Linear Algebra

Lecture 8 - Linear Endomorphisms

Oskar Kędzierski

2 December 2024

Endomorphism

Definition

Let V be a vector space and $\mathcal{A}$ its (ordered) basis. A linear transformation $\varphi : V \longrightarrow V$ is called a **linear endomorphism**. The matrix $M(\varphi)_{\mathcal{A}}^{\mathcal{A}}$ is called matrix of endomorphism relative to basis $\mathcal{A}$. It is denoted in short $M(\varphi)_{\mathcal{A}}$.

Example

The identity $id : V \longrightarrow V$ is a linear endomorphism and its matrix relative to any basis $\mathcal{A}$ is the identity matrix

$$M(\mathrm{id})_{\mathcal{A}} = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \in M(n \times n; \mathbb{R}),$$

where $n = \dim V$.

Example

Let

$$s : \mathbb{R}^2 \longrightarrow \mathbb{R}^2,$$

$$r : \mathbb{R}^2 \longrightarrow \mathbb{R}^2,$$

$$k : \mathbb{R}^2 \longrightarrow \mathbb{R}^2,$$

$$p : \mathbb{R}^2 \longrightarrow \mathbb{R}^2.$$

be linear endomorphisms of $\mathbb{R}^2$ defined as follows: s is a reflection of $\mathbb{R}^2$ about the x_1-axis, r rotation about the origin of $\mathbb{R}^2$ (i.e. $(0,0)$) by $\frac{\pi}{2}$ radians (i.e. 90 degrees) counter-clockwise, k is scaling by -2 in all directions (also called uniform scaling) and p is projection onto the x_2-axis.

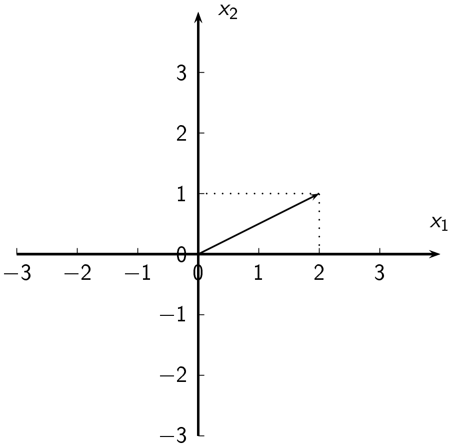
Example (continued)

For example, if $v = (2, 1)$ then
$s(v) = (2, -1)$, $r(v) = (-1, 2)$, $k(v) = (-4, -2)$, $p(v) = (0, 1)$.

Example (continued)

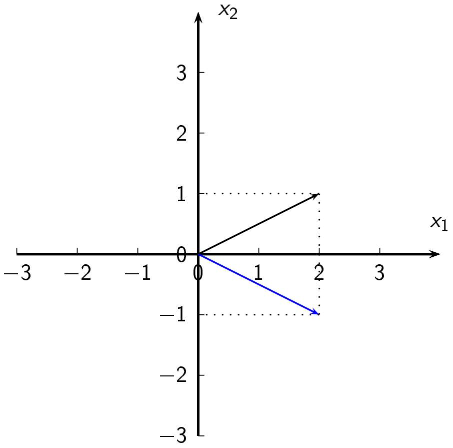
For example, if $v = (2, 1)$ then

$s(v) = (2, -1)$, $r(v) = (-1, 2)$, $k(v) = (-4, -2)$, $p(v) = (0, 1)$.



Example (continued)

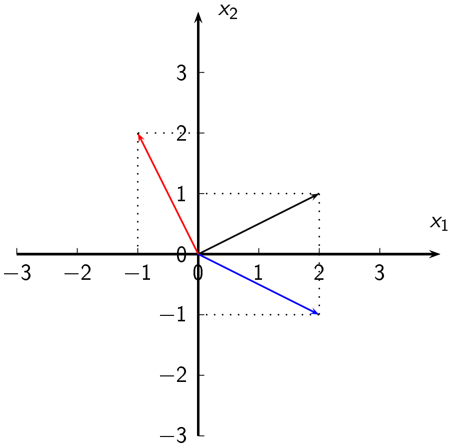
For example, if $v = (2, 1)$ then
$s(v) = (2, -1)$, $r(v) = (-1, 2)$, $k(v) = (-4, -2)$, $p(v) = (0, 1)$.



$s(x_1, x_2) = (x_1, -x_2)$

Example (continued)

For example, if $v = (2, 1)$ then
$s(v) = (2, -1)$, $r(v) = (-1, 2)$, $k(v) = (-4, -2)$, $p(v) = (0, 1)$.

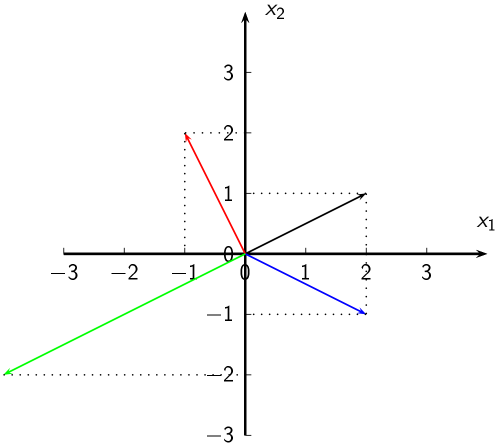


$s(x_1, x_2) = (x_1, -x_2)$

$r(x_1, x_2) = (-x_2, x_1)$

Example (continued)

For example, if $v = (2, 1)$ then
$s(v) = (2, -1)$, $r(v) = (-1, 2)$, $k(v) = (-4, -2)$, $p(v) = (0, 1)$.



