特征向量

3.11 Eigenvectors and eigenvalues

not eigenvector

if we take subspace W we find $T: W \to W$ or $T: W \to V|W$

We focus our attention on the study of linear transformations $T: V \to V$.

A very useful strategy is to determine the *T-invariant subspaces of V*, that is subspaces W such that $T(\mathbf{w}) \in W$ for all $\mathbf{w} \in W$.

Example 3.64. Consider

$$A = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$T: \mathbb{R}^4 \to \mathbb{R}^4$
$v \to Av$

$W_1 = \text{span}(e_1)$

if $w \in W_1$, then $w = \lambda e_1$ for some $\lambda \in \mathbb{R}$

$T(w_1) = Aw = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \lambda \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2\lambda \\ 0 \\ 0 \\ 0 \end{bmatrix} = (-2\lambda)e_1 \in W_1$

$T(e_1) = -2e_1$

$T(e_4) = 0\, e_1$

$\vec{0}$ not be an eigenvector

$W_2 = \text{span}(e_4)$, then T-invariant

$T(e_4) = \vec{0}$ so $T(\lambda e_4) = \vec{0}$

0 can be an eigenvector

$W_3 = \text{span}(e_2, e_3)$ $T(e_2) = e_2$ $T(e_3) = -e_3$

so $T(\lambda e_2 + \mu e_3) = T(\lambda e_2) + T(\mu e_3) = \lambda e_2 - \mu e_3 \in W_3$

It turns out that the simplest nontrivial case is extremely important: T-invariant lines (one-dimensional subspaces) L in V.

Let $\mathbf{v} \in L$ be any nonzero vector, then $\{\mathbf{v}\}$ is a basis for L. We want $T(\mathbf{v}) \in L$, so there exists $\lambda \in \mathbb{R}$ such that

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

The vector $\mathbf{v}$ is called an *eigenvector of T* and the scalar λ is called an *eigenvalue of T*.

Conversely, if a nonzero $\mathbf{v} \in V$ and $\lambda \in \mathbb{R}$ satisfy $T(\mathbf{v}) = \lambda \mathbf{v}$, then $\text{Span}\{\mathbf{v}\}$ is a one-dimensional T-invariant subspace of V.

Example 3.65. Consider the linear transformation $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

What is the effect on the vectors $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$?

$$T\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} = (-1)\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

3.11.1 Finding eigenvalues

Proposition 3.66. A scalar λ is an eigenvalue of the matrix A if and only if $x = \lambda$ is a solution of the equation

$$\det(A - xI) = 0.$$

Proof. λ is an eigenvalue of A $\xLeftrightarrow{\text{if and only if}}$ $T(v) = Av = \lambda v$

$T: v \to v$

$Av = \lambda v$

$\Leftrightarrow \exists v \neq 0$ such that

$\Leftrightarrow Av - \lambda v = 0$

$\Leftrightarrow (A - \lambda I)v = 0$

$\Leftrightarrow A - \lambda I$ is not invertible

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

Example 3.67. Find the eigenvalues of $\begin{bmatrix} 2 & 2 \\ -2 & 7 \end{bmatrix}$.

$$\det(A-\lambda I) = \det\left(\begin{bmatrix} 2 & 2 \\ -2 & 7 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right)$$

$$= \det\left(\begin{bmatrix} 2-\lambda & 2 \\ -2 & 7-\lambda \end{bmatrix}\right) = 0$$

$$(2-\lambda)(7-\lambda)+4=0$$

$$\lambda^2-9\lambda+18=0$$

$$(\lambda-6)(\lambda-3)=0$$ $\Rightarrow$ since A is 2×2 matrix so 2-degree polynomial

$\lambda=6$ or $\lambda=3$

so the eigenvalue of $\begin{bmatrix} 2 & 2 \\ -2 & 7 \end{bmatrix}$ is 3 or 6

If A is an $n \times n$ matrix, then the expression $\det(A - xI)$ is a polynomial of degree n in x, called the *characteristic polynomial of A*.

The equation $\det(A - xI) = 0$ is called the *characteristic equation of A*.

