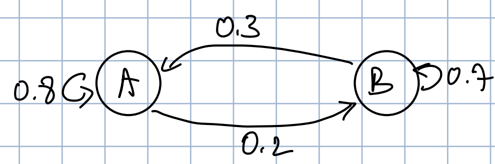
MODULE 7: MARKOV CHAINS AND EIGENVALUES

Linear Algebra 3:
Determinants & Eigenvalues

TOPIC 1: Markov Chains

Markov Chain and Steady States

Example:
A small town has 2 libraries, A and B.
After 1 month, among the books checked out of A,
↳ 80% returned to A.
↳ 20% returned to B.



After 1 month, among the books checked out of B,
↳ 30% returned to A
↳ 70% returned to B

If both libraries have 1000 books today, how many books does each library have after 1 month? After 1 year? After n months?

(assumption)
The books are equally divided by between the two branches, denoted by $\vec{x}_0 = \begin{bmatrix} .5 \\ .5 \end{bmatrix}$. What is the distribution after 1 month, call it $\vec{x}_1$? After two months?

After k months, the distribution is $\vec{x}_k$, which is what in terms of $\vec{x}_0$?

$$\vec{x}_1 = \begin{pmatrix} \text{proportion of books A, 1 month} \\ \text{proportion of books B, 1 month} \end{pmatrix} = \begin{pmatrix} .8 \times .5 + .3 \times .5 \\ .2 \times .5 + .7 \times .5 \end{pmatrix} = \begin{pmatrix} .8 & .3 \\ .2 & .7 \end{pmatrix} \begin{pmatrix} .5 \\ .5 \end{pmatrix}$$
$$= P\vec{x}_0 .$$

$$\left.\begin{aligned} \vec{x}_2 &= P\vec{x}_1 = P(P\vec{x}_0) = P^2\vec{x}_0 \\ \vec{x}_3 &= P\vec{x}_2 = P^3\vec{x}_0 \end{aligned}\right\} \vec{x}_k = P^k\vec{x}_0$$

Markov Chains

A few definitions:
↳ A <u>probability vector</u> is a vector, $\vec{x}$, with non-negative elements that sum to 1.
↳ A <u>stochastic matrix</u> is a square matrix, P, whose columns are probability vectors.
↳ A <u>Markov chain</u> is a sequence of probability vectors $\vec{x}_k$, and a stochastic matrix P, such that:

$$\vec{x}_{k+1} = P\vec{x}_k, \quad k = 0, 1, 2, \ldots$$

↳ A <u>steady-state vector</u> for P is a probability vector $\vec{q}$ such that $P\vec{q} = \vec{q}$.

Example: Identifying a Steady-State

Determine a steady-state vector for the stochastic matrix

$$P = \begin{pmatrix} .8 & .3 \\ .2 & .7 \end{pmatrix}$$

A steady-state would satisfy $P\vec{q} = \vec{q}$.

$$P\vec{q} - \vec{q} = \vec{0}$$
$$P\vec{q} - I\vec{q} = \vec{0}, \quad \because I\vec{q} = \vec{q}$$
$$(P-I)\vec{q} = \vec{0}$$
$$P - I = \begin{pmatrix} .8-1 & .3 \\ .2 & .7-1 \end{pmatrix} = \begin{pmatrix} -.2 & .3 \\ .2 & -.3 \end{pmatrix}, \left(\begin{array}{cc|c} -.2 & .3 & 0 \\ .2 & -.3 & 0 \end{array}\right)$$

$$\left.\begin{array}{l} -2x_1 + 3x_2 = 0 \\ \text{Set } x_2 = 2. \Rightarrow x_1 = 3. \end{array}\right\} \vec{q} = \frac{1}{\underbrace{5}_{3+2}}\begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

$3+2$ (such that $\frac{1}{3+2}(3) + \frac{1}{3+2}(2) = 1$).

Exercise:

Consider the following city migration problem, where there are two cities, X and Y. There are 12000 people among the two cities. Every year,

* 70 percent of the people from X stay in X, the remaining 30 percent move to Y.
* 40 percent of the people from Y stay in Y, the remaining 60 percent move to X.

Construct a Markov Chain for the above process and identify the steady state vector.

(a) If the Markov process is in a steady state, what is the population in the city X?

(b) If the Markov process is in a steady state, what is the population in the city Y?

Assuming the number of people are equally distributed among the two cities.

$$\vec{x}_0 = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}. \quad \vec{x}_1 = \begin{pmatrix} 0.7 \times 0.5 + 0.6 \times 0.5 \\ 0.3 \times 0.5 + 0.4 \times 0.5 \end{pmatrix} = \begin{pmatrix} 0.7 & 0.6 \\ 0.3 & 0.4 \end{pmatrix}\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$$
$$= P\vec{x}_0.$$

$$P = \begin{pmatrix} 0.7 & 0.6 \\ 0.3 & 0.4 \end{pmatrix}.$$

To get steady-state vector $\vec{q}$, $P\vec{q} = \vec{q}$.

