}\underline{w}_i$$

The ***change of basis matrix****, $P \in GL(n; R)$:*

$$P = (p_{ij}) = {}_{\mathcal{A}}[id_V]_{\mathcal{B}} \qquad \underline{w}_j = \sum_{i=1}^{n} p_{ij}\underline{v}_i$$

allowed us to define the ***trace*** *of an endomorphism, independent of a* ***basis****, since it allowed us to see A, B as* ***conjugate matrices****:*

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

and the trace of conjugate matrices is the same.
In a similar vein, the ***characteristic polynomial*** *of* ***conjugate matrices*** *is the same, so we can define the* ***characteristic polynomial of an endomorphism****, by using its representing matrix, with respect to* ***any*** *basis.https://www.overleaf.com/project/61e7e55b4eaaff498c0d6c2b*
To see this:

$$\mathcal{X}_B = det(xI_n - B) = det(xI_n - P^{-1}AP)$$

Notice that:

$$det(P^{-1}(xI_n - A)P) = det(xP^{-1}I_nP - P^{-1}AP) = det(xI_nP^{-1}AP)$$

So:

$$\mathcal{X}_B = det(P^{-1})det(xI_n - A)det(P) = det(xI_n - A) = \mathcal{X}_A$$

Thus, the ***eigenvalues*** *of f are the roots of $\mathcal{X}_f$, the* ***characteristic polynomial of*** *f. [Remark 4.5.9]*

1.3.1 Exercises (TODO)

1. **Show that every endomorphism of an odd dimensional real vector space has a real eigenvalue. Show furthermore that if the determinant of the endomorphism is a positive real number, then the endomorphism even has a positive real eigenvalue.**

1.4 Remark: Existence of Upper Triangular Representing Matrix

The following remark might look weird here, but it is useful in the next section, when triangularisation of matrices is discussed.

Consider the endomorphism of an n:-dimensional F-vector space:

$$f : V \to V$$

*Let W be a vector **subspace** of V, with $f(W) \subseteq W$. Then, define the following endomorphisms:*

$$g : W \to W \qquad \underline{w} \to f(\underline{w})$$

$$h : V/W \to V/W \qquad W + \underline{v} \to W + f(\underline{v})$$

We can construct a basis for V using a basis for W:

$$\mathcal{A} = (\underline{w}_1, \ldots, \underline{w}_m)$$

$$\mathcal{B} = (\underline{w}_1, \ldots, \underline{w}_m, \underline{v}_{m+1}, \ldots, \underline{v}_n)$$

*Moreover, the basis of V/W can be constructed by applying the **canonical map** $can : V \to V/W$ to the elements $\underline{v}_j$ of the basis $\mathcal{B}$ (we don't need to consider the $\underline{w}_1$, since $W + \underline{w}_i = W + \underline{0}$ under the canonical map):*

$$\mathcal{C} = (can(\underline{v}_{m+1}), \ldots, can(\underline{v}_n))$$

Now, turns out that we can write:

$$[f]_{\mathcal{B}} = \begin{pmatrix} {}_{\mathcal{A}}[g]_{\mathcal{A}} & {}_{\mathcal{A}}[e]_{\mathcal{C}} \\ 0 & {}_{\mathcal{C}}[h]_{\mathcal{C}} \end{pmatrix} = \begin{pmatrix} (a_{ij}) & (c_{ik}) \\ 0 & (b_{jk}) \end{pmatrix}$$

where the components a_{ij} of ${}_{\mathcal{A}}[g]_{\mathcal{A}}$ satisfy:

$$f(\underline{w}_j) = \sum_{i=1}^{m} a_{ij}\underline{w}_i$$

and we have coefficients b_{jk}, c_{ik} satisfying:

$$f(\underline{v}_k) = \sum_{i=1}^{m} c_{ik}\underline{w}_i + \sum_{j=m+1}^{n} b_{jk}\underline{v}_j$$

Moreover:

$$e : V/W \to W \qquad W + \underline{v}_k \to \sum_{i=1}^{m} c_{ik}\underline{w}_i$$

[Remark 4.5.10]

The above is also useful in telling us that:

$$\mathcal{X}_f = \mathcal{X}_g \mathcal{X}_h$$

(Informally, think that $V/W \cdot W \to V$).
To see why, notice that:

$$xI_n - {}_{\mathcal{B}}[f]_{\mathcal{B}} = \begin{pmatrix} xI_m - {}_{\mathcal{A}}[g]_{\mathcal{A}} & -{}_{\mathcal{A}}[e]_{\mathcal{C}} \\ 0 & xI_{n-m\mathcal{C}}[h]_{\mathcal{C}} \end{pmatrix}$$

and since the determinant of block diagonal matrices is the product of the determinants of the diagonal blocks:

$$\mathcal{X}_f(x) = det(xI_n - {}_{\mathcal{B}}[f]_{\mathcal{B}}) = det(xI_m - {}_{\mathcal{A}}[g]_{\mathcal{A}}) det(xI_{n-m\mathcal{C}}[h]_{\mathcal{C}}) = \mathcal{X}_g(x)\mathcal{X}_h(x)$$

1.5 Example of Remark

Consider the mapping:

$$f(x\underline{e}_1 + \underline{e}_2 y) = (2x + ay)\underline{e}_1 + y\underline{e}_2$$

where we are using $V = S(2)$.
We can pick $W = \langle \underline{e}_1 \rangle$, since:

$$f(\underline{e}_1) = 2\underline{e}_1 \implies f(W) \subseteq W$$

Then, we can define the basis of W as:

$$\mathcal{A} = \{\underline{e}_1\}$$

Moreover, define $g : W \to W$ as:

$$\forall w \in W, \qquad g(w) = f(w) = f(\lambda \underline{e}_1) = 2(\lambda \underline{e}_1) = 2w$$