When A has real entries, it is possible that some of its eigenvalues are complex (and not real). like $x^2+ax+b=0$ have no real solution but complex solution

When A has entries in $\mathbb{F}_2$, it is possible that some of its eigenvalues are not in $\mathbb{F}_2$ (exactly where they are is a matter for MAST30005 Algebra).

Example 3.68. Find the eigenvalues of $\begin{bmatrix} 3 & 2 & 1 \\ 0 & 3 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.

$$\det\left(\begin{bmatrix} 3 & 2 & 1 \\ 0 & 3 & 1 \\ 1 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \right)$$

$$= \det\left(\begin{bmatrix} 3-\lambda & 2 & 1 \\ 0 & 3-\lambda & 1 \\ 1 & 0 & 1-\lambda \end{bmatrix} \right)$$

$$\Rightarrow 3-\lambda \begin{bmatrix} 3-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix} - 2 \begin{bmatrix} 0 & 1 \\ 1 & 1-\lambda \end{bmatrix} + 1 \begin{bmatrix} 0 & 3-\lambda \\ 1 & 0 \end{bmatrix}$$

$$\Rightarrow 3-\lambda \left[(3-\lambda)(1-\lambda) \right] - 2(-1) \quad - 1(3-\lambda)$$

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

$$= (3-\lambda)^2 (1-\lambda) - (1-\lambda)$$

$$= \left[(3-\lambda)^2 - 1 \right] (1-\lambda) = 0$$

$\lambda_1 = 1$

$(3-\lambda)^2 - 1 = 0$

$3-\lambda = \pm 1$

$\lambda = 4$ or 2

$\boxed{\lambda = 1 \text{ or } 2 \text{ or } 4}$

3.11.2 Finding eigenvectors

If λ is an eigenvalue of A, then the corresponding eigenvectors are the solutions of the vector equation

$$(A - \lambda I)\mathbf{v} = \mathbf{0}.$$

In other words, these are the nonzero elements in the kernel of $A - \lambda I$. In this setting, the kernel of $A - \lambda I$ is called the *λ-eigenspace of A*.

We solve this equation as usual, via row reduction.

Example 3.69. Find eigenvectors for the eigenvalues obtained in Example 3.67.

from 3.67

$A=\begin{bmatrix}2&2\\-2&7\end{bmatrix}$ eigenvalue 3 & 6.

for $\lambda=3$

free parameter

$$A-3I=\begin{bmatrix}2&2\\-2&7\end{bmatrix}-\begin{bmatrix}3&0\\0&3\end{bmatrix}=\begin{bmatrix}-1&2\\-2&4\end{bmatrix}\xrightarrow{R_2\leftarrow R_2-2R_1}\begin{bmatrix}-1&2\\0&0\end{bmatrix}\xrightarrow{R_1\leftarrow(-R_1)}\begin{bmatrix}1&-2\\0&0\end{bmatrix}$$

$x_2=t$

$x_1=2t$

$\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}2\\1\end{bmatrix}t$ so $\begin{bmatrix}2\\1\end{bmatrix}$ is an eigenvector with eigenvalue 3

$A\begin{bmatrix}2\\1\end{bmatrix}=3\begin{bmatrix}2\\1\end{bmatrix}$

for $\lambda=6$

$$A-6I=\begin{bmatrix}2&2\\-2&7\end{bmatrix}-\begin{bmatrix}6&0\\0&6\end{bmatrix}=\begin{bmatrix}-4&2\\-2&1\end{bmatrix}\xrightarrow{R_1\leftarrow R_1-2R_2}\begin{bmatrix}0&0\\-2&1\end{bmatrix}\longrightarrow\begin{bmatrix}1&-\frac{1}{2}\\0&0\end{bmatrix}$$

$x_2=t$

$x_1=\frac{1}{2}t$

$\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}\frac{1}{2}\\1\end{bmatrix}t$, $t\in\mathbb{R}$ so $\begin{bmatrix}\frac{1}{2}\\1\end{bmatrix}$ is an eigenvector with eigenvalue 6

Example 3.70. Find the eigenspaces of $\begin{bmatrix}3&2&4\\2&0&2\\4&2&3\end{bmatrix}$.