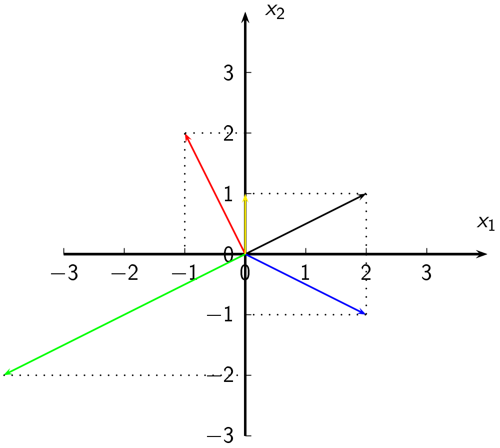
$$s(x_1, x_2) = (x_1, -x_2)$$

$$r(x_1, x_2) = (-x_2, x_1)$$

$$k(x_1, x_2) = (-2x_1, -2x_2)$$

Example (continued)

For example, if $v = (2,1)$ then
$s(v) = (2,-1)$, $r(v) = (-1,2)$, $k(v) = (-4,-2)$, $p(v) = (0,1)$.



$$s(x_1, x_2) = (x_1, -x_2)$$

$$r(x_1, x_2) = (-x_2, x_1)$$

$$k(x_1, x_2) = (-2x_1, -2x_2)$$

$$p(x_1, x_2) = (0, x_2)$$

Example (continued)

$$s(x_1, x_2) = (x_1, -x_2), \;\; r(x_1, x_2) = (-x_2, x_1),$$

$$k(x_1, x_2) = (-2x_1, -2x_2), \;\; p(x_1, x_2) = (0, x_2).$$

Example (continued)

$$s(x_1, x_2) = (x_1, -x_2), \;\; r(x_1, x_2) = (-x_2, x_1),$$
$$k(x_1, x_2) = (-2x_1, -2x_2), \;\; p(x_1, x_2) = (0, x_2).$$

The matrices of these endomorphisms relative to the standard basis $st = ((1, 0), (0, 1))$ look as follows:

$$M(s)_{st} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \; M(r)_{st} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

$$M(k)_{st} = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}, \; M(p)_{st} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Example (continued)

$$s(x_1, x_2) = (x_1, -x_2), \; r(x_1, x_2) = (-x_2, x_1),$$
$$k(x_1, x_2) = (-2x_1, -2x_2), \; p(x_1, x_2) = (0, x_2).$$

The matrices of these endomorphisms relative to the standard basis $st = ((1,0),(0,1))$ look as follows:

$$M(s)_{st} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \; M(r)_{st} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

$$M(k)_{st} = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}, \; M(p)_{st} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Take different basis, for example $\mathcal{A} = ((1,2),(1,1))$. The change-of-coordinate matrix is

$$M(\mathrm{id})^{\mathcal{A}}_{st} = (M(\mathrm{id})^{st}_{\mathcal{A}})^{-1} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}.$$

Example (continued)

Recall, $\mathcal{A} = ((1,2),(1,1))$ and $M(\mathrm{id})^{\mathcal{A}}_{st} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}$.

Example (continued)

Recall, $\mathcal{A} = ((1,2),(1,1))$ and $M(\mathrm{id})^{\mathcal{A}}_{st} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}$.

$$\begin{aligned}
s(1,2) &= (1,-2) &&= -3(1,2) + 4(1,1), \\
s(1,1) &= (1,-1) &&= -2(1,2) + 3(1,1), \\
r(1,2) &= (-2,1) &&= 3(1,2) - 5(1,1), \\
r(1,1) &= (-1,1) &&= 2(1,2) - 3(1,1), \\
k(1,2) &= (-2,-4) &&= -2(1,2) + 0(1,1), \\
k(1,1) &= (-2,-2) &&= 0(1,2) - 2(1,1), \\
p(1,2) &= (0,2) &&= 2(1,2) - 2(1,1), \\
p(1,1) &= (0,1) &&= 1(1,2) - 1(1,1).
\end{aligned}$$

Example (continued)

Recall, $\mathcal{A} = ((1,2),(1,1))$ and $M(\mathrm{id})_{st}^{\mathcal{A}} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}$.

$$\begin{aligned}
s(1,2) &= (1,-2) &&= -3(1,2) + 4(1,1),\\
s(1,1) &= (1,-1) &&= -2(1,2) + 3(1,1),\\
r(1,2) &= (-2,1) &&= 3(1,2) - 5(1,1),\\
r(1,1) &= (-1,1) &&= 2(1,2) - 3(1,1),\\
k(1,2) &= (-2,-4) &&= -2(1,2) + 0(1,1),\\
k(1,1) &= (-2,-2) &&= 0(1,2) - 2(1,1),\\
p(1,2) &= (0,2) &&= 2(1,2) - 2(1,1),\\
p(1,1) &= (0,1) &&= 1(1,2) - 1(1,1).
\end{aligned}$$

$$M(s)_{\mathcal{A}} = \begin{bmatrix} -3 & -2 \\ 4 & 3 \end{bmatrix}, \quad M(r)_{\mathcal{A}} = \begin{bmatrix} 3 & 2 \\ -5 & -3 \end{bmatrix},$$

$$M(k)_{\mathcal{A}} = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}, \quad M(p)_{\mathcal{A}} = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix}.$$

Example (continued)

We see that matrices of simple linear transformations look 'nice' relative to some bases and 'not-that-nice' relative to the others.

Example (continued)

We see that matrices of simple linear transformations look ‘nice’ relative to some bases and ‘not-that-nice’ relative to the others. That aim of this lecture is to find a way of computing those ‘nice’ ones in the general case. Note that determinants and the ranks of corresponding matrices did not change.

Matrix Similarity

Definition

Two matrices $A, B \in M(n \times n; \mathbb{R})$ are called **similar** if there exists an invertible matrix $C \in M(n \times n; \mathbb{R})$ such that

$$A = C^{-1}BC.$$

Matrix Similarity

Definition

Two matrices $A, B \in M(n \times n; \mathbb{R})$ are called **similar** if there exists an invertible matrix $C \in M(n \times n; \mathbb{R})$ such that

$$A = C^{-1}BC.$$

Proposition

Let $\varphi : V \longrightarrow V$ be a linear endomorphism of a finite dimensional vector space V. For any two bases $\mathcal{A}, \mathcal{B}$ of V the matrices $M(\varphi)_{\mathcal{A}}$ and $M(\varphi)_{\mathcal{B}}$ are similar.

Matrix Similarity

Definition

Two matrices $A, B \in M(n \times n; \mathbb{R})$ are called **similar** if there exists an invertible matrix $C \in M(n \times n; \mathbb{R})$ such that

$$A = C^{-1}BC.$$

Proposition

Let $\varphi : V \longrightarrow V$ be a linear endomorphism of a finite dimensional vector space V. For any two bases $\mathcal{A}, \mathcal{B}$ of V the matrices $M(\varphi)_{\mathcal{A}}$ and $M(\varphi)_{\mathcal{B}}$ are similar.

