Diagonalization

Bioinformatics Degree
Algebra

Departament de Matemàtiques

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Outline

Outline

Example: Population growth (Leslie model)

The Vollmar-Wasserman beetles (revisited)

- x_1 = number of youths (beetles 0 to 1 years old)
- x_2 = number of juveniles (beetles 1 to 2 year old)
- x_3 = number of adults (beetles 2 to 3 year old)

We put these numbers in a vector $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$.

We want to study the number of youths, juveniles and adults in a certain year k, assuming that this year is $k = 0$.
We write $x_i(k)$ for the quantity x_i in year k, and also write this information as a vector $x(k)$:

$$x(k) = \begin{pmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{pmatrix}.$$

We know that:

$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{pmatrix}}_{A} \begin{pmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{pmatrix} \text{ this is, } \mathbf{x}(k+1) = A\mathbf{x}(k)$$

Definition

Definition

A **homogeneous linear discrete dynamical system** is a matrix equation of the form

$$\mathtt{x}(k+1) = A\,\mathtt{x}(k), \quad k \in \mathbb{N},$$

where A is an $n \times n$ square matrix , and

$$\mathtt{x}(k) = \begin{pmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{pmatrix} \in \mathbb{R}^n.$$

The vector $\mathtt{x}(0)$ is called an **initial condition**.
A **solution** (or trajectory) is a collection of vectors $\{\mathtt{x}(k)\}_{k\geq 0}$ such that each $\mathtt{x}(k)$ satisfies the equation above.

Solutions

Lemma

The solutions to the system $\mathtt{x}(k+1) = A\,\mathtt{x}(k)$ are $\{\mathtt{x}(k)\}_{k\geq 0}$ with

$$\mathtt{x}(k) = A^k \mathtt{x}(0),\ k \geq 1.$$

- There's a unique solution with given initial condition $\mathtt{x}(0)$.
- The constant solutions $\mathtt{x}(k) = \mathtt{x}$ for all k are called steady states.
- If $\mathtt{x}$ is a steady state $\Rightarrow \mathtt{x} = A\mathtt{x}$, so $\mathtt{x}$ is either 0 or an eigenvector of A of eigenvalue 1.

Solutions

Lemma

The solutions to the system $\mathtt{x}(k+1) = A\,\mathtt{x}(k)$ are $\{\mathtt{x}(k)\}_{k\geq 0}$ with

$$\mathtt{x}(k) = A^k \mathtt{x}(0),\ k \geq 1.$$

- There's a unique solution with given initial condition $\mathtt{x}(0)$.
- The constant solutions $\mathtt{x}(k) = \mathtt{x}$ for all k are called steady states.
- If $\mathtt{x}$ is a steady state $\Rightarrow \mathtt{x} = A\mathtt{x}$, so $\mathtt{x}$ is either 0 or an eigenvector of A of eigenvalue 1.

Solutions

Lemma

The solutions to the system $\mathtt{x}(k+1) = A\,\mathtt{x}(k)$ are $\{\mathtt{x}(k)\}_{k\geq 0}$ with

$$\mathtt{x}(k) = A^k\mathtt{x}(0),\ k \geq 1.$$

- There's a unique solution with given initial condition $\mathtt{x}(0)$.
- The constant solutions $\mathtt{x}(k) = \mathtt{x}$ for all k are called steady states.
- If $\mathtt{x}$ is a steady state $\Rightarrow \mathtt{x} = A\mathtt{x}$, so $\mathtt{x}$ is either 0 or an eigenvector of A of eigenvalue 1.

Solutions

Lemma

The solutions to the system $\mathtt{x}(k+1) = A\,\mathtt{x}(k)$ are $\{\mathtt{x}(k)\}_{k\geq 0}$ with

$$\mathtt{x}(k) = A^k \mathtt{x}(0),\ k \geq 1.$$

- There's a unique solution with given initial condition $\mathtt{x}(0)$.
- The constant solutions $\mathtt{x}(k) = \mathtt{x}$ for all k are called steady states.
- If $\mathtt{x}$ is a steady state $\Rightarrow \mathtt{x} = A\mathtt{x}$, so $\mathtt{x}$ is either 0 or an eigenvector of A of eigenvalue 1.

Example

In the previous example, if $\mathrm{x}(0) = \begin{pmatrix} 40 \\ 40 \\ 20 \end{pmatrix}$, then:

$$\mathrm{x}(0)=\begin{pmatrix} 40 \\ 40 \\ 20 \end{pmatrix}, \mathrm{x}(1)=A\mathrm{x}(0)=\begin{pmatrix} 220 \\ 20 \\ 10 \end{pmatrix}, \mathrm{x}(2)=A^2\mathrm{x}(0)=\begin{pmatrix} 110 \\ 110 \\ 5 \end{pmatrix}, \mathrm{x}(3)=\begin{pmatrix} 455 \\ 55 \\ 27.5 \end{pmatrix}, \ldots$$

The eigenvalues of A are $1.5, -1.31, -0.19$ (0 is the only steady state in this case).
Will this population eventually survive?