In other words, the representing matrix ${}_{\mathcal{A}}[g]_{\mathcal{A}}$ which maps elements from basis $\mathcal{A}$ to basis $\mathcal{A}$ is simply:

$${}_{\mathcal{A}}[g]_{\mathcal{A}} = (2)$$

Now, we define the basis $\mathcal{B}$ for V, by extending $\mathcal{A}$:

$$\mathcal{B} = \{\underline{e}_1, \underline{e}_1 + \underline{e}_2\}$$

Similarly, we define the basis $\mathcal{C}$ as a basis for V/W:

$$\mathcal{C} = \{can(\underline{e}_1 + \underline{e}_2)\} = \{W + (\underline{e}_1 + \underline{e}_2)\}$$

Then the mapping:

$$h : V/W \to V/W \qquad h(W + v) = W + f(v) = W + f(\underline{e}_1 + \underline{e}_2) = W + (2 + a)\underline{e}_1 + \underline{e}_2 = W + \underline{e}_2$$

since $(2 + a)\underline{e}_1 \in W$. However, notice that:

$$W + \underline{e}_2 = (W + \underline{e}_2) + (W + 0) = (W + \underline{e}_2) + (W + \underline{e}_1) = W + (\underline{e}_1 + \underline{e}_2)$$

Hence:

$$h(W + v) = W + v$$

In particular, this means that:
$$_{\mathcal{C}}[h]_{\mathcal{C}} = (1)$$
(we are mapping from $\mathcal{C}$ to $\mathcal{C}$ by using the identity. The last step is to consider the map:
$$e : V/W \to W$$
via:
$$e(W + (\underline{e}_1 + \underline{e}_2)) = c_{11}\underline{e}_1$$
Here, c_{11} comes from:
$$f(\underline{e}_1 + \underline{e}_2) = c_{11}\underline{e}_1 + b_{11}(\underline{e}_1 + \underline{e}_2)$$
We know that:
$$f(\underline{e}_1 + \underline{e}_2) = (2 + a)\underline{e}_1 + \underline{e}_2$$
So we must have that:
$$c_{11}\underline{e}_1 + b_{11}(\underline{e}_1 + \underline{e}_2) = (2 + a)\underline{e}_1 + \underline{e}_2$$
In particular, $b_{11} = 1$ (so that the equality of $\underline{e}_2$ matches), so:
$$c_{11}\underline{e}_1 + \underline{e}_1 + \underline{e}_2 = (2 + a)\underline{e}_1 + \underline{e}_2 \implies c_{11} + 1 = 2 + a \implies c_{11} = 1 + a$$
Hence, once again, we must have that:
$$_{\mathcal{A}}[e]_{\mathcal{C}} = (1 + a)$$
The remark thus tells us that:
$$_{\mathcal{B}}[f]_{\mathcal{B}} = \begin{pmatrix} 2 & 1 + a \\ 0 & 1 \end{pmatrix}$$
We can verify this:
$$f(\underline{e}_1) = (2 + 0)\underline{e}_1 = 2\underline{e}_1 + 0(\underline{e}_1 + \underline{e}_2)$$
$$f(\underline{e}_1 + \underline{e}_2) = (2 + a)\underline{e}_1 + \underline{e}_2 = 2\underline{e}_1 + a\underline{e}_1 + \underline{e}_2 = (1 + a)\underline{e}_1 + (\underline{e}_1 + \underline{e}_2)$$
So
$$_{\mathcal{B}}[f]_{\mathcal{B}} = \begin{pmatrix} 2 & 1 + a \\ 0 & 1 \end{pmatrix}$$
as expected.

1.6 Theorem: Existence of Eigenvalues

Each ***endomorphism*** *of a* ***non-zero, finite dimensional*** *vector space over an* ***algebraically closed field*** *has an* ***eigenvalue****. [Theorem 4.5.4]*

Proof. Notice, given any endomorphism, its characteristic polynomial $\mathcal{X}_f$ won't be constant. This is a polynomial, and we are operating over an algebraically closed field, so in particular this means that any polynomial has at least one root. But then, by the remark above, the roots of a characteristic polynomial are precisely eigenvalues, as required.

The wording used is necessary:

- we require a non-zero vector space to ensure that an eigenvector will be non-zero

- we also require finite dimensionality, since, for example, the infinite dimensional polynomial space $V = \mathbb{C}[x]$ contains endomorphisms with no eigenvalues, such as:

$$f(P) = xP$$

□

2 Triangularisation

2.1 Proposition: Triangularisability

> *Let $f : V \to V$ be an **endomorphism** of a **finite dimensional** F-vector space V.*
> *The following are equivalent:*
>
> 1. *There exists an ordered basis:*
>
> $$\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$$
>
> *such that:*
>
> $$f(\underline{v}_j) = \sum_{i=1}^{j} a_{ij}\underline{v}_i, \qquad i \in [1, n]$$
>
> *In particular, this means that ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ will be a **triangular matrix**, with entries a_{ij}:*
>
> $$A = {}_{\mathcal{B}}[f]_{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix}$$
>
> *This means that f is **triangularisable**.*
>
> 2. *The **characteristic polynomial**, $\mathcal{X}_f$, decomposes into **linear factors** in $F[x]$*
>
> *[Proposition 4.6.1]*

Proof. 1. **Triangularisable Endomorphism Implies Characteristic Polynomial With linear Factors**

The determinant of an **upper triangular matrix** is the product of its diagonal entries. Hence:

$$det(xI_n - A) = \prod_{i=1}^{n}(x - a_{ii})$$

Thus, $\mathcal{X}_f$ decomposes into linear factors.