Proof.

$$M(\varphi)^{\mathcal{B}}_{\mathcal{B}} = M(\mathrm{id} \circ \varphi \circ \mathrm{id})^{\mathcal{B}}_{\mathcal{B}} = M(\mathrm{id})^{\mathcal{B}}_{\mathcal{A}} M(\varphi)^{\mathcal{A}}_{\mathcal{A}} M(\mathrm{id})^{\mathcal{A}}_{\mathcal{B}}.$$

Therefore

$$M(\varphi)_{\mathcal{B}} = C^{-1} M(\varphi)_{\mathcal{A}} C,$$

where $C = M(\mathrm{id})^{\mathcal{A}}_{\mathcal{B}}$. $\square$

Example

Let $\varphi((x_1, x_2)) = (x_1 + x_2, 2x_1 + 3x_2)$ be a linear endomorphism $\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$. Take $\mathcal{A} = st$ and $\mathcal{B} = ((-2, 1), (1, -1)$. Then

$$M(\varphi)_{\mathcal{A}} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} \text{ and } C = M(\mathrm{id})_{\mathcal{B}}^{\mathcal{A}} = \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}.$$

Example

Let $\varphi((x_1, x_2)) = (x_1 + x_2, 2x_1 + 3x_2)$ be a linear endomorphism $\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$. Take $\mathcal{A} = st$ and $\mathcal{B} = ((-2, 1), (1, -1))$. Then

$$M(\varphi)_{\mathcal{A}} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} \text{ and } C = M(\mathrm{id})^{\mathcal{A}}_{\mathcal{B}} = \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}.$$

Use $M(\varphi)_{\mathcal{B}} = C^{-1} M(\varphi)_{\mathcal{A}} C$ and compute $C^{-1} = \begin{bmatrix} -1 & -1 \\ -1 & -2 \end{bmatrix}$.

Example

Let $\varphi((x_1, x_2)) = (x_1 + x_2, 2x_1 + 3x_2)$ be a linear endomorphism $\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$. Take $\mathcal{A} = st$ and $\mathcal{B} = ((-2, 1), (1, -1)$. Then

$$M(\varphi)_{\mathcal{A}} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} \text{ and } C = M(\mathrm{id})^{\mathcal{A}}_{\mathcal{B}} = \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}.$$

Use $M(\varphi)_{\mathcal{B}} = C^{-1} M(\varphi)_{\mathcal{A}} C$ and compute $C^{-1} = \begin{bmatrix} -1 & -1 \\ -1 & -2 \end{bmatrix}$.

Then

$$M(\varphi)_{\mathcal{B}} = \begin{bmatrix} -1 & -1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}.$$

Example

Let $\varphi((x_1, x_2)) = (x_1 + x_2, 2x_1 + 3x_2)$ be a linear endomorphism $\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$. Take $\mathcal{A} = st$ and $\mathcal{B} = ((-2, 1), (1, -1)$. Then

$$M(\varphi)_{\mathcal{A}} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} \text{ and } C = M(\mathrm{id})^{\mathcal{A}}_{\mathcal{B}} = \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}.$$

Use $M(\varphi)_{\mathcal{B}} = C^{-1} M(\varphi)_{\mathcal{A}} C$ and compute $C^{-1} = \begin{bmatrix} -1 & -1 \\ -1 & -2 \end{bmatrix}$.

Then

$$M(\varphi)_{\mathcal{B}} = \begin{bmatrix} -1 & -1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}.$$

On the other hand,

$$\varphi((-2, 1)) = (-1, -1) = 2(-2, 1) + 3(1, -1),$$

$$\varphi((1, -1)) = (0, -1) = (-2, 1) + 2(1, -1).$$

Similar Matrices and Endomorphisms

Theorem

Let V be n-dimensional vector space and let $A, B \in M(n \times n; \mathbb{R})$. Then

$$A, B \text{ are similar} \Longleftrightarrow \text{there exists an endomorphism } \varphi : V \longrightarrow V$$
$$\text{and bases } \mathcal{A}, \mathcal{B} \text{ of } V \text{ such that } M(\varphi)_{\mathcal{A}} = A \text{ and } M(\varphi)_{\mathcal{B}} = B.$$

Similar Matrices and Endomorphisms

Theorem

Let V be n-dimensional vector space and let $A, B \in M(n \times n; \mathbb{R})$. Then

$$A, B \text{ are similar} \iff \text{there exists an endomorphism } \varphi : V \longrightarrow V \text{ and bases } \mathcal{A}, \mathcal{B} \text{ of } V \text{ such that } M(\varphi)_{\mathcal{A}} = A \text{ and } M(\varphi)_{\mathcal{B}} = B.$$

Proof.

$(\Leftarrow)$ was done before.

Similar Matrices and Endomorphisms

Theorem

Let V be n-dimensional vector space and let $A, B \in M(n \times n; \mathbb{R})$. Then

$$A, B \text{ are similar} \iff \text{there exists an endomorphism } \varphi : V \longrightarrow V$$
$$\text{and bases } \mathcal{A}, \mathcal{B} \text{ of } V \text{ such that } M(\varphi)_{\mathcal{A}} = A \text{ and } M(\varphi)_{\mathcal{B}} = B.$$

Proof.

($\Leftarrow$) was done before.
($\Rightarrow$) there exits an invertible matrix $C \in M(n \times n; \mathbb{R})$ such that $B = C^{-1}AC$. Let $\mathcal{A}$ be any basis of the vector space V and let φ be the unique linear endomorphism given by the condition $M(\varphi)^{\mathcal{A}}_{\mathcal{A}} = A$. If $\mathcal{B}$ is given by the condition $C = M(\mathrm{id})^{\mathcal{A}}_{\mathcal{B}}$ then $B = M(\varphi)_{\mathcal{B}}$. $\square$

Eigenvalues and Eigenvectors

Definition

Let $\varphi : V \longrightarrow V$ be a linear endomorphism of a finite dimensional vector space V. A constant $\lambda \in \mathbb{R}$ is called **eigenvalue** of φ if there exists a non-zero vector $v \in V$ such that

$$\varphi(v) = \lambda v.$$

Such vector v is called an **eigenvector** of φ associated to the eigenvalue λ.