$A=\begin{bmatrix}3&2&4\\2&0&2\\4&2&3\end{bmatrix}$

$$\det(A-\lambda I)=\cdots=(3-\lambda)(-3\lambda+\lambda^2-4)-2(6-2\lambda-8)+4(4+4\lambda)$$
$$=(3-\lambda)(\lambda-4)(\lambda+1)+4(\lambda+1)+16(\lambda+1)$$
$$=(-\lambda^2+7\lambda-12+20)(\lambda+1)$$
$$=(-\lambda^2+7\lambda+8)(\lambda+1)$$
$$=-(\lambda^2-7\lambda-8)(\lambda+1)=-(\lambda-8)(\lambda+1)^2=0$$

$\lambda=-1$ or $\lambda=8$

for $\lambda=-1$

2 free parameter

$A-(-I)=\cdots\begin{bmatrix}1&\frac{1}{2}&1\\0&0&0\\0&0&0\end{bmatrix}$

$x_1=-t-\frac{1}{2}s$

$x_2=s$

$x_3=t$

$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}-1\\0\\1\end{bmatrix}t+\begin{bmatrix}-\frac{1}{2}\\1\\0\end{bmatrix}s$

Then -1-eigenspace is span$\left\{\begin{bmatrix}-1\\0\\1\end{bmatrix},\begin{bmatrix}-\frac{1}{2}\\1\\0\end{bmatrix}\right\}$

$\mathbb{R}^3$

for $\lambda=8$ $A-8I=\begin{bmatrix}-5&2&4\\2&-8&2\\4&2&-3\end{bmatrix}\xrightarrow{\cdots}\begin{bmatrix}1&-4&1\\0&18&-9\\0&0&0\end{bmatrix}\to\begin{bmatrix}1&-4&1\\0&2&-1\\0&0&0\end{bmatrix}$

The 8-eigenspace is span$\left\{\begin{bmatrix}1\\\frac{1}{2}\\1\end{bmatrix}\right\}$

V_8 V_4

dim 1 dim 2.

$$\begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix} t \longrightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

Theorem 3.71. Let $T: V \to V$ be a linear transformation. Any set of eigenvectors of T with distinct eigenvalues is linearly independent.

Proof. In other words, if v_1 is eigenvector of λ_1

v_2 is eigenvector of λ_2

$\vdots$

v_n - - - - - λ_n

and $\lambda_1 \neq \lambda_2 \neq \cdots \neq \lambda_n$, then $\{v_1, v_2, \cdots, v_n\}$ is linear independent

Corollary 3.72. If A is an $n \times n$ matrix and the dimensions of the eigenspaces of A add up to n, then A has n linearly independent eigenvectors.

Proof.

Example 3.73. In Example 3.70,

$\lambda_1 = -1$ $\quad v_1 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} v_2$

$\lambda_2 = 8$ $\quad \begin{bmatrix} 1 \\ \frac{1}{2} \\ 1 \end{bmatrix} v_3$

v_1, v_2, v_3 linear independent

Example 3.74. How many linearly independent eigenvectors does $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ have?

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1-\lambda & 2 \\ 0 & 1-\lambda \end{bmatrix} = (1-\lambda)^2$$

$$\lambda = 1$$

1- eigenspan

$$A - I = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \to \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$x_1 = t$$

$$x_2 = 0$$

$$v = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} t$$

3.11.3 The Cayley–Hamilton Theorem

Theorem 3.75. Let A be an $n \times n$ matrix and let

$$\det(A - xI) = a_0 + a_1 x + \cdots + a_n x^n$$

be the characteristic polynomial of A. Then

$$a_0 I + a_1 A + \cdots + a_n A^n = \mathbf{0}.$$

Example 3.76. Consider $A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$.