2. **Characteristic Polynomial With Linear Factors Implies Triangularisable Endomorphism**
This part is a bit more involved, and will proceed by induction. In particular, we consider $n = dim(V)$

1 **Base Case: n = 1**

In this case, we will have a 1×1 matrix, which is automatically a diagonal matrix, with any basis $\mathcal{B} = \{\underline{v}_1\}$:

$$f(\underline{v}_1) = a_{11}\underline{v}_1$$

2 **Inductive Hypothesis: n = k**

Let's assume our claim is true. That is, if $dimV = k$, then we can find a basis:

$$\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_k\}$$

such that:

$$f(\underline{v}_j) = \sum_{i=1}^{j} a_{ij}\underline{v}_i, \qquad j \in [1, k]$$

and so that the matrix ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ is upper triangular:

$${}_{\mathcal{B}}[f]_{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ 0 & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{kk} \end{pmatrix}$$

3 **Inductive Step: n = k+1**

(a) **Recapping Remark 4.5.10**

We now consider what happens when $n = k + 1$. For this we want to use Remark 4.5.10 (1.4).

As a recap, to avoid having to scroll, this lovely remark tells us that there exists a basis, which allows us to write a representing matrix as a upper triangular matrix, which is conveniently what we try to prove.

To do this, it employs a subspace $W \subseteq V$, satisfying:

$$f(W) \subseteq W$$

and produces homomorphisms:

$$g : W \to W \qquad \underline{w} \to f(\underline{w})$$

$$h : V/W \to V/W \qquad W + \underline{v} \to W + f(\underline{v})$$

It also constructs a basis for V by extending the basis for W. Lastly, it tells us that we can decompose the characteristic polynomial:

$$\mathcal{X}_f(x) = \mathcal{X}_g(x)\mathcal{X}_h(x)$$

The assumption that $\mathcal{X}_f(x)$ decomposes into linear factors thus means that $\mathcal{X}_g(x), \mathcal{X}_h(x)$ also decompose into linear factors, which are precisely eigenvectors.

(b) **Choosing a Basis W**

To construct the W, we are told that there are 3 "obvious". The choices rely on the existence of at least one eigenvector (which we can assume, since $\mathcal{X}_f$ decomposes into linear factors), $\underline{v}_1$, which satisfies $f(\underline{v}_1) = \lambda \underline{v}_1 \in V$, and on the fact that we seek $W, V/W$ with $dim(W) < dim(V) = k+1$ and $dim(V/W) < dim(V) = k+1$, since this then allows us to employ the inductive hypothesis. Using this, the "obvious" choices are:

i. W as the span of $\underline{v}_1$:

$$W = \langle \underline{v}_1 \rangle = \{\mu \underline{v}_1 \mid \mu \in F\} \subseteq V$$

- We have that $dim(W) = 1$, so by (1.4), we must have that $dim(V/W) = (k+1) - 1 = k$
  - in (1.4) a basis for $dim(V/W)$ is constructed by applying the canonical map to the vectors used to extend the basis for W
  - this then means that $dim(V/W) = dim(V) - dim(W)$
- since $\underline{v}_1$ is an eigenvector, $f(W) = \{\lambda\mu \underline{v}_1 \mid \mu \in F\} \subseteq W$
- this choice of W then satisfies our 2 requirements ((1.4) and the inductive hypothesis can be applied)

ii. W as:

$$W = ker(f - \lambda 1_V)$$

Thus, W contains all those eigenvectors with the same eigenvalue λ as $\underline{v}_1$ (since $(f - \lambda 1_V)(\underline{v}) = f(\underline{v}) - \lambda \underline{v})$. Using similar arguments as above, we can see that $f(W) \subseteq W$, and that $dim(W) \geq 1, dim(V/W) \leq k$.

iii. W as:

$$W = im(\lambda 1_V - f)$$

- to see why $f(W) \subseteq W$:

$$\underline{w} \in W \implies f(\underline{w}) = (\lambda 1_v - f)(\underline{w}) + \lambda \underline{w}$$

  Clearly, $(\lambda 1_v - f)(\underline{w}) \in W$, and by hypothese $\underline{w} \in W$, so $f(\underline{w}) \in W \implies f(W) \subseteq W$
- the mapping is:

$$\lambda 1_V - f : V \to V$$

  so by rank nullity:

$$dim(V) = dim(W) + dim(ker(\lambda 1_V - f)) \implies dim(W) = k + 1 - dim(ker(\lambda 1_V - f))$$

  The kernel is non-zero (for example, $\underline{v}_1$ is in the kernel), so it follows that $dim(W) \leq k$. Similarly, $dim(V/W) \leq k$.

(c) **Applying the Inductive Step**

For this, we will make use of:

$$W = \langle \underline{v}_1 \rangle$$

with basis:

$$\mathcal{A} = \{\underline{v}_1\}$$

where we shall relabel its eigenvalue (conveniently):

$$\lambda = a_{11}$$

such that:

$$f(\underline{v}_1) = a_{11} \underline{v}_1$$

Now, as we defined $g : W \to W$ via:

$$\underline{w} \in \mathcal{A} \implies g(\underline{w}) = f(\underline{w}) = f(\underline{v}_1) = a_{11} \underline{v}_1 = a_{11} \underline{w}$$