Eigenvalues and Eigenvectors

Definition

Let $\varphi : V \longrightarrow V$ be a linear endomorphism of a finite dimensional vector space V. A constant $\lambda \in \mathbb{R}$ is called **eigenvalue** of φ if there exists a non-zero vector $v \in V$ such that

$$\varphi(v) = \lambda v.$$

Such vector v is called an **eigenvector** of φ associated to the eigenvalue λ.

Remark (geometric interpretation)

A vector $v \in V$ is an eigenvector of φ if and only if $\varphi(\operatorname{lin}(v)) \subset \operatorname{lin}(v)$ and $\operatorname{lin}(v) \neq \{0\}$, i.e. v is a non-zero vector and the line spanned by v is mapped into itself.

Eigenvalues and Eigenvectors (continued)

Let $\varphi : V \longrightarrow V$ be a linear endomorphism. For any eigenvalue λ of φ let $V_{(\lambda)}$ denote the set of all eigenvectors associated to λ together with the zero vector, i.e.

$$V_{(\lambda)} = \{v \in V \mid \varphi(v) = \lambda v.\}$$

Proposition

The subset $V_{(\lambda)} \subset V$ is a subspace of V.

Eigenvalues and Eigenvectors (continued)

Let $\varphi : V \longrightarrow V$ be a linear endomorphism. For any eigenvalue λ of φ let $V_{(\lambda)}$ denote the set of all eigenvectors associated to λ together with the zero vector, i.e.

$$V_{(\lambda)} = \{v \in V \mid \varphi(v) = \lambda v.\}$$

Proposition

The subset $V_{(\lambda)} \subset V$ is a subspace of V.

Proof.

Let $v, w \in V_{(\lambda)}$. Then $\varphi(v + w) = \varphi(v) + \varphi(w) = \lambda v + \lambda w = \lambda(v + w)$. Hence $v + w \in V_{(\lambda)}$. For any $\alpha \in \mathbb{R}$ we have $\varphi(\alpha v) = \alpha\varphi(v) = \lambda(\alpha v)$. Hence $\alpha v \in V_{(\lambda)}$. $\square$

Eigenvalues and Eigenvectors (continued)

Let $\varphi : V \longrightarrow V$ be a linear endomorphism. For any eigenvalue λ of φ let $V_{(\lambda)}$ denote the set of all eigenvectors associated to λ together with the zero vector, i.e.

$$V_{(\lambda)} = \{v \in V \mid \varphi(v) = \lambda v.\}$$

Proposition

The subset $V_{(\lambda)} \subset V$ is a subspace of V.

Proof.

Let $v, w \in V_{(\lambda)}$. Then
$\varphi(v + w) = \varphi(v) + \varphi(w) = \lambda v + \lambda w = \lambda(v + w)$. Hence
$v + w \in V_{(\lambda)}$. For any $\alpha \in \mathbb{R}$ we have $\varphi(\alpha v) = \alpha\varphi(v) = \lambda(\alpha v)$.
Hence $\alpha v \in V_{(\lambda)}$. $\square$

For any eigenvalue λ of φ the subspace $V_{(\lambda)}$ is called **the eigenspace** associated to λ.

Eigenvalues and Eigenvectors (continued)

Let $\varphi : V \longrightarrow V$ be a linear endomorphism. For any eigenvalue λ of φ let $V_{(\lambda)}$ denote the set of all eigenvectors associated to λ together with the zero vector, i.e.

$$V_{(\lambda)} = \{v \in V \mid \varphi(v) = \lambda v.\}$$

Proposition

The subset $V_{(\lambda)} \subset V$ is a subspace of V.

Proof.

Let $v, w \in V_{(\lambda)}$. Then $\varphi(v + w) = \varphi(v) + \varphi(w) = \lambda v + \lambda w = \lambda(v + w)$. Hence $v + w \in V_{(\lambda)}$. For any $\alpha \in \mathbb{R}$ we have $\varphi(\alpha v) = \alpha\varphi(v) = \lambda(\alpha v)$. Hence $\alpha v \in V_{(\lambda)}$. $\square$

For any eigenvalue λ of φ the subspace $V_{(\lambda)}$ is called **the eigenspace** associated to λ. It is straightforward that $\varphi(V_{(\lambda)}) \subset V_{(\lambda)}$.

Example

Let $s : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a reflection of $\mathbb{R}^2$ about the x_1-axis. Then $V_{(1)} = \text{lin}((1, 0))$ and $V_{(-1)} = \text{lin}((0, 1))$.

Example

Let $s : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a reflection of $\mathbb{R}^2$ about the x_1-axis. Then $V_{(1)} = \text{lin}((1,0))$ and $V_{(-1)} = \text{lin}((0,1))$. The rotation r about the origin of $\mathbb{R}^2$ by $\frac{\pi}{2}$ radians counter-clockwise has no eigenvalues (no line is mapped into itself).

Example

Let $s : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a reflection of $\mathbb{R}^2$ about the x_1-axis. Then $V_{(1)} = \mathrm{lin}((1,0))$ and $V_{(-1)} = \mathrm{lin}((0,1))$. The rotation r about the origin of $\mathbb{R}^2$ by $\frac{\pi}{2}$ radians counter-clockwise has no eigenvalues (no line is mapped into itself). In the case of uniform scaling k by -2 in all directions any non-zero vector is eigenvector associated to -2, i.e. $V_{(-2)} = \mathbb{R}^2$.

Example

Let $s : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a reflection of $\mathbb{R}^2$ about the x_1-axis. Then $V_{(1)} = \mathrm{lin}((1,0))$ and $V_{(-1)} = \mathrm{lin}((0,1))$. The rotation r about the origin of $\mathbb{R}^2$ by $\frac{\pi}{2}$ radians counter-clockwise has no eigenvalues (no line is mapped into itself). In the case of uniform scaling k by -2 in all directions any non-zero vector is eigenvector associated to -2, i.e. $V_{(-2)} = \mathbb{R}^2$.The projection p onto the x_2-axis has two eigenspaces: $V_{(0)} = \mathrm{lin}((1,0))$ and $V_{(1)} = \mathrm{lin}((0,1))$.