$$(P-I)\vec{q} = \vec{0}.$$

$$P - I = \begin{pmatrix} 0.7-1 & 0.6 \\ 0.3 & 0.4-1 \end{pmatrix} = \begin{pmatrix} -0.3 & 0.6 \\ 0.3 & -0.6 \end{pmatrix}, \left(\begin{array}{cc|c} -0.3 & 0.6 & 0 \\ 0.3 & -0.6 & 0 \end{array}\right)$$

$-0.3x_1 + 0.6x_2 = 0$. Let $x_2 = 1. \Rightarrow x_1 = 2$.

$$\Rightarrow \text{steady-state vector } \vec{q} = \frac{1}{2+1}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Population of the two cities:

$$\frac{1}{3}\begin{pmatrix} 2 \\ 1 \end{pmatrix} \times 12000 = \begin{pmatrix} 8000 \\ 4000 \end{pmatrix}.$$

$\therefore$ Population in X = 8000 #

Population in Y = 4000 #

Markov Chain Convergence

We often want to know what happens to a process,

$$\vec{x}_{k+1} = P\vec{x}_k, \quad k = 0, 1, 2, \ldots \quad \text{——} \ (1)$$

as $k \to \infty$.

We may want to know, for example, if the sequence generated by (1),

$$\vec{x}_1, \vec{x}_2, \vec{x}_3, \ldots$$

will converge to a steady-state, and if so, what those steady-state vectors are.

Regular Stochastic Matrices

DEFINITION — Regular Stochastic Matrices

↳ A stochastic matrix P is regular if there is some k such that P^k only contains strictly positive entries.

This matrix is regular stochastic:

$$A = \begin{pmatrix} 0.1 & 0.7 \\ 0.9 & 0.3 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 1 & 7 \\ 9 & 3 \end{pmatrix}$$

Another example of a regular stochastic matrix:

$$B = \frac{1}{10}\begin{pmatrix} 0 & 1 & 2 \\ 2 & 8 & 7 \\ 8 & 1 & 1 \end{pmatrix}$$

Note that

$$B^2 = \frac{1}{10^2}\begin{pmatrix} 0 & 1 & 2 \\ 2 & 8 & 7 \\ 8 & 1 & 1 \end{pmatrix}^2 = \frac{1}{100}\begin{pmatrix} 18 & 10 & 9 \\ 72 & 73 & 67 \\ 10 & 17 & 24 \end{pmatrix}$$

Because B^2 has strictly positive entries for $k=2$, B is stochastic.
It can be very difficult to determine whether a matrix is regular stochastic.

Convergence and Regular Stochastic Matrices

Theorem:
If P is a regular stochastic matrix, then P has a unique steady-state vector $\vec{q}$, and $\vec{x}_{k+1} = P\vec{x}_k$ converges to $\vec{q}$ as $k \to \infty$.

Example: Car Rental company

A car rental company has 3 rental locations, A, B and C. Cars can be returned at any location. The table below gives the pattern of rental and returns for a given week.

		rented from		
		A	B	C
returned to	A	0.8	0.1	0.2
	B	0.2	0.6	0.3
	C	0	0.3	0.5

There are 1000 cars at each location today.
(a) Construct a stochastic matrix, P, for the given problem.
(b) What happens to the distribution of cars after a long time? You may assume that P is regular.

(a) If $x_{A,k}, x_{B,k}, x_{C,k}$ are the number of cars in week k at locations A, B, C respectively, then after one week,

$$\begin{aligned} x_{A,1} &= 0.8x_{A,0} + 0.1x_{B,0} + 0.1x_{C,0} \\ x_{B,1} &= 0.2x_{A,0} + 0.6x_{B,0} + 0.3x_{C,0} \\ x_{C,1} &= \phantom{0.2x_{A,0} +{}} 0.3x_{B,0} + 0.5x_{C,0} \end{aligned} \quad \Rightarrow \vec{x}_1 = P\vec{x}_0.$$

where $P = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.2 & 0.6 & 0.3 \\ 0 & 0.3 & 0.5 \end{pmatrix}$ *

(b) P is regular $\Rightarrow$ can assume that P has a unique steady-state vector $\vec{q}$, and $\vec{x}_{k+1} = P\vec{x}_k$ converges to $\vec{q}$ as $k \to \infty$.

The steady-state vector, $\vec{q}$ is given by:

$$\begin{aligned} P\vec{q} &= \vec{q} \\ P\vec{q} - \vec{q} &= \vec{0} \\ (P - I)\vec{q} &= \vec{0} \end{aligned}$$

$\vec{q}$ is a probability vector in Null $(P-I)$. We need to calculate $P-I$ and $\vec{q}$...

$$P - I = \begin{pmatrix} 0.8-1 & 0.1 & 0.2 \\ 0.2 & 0.6-1 & 0.3 \\ 0 & 0.3 & 0.5-1 \end{pmatrix} = \begin{pmatrix} -0.2 & 0.1 & 0.2 \\ 0.2 & -0.4 & 0.3 \\ 0 & 0.3 & -0.5 \end{pmatrix}$$

$\vec{q}$ is a vector in the null space of the above matrix. We apply the usual process for finding a vector in the null space of a matrix.

$$\left(\begin{array}{ccc|c} -0.2 & 0.1 & 0.2 & 0 \\ 0.2 & -0.4 & 0.3 & 0 \\ 0 & 0.3 & -0.5 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} -2 & 1 & 2 & 0 \\ 0 & -3 & 5 & 0 \\ 0 & 3 & -5 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 6 & 0 & 11 & 0 \\ 0 & -3 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

x_3 is free. Let $x_3 = 6$. $\Rightarrow$ $6x_1 + 11x_3 = 0 \Rightarrow x_1 = -11$.
$-3x_2 + 5x_3 = 0 \Rightarrow x_2 = 10$.