- Study $\mathrm{x}(k)$ when k tends to infinity; this is called the **long-term behavior** (or asymptotic behavior) of the system.
- As $\mathrm{x}(k) = A^k\mathrm{x}(0)$, we need to compute powers of matrices and study its limit when k tends to infinite.

Example

In the previous example, if $\mathtt{x}(0) = \begin{pmatrix} 40 \\ 40 \\ 20 \end{pmatrix}$, then:

$$\mathtt{x}(0)=\begin{pmatrix} 40 \\ 40 \\ 20 \end{pmatrix}, \mathtt{x}(1)=A\mathtt{x}(0)=\begin{pmatrix} 220 \\ 20 \\ 10 \end{pmatrix}, \mathtt{x}(2)=A^2\mathtt{x}(0)=\begin{pmatrix} 110 \\ 110 \\ 5 \end{pmatrix}, \mathtt{x}(3)=\begin{pmatrix} 455 \\ 55 \\ 27.5 \end{pmatrix}, \ldots$$

The eigenvalues of A are $1.5, -1.31, -0.19$ (0 is the only steady state in this case).
Will this population eventually survive?

- Study $\mathtt{x}(k)$ when k tends to infinity; this is called the **long-term behavior** (or asymptotic behavior) of the system.
- As $\mathtt{x}(k) = A^k\mathtt{x}(0)$, we need to compute powers of matrices and study its limit when k tends to infinite.

Example (cont.)

In the previous example, if $\mathtt{x}(0) = \begin{pmatrix} 40 \\ 40 \\ 20 \end{pmatrix}$, then:

$$\mathtt{x}(10) = A^{10}\mathtt{x}(0) = \begin{pmatrix} 4571.91 \\ 2162.37 \\ 238.50 \end{pmatrix}, \mathtt{x}(20) = \begin{pmatrix} 301860.73 \\ 110036.15 \\ 16541.80 \end{pmatrix}, \mathtt{x}(30) = \begin{pmatrix} 17971431.25 \\ 6129573.17 \\ 995030.54 \end{pmatrix}, \dots$$

it seems to go to infinite. But, the proportion between populations seems to stabilize:

$$s := x_1(10)+x_2(10)+x_3(10) \Rightarrow \quad \frac{x_1(10)}{s} = 0.6558 \quad \frac{x_2(10)}{s} = 0.3100 \quad \frac{x_3(10)}{s} = 0.0342$$

$$s := x_1(20)+x_2(20)+x_3(20) \Rightarrow \quad \frac{x_1(20)}{s} = 0.7046 \quad \frac{x_2(20)}{s} = 0.2568 \quad \frac{x_3(20)}{s} = 0.0386$$

$$s := x_1(30)+x_2(30)+x_3(30) \Rightarrow \quad \frac{x_1(30)}{s} = 0.7161 \quad \frac{x_2(30)}{s} = 0.2442 \quad \frac{x_3(30)}{s} = 0.0397$$

Example (cont.)

Also the rate between $x(k)$ and $x(k+1)$ (the "growth rate") seems to have a tendency:

$$\frac{x_1(31)}{x_1(30)} = 1.623 \quad \frac{x_2(31)}{x_2(30)} = 1.372 \quad \frac{x_3(31)}{x_3(30)} = 1.663$$
$$\frac{x_1(41)}{x_1(40)} = 1.507 \quad \frac{x_2(41)}{x_2(40)} = 1.491 \quad \frac{x_3(41)}{x_3(40)} = 1.510$$

This is all related to eigenvectors and eigenvalues!

Outline

Diagonalizable matrices: long-term behavior.

Example: $\mathrm{x}(k+1) = A\mathrm{x}(k)$ where

$$A = \begin{pmatrix} 0.65 & -0.15 \\ -0.15 & 0.65 \end{pmatrix}.$$

Solutions: $\mathrm{x}(k) = A^k \mathrm{x}(0)$. To compute A^k: diagonalize A.

- The eigenvalues of A are 0.8 and 0.5 with respective eigenvectors and $v_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
- We write $\mathrm{x}(0)$ in the basis v_1, v_2: $\mathrm{x}(0) = c_1 v_1 + c_2 v_2$. For ex., if $\mathrm{x}(0) = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$, then $c_1 = 2$, $c_2 = 1$,$\mathrm{x}(0) = 2v_1 + 1v_2$.
- As $Av_i = \lambda_i v_i$, we have:

$$\mathrm{x}(k) = A^k \mathrm{x}(0) = A^k(c_1 v_1 + c_2 v_2) = c_1 A^k v_1 + c_2 A^k v_2 = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k$$

$$\mathrm{x}(k) = c_1 0.8^k \begin{pmatrix} -1 \\ 1 \end{pmatrix} + c_2 0.5^k \begin{pmatrix} 1 \\ 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Diagonalizable matrices: long-term behavior.