Example

Let $s : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a reflection of $\mathbb{R}^2$ about the x_1-axis. Then $V_{(1)} = \mathrm{lin}((1,0))$ and $V_{(-1)} = \mathrm{lin}((0,1))$. The rotation r about the origin of $\mathbb{R}^2$ by $\frac{\pi}{2}$ radians counter-clockwise has no eigenvalues (no line is mapped into itself). In the case of uniform scaling k by -2 in all directions any non-zero vector is eigenvector associated to -2, i.e. $V_{(-2)} = \mathbb{R}^2$.The projection p onto the x_2-axis has two eigenspaces: $V_{(0)} = \mathrm{lin}((1,0))$ and $V_{(1)} = \mathrm{lin}((0,1))$.
Note that for s, k and p there exist a basis (the standard one) consisting of eigenvectors. The matrices of those endomorphisms in the standard basis are diagonal.

Example

Let $s : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a reflection of $\mathbb{R}^2$ about the x_1-axis. Then $V_{(1)} = \mathrm{lin}((1,0))$ and $V_{(-1)} = \mathrm{lin}((0,1))$. The rotation r about the origin of $\mathbb{R}^2$ by $\frac{\pi}{2}$ radians counter-clockwise has no eigenvalues (no line is mapped into itself). In the case of uniform scaling k by -2 in all directions any non-zero vector is eigenvector associated to -2, i.e. $V_{(-2)} = \mathbb{R}^2$.The projection p onto the x_2-axis has two eigenspaces: $V_{(0)} = \mathrm{lin}((1,0))$ and $V_{(1)} = \mathrm{lin}((0,1))$.
Note that for s, k and p there exist a basis (the standard one) consisting of eigenvectors. The matrices of those endomorphisms in the standard basis are diagonal.

$$M(s)_{st} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \; M(k)_{st} = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix},$$

$$M(p)_{st} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Characteristic Polynomial

Definition

Let $A \in M(n \times n; \mathbb{R})$. The polynomial $w_A(\lambda) = \det(A - \lambda I_n)$ is called the characteristic polynomial of A.

Characteristic Polynomial

Definition

Let $A \in M(n \times n; \mathbb{R})$. The polynomial $w_A(\lambda) = \det(A - \lambda I_n)$ is called the characteristic polynomial of A.

The degree of $w_A(\lambda)$ is equal to n.

Characteristic Polynomial

Definition

Let $A \in M(n \times n; \mathbb{R})$. The polynomial $w_A(\lambda) = \det(A - \lambda I_n)$ is called the characteristic polynomial of A.

The degree of $w_A(\lambda)$ is equal to n.

Example

Let $A = \begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix}$. Then

$$w_A(\lambda) = \det \begin{bmatrix} 4-\lambda & 2 \\ 3 & 3-\lambda \end{bmatrix} = (4-\lambda)(3-\lambda) - 6 = \lambda^2 - 7\lambda + 6.$$

Characteristic Polynomial

Definition

Let $A \in M(n \times n; \mathbb{R})$. The polynomial $w_A(\lambda) = \det(A - \lambda I_n)$ is called the characteristic polynomial of A.

The degree of $w_A(\lambda)$ is equal to n.

Example

Let $A = \begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix}$. Then

$$w_A(\lambda) = \det \begin{bmatrix} 4 - \lambda & 2 \\ 3 & 3 - \lambda \end{bmatrix} = (4 - \lambda)(3 - \lambda) - 6 = \lambda^2 - 7\lambda + 6.$$

Proposition

Let $A, B \in M(n \times n; \mathbb{R})$ be similar matrices. Then $w_A = w_B$.

Characteristic Polynomial

Definition

Let $A \in M(n \times n; \mathbb{R})$. The polynomial $w_A(\lambda) = \det(A - \lambda I_n)$ is called the characteristic polynomial of A.

The degree of $w_A(\lambda)$ is equal to n.

Example

Let $A = \begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix}$. Then

$$w_A(\lambda) = \det \begin{bmatrix} 4-\lambda & 2 \\ 3 & 3-\lambda \end{bmatrix} = (4-\lambda)(3-\lambda) - 6 = \lambda^2 - 7\lambda + 6.$$

Proposition

Let $A, B \in M(n \times n; \mathbb{R})$ be similar matrices. Then $w_A = w_B$.

Proof.

There exists an invertible matrix C such that $A = C^{-1}BC$. But
$w_A(\lambda) = \det(A - \lambda I_n) = \det\big(C^{-1}BC - C^{-1}\lambda I_n C\big) =$
$\det\big(C^{-1}(B - \lambda I_n)C\big) = (\det C)^{-1} \det(B - \lambda I_n) \det C =$
$w_B(\lambda)$. $\square$

Characteristic Polynomial (continued)

Definition

Let $\varphi : V \longrightarrow V$ be a linear endomorphism of a finite dimensional vector space V. The characteristic polynomial w_φ of φ is the characteristic polynomial of matrix $M(\varphi)_\mathcal{A}$ where $\mathcal{A}$ is a basis of V.

Characteristic Polynomial (continued)

Definition

Let $\varphi : V \longrightarrow V$ be a linear endomorphism of a finite dimensional vector space V. The characteristic polynomial w_φ of φ is the characteristic polynomial of matrix $M(\varphi)_\mathcal{A}$ where $\mathcal{A}$ is a basis of V.

By the previous proposition the characteristic polynomial of φ does not depend on the basis $\mathcal{A}$.

Finding Eigenvalues and Eigenvectors

Theorem

Let $\varphi : V \longrightarrow V$ *be a linear endomorphism of a finite dimensional vector space* V*.*

Finding Eigenvalues and Eigenvectors

Theorem

Let $\varphi : V \longrightarrow V$ be a linear endomorphism of a finite dimensional vector space V.

i) *$\alpha \in \mathbb{R}$ is an eigenvalue of $\varphi \iff \alpha$ is a root the characteristic polynomial of φ,*

Finding Eigenvalues and Eigenvectors

Theorem

Let $\varphi : V \longrightarrow V$ be a linear endomorphism of a finite dimensional vector space V.

i) *$\alpha \in \mathbb{R}$ is an eigenvalue of $\varphi \iff \alpha$ is a root the characteristic polynomial of φ,*

ii) *let $\mathcal{A} = (v_1, \dots, v_n)$ and $A = M(\varphi)_{\mathcal{A}}$. The vector $v = x_1 v_1 + \ldots + x_n v_n$ is an eigenvector of φ associated to α if and only if*

$$(A - \alpha I_n) \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Finding Eigenvalues and Eigenvectors (continued)

Proof.