(recall, we can fully define a map based on the value it takes on the basis elements). Thus, it follows that:

$$_{\mathcal{A}}[g]_{\mathcal{A}} = \begin{pmatrix} a_{11} \end{pmatrix}$$

We also had a mapping:

$$h : V/W \to V/W$$

By construction, $dim(V/W) = k$, so the inductive hypothesis applies; in particular, V/W has an ordered basis:

$$\mathcal{D} = \{\underline{u}_2, \dots, \underline{u}_{k+1}\}$$

(we have conveniently labelled the basis elements), such that:

$$f(\underline{u}_j) = \sum_{i=2}^{j} a_{ij}\underline{u}_i, \quad j \in [2, k+1]$$

and so:

$$_{\mathcal{D}}[h]_{\mathcal{D}} = \begin{pmatrix} a_{22} & a_{23} & \dots & a_{2(k+1)} \\ 0 & a_{33} & \dots & a_{3(k+1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{(k+1)(k+1)} \end{pmatrix}$$

From (1.4), we have h defined as:

$$h(W + \underline{v}) = W + f(\underline{v})$$

where $W + \underline{v}$ is nothing but the effect of applying the canonical map to $\underline{v}$.
Now, for each $\underline{u}_j, j \in [2, k+1]$, let $\underline{v}_j \in V$ such that:

$$can(\underline{v}_j) = W + \underline{v}_j = \underline{u_j}$$

Then:

$$h(\underline{u}_j) = h(W + \underline{v}_j) = W + f(\underline{v}_j) = W + \sum_{i=2}^{j} a_{ij}\underline{v}_i = \sum_{i=2}^{j} a_{ij}(W + \underline{v}_i) = \sum_{i=2}^{j} a_{ij}\underline{u}_i$$

so it follows that:

$$h(\underline{u}_j) - \sum_{i=2}^{j} a_{ij}\underline{u}_i = \underline{0}_{V/W} \in V/W$$

Now consider:

$$\begin{aligned} can\left(f(\underline{v}_j) - \sum_{i=2}^{j} a_{ij}\underline{v}_i\right) &= can(f(\underline{v}_j) - \sum_{i=2}^{j} a_{ij}can(\underline{v}_i) \\ &= (W + f(\underline{v}_j) - \sum_{i=2}^{j} \underline{u}_i \\ &= h(W + \underline{v}_j) - \sum_{i=2}^{j} \underline{u}_i \\ &= h(\underline{u}_j) - \sum_{i=2}^{j} \underline{u}_i \\ &= \underline{0}_{V/W} \in V/W \end{aligned}$$

But then it must be the case that, from the definition of the canonical mapping:

$$f(\underline{v}_j) - \sum_{i=2}^{j} a_{ij}\underline{v}_i \in W \implies \exists a_{ij} \in F \ : \ f(\underline{v}_j) - \sum_{i=2}^{j} a_{ij}\underline{v}_i = a_{1j}\underline{v}_1, \quad j \in [2, k+1]$$

So rearranging:

$$f(\underline{v}_j) = \sum_{i=1}^{j} a_{ij}\underline{v}_i, \qquad j \in [1, k+1]$$

Thus, we have proven the inductive step, and it follows that the ordered basis $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ leads to the matrix:

$$_{\mathcal{B}}[f]_{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1(k+1)} \\ 0 & a_{22} & \dots & a_{2(k+1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{(k+1)(k+1)} \end{pmatrix}$$

In terms of the matrix in (1.4):

$$_{\mathcal{B}}[f]_{\mathcal{B}} = \begin{pmatrix} _{\mathcal{A}}[g]_{\mathcal{A}} & _{\mathcal{A}}[e]_{\mathcal{C}} \\ 0 & _{\mathcal{C}}[h]_{\mathcal{C}} \end{pmatrix}$$

where:

$$e : V/W \to W \qquad e(\underline{u}_j) = a_{ij}\underline{v}_1$$

$\square$

2.2 Remark: Triangularisability and Conjugacy

In general, an endomorphism:

$$A : F^n \to F^n$$

is ***triangularisable*** *if and only if A is conjugate to an* ***upper triangular matrix*** *B, such that for invertible P:*

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

[Remark 4.6.3.1]

2.3 Remark: Triangularisability and Algebraic Closure

Algebraic closure *of* $\mathbb{C}$ *means that any endomorphism of a finite dimensional* $\mathbb{C}$*-vector space can be decomposed into* ***linear factors****, so in particular, any such endomorphism is* ***triangularisable****.*
On the other hand, $\mathbb{R}$*-vector fields are not all triangularisable. For example, the endomorphism representing a* θ º *anticlockwise rotation is given by:*

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

with characteristic polynomial:

$$\mathcal{X}(x) = x^2 - 2x\cos\theta + 1$$

which has no real roots (except if $\theta = n\pi$*), so this endomorphism isn't triangularisable (unless* $\theta = n\pi$*). [Remark 4.6.3.2]*

2.4 Remark: Triangularisability and Embedded Subspaces

An ***endomorphism*** $f : V \to V$ *of an n-dimensional* F*-vector space* V *is* ***triangularisable*** *if and only if there is a* ***sequence of subspaces****:*

$$V_0 = \{\underline{0}\} \subset V_1 \subset \ldots \subset V_n = V$$

such that $dim(V_i) = i$ *and* $f(V_i) \subseteq V_i$ *[Remark 4.6.3.3]*

2.5 Remark: Importance of Triangular Matrices

Given an ***invertible*** *upper triangular matrix, it is straightforward to compute the solution to a system of simultaneous equations, where we can solve for* x_i *by using* $x_1, \ldots, x_{i-1}$*, and* x_1 *can be "read off". [Remark 4.6.3.4]*