Example: $\mathtt{x}(k+1) = A\mathtt{x}(k)$ where

$$A = \begin{pmatrix} 0.65 & -0.15 \\ -0.15 & 0.65 \end{pmatrix}.$$

Solutions: $\mathtt{x}(k) = A^k \mathtt{x}(0)$. To compute A^k: diagonalize A.

- The eigenvalues of A are 0.8 and 0.5 with respective eigenvectors and $v_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
- We write $\mathtt{x}(0)$ in the basis v_1, v_2: $\mathtt{x}(0) = c_1 v_1 + c_2 v_2$. For ex., if $\mathtt{x}(0) = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$, then $c_1 = 2$, $c_2 = 1$,$\mathtt{x}(0) = 2v_1 + 1v_2$.
- As $Av_i = \lambda_i v_i$, we have:

$$\mathtt{x}(k) = A^k \mathtt{x}(0) = A^k(c_1 v_1 + c_2 v_2) = c_1 A^k v_1 + c_2 A^k v_2 = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k$$

$$\mathtt{x}(k) = c_1 0.8^k \begin{pmatrix} -1 \\ 1 \end{pmatrix} + c_2 0.5^k \begin{pmatrix} 1 \\ 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Diagonalizable matrices: long-term behavior.

Example: $\mathrm{x}(k+1) = A\mathrm{x}(k)$ where

$$A = \begin{pmatrix} 0.65 & -0.15 \\ -0.15 & 0.65 \end{pmatrix}.$$

Solutions: $\mathrm{x}(k) = A^k \mathrm{x}(0)$. To compute A^k: diagonalize A.

- The eigenvalues of A are 0.8 and 0.5 with respective eigenvectors and $v_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
- We write $\mathrm{x}(0)$ in the basis v_1, v_2: $\mathrm{x}(0) = c_1 v_1 + c_2 v_2$. For ex., if $\mathrm{x}(0) = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$, then $c_1 = 2$, $c_2 = 1$,$\mathrm{x}(0) = 2v_1 + 1v_2$.
- As $Av_i = \lambda_i v_i$, we have:

$$\mathrm{x}(k) = A^k \mathrm{x}(0) = A^k (c_1 v_1 + c_2 v_2) = c_1 A^k v_1 + c_2 A^k v_2 = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k$$

$$\mathrm{x}(k) = c_1 0.8^k \begin{pmatrix} -1 \\ 1 \end{pmatrix} + c_2 0.5^k \begin{pmatrix} 1 \\ 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Diagonalizable matrices: long-term behavior

If the system is $\mathtt{x}(k+1) = A\mathtt{x}(k)$ and A diagonalizes, to study the long-term behavior of the solutions $\mathtt{x}(k) = A^k \mathtt{x}(0)$ depending on $\mathtt{x}(0)$ we do:

- Compute and order the eigenvalues such that $|\lambda_1| \geq |\lambda_2| \geq \dots |\lambda_n|$.
- Compute the corresponding basis of eigenvectors $\mathbf{v} = \{v_1, \dots, v_n\}$.
- Compute the coordinates of $\mathtt{x}(0)$ in the basis $v_1, \dots, v_n$: $\mathtt{x}(0) = c_1 v_1 + \dots + c_n v_n$.
- Then, as $Av_i = \lambda_i v_i$, the solutions $\mathtt{x}(k)$ are:

$$\mathtt{x}(k) = A^k \mathtt{x}(0) = A^k(c_1 v_1 + \dots + c_n v_n) = c_1 A^k v_1 + \dots + c_n A^k v_n =$$

$$= c_1 \lambda_1^k v_1 + \dots + c_n \lambda_n^k v_n.$$

Diagonalizable matrices: long-term behavior

If the system is $\mathrm{x}(k+1) = A\mathrm{x}(k)$ and A diagonalizes, to study the long-term behavior of the solutions $\mathrm{x}(k) = A^k\mathrm{x}(0)$ depending on $\mathrm{x}(0)$ we do:

- Compute and order the eigenvalues such that $|\lambda_1| \geq |\lambda_2| \geq \dots |\lambda_n|$.
- Compute the corresponding basis of eigenvectors $\mathbf{v} = \{v_1, \dots, v_n\}$.
- Compute the coordinates of $\mathrm{x}(0)$ in the basis $v_1, \dots, v_n$: $\mathrm{x}(0) = c_1 v_1 + \cdots + c_n v_n$.
- Then, as $Av_i = \lambda_i v_i$, the solutions $\mathrm{x}(k)$ are:

$$\mathrm{x}(k) = A^k\mathrm{x}(0) = A^k(c_1 v_1 + \cdots + c_n v_n) = c_1 A^k v_1 + \cdots + c_n A^k v_n =$$

$$= c_1 \lambda_1^k v_1 + \cdots + c_n \lambda_n^k v_n.$$

Diagonalizable matrices: long-term behavior

If the system is $\mathrm{x}(k+1) = A\mathrm{x}(k)$ and A diagonalizes, to study the long-term behavior of the solutions $\mathrm{x}(k) = A^k\mathrm{x}(0)$ depending on $\mathrm{x}(0)$ we do:

- Compute and order the eigenvalues such that $|\lambda_1| \geq |\lambda_2| \geq \ldots |\lambda_n|$.
- Compute the corresponding basis of eigenvectors $\mathbf{v} = \{v_1, \ldots, v_n\}$.
- Compute the coordinates of $\mathrm{x}(0)$ in the basis $v_1, \ldots, v_n$: $\mathrm{x}(0) = c_1 v_1 + \cdots + c_n v_n$.
- Then, as $Av_i = \lambda_i v_i$, the solutions $\mathrm{x}(k)$ are:

$$\mathrm{x}(k) = A^k\mathrm{x}(0) = A^k(c_1 v_1 + \cdots + c_n v_n) = c_1 A^k v_1 + \cdots + c_n A^k v_n =$$

$$= c_1 \lambda_1^k v_1 + \cdots + c_n \lambda_n^k v_n.$$

Diagonalizable matrices: long-term behavior

If the system is $\mathtt{x}(k+1) = A\mathtt{x}(k)$ and A diagonalizes, to study the long-term behavior of the solutions $\mathtt{x}(k) = A^k\mathtt{x}(0)$ depending on $\mathtt{x}(0)$ we do:

- Compute and order the eigenvalues such that $|\lambda_1| \geq |\lambda_2| \geq \ldots |\lambda_n|$.
- Compute the corresponding basis of eigenvectors $\mathbf{v} = \{v_1, \ldots, v_n\}$.
- Compute the coordinates of $\mathtt{x}(0)$ in the basis $v_1, \ldots, v_n$: $\mathtt{x}(0) = c_1 v_1 + \cdots + c_n v_n$.
- Then, as $Av_i = \lambda_i v_i$, the solutions $\mathtt{x}(k)$ are:

$$\mathtt{x}(k) = A^k\mathtt{x}(0) = A^k(c_1 v_1 + \cdots + c_n v_n) = c_1 A^k v_1 + \cdots + c_n A^k v_n =$$

$$= c_1 \lambda_1^k v_1 + \cdots + c_n \lambda_n^k v_n.$$

As matrices

If $A = PDP^{-1}$, $D = diag(\lambda_1, \ldots, \lambda_n)$, $P = A_{\mathbf{v}\to e}$

$$P = \begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix},$$

then $A^k = PD^kP^{-1}$. Given $\mathrm{x}(0)$, if $P^{-1}\mathrm{x}(0) = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$

$$\Rightarrow \mathrm{x}(k) = PD^kP^{-1}\mathrm{x}(0) = P \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = P \begin{pmatrix} c_1\lambda_1^k \\ \vdots \\ c_n\lambda_n^k \end{pmatrix},$$

$$\mathrm{x}(k) = \begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix} \begin{pmatrix} c_1\lambda_1^k \\ \vdots \\ c_n\lambda_n^k \end{pmatrix} = c_1\lambda_1^k \begin{pmatrix} v_1 \end{pmatrix} + \cdots + c_n\lambda_n^k \begin{pmatrix} v_n \end{pmatrix}.$$

Long-term behavior

We have $|\lambda_1| \geq |\lambda_2| \geq \ldots |\lambda_n|$ and

$$\mathtt{x}(k) = c_1 \lambda_1^k v_1 + \cdots + c_n \lambda_n^k v_n,$$

note that:

- if $\mathtt{x}(0) = v_i$, then $\mathtt{x}(k) = \lambda_i^k v_i$.
- when $|\lambda_1| > |\lambda_2|$, we'll see that λ_1 and v_1 determine the long-term behaviour

Long-term behavior

We have $|\lambda_1| \geq |\lambda_2| \geq \ldots |\lambda_n|$ and

$$\mathtt{x}(k) = c_1 \lambda_1^k v_1 + \cdots + c_n \lambda_n^k v_n,$$

note that:

- if $\mathtt{x}(0) = v_i$, then $\mathtt{x}(k) = \lambda_i^k v_i$.
- when $|\lambda_1| > |\lambda_2|$, we'll see that λ_1 and v_1 determine the long-term behaviour

Definition

If there is an eigenvalue λ_1 that satisfies $|\lambda_1| > |\lambda_i|$, then λ_1 is real and is called the **dominant** eigenvalue and the corresponding eigenvalue is called the **dominant eigenvector**.