Let $v = x_1 v_1 + \ldots + x_n v_n$. Then $\varphi(v) = \alpha v$ if and only if

$$A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \alpha \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \iff (A - \alpha I_n) \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Finding Eigenvalues and Eigenvectors (continued)

Proof.

Let $v = x_1 v_1 + \ldots + x_n v_n$. Then $\varphi(v) = \alpha v$ if and only if

$$A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \alpha \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \iff (A - \alpha I_n) \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

From the previous lecture we know that there exists a non-zero solution of the latter if and only if $\det(A - \alpha I_n) = 0$, i.e. $w_A(\alpha) = 0$.

$\square$

Example

Let $\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be an endomorphism of $\mathbb{R}^3$ given by $\varphi(x_1, x_2, x_3) = (4x_1 + 4x_2, -x_1, x_1 + 3x_2 + 3x_3)$. Its matrix in the standard basis is $A = M(\varphi)_{st} = \begin{bmatrix} 4 & 4 & 0 \\ -1 & 0 & 0 \\ 1 & 3 & 3 \end{bmatrix}$.

$$A - \lambda I = \begin{bmatrix} 4-\lambda & 4 & 0 \\ -1 & -\lambda & 0 \\ 1 & 3 & 3-\lambda \end{bmatrix}.$$

Hence $w_\varphi(\lambda) = \det(A - \lambda I) = (3-\lambda)((4-\lambda)(-\lambda) + 4)) = (3-\lambda)(\lambda^2 - 4\lambda + 4) = (3-\lambda)(2-\lambda)^2$.

Example

Let $\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be an endomorphism of $\mathbb{R}^3$ given by $\varphi(x_1, x_2, x_3) = (4x_1 + 4x_2, -x_1, x_1 + 3x_2 + 3x_3)$. Its matrix in the standard basis is $A = M(\varphi)_{st} = \begin{bmatrix} 4 & 4 & 0 \\ -1 & 0 & 0 \\ 1 & 3 & 3 \end{bmatrix}$.

$$A - \lambda I = \begin{bmatrix} 4-\lambda & 4 & 0 \\ -1 & -\lambda & 0 \\ 1 & 3 & 3-\lambda \end{bmatrix}.$$

Hence $w_\varphi(\lambda) = \det(A - \lambda I) = (3-\lambda)((4-\lambda)(-\lambda)+4) = (3-\lambda)(\lambda^2 - 4\lambda + 4) = (3-\lambda)(2-\lambda)^2$. There are two eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 3$. To find $V_{(2)}$ we solve a system of linear equations:

$$V_{(2)} : \begin{bmatrix} 2 & 4 & 0 \\ -1 & -2 & 0 \\ 1 & 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Example (continued)

$$\begin{bmatrix} 2 & 4 & 0 \\ -1 & -2 & 0 \\ 1 & 3 & 1 \end{bmatrix} \xrightarrow[r_3+r_2]{r_1+2r_2} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{r_1-2r_2} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Example (continued)

$$\begin{bmatrix} 2 & 4 & 0 \\ -1 & -2 & 0 \\ 1 & 3 & 1 \end{bmatrix} \xrightarrow[r_3+r_2]{r_1+2r_2} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{r_1-2r_2} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore $x_1 = 2x_3, \; x_2 = -x_3, \; x_3 \in \mathbb{R}$, i.e.

$$V_{(2)} = \{(2x_3, -x_3, x_3) \mid x_3 \in \mathbb{R}\} = \mathrm{lin}((2, -1, 1)).$$

Example (continued)

$$\begin{bmatrix} 2 & 4 & 0 \\ -1 & -2 & 0 \\ 1 & 3 & 1 \end{bmatrix} \xrightarrow[r_3+r_2]{r_1+2r_2} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{r_1-2r_2} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore $x_1 = 2x_3,\ x_2 = -x_3,\ x_3 \in \mathbb{R}$, i.e.

$$V_{(2)} = \{(2x_3, -x_3, x_3) \mid x_3 \in \mathbb{R}\} = \mathrm{lin}((2, -1, 1)).$$

$$V_{(3)} : \begin{bmatrix} 1 & 4 & 0 \\ -1 & -3 & 0 \\ 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Example (continued)

$$\begin{bmatrix} 1 & 4 & 0 \\ -1 & -3 & 0 \\ 1 & 3 & 0 \end{bmatrix} \xrightarrow[]{\substack{r_1+r_2 \\ r_3+r_2}} \begin{bmatrix} 0 & 1 & 0 \\ -1 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{r_2+3r_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Example (continued)

$$\begin{bmatrix} 1 & 4 & 0 \\ -1 & -3 & 0 \\ 1 & 3 & 0 \end{bmatrix} \xrightarrow[r_3+r_2]{r_1+r_2} \begin{bmatrix} 0 & 1 & 0 \\ -1 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{r_2+3r_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore $x_1 = x_2 = 0,\ x_3 \in \mathbb{R}$, i.e.

$$V_{(3)} = \{(0, 0, x_3) \mid x_3 \in \mathbb{R}\} = \mathrm{lin}((0, 0, 1)).$$

Example (continued)

Recall that

$$\varphi(x_1, x_2, x_3) = (4x_1 + 4x_2, -x_1, x_1 + 3x_2 + 3x_3),$$

$$V_{(2)} = \mathrm{lin}((2, -1, 1)),$$

$$V_{(3)} = \mathrm{lin}((0, 0, 1)),$$

and check those directly

$$\varphi(2, -1, 1) =$$

Example (continued)

Recall that

$$\varphi(x_1, x_2, x_3) = (4x_1 + 4x_2, -x_1, x_1 + 3x_2 + 3x_3),$$

$$V_{(2)} = \mathrm{lin}((2, -1, 1)),$$

$$V_{(3)} = \mathrm{lin}((0, 0, 1)),$$

and check those directly

$$\varphi(2, -1, 1) = (4, -2, 2) = 2(2, -1, 1),$$

Example (continued)