2.5.1 Worked Example: Triangularisation

Consider the endomorphism $f : \mathbb{R}^3 \to \mathbb{R}^3$ represented by:

$$[f] = \begin{pmatrix} 14 & 8 & 3 \\ -17 & -9 & -3 \\ 1 & 0 & 0 \end{pmatrix}$$

We compute its characteristic polynomial:

$$\begin{aligned}
\mathcal{X}_f(x) &= 1 \times \begin{vmatrix} 8 & 3 \\ -9-x & -3 \end{vmatrix} - 0 - x \times \begin{vmatrix} 14-x & 8 \\ -17 & -9-x \end{vmatrix} \\
&= -24 + 3(9+x) + x(14-x)(9+x) - x(17)(8) \\
&= x(14-x)(9+x) + 27 + 3x - 136x - 24 \\
&= x(14-x)(9+x) - 133x + 3 \\
&= x(14(9) + 5x - x^2) - 133x + 3 \\
&= -x^3 + 5x^2 + 126x - 133x + 3 \\
&= -x^3 + 5x^2 - 7x + 3 \\
&= (x-1)^2(x-3)
\end{aligned}$$

In particular, notice that since $\mathcal{X}_f(x)$ decomposes into linear factors, it is triangularisable.

To find a basis for this triangularisation, we can use example 3 of the proof of Proposition 4.6.1. That is, we use:

$$W = im([f] - \lambda I_3)$$

With $\lambda = 3$, we want to consider the image of:

$$[f] - 3I_3 = \begin{pmatrix} 11 & 8 & 3 \\ -17 & -12 & -3 \\ 1 & 0 & -3 \end{pmatrix}$$

Thinking about it, the image is just the **column space** of the matrix. This is because when we apply matrix multiplication, the element v_1 gets multiplied by the first column, v_2 gets multiplied by the second column and so on. Hence, the result of matrix multiplication is a linear combination of the column vectors of the matrix. As a concrete example:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 \\ a_{21}v_1 + a_{22}v_2 \end{pmatrix} = v_1 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + v_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$$

Hence, $dim(W)$ is just the number of linearly independent columns. We can see that:

$$-\frac{1}{3} \begin{pmatrix} 3 \\ -3 \\ -3 \end{pmatrix} + \frac{6}{4} \begin{pmatrix} 8 \\ -12 \\ 0 \end{pmatrix} = \begin{pmatrix} 11 \\ -17 \\ 1 \end{pmatrix}$$

so we can pick as a basis of W:

$$\mathcal{A} = \left\{ \underline{w}_1 = \begin{pmatrix} 2 \\ -3 \\ 0 \end{pmatrix}, \underline{w}_2 = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \right\}$$

Now, to get ${}_{\mathcal{A}}[g]_{\mathcal{A}}$, we need to compute $f(\underline{w}_1), f(\underline{w}_2)$ (which are used to compute the columns of the representing matrix):

$$f(\underline{w}_1) = \begin{pmatrix} 14 & 8 & 3 \\ -17 & -9 & -3 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -3 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ -7 \\ 2 \end{pmatrix} = 3\underline{w}_1 - 2\underline{w}_2$$

$$f(\underline{w}_2) = \begin{pmatrix} 14 & 8 & 3 \\ -17 & -9 & -3 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ -5 \\ 1 \end{pmatrix} = 2\underline{w}_1 - \underline{w}_2$$

Thus:

$$_{\mathcal{A}}[g]_{\mathcal{A}} = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix}$$

Moreover, we can extend $\mathcal{A}$ to a basis of V, say:

$$\mathcal{B} = \{\underline{w}_1, \underline{w}_2, \underline{e}_1\}$$

Then

3 Diagonalisability

3.1 Diagonalisable Endomorphisms

- **What is a diagonalisable endomorphism?**
  - let $f : V \to V$ be an **endomorphism**
  - f is **diagonalisable** if and only if a basis for V is composed entirely of **eigenvectors** of f
- **How can we think of diagonalisable endomorphisms, in terms of matrices?**
  - consider a finite dimensional V
  - then, f is **diagonalisable** if and only if V has an ordered basis $\mathcal{B}$, with $_{\mathcal{B}}[f]_{\mathcal{B}}$ a **diagonal matrix**
  - in particular, the diagonal entries of $_{\mathcal{B}}[f]_{\mathcal{B}}$ will be the the **eigenvalues** corresponding to the **eigenvectors** in $\mathcal{B}$
- **When is a square matrix diagonalisable?**
  - a **square matrix** $A \in Mat(n; F)$ is **diagonalisable** if and only if its corresponding **linear mapping** $A\circ : F^n \to F^n$ is **diagonalisable**
  - in particular, this means that A will be diagonalisable **if and only if** it is **conjugate** to a diagonal matrix; in fact, $\exists P \in GL(n; F)$ such that:

    $$P^{-1}AP = diag(\lambda_1, \ldots, \lambda_n)$$

    where P is a matrix of the eigenvectors of A, and λ_i are the corresponding eigenvalues

3.1.1 Examples

- the only **diagonalisable** nilpotent matrix is the 0 matrix. If A is nilpotent, then $\exists k$ such that:

  $$A^k = 0$$

  Assuming that A is a diagonalisable, then:

  $$P^{-1}AP = D \implies (P^{-1}AP)^k = D^k \implies D^k = 0$$

  Since D is diagonal, D^k is a diagonal matrix with the entries of D to the power of k. Hence, $(D_{ii})^k = 0 \iff D_{ii} = 0$. Thus, D must be the 0 matrix. But then:

  $$P^{-1}AP = 0 \iff A = P0P^{-1} \implies A = 0$$

- consider the matrix:

$$A = \begin{pmatrix} 7 & 2 \\ -18 & -6 \end{pmatrix}$$

with characteristic polynomial:

$$(7-x)(-6-x)+36 = x^2 - x - 6 = (x-3)(x+2)$$

A has eigenvalues $\lambda = 3, -2$. If $\lambda_1 = -3$, then consider:

$$A - 3I_n = \begin{pmatrix} 4 & 2 \\ -18 & -9 \end{pmatrix}$$

This means that there is an eigenvector:

$$\underline{v}_1 = \langle -1, 2 \rangle$$

Similarly, for $\lambda_2 = 2$, an eigenvector is $\langle 2, -9 \rangle$. We can then construct the matrices:

$$P = \begin{pmatrix} -1 & 2 \\ 2 & -9 \end{pmatrix} \qquad P^{-1} = \frac{1}{5}\begin{pmatrix} -9 & -2 \\ -2 & -1 \end{pmatrix}$$