$$\det(A - \underset{A}{\underset{\uparrow}{\lambda}} I) = \begin{vmatrix} 3-\lambda & 2 \\ 1 & 4-\lambda \end{vmatrix}$$

$$= \lambda^2 - 7\lambda + 10$$

$$10I - 7A + A^2 = \vec{0}$$

We can make use of the characteristic equation to compute powers of A:

from 3.7b

$$10I - 7A + A^2 = 0$$

$$-7A + A^2 = 10I$$

$$(-7I + A)A = -10I$$

$$\underbrace{\frac{-1}{10}\left(\frac{-1}{7}I + A\right)}_{A^{-1}}A = I$$

not all matrix is invertible

for matrix is not invertible

$\Rightarrow$ the polynomial has 0 root

kernal of A has non-zero solution

227 $Ax = 0$ has non-zero solution

3.11.4 Diagonalisation

Suppose you are given a linear transformation $T\colon V \to V$. We have already seen that some choices of ordered basis of V yield simpler matrices than other choices.

Two natural questions present themselves:

(a) How simple can we make the matrix representation?

(b) How can we compute an ordered basis that achieves such a simple matrix representation?

The ultimate answer to these questions depends on the field of scalars of V, and takes us to MAST20022 *Group Theory and Linear Algebra*, MAST30005 *Algebra*, and beyond.

We will be focussing on a special case that is very useful in practice, that where the matrix representation is diagonal.

We say that a linear transformation $T\colon V \to V$ is *diagonalisable* if there exists an ordered basis $\mathcal{B}$ such that the matrix $[T]_{\mathcal{B}}$ is diagonal.

We say that an $n \times n$ matrix A is *diagonalisable* if the associated linear transformation $\mathbb{R}^n \to \mathbb{R}^n$ given by $\mathbf{v} \mapsto A\mathbf{v}$ is diagonalisable.

Theorem 3.77. A linear transformation $T\colon V \to V$ is diagonalisable if and only if V has a basis $\mathcal{B}$ consisting of eigenvectors for T.

Proof.

$$[T]_B = \begin{bmatrix} | & | & & | \\ [T(v_1)]_B & [T(v_2)]_B & \cdots & [T(v_n)]_B \\ | & | & & | \end{bmatrix} \quad \text{where } B=\{v_1, \cdots v_n\}$$

$$\Longleftrightarrow [T(v_1)]_B = \begin{bmatrix} d_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad [T(v_n)]_B = \begin{bmatrix} 0 \\ \vdots \\ d_n \end{bmatrix}$$

$$\Longleftrightarrow [T(v_1)]_B = d_1 v_1 \qquad \cdots [T(v_n)]_B = d_n v_n$$

$$\Longleftrightarrow v_1, \cdots v_n \text{ are eigenvector of } T$$

Example 3.78. Consider $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$.

Example 3.79. Consider $\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$.

seen $\lambda = -1$, $\lambda = 8$ eigenvalue

-1-eigenspace dim $= 2$ span$\left\{ \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$

8-eigenspace, span$\left\{ \begin{bmatrix} 1 \\ \frac{1}{2} \\ 1 \end{bmatrix} \right\}$

Get 3 linearly indep eigenvectors for the basis of $\mathbb{R}^3$

so ... is diagonalisable

Corollary 3.80. An $n \times n$ matrix A is diagonalisable if and only if there exist a diagonal $n \times n$ matrix D and an invertible $n \times n$ matrix P such that

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

Moreover, the columns of P are linearly independent eigenvectors of A and the diagonal entries of D are the corresponding eigenvalues.

Proof.

by def: A is diagonalisable $\Leftrightarrow$ $T: \mathbb{R}^n \to \mathbb{R}^n$ given by $T(v) = Av$ is diagonalisable

change of basis

$$\underset{P^{-1}}{P_{B \leftarrow S}} \; \underset{A}{[T]_S} \; \underset{P}{P_{S \leftarrow B}} = \underbrace{[T]_B}_{D}$$

Let $D = [T]_B$ $\quad P = \begin{bmatrix} | & | & | & & | \\ v_1 & v_2 & v_3 & \cdots & v_n \\ | & | & | & & | \end{bmatrix}$

Example 3.81. Consider $\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$.