Therefore, a vector in the null space of $P - I$ is $\begin{pmatrix} 11 \\ 10 \\ 6 \end{pmatrix}$.

A probability vector in the null space is $\vec{q} = \frac{1}{27}\begin{pmatrix} 11 \\ 10 \\ 6 \end{pmatrix}$. This is our steady-state vector.

No matter what the initial distribution of cars happen to be, after a long period of time, the distribution of cars is given by $\vec{q}$.

Exercise:
Consider the city migration problem used earlier, where there are two cities, X and Y, and every year;

* 70% of the people from X stay in X, the remaining 30% percent move to Y.
* 40% of the people from Y stay in Y, the remaining 60% percent move to X.

The initial population of X is 10000. The initial population of Y is 20000.

After a long period of time, what is the population in city X?
(Hint: first construct a stochastic matrix, P, so that $x_{k+1} = Px_k$ gives the distribution of people at month $k+1$).

Total Population = 10000 + 20000 = 30000.

$\vec{x}_0 = \begin{pmatrix} 1/3 \\ 2/3 \end{pmatrix}$. $\vec{x}_1 = \begin{pmatrix} 1/3 \times 0.7 + 2/3 \times 0.6 \\ 1/3 \times 0.3 + 2/3 \times 0.4 \end{pmatrix} = \begin{pmatrix} 0.7 & 0.6 \\ 0.3 & 0.4 \end{pmatrix}\begin{pmatrix} 1/3 \\ 2/3 \end{pmatrix} = P\vec{x}_0$.

$\Rightarrow P = \begin{pmatrix} 0.7 & 0.6 \\ 0.3 & 0.4 \end{pmatrix}$

$P\vec{q} = \vec{q} \Rightarrow P\vec{q} - \vec{q} = \vec{0} \Rightarrow (P-I)\vec{q} = \vec{0}$

$P - I = \begin{pmatrix} -0.3 & 0.6 \\ 0.3 & -0.6 \end{pmatrix}$, $\left(\begin{array}{cc|c} -0.3 & 0.6 & 0 \\ 0.3 & -0.6 & 0 \end{array}\right)$

$-0.3x_1 + 0.6x_2 = 0$. Let $x_2 = 1$. $\Rightarrow x_1 = 2$.

$\Rightarrow \vec{q} = \frac{1}{3}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Population, $\frac{1}{3}\begin{pmatrix} 2 \\ 1 \end{pmatrix} \times 30000 = \begin{pmatrix} 20000 \\ 10000 \end{pmatrix}$. $\therefore$ Population of X = 20000 #

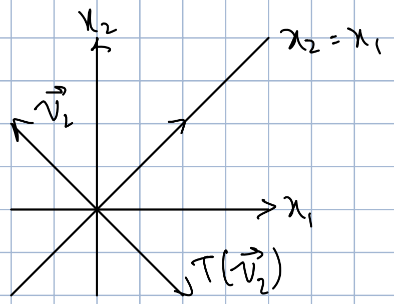
TOPIC 2: Eigenvalues and Eigenvectors

Motivating Problem: Consider the linear transform

$$T_A(\vec{x}) = A\vec{x} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Can you state a nonzero vector, $\vec{v} \in \mathbb{R}^2$ that satisfies the following equation?

$$A\vec{v} = \lambda\vec{v}, \ \lambda \in \mathbb{R}$$



$\vec{v}_1 = \begin{pmatrix}1\\1\end{pmatrix}$, $T(\vec{v}_1) = A\vec{v}_1 = \begin{pmatrix}1\\1\end{pmatrix} = \lambda\vec{v}_1$, $\lambda = 1$.

or: $\vec{v}_2 = \begin{pmatrix}-1\\1\end{pmatrix}$, $T(\vec{v}_2) = A\vec{v}_2 = \begin{pmatrix}-1\\1\end{pmatrix} = -1\vec{v}_2$

Eigenvalues and Eigenvectors

If $A \in \mathbb{R}^{n\times n}$, and there is a $\vec{v} \neq \vec{0}$ in $\mathbb{R}^n$, and

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

then $\vec{v}$ is an eigenvector for A, and $\lambda \in \mathbb{C}$ is the corresponding eigenvalue.

Note that
* we will only consider the case where A is square
* even when all entries of A and $\vec{v}$ are real, λ can be complex (a rotation of the plane has no real eigenvalues)
* if $\lambda \in \mathbb{R}$, then → $\lambda > 0 \Rightarrow A\vec{v}$ and $\vec{v}$ point in the same direction
  ↳ $\lambda < 0 \Rightarrow A\vec{v}$ and $\vec{v}$ point in opposite directions

Determining Whether a Vector is an Eigenvector

Which of the following vectors are eigenvectors of $A = \begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}$?