So computing:

$$P^{-1}AP = \frac{1}{5}\begin{pmatrix} -9 & -2 \\ -2 & -1 \end{pmatrix}\begin{pmatrix} 7 & 2 \\ -18 & -6 \end{pmatrix}\begin{pmatrix} -1 & 2 \\ 2 & -9 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -27 & -6 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} -1 & 2 \\ 2 & -9 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 15 & 0 \\ 0 & -10 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$$

as expected.

3.2 Lemma: Linear Independence of Eigenvectors

Let $f : V \to V$ be an ***endomorphism*** *of a* ***vector space*** *V, and let:*

$$\underline{v}_1, \ldots, \underline{v}_n$$

be ***eigenvectors*** *of f with* ***pairwise different eigenvalues****:*

$$\lambda_1, \ldots, \lambda_n$$

Then, the ***eigenvectors*** *are* ***linearly independent****. [Lemma 4.6.9]*

Proof. Suppose that:

$$\sum_{i=1}^{n} \alpha_i \underline{v}_i = 0$$

Consider the endomorphism:

$$(f - \lambda_2 id_V) \circ \ldots \circ (f - \lambda_n id_V)$$

In particular:

$$\begin{aligned}
&(f-\lambda_2 id_V)\circ\ldots\circ(f-\lambda_n id_V)(\underline{v}_i)\\
=&(f-\lambda_2 id_V)\circ\ldots\circ(f-\lambda_{n-1} id_V)((f-\lambda_n id_v)(\underline{v}_i))\\
=&(f-\lambda_2 id_V)\circ\ldots\circ(f-\lambda_{n-1} id_V)(f(\underline{v}_i)-\lambda_n\underline{v}_i)\\
=&(f-\lambda_2 id_V)\circ\ldots\circ(f-\lambda_{n-1} id_V)(\lambda_i\underline{v}_i-\lambda_n\underline{v}_i)\\
=&(\lambda_i-\lambda_n)(f-\lambda_2 id_V)\circ\ldots\circ(f-\lambda_{n-1} id_V)(\underline{v}_i)\\
=&\prod_{j=2}^{n}(\lambda_i-\lambda_j)\underline{v}_i
\end{aligned}$$

In particular, if $i \neq 1$, the product becomes 0, whilst if $i = 1$ it will be non-zero. Thus, if we apply the endomorphism to $\sum_{i=1}^{n}\alpha_i\underline{v}_i = 0$, it results in:

$$\alpha_1\prod_{j=2}^{n}(\lambda_1-\lambda_j)\underline{v}_1 = \underline{0}$$

Since the eignevalues are all distinct, this will only be true if $\alpha_1 = 0$. Employing similar endomorphisms (with the term $(f - \lambda_k id_V)$ missing from the composition), will imply that $a_k = 0, \forall k \in [1, n]$. Thus, the vectors will be linearly independent.

□

4 The Cayley-Hamilton Theorem

4.1 Theorem: The Cayley-Hamilton Theorem

Let $A \in Mat(n; R)$ be a ***square matrix****, with entries in a* ***commutative ring*** *R.*
Then, evaluating its ***characteristic polynomial $\mathcal{X}_A(x)$ at the matrix A results in 0.***

4.1.1 Examples

Consider the matrix:

$$A = \begin{pmatrix} 14 & 8 & 3 \\ -17 & -9 & -3 \\ 1 & 0 & 0 \end{pmatrix}$$

This has characteristic polynomial:

$$\mathcal{X}_A(x) = x^3 - 5x^2 + 7x - 3$$

Calculating:

$$A^3 - 5A^2 + 7A - 3I = 0$$

5 Workshop

1. **True or false. Any $\mathbb{R}$-linear mapping $f : \mathbb{R}^5 \to \mathbb{R}^5$ has an eigenvalue.**

   This is true. This is because the characteristic polynomial of f will be of degree 5, so over $\mathbb{R}$, it can get factored into linear terms and irreducible quadratics. The linear terms always have a root in $\mathbb{R}$, so f will have at least one real eigenvalue. (Exercise 78 of the Notes)

(a) **Show that the matrices:**

$$\begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

are similar/conjugate.
Recall, 2 matrices are similar if we can find an (invertible) matrix T such that:

$$A = T^{-1}BT \implies TA = BT$$

For this case, notice, the first column is common between the 2, and the 2 are upper triangular. This means we seek:

$$T = \begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix}$$

We compute:

$$TA = \begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 3+2a \\ 0 & 2b \end{pmatrix}$$

$$BT = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix} = \begin{pmatrix} 2 & 2a+b \\ 0 & 2b \end{pmatrix}$$

Hence, we require:

$$3 + 2a = 2a + b \implies b = 3$$

and a is free. We can thus pick:

$$\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

and T is invertible, since $det(T) = 3$.