$-1 \qquad 8$

$$B = \left\{ \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ \frac{1}{2} \\ 1 \end{bmatrix} \right\}$$

$$P = \begin{bmatrix} -\frac{1}{2} & -1 & 1 \\ 1 & 0 & \frac{1}{2} \\ 0 & 1 & 1 \end{bmatrix}$$

Then $P^{-1}AP = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 8 \end{bmatrix}$

Proposition 3.82. If an $n \times n$ matrix A has $\underbrace{n \text{ distinct eigenvalues}}_{\text{no repeation}}$, then A is diagonalisable.

Proof.

$\lambda_1, \cdots \lambda_n$ distinct

$\downarrow$

$v_1 \cdots v_n$ eigenvector $\rightarrow$ independent

$\downarrow$

basis

Example 3.83. Consider $\begin{bmatrix} 2 & 2 \\ -2 & 7 \end{bmatrix}$. $= A$

eigenvalue 3 & 6

so A in diagonalisable

$P^{-1}AP = \begin{bmatrix} 3 & 0 \\ 0 & 6 \end{bmatrix}$

3.11.5 Matrix powers and Markov chains

Example 3.84. Suppose that 2% of Victorians move to Queensland every year, 1% of Queenslanders move to Victoria every year, and everyone else stays put.

Starting with populations of 5.1 million Queenslanders and 6.3 million Victorians, what will happen after 20 years?

$v_0 = \begin{bmatrix} 5.1 \\ 6.3 \end{bmatrix} \longrightarrow v_1 = \begin{bmatrix} 0.99 \times 5.1 + 0.02 \times 6.3 \\ 0.98 \times 6.3 + 0.01 \times 5.1 \end{bmatrix}$

$= \underbrace{\begin{array}{cc} Q & V \\ \left[\begin{array}{cc} 0.99 & 0.02 \\ 0.01 & 0.98 \end{array}\right] & \end{array}}_{A} \begin{bmatrix} 5.1 \\ 6.3 \end{bmatrix}$

$v_{20} = A^{20} v_0$

$\Downarrow$

$v_n = A^n v_0$

So we have to compute a high power of a matrix A. There is one situation when this is very easy, namely for a diagonal matrix

$$D = \begin{bmatrix} d_1 & & 0 \\ & d_2 & \\ 0 & & \ddots & \\ & & & d_n \end{bmatrix} \quad \Rightarrow \quad D^k = \begin{bmatrix} d_1^k & & & \\ & d_2^k & & \\ & & \ddots & \\ & & & d_n^k \end{bmatrix}$$

The next best thing is when A is diagonalisable:

$$P^{-1}AP = D \quad \Rightarrow$$

$$P(P^{-1}AP) = PD$$
$$(PP^{-1})AP = PD$$
$$AP = PD$$
$$APP^{-1} = PDP^{-1}$$
$$A = PDP^{-1}$$

$$A^k = (PDP^{-1})^k$$
$$= (PD\cancel{P^{-1})(P}D\cancel{P^{-1})(P}D\cancel{P^{-1}}) \cdots (\cancel{P}DP^{-1})$$
$$= P\underbrace{DDDD\cdots}_{k}P^{-1}$$
$$= PD^kP^{-1}$$

Going back to our population example:

eigenvector

$$\det(A-\lambda I) = \begin{vmatrix} 0.99-\lambda & 0.02 \\ 0.01 & 0.98-\lambda \end{vmatrix} = (0.99-\lambda)(0.98-\lambda) - 0.02 \cdot 0.01$$
$$= 0.9702 - 1.97\lambda + \lambda^2 - 0.0002$$
$$= \lambda^2 - 1.97\lambda + 0.97$$
$$= (\lambda - 0.97)(\lambda - 1)$$

$$\lambda_1 = 0.97 \quad \lambda_2 = 1$$

for $\lambda = 0.97$

free

$$A - \lambda I = \begin{bmatrix} -0.01 & 0.02 \\ 0.01 & -0.02 \end{bmatrix} \to \begin{bmatrix} 0.01 & -0.02 \\ 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}$$

$$u_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$\lambda = 1$ $$A - \lambda I = \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \to \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad u_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

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

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 0.97 \end{bmatrix}$$

Note that the matrix $A = \begin{bmatrix} 0.98 & 0.01 \\ 0.02 & 0.99 \end{bmatrix}$ has nonnegative entries, and the sum of each column is 1.