(a) $\vec{v}_1 = \begin{pmatrix}1\\1\end{pmatrix}$ (b) $\vec{v}_2 = \begin{pmatrix}2\\2\end{pmatrix}$ (c) $\vec{v}_3 = \begin{pmatrix}1\\-1\end{pmatrix}$ (d) $\vec{v}_4 = \begin{pmatrix}2\\1\end{pmatrix}$ (e) $\vec{v}_5 = \begin{pmatrix}0\\0\end{pmatrix}$

(a) $A\vec{v}_1 = \lambda\vec{v}_1$?
$A\vec{v}_1 = \begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}2\\2\end{pmatrix} = 2\vec{v}_1$, $\lambda = 2$
$\therefore v_1$ is an eigenvector of A.

(b) $A\vec{v}_2 = \lambda\vec{v}_2$?
$A\vec{v}_2 = \begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}2\\2\end{pmatrix} = \begin{pmatrix}4\\4\end{pmatrix} = 2\vec{v}_2$, $\lambda = 2$.
$\therefore \vec{v}_2$ is an eigenvector of A.

(c) $A\vec{v}_3 = \lambda\vec{v}_3$?
$A\vec{v}_3 = \begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} = 0\vec{v}_3$, $\lambda = 0$.
$\therefore \vec{v}_3$ is an eigenvector of A.

(d) $A\vec{v}_4 = \lambda\vec{v}_4$?
$A\vec{v}_4 = \begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}3\\3\end{pmatrix} \neq \lambda v^4$.
$\therefore \vec{v}_4$ is not an eigenvector of A.

(e) $A\vec{v}_5 = \lambda\vec{v}_5$?
By definition, $\vec{v}_5$ (a zero vector) cannot be an eigenvector.

Determining whether a Number is an Eigenvalue

Determine whether $\lambda = 3$ is an eigenvalue of $A = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix}$.

$$A\vec{v} = \lambda\vec{v} = 3\vec{v}$$
$$A\vec{v} - 3\vec{v} = \vec{0} \Rightarrow A\vec{v} - 3I\vec{v} = \vec{0}$$
$$(A - 3I)\vec{v} = \vec{0}$$

* A - 3I needs to be singular

$$A - 3I = \begin{pmatrix} 2-3 & -4 \\ -1 & -1-3 \end{pmatrix} = \begin{pmatrix} -1 & -4 \\ -1 & -4 \end{pmatrix} \Rightarrow \text{singular matrix.}$$

$\therefore \lambda = 3$ is an eigenvalue of A.

Another Interpretation of What Eigenvalues Are

From the previous examples, we saw how an eigenvalue of a matrix is a number, λ, that
↳ satisfies $A\vec{v} = \lambda\vec{v}$ for eigenvector $\vec{v}$
↳ makes $A - \lambda I$ singular

Recall: a singular matrix is not invertible!

Exercise: Suppose A is the matrix below.

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

Which of the following are eigenvectors of matrix A that correspond to eigenvalue $\lambda = 1$?

(a) $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ By definition, zero vectors cannot be eigenvectors of A.

(b) $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $A\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\lambda = 3$. $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector of A.

(c) $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ $A\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\lambda = 1$. $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is an eigenvector of A. ✓

Eigenspaces

DEFINITION — Eigenspaces

↳ Suppose $A \in \mathbb{R}^{n \times n}$. The eigenvectors for a given λ span a subspace of $\mathbb{R}^n$ called the λ-eigenspace of A.

Note: the λ-eigenspace for matrix A is $\text{Nul}(A - \lambda I)$, because:

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

If $\vec{v} \neq \vec{0}$, we must have that $A - \lambda I$ is singular. Thus, eigenvectors span the null space of $A - \lambda I$.

Eigenspaces in $\mathbb{R}^2$

Construct a basis for the eigenspaces for the matrix whose eigenvalues are given. Sketch the eigenvectors and eigenspaces.

$$A = \begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix}, \quad \lambda = -1, 2$$

$\lambda = -1$: $(A - (-1)I \mid \vec{0})$

$$= \left(\begin{array}{cc|c} 6 & -6 & 0 \\ 3 & -3 & 0 \end{array}\right)$$

x_2 is free, $x_1 = x_2$.

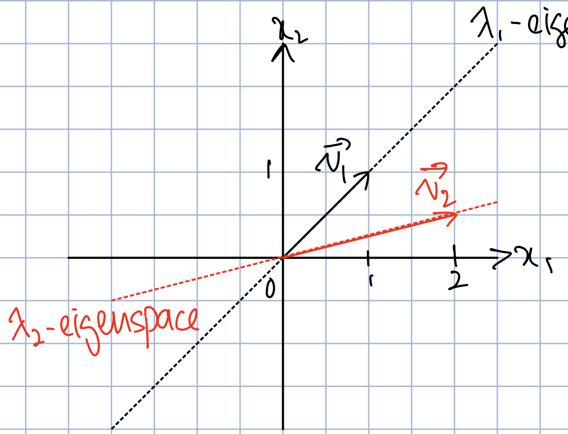
Let $\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

$\lambda = 2$: $(A - 2I \mid \vec{0})$

$$= \left(\begin{array}{cc|c} 3 & -6 & 0 \\ 3 & -6 & 0 \end{array}\right)$$

x_2 is free, $x_1 = 2x_2$.