(b) **Show that the matrices**

$$\begin{pmatrix} 2 & 1 & 2 \\ 0 & 2 & -4 \\ 0 & 0 & 3 \end{pmatrix} \qquad \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

are similar/conjugate.
This time, the first column, second column and the last row are the same. Both of them are upper triangular. Hence, we seek:

$$T = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}$$

We compute:

$$TA = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 2 \\ 0 & 2 & -4 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 2+3a \\ 0 & 2 & 3b-4 \\ 0 & 0 & 3 \end{pmatrix}$$

$$BT = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 2a+b \\ 0 & 2 & 2b \\ 0 & 0 & 3 \end{pmatrix}$$

Hence, we require:

$$3b - 4 = 2b \implies b = 4$$
$$2 + 3a = 2a + b \implies a = 2$$

Hence:

$$T = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}$$

which is invertible, since $det(T) = 1$.

(c) **Put:**

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 3 & 0 \\ 5 & 7 & 0 \end{pmatrix}$$

into triangular form. Show that by conjugating your answer, if necessary, this may be put in the form:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

For this one, the solutions and I diverge in result, but we both produce a triangular matrix which is similar to the one provided.
I'll include both (they both apparently follow Example 4.6.4, but that is very weirdly written - or at least I find it hard to understand).

We need to compute the characteristic polynomial:

$$\begin{aligned} |A - xI_3| &= \begin{vmatrix} 1-x & -1 & 0 \\ 1 & 3-x & 0 \\ 5 & 7 & -x \end{vmatrix} \\ &= -x[(1-x)(3-x)+1] \\ &= -x[x^2 - 4x + 3 + 1] \\ &= -x(x-2)^2 \end{aligned}$$

From solutions: we now compute the eigenvectors. If $\lambda = 0$, then the asociated eigenvector:

$$\underline{v}_1 = (0, 0, 1)^T$$

If $\lambda = 2$ then:

$$\underline{v}_2 = (-1, 1, 1)^T$$

We now construct a basis for $im(A - \lambda I_3)$. If we pick $\lambda = 2$, then:

$$A - 2I_3 = \begin{pmatrix} -1 & -1 & 0 \\ 1 & 1 & 0 \\ 5 & 7 & -2 \end{pmatrix}$$

$im(A - \lambda I_3)$ is spanned by the LiD columns of $A - 2I_3$. In the solutions they pick the eigenvectors as a basis, so that works. Then:

$$A\underline{v}_1 = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 3 & 0 \\ 5 & 7 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \underline{0} = 0\underline{v}_1 + 0\underline{v}_2$$

$$A\underline{v}_2 = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 3 & 0 \\ 5 & 7 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \\ 2 \end{pmatrix} = 0\underline{v}_1 + 2\underline{v}_2$$

We need to extend this basis to a basis for $\mathbb{R}^3$. For this we can just use $\underline{e}_1$. We compute:

$$A\underline{e}_1 = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 3 & 0 \\ 5 & 7 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 5 \end{pmatrix} = 4\underline{v}_1 + \underline{v}_2 + 2\underline{e}_1$$

Using this basis, we construct the representing matrix B:

$$\begin{pmatrix} 0 & 0 & 4 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

which is indeed upper triangular.
We can show that this is similar to the matrix that they give us. Notice, they are practically identical, except for the first row. This tells us to seek:

$$T = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and:

$$\begin{pmatrix} 1 & 0 & a \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 4 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 4+2a \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

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

Thus, we pick $a = -2$:

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

which is invertible, since $det(T) = 1$

As a basis, I just picked the first and third columns of the matrix:

$$\underline{v}_1 = (0, 0, -2)^T \qquad \underline{v}_2 = (-1, 1, 5)^T$$

Then:

$$A\underline{v}_1 = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 3 & 0 \\ 5 & 7 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix} = \underline{0} = 0\underline{v}_1 + 0\underline{v}_2$$

$$A\underline{v}_2 = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 3 & 0 \\ 5 & 7 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \\ 2 \end{pmatrix} = 4\underline{v}_1 + 2\underline{v}_2$$

To extend to a basis of $\mathbb{R}^3$, we can pick $\underline{e}_2$, so we compute:

$$A\underline{e}_2 = \begin{pmatrix} -1 \\ 3 \\ 7 \end{pmatrix} = -\underline{v}_1 + \underline{v}_2 + 2\underline{e}_2$$

Using this basis, we obtain the representing matrix:

$$\begin{pmatrix} 0 & 4 & -2 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

This is identical to the matrix they give us, except for the 2 top-right entries. We thus look for:

$$T = \begin{pmatrix} 1 & a & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

So:

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 4 & -2 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 4+2a & -2+a+2b \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & a & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

Hence, we require:

$$4 + 2a = 0 \implies a = -2$$

and:

$$-2 + a + 2b = 0 \implies 2b = 4 \implies b = 2$$

So we can pick:

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

2. **There are 215 students registered for Honours Algebra. Let $M \in Mat(215; \mathbb{R})$ be the matrix with $M_{ij} = 1$ if student i and student j have met each other, and 0 if they haven't. Here we assume that $M_{ii} = 0$.**

   (a) **Let $\underline{u} \in \mathbb{R}^{215}$ be the vector each of whose entries is 1. What does the vector $M\underline{u}$ represent?**

   The ith component of $M\underline{u}$ represents the number of people which student i has met.

   (b) **What information is contained in M^2?**

   If $i \neq j$, $(M^2)_{ij}$ denotes the number of people which are known by both i and j. Otherwise, if $i = j$, $(M^2)_{ii}$ denotes the total number of people which student i has met.

   To formalise this:

   $$(M^2)_{ij} = \sum_{k=1}^{215} M_{ik} M_{kj}$$

   Now, $M_{ik}M_{kj}$ is non-zero only if i knows k and k knows j. Hence, $\sum_{k=1}^{215} M_{ik}M_{kj}$ counts all the k which are known by both i and j.

   $$(M^2)_{ii} = \sum_{k=1}^{215} M_{ik} M_{ki}$$

   Now, $M_{ik}M_{ki}$ is non-zero only if i knows k, so $\sum_{k=1}^{215} M_{ik}M_{ki}$ is all the students k which i knows.