- Example: if the eigenvalues of a 4×4 matrix are $-4, -3, 1, 2$, then $\lambda_1 = -4$ is the dominant eigenvalue.
- Example: if the eigenvalues of a 4×4 matrix are $-4, -3, 1, 4$, then there is no dominant eigenvalue.
  Note:
- If A has complex eigenvalues, $|\lambda|$ refers to the modulus (or absolute value) of the complex number: $|a + bi| = \sqrt{a^2 + b^2}$
- As A is real, non-real eigenvalues appear in conjugate pairs $(a + bi, a - bi)$ and have the same modulus
- Example: if the eigenvalues of a 3×3 matrix are 6, $-1 + 2i$, $-1 - 2i$, then 6 is the dominant eigenvalue.
- Example: if the eigenvalues of a 2×2 matrix are $-1 + 2i$, $-1 - 2i$, then there is no dominant eigenvalue.

Definition

If there is an eigenvalue λ_1 that satisfies $|\lambda_1| > |\lambda_i|$, then λ_1 is real and is called the **dominant** eigenvalue and the corresponding eigenvalue is called the **dominant eigenvector**.

- Example: if the eigenvalues of a 4×4 matrix are $-4, -3, 1, 2$, then $\lambda_1 = -4$ is the dominant eigenvalue.
- Example: if the eigenvalues of a 4×4 matrix are $-4, -3, 1, 4$, then there is no dominant eigenvalue.
  Note:
- If A has complex eigenvalues, $|\lambda|$ refers to the modulus (or absolute value) of the complex number: $|a + bi| = \sqrt{a^2 + b^2}$
- As A is real, non-real eigenvalues appear in conjugate pairs $(a + bi, a - bi)$ and have the same modulus
- Example: if the eigenvalues of a 3×3 matrix are 6, $-1 + 2i$, $-1 - 2i$, then 6 is the dominant eigenvalue.
- Example: if the eigenvalues of a 2×2 matrix are $-1 + 2i$, $-1 - 2i$, then there is no dominant eigenvalue.

Definition

If there is an eigenvalue λ_1 that satisfies $|\lambda_1| > |\lambda_i|$, then λ_1 is real and is called the **dominant** eigenvalue and the corresponding eigenvalue is called the **dominant eigenvector**.

- Example: if the eigenvalues of a 4×4 matrix are $-4, -3, 1, 2$, then $\lambda_1 = -4$ is the dominant eigenvalue.
- Example: if the eigenvalues of a 4×4 matrix are $-4, -3, 1, 4$, then there is no dominant eigenvalue.
  Note:
- If A has complex eigenvalues, $|\lambda|$ refers to the modulus (or absolute value) of the complex number: $|a + bi| = \sqrt{a^2 + b^2}$
- As A is real, non-real eigenvalues appear in conjugate pairs $(a + bi, a - bi)$ and have the same modulus
- Example: if the eigenvalues of a 3×3 matrix are $6, -1 + 2i$, $-1 - 2i$, then 6 is the dominant eigenvalue.
- Example: if the eigenvalues of a 2×2 matrix are $-1 + 2i$, $-1 - 2i$, then there is no dominant eigenvalue.

Definition

If there is an eigenvalue λ_1 that satisfies $|\lambda_1| > |\lambda_i|$, then λ_1 is real and is called the **dominant** eigenvalue and the corresponding eigenvalue is called the **dominant eigenvector**.

- Example: if the eigenvalues of a 4×4 matrix are $-4, -3, 1, 2$, then $\lambda_1 = -4$ is the dominant eigenvalue.
- Example: if the eigenvalues of a 4×4 matrix are $-4, -3, 1, 4$, then there is no dominant eigenvalue.
  Note:
- If A has complex eigenvalues, $|\lambda|$ refers to the modulus (or absolute value) of the complex number: $|a + bi| = \sqrt{a^2 + b^2}$
- As A is real, non-real eigenvalues appear in conjugate pairs $(a + bi, a - bi)$ and have the same modulus
- Example: if the eigenvalues of a 3×3 matrix are 6, $-1 + 2i$, $-1 - 2i$, then 6 is the dominant eigenvalue.
- Example: if the eigenvalues of a 2×2 matrix are $-1 + 2i$, $-1 - 2i$, then there is no dominant eigenvalue.

Definition

If there is an eigenvalue λ_1 that satisfies $|\lambda_1| > |\lambda_i|$, then λ_1 is real and is called the **dominant** eigenvalue and the corresponding eigenvalue is called the **dominant eigenvector**.

- Example: if the eigenvalues of a 4×4 matrix are $-4, -3, 1, 2$, then $\lambda_1 = -4$ is the dominant eigenvalue.
- Example: if the eigenvalues of a 4×4 matrix are $-4, -3, 1, 4$, then there is no dominant eigenvalue.

  Note:
- If A has complex eigenvalues, $|\lambda|$ refers to the modulus (or absolute value) of the complex number: $|a + bi| = \sqrt{a^2 + b^2}$
- As A is real, non-real eigenvalues appear in conjugate pairs $(a + bi, a - bi)$ and have the same modulus
- Example: if the eigenvalues of a 3×3 matrix are 6, $-1 + 2i$, $-1 - 2i$, then 6 is the dominant eigenvalue.
- Example: if the eigenvalues of a 2×2 matrix are $-1 + 2i$, $-1 - 2i$, then there is no dominant eigenvalue.