The model used in the population study is an example of a *Markov chain*: For us, this will be a system that has a finite number of states, and follows a sequence of events in which the probability of passing from state i to state j depends only on i and j (and not on the past events).

The population example can be thought of as having two states: any person can be either in Victoria or in Queensland.

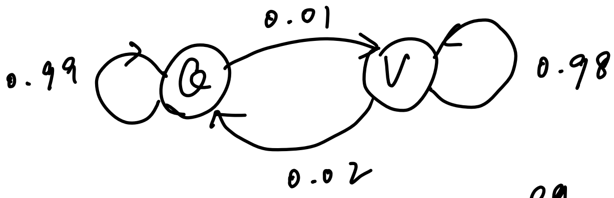
$$\lim_{n\to\infty} A^n = \lim_{n\to\infty} (P D^n P^{-1}) = P \cdot \lim_{n\to\infty} D^n \cdot P^{-1}$$

$$= P \lim_{n\to\infty} \begin{bmatrix} 1^n & 0 \\ 0 & 0.97^n \end{bmatrix} P^{-1}$$

$$= P \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot P^{-1}$$

$$= \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

More generally, to a Markov chain we associate a *transition matrix* A as follows: The entry in column i and row j is the probability of going from state i to state j.



in one step to get from a to b (possibility).

A^2 in two step ... (possibility)

eg

$$A = \begin{array}{c} \text{from} \ldots \\ \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \\ P_{41} & P_{42} & P_{43} & P_{44} \end{bmatrix} \end{array} \text{to} \ldots$$

So the entries of the transition matrix are always between 0 and 1 (being probabilities). Also, each column keeps track of the probabilities of all possible outcomes from a given state, so the entries must add up to 1.

A matrix with these properties is called a *stochastic matrix*.

Theorem 3.85. Let A be a stochastic matrix that is regular, that is there exists $n \in \mathbb{N}$ such that all entries of A^n are positive. Then

(a) 1 is an eigenvalue of A.

(b) $\lim_{k\to\infty} A^k$ exists and all its columns are equal to the same vector $\mathbf{v}$.

(c) $A\mathbf{v} = \mathbf{v}$. (This $\mathbf{v}$ is called the *steady state vector*.)

Proof.

$$A = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ \vdots & & & \\ p_{m1} & \cdots & & p_{mn} \end{bmatrix}$$

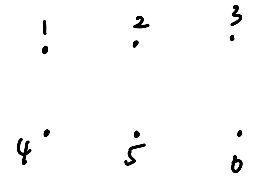
$$A^T = \begin{bmatrix} p_{11} & \cdots & p_{m1} \\ p_{12} & \ddots & \vdots \\ \vdots & & \\ p_{1n} & \cdots & p_{mn} \end{bmatrix} \qquad A^T \underbrace{\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}}_{\text{eigenvector}} = \begin{bmatrix} p_{11} + \cdots + p_{m1} \\ p_{12} \cdots p_{m2} \\ \vdots \\ p_{1n} \cdots p_{mn} \end{bmatrix} = \underbrace{1}_{\text{eigenvalue}} \times \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix},$$

it turns out that A^T and A has the same eigenvalue

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

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

Example 3.86 (A random walk)**.** Consider the following arrangement of 6 dots:



A walker is allowed to move up, down, left, or right from a dot to a neighbouring dot. The probability is the same for all the neighbours of a dot.

The transition matrix is

$$A = \begin{bmatrix} 0 & \frac{1}{3} \\ \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \\ \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \\ 0 & 0 \end{bmatrix}$$

Example 3.87 (Google's PageRank algorithm)**.** Google's original search algorithm models web browsing as a random walk on the N webpages making up the Internet.

If page j has $\ell_j \neq 0$ links then the probability a user will move to page i from page j is

$$p_{ij} = \begin{cases} \frac{1-q}{\ell_j} + \frac{q}{N} & \text{if page } j \text{ links to page } i \\ \frac{q}{N} & \text{if page } j \text{ does not link to page } i \end{cases}$$

Usually $q = 0.15$.