Recall that

$$\varphi(x_1, x_2, x_3) = (4x_1 + 4x_2, -x_1, x_1 + 3x_2 + 3x_3),$$

$$V_{(2)} = \operatorname{lin}((2, -1, 1)),$$

$$V_{(3)} = \operatorname{lin}((0, 0, 1)),$$

and check those directly

$$\varphi(2, -1, 1) = (4, -2, 2) = 2(2, -1, 1),$$

$$\varphi(0, 0, 1) =$$

Example (continued)

Recall that

$$\varphi(x_1, x_2, x_3) = (4x_1 + 4x_2, -x_1, x_1 + 3x_2 + 3x_3),$$

$$V_{(2)} = \mathrm{lin}((2, -1, 1)),$$
$$V_{(3)} = \mathrm{lin}((0, 0, 1)),$$

and check those directly

$$\varphi(2, -1, 1) = (4, -2, 2) = 2(2, -1, 1),$$

$$\varphi(0, 0, 1) = (0, 0, 3) = 3(0, 0, 1).$$

Remarks

i) if $\varphi : V \longrightarrow V$ and $\dim V$ is odd then the degree of w_φ is odd therefore it has at least one real root so there exists an eigenvector of φ,

Remarks

i) if $\varphi : V \longrightarrow V$ and $\dim V$ is odd then the degree of w_φ is odd therefore it has at least one real root so there exists an eigenvector of φ,

ii) $\dim V_{(\alpha)} \leqslant$ multiplicity of the root α in w_φ, cf. the last example (2 is a root of multiplicity 2 but $\dim V_{(2)} = 1$),

Remarks

i) if $\varphi : V \longrightarrow V$ and $\dim V$ is odd then the degree of w_φ is odd therefore it has at least one real root so there exists an eigenvector of φ,

ii) $\dim V_{(\alpha)} \leqslant$ multiplicity of the root α in w_φ, cf. the last example (2 is a root of multiplicity 2 but $\dim V_{(2)} = 1$),

iii) if $A \in M(n \times n; \mathbb{R})$ then $w_A(A) = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix}$, i.e. matrix A substituted to its characteristic polynomial gives the zero matrix (Cayley-Hamilton theorem).

Example

Let $A = \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix}$ and $w_A(\lambda) = \lambda^2 - 2\lambda - 2$. Then

$$w_A(A) = \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix}^2 - 2\begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix} - 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 4 & 6 \\ 2 & 4 \end{bmatrix} + \begin{bmatrix} -2 & -6 \\ -2 & -2 \end{bmatrix} + \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Cayley–Hamilton Theorem

Theorem

For any $A \in M(n \times n; \mathbb{R})$ *and* $w_A(\lambda) = \det(A - \lambda I_n)$

$$w_A(A) = 0.$$

Cayley–Hamilton Theorem

Theorem

For any $A \in M(n \times n; \mathbb{R})$ *and* $w_A(\lambda) = \det(A - \lambda I_n)$

$$w_A(A) = 0.$$

Proof.

Let $B = \operatorname{adj}(A - \lambda I_n)$ be the adjugate matrix of the matrix $A - \lambda I_n$.

Cayley–Hamilton Theorem

Theorem

For any $A \in M(n \times n; \mathbb{R})$ *and* $w_A(\lambda) = \det(A - \lambda I_n)$

$$w_A(A) = 0.$$

Proof.

Let $B = \mathrm{adj}(A - \lambda I_n)$ be the adjugate matrix of the matrix $A - \lambda I_n$. The entries of B are polynomials of degree at most $n-1$. By separating monomials of the same degree one can write

$$B = \lambda^{n-1} B_{n-1} + \lambda^{n-2} B_{n-2} + \ldots + \lambda B_1 + B_0,$$

where $B_i \in M(n \times n; \mathbb{R})$ for $i = 0, \ldots, n-1$.

Cayley–Hamilton Theorem (continued)

Proof.

By the matrix inverse formula

$$B(A - \lambda I_n) = w_A(\lambda) I_n = \lambda^n a_n I_n + \lambda^{n-1} a_{n-1} I_n + \ldots + \lambda a_1 I_n + a_0 I_n,$$

where

$$w_A(\lambda) = \lambda^n a_n + \lambda^{n-1} a_{n-1} + \ldots + \lambda a_1 + a_0,$$

is the characteristic polynomial of matrix A.

Cayley–Hamilton Theorem (continued)

Proof.

By the matrix inverse formula

$$B(A-\lambda I_n) = w_A(\lambda) I_n = \lambda^n a_n I_n + \lambda^{n-1} a_{n-1} I_n + \ldots + \lambda a_1 I_n + a_0 I_n,$$

where

$$w_A(\lambda) = \lambda^n a_n + \lambda^{n-1} a_{n-1} + \ldots + \lambda a_1 + a_0,$$

is the characteristic polynomial of matrix A. Hence

$$B(A-\lambda I_n) = \lambda^{n-1} B_{n-1} A + \ldots + \lambda^2 B_2 A + \lambda B_1 A + B_0 A +$$

$$-\lambda^n B_{n-1} - \lambda^{n-1} B_{n-2} - \ldots - \lambda^2 B_1 - \lambda B_0 =$$

$$= -\lambda^n B_{n-1} + \lambda^{n-1}(B_{n-1}A - B_{n-2}) + \lambda^{n-2}(B_{n-2}A - B_{n-3}) + \ldots +$$

$$+\lambda^2(B_2 A - B_1) + \lambda(B_1 A - B_0) + B_0 A.$$

Two polynomials with real coefficients are equal if and only if they have the same coefficients, therefore,

Cayley–Hamilton Theorem (continued)

Proof.

$$\begin{aligned} -B_{n-1} &= a_n I_n, \\ B_{n-1}A - B_{n-2} &= a_{n-1} I_n, \\ &\vdots \\ B_1 A - B_0 &= a_1 I_n, \\ B_0 A &= a_0 I_n. \end{aligned}$$

Multiplying those equations on the right by $A^n, A^{n-1}, \ldots, A, A^0 = I_n$ respectively one gets

$$\begin{aligned} -B_{n-1}A^n &= a_n A^n, \\ B_{n-1}A^n - B_{n-2}A^{n-1} &= a_{n-1}A^{n-1}, \\ &\vdots \\ B_1 A^2 - B_0 A &= a_1 A, \\ B_0 A &= a_0 I_n. \end{aligned}$$

Cayley–Hamilton Theorem – Proof

Proof.