Definition

If there is an eigenvalue λ_1 that satisfies $|\lambda_1| > |\lambda_i|$, then λ_1 is real and is called the **dominant** eigenvalue and the corresponding eigenvalue is called the **dominant eigenvector**.

- Example: if the eigenvalues of a 4×4 matrix are $-4, -3, 1, 2$, then $\lambda_1 = -4$ is the dominant eigenvalue.
- Example: if the eigenvalues of a 4×4 matrix are $-4, -3, 1, 4$, then there is no dominant eigenvalue.
  Note:
- If A has complex eigenvalues, $|\lambda|$ refers to the modulus (or absolute value) of the complex number: $|a + bi| = \sqrt{a^2 + b^2}$
- As A is real, non-real eigenvalues appear in conjugate pairs $(a + bi, a - bi)$ and have the same modulus
- Example: if the eigenvalues of a 3×3 matrix are 6, $-1 + 2i$, $-1 - 2i$, then 6 is the dominant eigenvalue.
- Example: if the eigenvalues of a 2×2 matrix are $-1 + 2i$, $-1 - 2i$, then there is no dominant eigenvalue.

Definition

If there is an eigenvalue λ_1 that satisfies $|\lambda_1| > |\lambda_i|$, then λ_1 is real and is called the **dominant** eigenvalue and the corresponding eigenvalue is called the **dominant eigenvector**.

- Example: if the eigenvalues of a 4×4 matrix are $-4, -3, 1, 2$, then $\lambda_1 = -4$ is the dominant eigenvalue.
- Example: if the eigenvalues of a 4×4 matrix are $-4, -3, 1, 4$, then there is no dominant eigenvalue.
  Note:
- If A has complex eigenvalues, $|\lambda|$ refers to the modulus (or absolute value) of the complex number: $|a + bi| = \sqrt{a^2 + b^2}$
- As A is real, non-real eigenvalues appear in conjugate pairs $(a + bi, a - bi)$ and have the same modulus
- Example: if the eigenvalues of a 3×3 matrix are 6, $-1 + 2i$, $-1 - 2i$, then 6 is the dominant eigenvalue.
- Example: if the eigenvalues of a 2×2 matrix are $-1 + 2i$, $-1 - 2i$, then there is no dominant eigenvalue.

Long-term behavior

If we have $|\lambda_1| > |\lambda_2| \geq \ldots |\lambda_n|$ (λ_1 dominant) and

$$\mathrm{x}(k) = c_1 \lambda_1^k v_1 + \cdots + c_n \lambda_n^k v_n,$$

then when k is large:

- λ_1^k grows faster than λ_i^k so, if $c_1 \neq 0$,

$$\mathrm{x}(k) \sim c_1 \lambda_1^k v_1 \quad \text{for } k \text{ big, and}$$

  - If $|\lambda_1| < 1$, $\mathrm{x}(k) \to 0$ when $k \to \infty$.
  - If $\lambda_1 = 1$ $c_1 \neq 0$, then $\mathrm{x}(k) \to c_1 v_1$ when $k \to \infty$.
  - If $|\lambda_1| > 1$ and $c_1 \neq 0$, then $\mathrm{x}(k)$ tends to a vector with infinite components in the direction of v_1.

- the **growth rate** is given by λ_1: $\frac{\mathrm{x}_j(k+1)}{\mathrm{x}_j(k)} \sim \lambda_1$, so $\mathrm{x}(k+1) \sim \lambda_1 \mathrm{x}(k)$

Special things happen with initial conditions that have $c_1 = 0$.

Long-term behavior

If we have $|\lambda_1| > |\lambda_2| \geq \dots |\lambda_n|$ (λ_1 dominant) and

$$\mathrm{x}(k) = c_1 \lambda_1^k v_1 + \cdots + c_n \lambda_n^k v_n,$$

then when k is large:

- λ_1^k grows faster than λ_i^k so, if $c_1 \neq 0$,

  $$\mathrm{x}(k) \sim c_1 \lambda_1^k v_1 \quad \text{for } k \text{ big, and}$$

  - If $|\lambda_1| < 1$, $\mathrm{x}(k) \to 0$ when $k \to \infty$.
  - If $\lambda_1 = 1$ $c_1 \neq 0$, then $\mathrm{x}(k) \to c_1 v_1$ when $k \to \infty$.
  - If $|\lambda_1| > 1$ and $c_1 \neq 0$, then $\mathrm{x}(k)$ tends to a vector with infinite components in the direction of v_1.
- the **growth rate** is given by λ_1: $\frac{\mathrm{x}_j(k+1)}{\mathrm{x}_j(k)} \sim \lambda_1$, so

  $\mathrm{x}(k+1) \sim \lambda_1 \mathrm{x}(k)$

Special things happen with initial conditions that have $c_1 = 0$.