If page j has no links, then

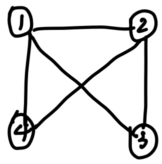
$$p_{ij} = \frac{1}{N} \quad \text{for all } i.$$

This determines a regular stochastic matrix, which has a steady state vector. This is the PageRank; it has nonnegative entries that add to 1. Its j-th entry is the proportion of time that page j is visited.

Suppose we consider the PageRank on an intranet with 4 pages linked as follows:

内部网

suppose	page	Link
	1	2 3 4
	2	3 4
	3	–
	4	2



The transition matrix is

To get the steady state vector we compute

$$\lim_{k\to\infty} A^k =$$

4 Inner product spaces

We go back to the notion of dot product on $\mathbb{R}^n$. This operation allowed us to define the length of a vector, the angle between two vectors (including the concept of orthogonality), and the distance between two vectors. As these concepts are very useful, we explore the possibility of extending them to general vector spaces.

To do so, recall the following properties of the dot product:

1. $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$
2. $(\lambda\mathbf{u})\cdot\mathbf{v} = \lambda(\mathbf{u}\cdot\mathbf{v})$
3. $(\mathbf{u}+\mathbf{v})\cdot\mathbf{w} = \mathbf{u}\cdot\mathbf{w} + \mathbf{v}\cdot\mathbf{w}$
4. (a) $\mathbf{u}\cdot\mathbf{u} \geq 0$

   (b) $\mathbf{u}\cdot\mathbf{u} = 0 \Rightarrow \mathbf{u} = \mathbf{0}$.

4.1 Inner products

Let V be a vector space with field of scalars $\mathbb{R}$.

An *inner product on V* is a function

$$\langle\, , \rangle : V \times V \to \mathbb{R}$$

<u,v>

satisfying

Prototype

eg. in $\mathbb{R}^2$ $\langle u, v\rangle = u \cdot v = u_1v_1 + u_2v_2$

1. $\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{u}\rangle$
2. $\langle \lambda\mathbf{u}, \mathbf{v}\rangle = \lambda\langle \mathbf{u}, \mathbf{v}\rangle$
3. $\langle \mathbf{u}+\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle$
4. (a) $\langle \mathbf{u}, \mathbf{u}\rangle \geq 0$

   (b) $\langle \mathbf{u}, \mathbf{u}\rangle = 0 \Rightarrow \mathbf{u} = \mathbf{0}$.

Example 4.1. Consider

$$\langle \mathbf{u}, \mathbf{v}\rangle = u_1v_1 + 2u_2v_2$$

on $\mathbb{R}^2$.

① $\langle u, v\rangle = u_1v_1 + 2u_2v_2 = v_1u_1 + 2v_2u_2 = \langle v, u\rangle$

② $\langle \lambda u, v\rangle = \lambda u_1v_1 + 2\lambda u_2v_2 = \lambda\langle u, v\rangle$

$\lambda u = (\lambda u_1, \lambda u_2)$

③ $\langle u+v, w\rangle = \langle (u_1+v_1, u_2+v_2), w\rangle$

$= (u_1+v_1)w_1 + (u_2+v_2)w_2$

$= (u_1w_1 + u_2w_2) + (v_1w_1 + v_2w_2)$

$= \langle u, w\rangle + \langle v, w\rangle$

④ $\langle u, u\rangle = u_1^2 + 2u_2^2 \geq 0$

⑤ $\langle u, u\rangle = 0 \Rightarrow u_1^2 = 2u_2^2 = 0$

$\Rightarrow u_1 = u_2 = 0$

$u = 0$

Example 4.2. Consider

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + u_1v_2 + u_2v_2$$

on $\mathbb{R}^2$.

Consider $u = (1,0)$ & $v = (2,1)$

$<u, v> = 2+1+0 = 3$

$<v, u> = v_1u_1 + v_1u_2 + v_2u_2$

$= 2 + 0 + 0 = 2$

not an inner product

Example 4.3. Consider

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 - u_2v_2 + u_3v_3$$

on $\mathbb{R}^3$.

$u = (0,1,0)$

$<u, u> = 0 - 1 + 0 = -1 \quad < 0$

not an inner product