Let $\vec{v}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.



Basis for λ_1-eigenspace is $\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$.

Basis for λ_2-eigenspace is $\left\{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}$.

Eigenspaces in $\mathbb{R}^3$

Construct a basis for the eigenspaces for the matrix whose eigenvalues are given.

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

$$\lambda_1 = 1: \ (A - \lambda_1 I \mid 0) = \left(\begin{array}{ccc|c} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right).$$

$x_1 = 0$.
$x_2 = -x_3$
x_3 is free.

$$\vec{v}_1 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 3: \ (A - \lambda_2 I \mid 0) = \left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right).$$

x_1, x_3 are free
$x_2 = x_3$.

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

A basis for λ_1-eigenspace is $\left\{\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}\right\}$; a basis for λ_2-eigenspace is $\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right\}$.

Exercise:

Suppose A is the matrix below.

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

$\lambda = -2$ is an eigenvalue of A.

What is the dimension of the eigenspace for $\lambda = -2$?

$$\lambda = -2 : (A + 2I \mid 0) = \left(\begin{array}{ccc|c} -2 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ -2 & 1 & 3 & 0 \end{array}\right).$$

$A - \lambda I$ has only one pivot, so there are 2 non-pivot columns.
nonpivotal columns = dimension of eigenspace = 2. *

x_2 and x_3 are free.
$2x_1 = x_2 + x_3$
Let $x_2 = x_3 = 1$. $x_1 = 1$.
Let $x_2 = 0$, $x_3 = 1$. $x_1 = \frac{1}{2}$

Let $\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 1/2 \\ 0 \\ 1 \end{pmatrix}$.

Basis for λ-eigenspace
$: \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 0 \\ 1 \end{pmatrix} \right\}$

What is the dimension of the eigenspace for $\lambda = -1$?

$$\lambda = -1 : (A + I \mid 0) = \left(\begin{array}{ccc|c} -3 & 1 & 3 & 0 \\ 0 & -1 & 0 & 0 \\ -2 & 1 & 2 & 0 \end{array}\right).$$

$A - \lambda I$ has 2 pivots, so there is one non-pivot column.
non-pivotal columns = dimension of eigenspace = 1. *

$x_2 = 0$.
$x_1 = x_3$. Let $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Dimension of eigenspace = 3.

Eigenvalue Theorems

Motivation: Suppose A is a real $n \times n$ matrix?
↳ How can we determine the eigenvalues of A?
↳ If A has some special structure (eg. singular, stochastic, triangular), what can be said about the eigenvalues of A?

Theorems that deal with eigenvalues of a matrix can help us calculate eigenvalues.

The following theorems can help us identify eigenvalues or eigenvectors of a matrix.

* The diagonal elements of a triangular matrix are its eigenvalues.

* A is not invertible $\Leftrightarrow$ 0 is an eigenvalue of A.

* Stochastic matrices have an eigenvalue equal to 1.

* If $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ are eigenvectors that correspond to distinct eigenvalues, then $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ are linearly independent.

Proofs of these theorems are relatively short. There are many other eigenvalue theorems that we could explore!

Determining Eigenvalues by Inspection

By inspection, give two eigenvalues for each of the following matrices.

① $A = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$, singular $\Rightarrow$ 0 is an eigenvalue. $\lambda_1 = 0$.
stochastic $\Rightarrow \lambda_2 = 1/2$.

② $B = \begin{pmatrix} 2 & 3 \\ 0 & 5 \end{pmatrix}$ B is triangular $\Rightarrow \lambda_1 = 2, \lambda_2 = 5$.

Warning!
We cannot determine the eigenvalues of a matrix from its reduced form.
↳ row reductions can change the eigenvalues of a matrix

Suppose $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. The eigenvalues and corresponding eigenvectors are

$$\lambda_1 = 2,\ \vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\quad \lambda_2 = 0,\ \vec{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

We can verify this:

$$A\vec{v}_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2\vec{v}_1$$

$$A\vec{v}_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0\vec{v}_2$$

But the RREF of A is $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, whose eigenvalues are 1 and 0.

Exercise:
Suppose vectors $\vec{v}_1$ and $\vec{v}_2$ are eigenvectors of A, and $\vec{v}_1 \neq \vec{v}_2$.
Which of the following situations are possible?

✗ $\vec{v}_1$ and $\vec{v}_2$ correspond to different eigenvalues, and are linearly dependent.

Eigenvectors corresponding to different eigenvalues are always independent.

✓ $\vec{v}_1$ and $\vec{v}_2$ correspond to the same eigenvalue, and are linearly dependent.

If $\vec{x}$ is an eigenvector of A for eigenvalue λ, then $k\vec{x}$ is also an eigenvector of A.