Long-term behavior

If we have $|\lambda_1| > |\lambda_2| \geq \dots |\lambda_n|$ (λ_1 dominant) and

$$\mathrm{x}(k) = c_1 \lambda_1^k v_1 + \cdots + c_n \lambda_n^k v_n,$$

then when k is large:

- λ_1^k grows faster than λ_i^k so, if $c_1 \neq 0$,

  $$\mathrm{x}(k) \sim c_1 \lambda_1^k v_1 \quad \text{for } k \text{ big, and}$$

  - If $|\lambda_1| < 1$, $\mathrm{x}(k) \to 0$ when $k \to \infty$.
  - If $\lambda_1 = 1$ $c_1 \neq 0$, then $\mathrm{x}(k) \to c_1 v_1$ when $k \to \infty$.
  - If $|\lambda_1| > 1$ and $c_1 \neq 0$, then $\mathrm{x}(k)$ tends to a vector with infinite components in the direction of v_1.
- the **growth rate** is given by λ_1: $\frac{\mathrm{x}_j(k+1)}{\mathrm{x}_j(k)} \sim \lambda_1$, so $\mathrm{x}(k+1) \sim \lambda_1 \mathrm{x}(k)$

Special things happen with initial conditions that have $c_1 = 0$.

Long-term behavior

If we have $|\lambda_1| > |\lambda_2| \geq \ldots |\lambda_n|$ (λ_1 dominant) and

$$\mathrm{x}(k) = c_1\lambda_1^k v_1 + \cdots + c_n\lambda_n^k v_n,$$

then when k is large:

- λ_1^k grows faster than λ_i^k so, if $c_1 \neq 0$,

$$\mathrm{x}(k) \sim c_1\lambda_1^k v_1 \quad \text{for } k \text{ big, and}$$

  - If $|\lambda_1| < 1$, $\mathrm{x}(k) \to 0$ when $k \to \infty$.
  - If $\lambda_1 = 1$ $c_1 \neq 0$, then $\mathrm{x}(k) \to c_1 v_1$ when $k \to \infty$.
  - If $|\lambda_1| > 1$ and $c_1 \neq 0$, then $\mathrm{x}(k)$ tends to a vector with infinite components in the direction of v_1.
- the **growth rate** is given by λ_1: $\frac{\mathrm{x}_j(k+1)}{\mathrm{x}_j(k)} \sim \lambda_1$, so $\mathrm{x}(k+1) \sim \lambda_1 \mathrm{x}(k)$

Special things happen with initial conditions that have $c_1 = 0$.

Long-term behavior

If we have $|\lambda_1| > |\lambda_2| \geq \dots |\lambda_n|$ (λ_1 dominant) and

$$\mathrm{x}(k) = c_1 \lambda_1^k v_1 + \cdots + c_n \lambda_n^k v_n,$$

then when k is large:

- λ_1^k grows faster than λ_i^k so, if $c_1 \neq 0$,

$$\mathrm{x}(k) \sim c_1 \lambda_1^k v_1 \quad \text{for } k \text{ big, and}$$

  - If $|\lambda_1| < 1$, $\mathrm{x}(k) \to 0$ when $k \to \infty$.
  - If $\lambda_1 = 1$ $c_1 \neq 0$, then $\mathrm{x}(k) \to c_1 v_1$ when $k \to \infty$.
  - If $|\lambda_1| > 1$ and $c_1 \neq 0$, then $\mathrm{x}(k)$ tends to a vector with infinite components in the direction of v_1.
- the **growth rate** is given by λ_1: $\frac{\mathrm{x}_j(k+1)}{\mathrm{x}_j(k)} \sim \lambda_1$, so $\mathrm{x}(k+1) \sim \lambda_1 \mathrm{x}(k)$

Special things happen with initial conditions that have $c_1 = 0$.

- Which matrices have a dominant eigenvalue?
- Which matrices have a steady state (x such that $A\mathrm{x} = \mathrm{x}$) different from 0? This is, which matrices have eigenvalue 1?

- Which matrices have a dominant eigenvalue?
- Which matrices have a steady state ($\mathtt{x}$ such that $A\mathtt{x} = \mathtt{x}$) different from 0? This is, which matrices have eigenvalue 1?

Outline

Stochastic matrices

Definition

A (column) **stochastic matrix** is a non-negative $n \times n$ matrix whose columns sum to 1.

A similar definition can be made for rows.
As columns sum to 1, if A is a stochastic matrix we have:

$$(11\ldots1)A = (11\ldots1)$$

$$A^t \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Thus, 1 is an eigenvalue of A^t and $(11\ldots1)^t$ is a positive eigenvector for A^t.

Properties of stochastic matrices

- 1 is an eigenvalue of A
- If x sums to 1, then Ax still sums to 1.

Non-negative vectors that sum to 1 are called **probability vectors or distributions**.