3. **Let:**

   $$A = \begin{pmatrix} 7 & 2 \\ 1 & 6 \end{pmatrix}$$

   (a) **Find the 2 eigenvalues of A.**

   We compute the characteristic polynomial:

   $$\begin{vmatrix} 7 - x & 2 \\ 1 & 6 - x \end{vmatrix} = (7 - x)(6 - x) - 2 = x^2 - 13x + 40 = (x - 5)(x - 8)$$

   Hence, the 2 eigenvalues are:

   $$\lambda_1 = 5 \qquad \lambda_2 = 8$$

(b) **Find eigenvectors for both eigenvalues.**

For $\lambda_1 = 5$:

$$\begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} \underline{v}_1 = \underline{0} \implies \underline{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

For $\lambda_2 = 8$:

$$\begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix} \underline{v}_2 = \underline{0} \implies \underline{v}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

(c) **Find an invertible matrix P such that $P^{-1}AP$**

We can pick P to be the matrix with eigenvectors as columns. Then:

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

means that:

$$P^{-1}AP = D$$

where:

$$D = \begin{pmatrix} 5 & 0 \\ 0 & 8 \end{pmatrix}$$

4. **Each year $\frac{1}{4}$ of the haggises outside Scotland move in, and $\frac{1}{8}$ of the haggises inside Scholand move out. Let h_k be the number of haggis in Scotland in year k, and let g_k be the number of haggis outside of Scotland in year k, where $k \geq 0$.**

(a) **Write down a matrix equation that describes the number of haggises inside and outside in year 1, in terms of h_0 and g_0. The matrix you write should have the following 2 properties:**

- **each column sums to 1**
- **each entry is non-negative**

Notice, we have 2 equations:

$$h_1 = \frac{7}{8}h_0 + \frac{1}{4}g_0$$
$$g_1 = \frac{1}{8}h_0 + \frac{3}{4}g_0$$

So we consider the matrix equation:

$$\frac{1}{8}\begin{pmatrix} 7 & 2 \\ 1 & 6 \end{pmatrix}\begin{pmatrix} h_0 \\ g_0 \end{pmatrix} = \begin{pmatrix} h_1 \\ g_1 \end{pmatrix}$$

Notice, the columns must add to 1 and be positive because the represent the proportions of haggises which enter or leave Scotland.

(b) **Write down a matrix equation that describes the number of haggises inside and outside in year k, in terms of h_0 and g_0. This matrix should satisfy the same properties as the matrix above.**

For h_k, g_k, we just need to consider powers of the matrix:

$$\frac{1}{8^k}\begin{pmatrix} 7 & 2 \\ 1 & 6 \end{pmatrix}^k \begin{pmatrix} h_0 \\ g_0 \end{pmatrix} = \begin{pmatrix} h_k \\ g_k \end{pmatrix}$$

The columns still represent proportions, so they will still add up to 1, and be positive.

(c) **Use a variation on your solution to Exercise 3 to solve this.**

We notice that our equation matrix is $\frac{1}{8}A$ from exercise 3, so in particular it is diagonalisable and:

$$\begin{aligned}
\frac{1}{8^k}\begin{pmatrix} 7 & 2 \\ 1 & 6 \end{pmatrix}^k &= \frac{1}{8^k}(P^{-1}DP)^k \\
&= \frac{1}{8^k}P^{-1}D^kP \\
&= \frac{1}{3\times 8^k}\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 5^k & 0 \\ 0 & 8^k \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \\
&= \frac{1}{3\times 8^k}\begin{pmatrix} 5^k & -2\times 8^k \\ 5^k & 8^k \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \\
&= \frac{1}{3\times 8^k}\begin{pmatrix} 5^k + 2\times 8^k & 2\times(5^k - 8^k) \\ 5^k - 8^k & 2\times 5^k + 8^k \end{pmatrix}
\end{aligned}$$

So then:

$$\begin{aligned}
\begin{pmatrix} h_k \\ g_k \end{pmatrix} &= \frac{1}{3\times 8^k}\begin{pmatrix} 5^k + 2\times 8^k & 2\times(5^k - 8^k) \\ 5^k - 8^k & 2\times 5^k + 8^k \end{pmatrix}\begin{pmatrix} h_0 \\ g_0 \end{pmatrix} \\
&= \frac{1}{3\times 8^k}\begin{pmatrix} h_0(5^k + 2\times 8^k) + g_0(2\times(5^k - 8^k)) \\ h_0(5^k - 8^k) + g_0(2\times 5^k + 8^k) \end{pmatrix} \\
&= \begin{pmatrix} \frac{h_0}{3}\left(\left(\frac{5}{8}\right)^k + 2\right) + \frac{2g_0}{3}\left(\left(\frac{5}{8}\right)^k - 1\right) \\ \frac{h_0}{3}\left(\left(\frac{5}{8}\right)^k - 1\right) + \frac{g_0}{3}\left(2\left(\frac{5}{8}\right)^k + 1\right) \end{pmatrix}
\end{aligned}$$

(d) **What do you expect the proportion of haggises inside and outside Scotland to be in the long run?**

As $k \to \infty$, we get that:

$$\begin{pmatrix} h_k \\ g_k \end{pmatrix} \to \begin{pmatrix} \frac{2}{3}(h_0 + g_0) \\ \frac{1}{3}(h_0 + g_0) \end{pmatrix}$$

Since $h_0 + g_0$ is the total number of haggis, we expect that $\frac{2}{3}$ of the total haggis stays in Scotland, and that $\frac{1}{3}$ leaves Scotland.