✓ $\vec{v}_1$ and $\vec{v}_2$ correspond to the same eigenvalue, and are linearly independent.

Eigenvectors corresponding to the same eigenvalue can be independent.

TOPIC 3: The Characteristic Equation

The Characteristic Polynomial

Recall: λ is an eigenvalue of $A \Leftrightarrow (A-\lambda I)$ is not invertible.

Therefore, to calculate the eigenvalues of A, we can solve

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

* $\det(A-\lambda I)$ is the characteristic polynomial of A
* $\det(A-\lambda I)=0$ is the characteristic equation of A
* the roots of the characteristic polynomial are the eigenvalues of A

Example: Calculating the Eigenvalues of a 2x2 Matrix

The characteristic polynomial of $A = \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}$ is:

$$\begin{aligned} \det(A-\lambda I) &= \det\begin{pmatrix} 4-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix} \\ &= (4-\lambda)(1-\lambda) - 4 \\ &= \lambda^2 - 5\lambda + 4 - 4 \\ &= \lambda^2 - 5\lambda = \lambda(\lambda-5) \end{aligned}$$

So, the eigenvalues of A are: $\lambda = 0, 5$

Characteristic Polynomial of 2x2 Matrices

The trace of a matrix is the sum of its diagonal elements.

Example:
Express the characteristic equation of $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in terms of its determinant and trace.

$$\begin{aligned} 0 &= \det(M-\lambda I) \\ &= (a-\lambda)(d-\lambda) - bc \\ &= ad - \lambda(a+d) + \lambda^2 - bc \\ &= \lambda^2 - \lambda(a+d) + (ad-bc) \\ &= \lambda^2 - \lambda(\text{trace } M) + \det(M) \end{aligned}$$

Using the Trace to Identify Eigenvalues

Although the characteristic polynomial can always be used to determine eigenvalues, sometimes we can identify eigenvalues by inspection.

Example: By inspection, what are the eigenvalues of A?

$$A = \begin{pmatrix} 6 & 18 \\ 3 & 9 \end{pmatrix}$$

singular: $\lambda_1 = 0$
trace $M = 6+9 = 15 = \lambda_2$

Numerical Notes:

* The eigenvalues of any matrix larger than 2x2 should be found using a computer, unless the matrix has a special structure.

* Software for computing eigenvalues tend to avoid the characteristic polynomial.

* Nevertheless, the characteristic polynomial is important for theoretical purposes.

Exercise:
Suppose A is the matrix below:

$$A = \begin{pmatrix} 2 & 4 \\ 3 & 6 \end{pmatrix}$$

Characteristic polynomial of A
$= \det(A - \lambda I)$
$= (2-\lambda)(6-\lambda) - 4(3)$
$= 12 - 8\lambda + \lambda^2 - 12$
$= \lambda^2 - 8\lambda = \lambda^2 + b\lambda + c.$

The characteristic polynomial has the form $\lambda^2 + b\lambda + c$.

What is b equal to? $b = -\text{trace}(M) = 2+6 = 8$
What is c equal to? $c = \det(M) = 2(6) - 4(3) = 0$.

Algebraic Multiplicity

DEFINITION — Algebraic Multiplicity

↳ The algebraic multiplicity of an eigenvalue is its multiplicity as a root of the characteristic polynomial.

Example:
Compute the algebraic multiplicities of the eigenvalues for the matrix

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

λ	Algebraic multiplicities
1	1
0	2
-1	1

Characteristic polynomial $= (1-\lambda)(0-\lambda)(-1-\lambda)(0-\lambda)$

Geometric Multiplicity

DEFINITION — Geometric Multiplicity

↳ The geometric multiplicity of an eigenvalue λ is the dimension of $\text{Null}(A - \lambda I)$.

* Geometric multiplicity is always at least 1. It can be smaller than algebraic multiplicity.
* Here is the basic example:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{Characteristic polynomial} = (0-\lambda)(0-\lambda)$$

$\lambda = 0$ is the only eigenvalue. Its algebraic multiplicity is 2, but the geometric multiplicity is 1.

$$\left(\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right), \quad \begin{array}{l} x_2 = 0, \\ x_1 \text{ is free} \end{array}. \quad \vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Properties of Algebraic and Geometric Multiplicities

Suppose that, for an $n \times n$ matrix A,
* a_i is the algebraic multiplicity of λ_i
* g_i is the geometric multiplicity of λ_i

The algebraic and geometric multiplicities have the following properties.
* $1 \leq a_i \leq n$
* $1 \leq g_i \leq a_i$

Example:
Give an example of a 4×4 matrix with $\lambda = 0$ the only eigenvalue, but the geometric multiplicity of $\lambda = 0$ is one.