Properties of stochastic matrices

- 1 is an eigenvalue of A
- If $\mathtt{x}$ sums to 1, then $A\mathtt{x}$ still sums to 1.

Non-negative vectors that sum to 1 are called **probability vectors or distributions**.

Stochastic matrices

Theorem (Perron-Frobenius)

If A is a stochastic matrix, then 1 is an eigenvalue and $|\lambda| \leq 1$ for any other eigenvalue λ. Moreover, if A is positive,

- *1 is the dominant eigenvalue*
- *1 has a positive eigenvector v (a steady state)*
- *no other eigenvalue has positive eigenvectors.*
- *If we take v to sum to 1, then v is called the* **stationary distribution** *and*

$$\lim A^k = (v\, v \ldots v)$$

$$and \qquad \lim A^k \mathrm{x} = v$$

for any probability vector x,

- *and the distribution of $\mathrm{x}(k)$ tends to v: $\frac{\mathrm{x}(k)}{\sum_i x_i(k)} \sim v$.*

Stochastic matrices

Theorem (Perron-Frobenius)

If A is a stochastic matrix, then 1 is an eigenvalue and $|\lambda| \leq 1$ for any other eigenvalue λ. Moreover, if A is positive,

- *1 is the dominant eigenvalue*
- *1 has a positive eigenvector v (a steady state)*
- *no other eigenvalue has positive eigenvectors.*
- *If we take v to sum to 1, then v is called the* **stationary distribution** *and*

$$\lim A^k = (v\, v \ldots v)$$

$$\text{and} \qquad \lim A^k \mathrm{x} = v$$

for any probability vector x,

- *and the distribution of $\mathrm{x}(k)$ tends to v: $\frac{\mathrm{x}(k)}{\sum_i x_i(k)} \sim v$.*

Stochastic matrices

Theorem (Perron-Frobenius)

If A is a stochastic matrix, then 1 is an eigenvalue and $|\lambda| \leq 1$ for any other eigenvalue λ. Moreover, if A is positive,

- *1 is the dominant eigenvalue*
- *1 has a positive eigenvector v (a steady state)*
- *no other eigenvalue has positive eigenvectors.*
- *If we take v to sum to 1, then v is called the* **stationary distribution** *and*

$$\lim A^k = (v\, v \ldots v)$$

$$and \qquad \lim A^k \mathrm{x} = v$$

for any probability vector x,

- *and the distribution of $\mathrm{x}(k)$ tends to v: $\frac{\mathrm{x}(k)}{\sum_i x_i(k)} \sim v$.*

Stochastic matrices

Theorem (Perron-Frobenius)

If A is a stochastic matrix, then 1 is an eigenvalue and $|\lambda| \leq 1$ for any other eigenvalue λ. Moreover, if A is positive,

- *1 is the dominant eigenvalue*
- *1 has a positive eigenvector v (a steady state)*
- *no other eigenvalue has positive eigenvectors.*
- *If we take v to sum to 1, then v is called the* **stationary distribution** *and*

$$\lim A^k = (v\, v \ldots v)$$

$$\textit{and} \qquad \lim A^k \mathrm{x} = v$$

  for any probability vector x,
- *and the distribution of* $\mathrm{x}(k)$ *tends to v:* $\frac{\mathrm{x}(k)}{\sum_i x_i(k)} \sim v$.

Stochastic matrices

Theorem (Perron-Frobenius)

If A is a stochastic matrix, then 1 is an eigenvalue and $|\lambda| \leq 1$ for any other eigenvalue λ. Moreover, if A is positive,

- *1 is the dominant eigenvalue*
- *1 has a positive eigenvector v (a steady state)*
- *no other eigenvalue has positive eigenvectors.*
- *If we take v to sum to 1, then v is called the* **stationary distribution** *and*

$$\lim A^k = (v\, v \ldots v)$$

$$and \qquad \lim A^k \mathtt{x} = v$$

for any probability vector $\mathtt{x}$,

- *and the distribution of $\mathtt{x}(k)$ tends to v: $\frac{\mathtt{x}(k)}{\sum_i x_i(k)} \sim v$.*

Stochastic matrices

Theorem (Perron-Frobenius)

If A is a stochastic matrix, then 1 is an eigenvalue and $|\lambda| \leq 1$ for any other eigenvalue λ. Moreover, if A is positive,

- *1 is the dominant eigenvalue*
- *1 has a positive eigenvector v (a steady state)*
- *no other eigenvalue has positive eigenvectors.*
- *If we take v to sum to 1, then v is called the* **stationary distribution** *and*

$$\lim A^k = (v\, v \ldots v)$$

$$and \qquad \lim A^k \mathtt{x} = v$$

  for any probability vector $\mathtt{x}$,
- *and the distribution of $\mathtt{x}(k)$ tends to v: $\frac{\mathtt{x}(k)}{\sum_i x_i(k)} \sim v$.*