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \lambda = 0,\ a = 2,\ g = 1.$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \lambda = 0,\ a = 4,\ g = 1$$

Example: Algebraic Multiplicity
For what values of k does the matrix have one eigenvalue with algebraic multiplicity 2?

$$\begin{pmatrix} -3 & k \\ 2 & -6 \end{pmatrix}$$

$$0 = \det\begin{pmatrix} -3-\lambda & k \\ 2 & -6-\lambda \end{pmatrix}$$
$$= (-3-\lambda)(-6-\lambda) - 2k$$
$$0 = \lambda^2 + 9\lambda + 18 - 2k$$
$$\lambda = \frac{-9 \pm \sqrt{81 - 4(1)(18-2k)}}{2(1)}$$
$$\lambda = -\frac{9}{2} \pm \frac{\sqrt{81 - 8(9-k)}}{2}$$

To obtain repeated eigenvalues, $\sqrt{81 - 8(9-k)} = 0$. $\Rightarrow 9 - k = \frac{81}{8}$
$\therefore k = -9/8$ ✱

Exercise:
Suppose A is the matrix below.

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

For what value of k does A have one eigenvalue with algebraic multiplicity 2 and geometric multiplicity 1?

$$0 = (1-\lambda)(3-\lambda) - k$$
$$= \lambda^2 - 4\lambda + 3 - k.$$
$$\lambda = \frac{4 \pm \sqrt{16 - 4(3-k)}}{2(1)}$$
$$= 2 \pm \sqrt{4-(3-k)}$$
$$\lambda = 2 \pm \sqrt{1+k}.$$

To obtain repeated eigenvalues, $\sqrt{1+k} = 0. \Rightarrow k = -1.$

The Long Term Behavior of Markov Chains

Recall that we often want to know what happens to a Markov chain

$$\vec{x}_{k+1} = P\vec{x}_k, \quad k = 0, 1, 2, \ldots$$

as $k \to \infty$.

If P is a regular stochastic matrix there will be a unique steady state

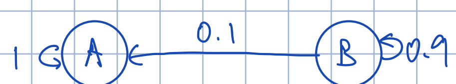
Now we can explore the following questions.
* if we do not know whether P is regular, what else might we do to describe the long-term behavior of the system?
* what can eigenvalues tell us about the behavior of these systems?

Example: Eigenvalues and Markov Chains

Consider the Markov chain:

$$\vec{x}_{k+1} = P\vec{x}_k = \begin{pmatrix} 1 & 0.1 \\ 0 & 0.9 \end{pmatrix}\vec{x}_k, \quad k = 0, 1, 2, 3, \ldots, \quad \vec{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

This system can be represented schematically with two nodes, A and B:



Goal: use eigenvalues to describe the long-term behavior of our system.

Use the eigenvalues and eigenvectors of P to determine what $\vec{x}_k$ tends to as $k \to \infty$. The eigenvalues and eigenvectors of P are

$$\lambda_1 = 1, \ \vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \ \lambda_2 = 0.9, \ \vec{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$x_0 = c_1\vec{v}_1 + c_2\vec{v}_2$$
$$x_1 = Px_0 = P(c_1\vec{v}_1 + c_2\vec{v}_2) = c_1P\vec{v}_1 + c_2P_2\vec{v}_2$$
$$= c_1\lambda_1\vec{v}_1 + c_2\lambda_2\vec{v}_2$$

$$x_2 = Px_1 = P(c_1\lambda_1\vec{v}_1 + c_2\lambda_2\vec{v}_2)$$
$$= c_1\lambda_1^2\vec{v}_1 + c_2\lambda_2^2\vec{v}_2$$
$$\vdots$$
$$x_k = Px_k = c_1\lambda_1^k\vec{v}_1 + c_2\lambda_2^k\vec{v}_2$$

$\lambda_1 = 1,\ \lambda_2 = 0.9$, so as $k\to\infty$, $x_k \to c_1\vec{v}_1$

What is c_1?

$$x_0 = \begin{pmatrix}1\\0\end{pmatrix} = c_1\vec{v}_1 + c_2\vec{v}_2 = c_1\begin{pmatrix}1\\0\end{pmatrix} + c_2\begin{pmatrix}1\\-1\end{pmatrix}. \quad \left(\begin{array}{cc|c}1 & 1 & 1\\ 0 & -1 & 0\end{array}\right)$$
$$\Rightarrow c_2 = 0,\ c_1 = 1.$$
$$\therefore x_k \to v_1$$

A More General Example:

The eigenvalues of a 3×3 stochastic matrix A are

$$\lambda_1 = 1,\ \lambda_2 = \frac{1}{4},\ \lambda_3 = \frac{1}{8}$$

The respective eigenvectors corresponding to these eigenvalues are $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$.

If $\vec{p}$ is a probability vector in $\mathbb{R}^3$, what does $A^k\vec{p}$ tend to as $k\to\infty$?

$$A^k\vec{p} = A^k(c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3)$$
$$= c_1A^k\vec{v}_1 + c_2A^k\vec{v}_2 + c_3A^k\vec{v}_3$$
$$A^k\vec{p} = c_1\lambda_1^k\vec{v}_1 + c_2\lambda_2^k\vec{v}_2 + c_3\lambda_3^k\vec{v}_3$$

Since $\lambda_2 < 1$ and $\lambda_3 < 1$, $c_2\lambda_2^k\vec{v}_2 \to 0$ and $c_3\lambda_3^k\vec{v}_3 \to 0$.

Also, $\lambda_1 = 1 \Rightarrow c_1\lambda_1^k\vec{v}_1 \to c_1\vec{v}_1$ as $k\to\infty$.

$$\therefore A^k\vec{p} \to c_1\vec{v}_1 \quad \text{as } k\to\infty.$$

Exercise:

Suppose $\vec{v}_1, \vec{v}_2$ are eigenvectors of a 3×3 matrix A that correspond to eigenvalues λ_1, λ_2.

$$\vec{v}_1 = \begin{pmatrix}3\\5\\0\end{pmatrix},\ \vec{v}_2 = \begin{pmatrix}5\\0\\7\end{pmatrix}.\ \lambda_1 = 1,\ \lambda_2 = \frac{1}{10}$$

Vector $\vec{p}$ is such that $\vec{p} = \vec{v}_1 - 13\vec{v}_2$.

Suppose that as $k\to\infty$, we have that $A^k\vec{p} \to \vec{q}$, where

$$\vec{q} = \begin{pmatrix}a\\b\\c\end{pmatrix}.$$

What is a equal to?

$$A\vec{p} = A(\vec{v}_1 - 13\vec{v}_2)$$
$$= \lambda_1\vec{v}_1 - 13\lambda_2\vec{v}_2 = \vec{v}_1 - \frac{13}{10}\vec{v}_2$$
$$A^2\vec{p} = A\left(\vec{v}_1 - \frac{13}{10}\vec{v}_2\right) = \vec{v}_1 - \frac{13}{10^2}\vec{v}_2$$
$$A^3\vec{p} = A\left(\vec{v}_1 - \frac{13}{10^2}\vec{v}_2\right) = \vec{v}_1 - \frac{13}{10^3}\vec{v}_3$$
$$\vdots$$
$$A^k\vec{p} = \vec{v}_1 - \frac{13}{10^k}\vec{v}_2.$$

As $k\to\infty$, $A^k\vec{p} \to \vec{v}_1 + 0 = \vec{v}_1 = \begin{pmatrix}3\\5\\0\end{pmatrix}$. $\therefore a = 3$.

Similar Matrices

Motivation: suppose A is an $n \times n$ matrix.

* in some applications, we need to compute A^k for large k.
* computing A^k directly could require many computations, especially if n is large and many of the elements in A are nonzero

Using the concept of similar matrices, we can obtain a more efficient approach.

DEFINITION — Similar Matrices

↳ $n \times n$ matrices A and B are similar if there is a P so that $A = PBP^{-1}$.

Example: if $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$, then $P^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ and

$$PBP^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = A.$$

By construction, A is similar to B.

Theorem (Similar Matrices and the Characteristic Polynomial)
If A and B are similar, then they have the same characteristic polynomial.

Proof: The characteristic polynomial of A is $\det(A - \lambda I)$, and if $A = PBP^{-1}$, then

$$\begin{aligned}
A - \lambda I &= PBP^{-1} - \lambda I \\
&= PBP^{-1} - \lambda PP^{-1} \\
&= (PB - \lambda P)P^{-1} \\
A - \lambda I &= P(B - \lambda)P^{-1} \\
\det(A - \lambda I) &= \det(P(B-\lambda)P^{-1}) \\
&= \det(P)\det(B - \lambda)\det(P^{-1}) \\
&= \det(P)\det(P^{-1})\det(B - \lambda) \\
\det(A - \lambda I) &= \det(B - \lambda) \qquad \because \det(P)\det(P^{-1}) = 1.
\end{aligned}$$

The characteristic polynomials of A and B are the same. (QED)

Note:

* If two matrices have the same characteristic polynomial, then they have the same eigenvalues.
* The converse is not always true: two matrices can have the same eigenvalues but not be similar.

Consider:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad \lambda = 0, 0$$

Can A be similar? If A and B are similar, then $A = PBP^{-1}$, but

$$PBP^{-1} = P\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}P^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq A$$

True or False:

(1) If A is similar to the identity matrix, then $A=I$.

$A = PIP^{-1} = PP^{-1} = I \Rightarrow$ true.

(2) If A is similar to B, and $A = PBP^{-1}$, then $A^2 = PB^2P^{-1}$.

$$A^2 = (PBP^{-1})^2 = PBP^{-1}PBP^{-1}$$
$$= PBBP^{-1} = PB^2P^{-1}. \Rightarrow \text{true}.$$

(3) If A and B have the same eigenvalues, then A and B are similar. False

(4) If A is similar to B, then we can find an invertible matrix P so that $A^k = PB^kP^{-1}$.

If the matrices are similar, then by definition of similar matrices we can find an invertible P so that

$$A = PBP^{-1}.$$

Note also that

$$A^2 = (PBP^{-1})(PBP^{-1}) = PB^2P^{-1}.$$

We used that $P^{-1}P = I$. Likewise,

$$A^3 = (PBP^{-1})(PBP^{-1})(PBP^{-1}) = PB^3P^{-1}.$$

In general,

$$A^k = PB^kP^{-1}.$$

Note that one of the reasons we introduce similar matrices is so that we can have an efficient way of computing A^k.