This sums to

$$w_A(A) = a_n A^n + a_{n-1} A^{n-1} + \ldots + a_1 A + a_0 I_n = 0.$$

$\square$

Remark

There exist other conceptual proofs of the Cayley-Hamilton theorem (using abstract algebra of Schur decomposition).

Schur Decomposition

Proposition

For any matrix $A \in M(n \times n; \mathbb{C})$ *there exists a unitary matrix* $U \in M(n \times n; \mathbb{C})$ *(i.e.* $U^*U = UU^* = I$, *where* $U^* = \overline{U}^{\mathsf{T}}$*) and an upper triangular matrix* $T = [t_{ij}] \in M(n \times n; \mathbb{C})$ *(i.e.* $t_{ij} = 0$ *for* $i > j$*) such that*

$$A = UTU^*.$$

The decomposition is not unique and the diagonal entries of matrix T *are exactly (complex) eigenvalues of matrix* A*.*

Schur Decomposition

Proposition

For any matrix $A \in M(n \times n; \mathbb{C})$ *there exists a unitary matrix* $U \in M(n \times n; \mathbb{C})$ *(i.e.* $U^*U = UU^* = I$, *where* $U^* = \overline{U}^\mathsf{T}$*) and an upper triangular matrix* $T = [t_{ij}] \in M(n \times n; \mathbb{C})$ *(i.e.* $t_{ij} = 0$ *for* $i > j$*) such that*

$$A = UTU^*.$$

The decomposition is not unique and the diagonal entries of matrix T *are exactly (complex) eigenvalues of matrix* A*.*

Proof.

Omitted. □

Cayley–Hamilton Theorem Alternative Proof via Schur Decomposition

Proof.

Let $UTU^* = A$. Then

$$w_A(A) = Uw_A(T)U^*.$$

Moreover, if

$$w_A(\lambda) = (\lambda - \lambda_1) \cdot \ldots \cdot (\lambda - \lambda_n),$$

then

$$w_A(T) = (T - \lambda_1 I) \cdot \ldots \cdot (T - \lambda_n I) = 0,$$

that is, the first k columns of the product

$$(T - \lambda_1 I) \cdot \ldots \cdot (T - \lambda_k I),$$

are zero. $\square$

Characteristic Polynomials of AB and BA

Proposition

Let $A \in M(m \times n; \mathbb{R})$ and let $B \in M(n \times m; \mathbb{R})$ where $m \geqslant n$. Then $AB \in M(m \times m; \mathbb{R})$, $BA \in M(n \times n; \mathbb{R})$ and

$$w_{AB}(\lambda) = \lambda^{m-n} w_{BA}(\lambda),$$

that is eigenvalues of AB and BA (up to $m - n$ zeroes) are the same. Moreover, the dimensions of eigenspaces corresponding to non–zero eigenvalues are the same.

Characteristic Polynomials of AB and BA

Proposition

Let $A \in M(m \times n; \mathbb{R})$ and let $B \in M(n \times m; \mathbb{R})$ where $m \geqslant n$. Then $AB \in M(m \times m; \mathbb{R})$, $BA \in M(n \times n; \mathbb{R})$ and

$$w_{AB}(\lambda) = \lambda^{m-n} w_{BA}(\lambda),$$

that is eigenvalues of AB and BA (up to $m - n$ zeroes) are the same. Moreover, the dimensions of eigenspaces corresponding to non–zero eigenvalues are the same.

Proof.

Let

$$M = \left[\begin{array}{c|c} AB & 0 \\ \hline B & 0 \end{array}\right], \quad N = \left[\begin{array}{c|c} 0 & 0 \\ \hline B & BA \end{array}\right], \quad C = \left[\begin{array}{c|c} I_m & A \\ \hline 0 & I_n \end{array}\right],$$

be $(m+n) \times (m+n)$ matrices.

Characteristic Polynomials of AB and BA

Proof.

Then

$$C^{-1} = \left[\begin{array}{c|c} I_m & -A \\ \hline 0 & I_n \end{array}\right], \quad C^{-1}MC = N,$$

i.e. the matrices are similar hence they have the same eigenvalues.

This holds as

$$MC = CN,$$

$$\left[\begin{array}{c|c} AB & 0 \\ \hline B & 0 \end{array}\right] \left[\begin{array}{c|c} I_m & A \\ \hline 0 & I_n \end{array}\right] = \left[\begin{array}{c|c} I_m & A \\ \hline 0 & I_n \end{array}\right] \left[\begin{array}{c|c} 0 & 0 \\ \hline B & BA \end{array}\right] = \left[\begin{array}{c|c} AB & ABA \\ \hline B & BA \end{array}\right].$$

Characteristic Polynomials of AB and BA

Proof.

Alternatively, for $\lambda \neq 0$ the following linear transfomations are inverse to each other hence invertible

$$\ker(AB - \lambda I) \ni v \mapsto \frac{1}{\lambda} Bv \in \ker(BA - \lambda I),$$

$$\ker(BA - \lambda I) \ni v \mapsto \frac{1}{\lambda} Av \in \ker(AB - \lambda I).$$

In particular $\ker(AB - \lambda I) \neq \{0\}$ if and only if $\ker(BA - \lambda I) \neq \{0\}$. $\square$

Nilpotent Matrix

Definition

Let $A \in M(n \times n; \mathbb{R})$. Matrix A is nilpotent if there exists $k \geqslant 1$ such that

$$A^k = 0.$$

Nilpotent Matrix

Definition

Let $A \in M(n \times n; \mathbb{R})$. Matrix A is nilpotent if there exists $k \geqslant 1$ such that

$$A^k = 0.$$

Proposition

If matrix $A \in M(n \times n; \mathbb{R})$ is nilpotent and $\lambda \in \mathbb{R}$ is an eigenvalue of A then $\lambda = 0$, i.e. all eigenvalues are equal to 0.

Nilpotent Matrix

Definition

Let $A \in M(n \times n; \mathbb{R})$. Matrix A is nilpotent if there exists $k \geqslant 1$ such that

$$A^k = 0.$$

Proposition

If matrix $A \in M(n \times n; \mathbb{R})$ is nilpotent and $\lambda \in \mathbb{R}$ is an eigenvalue of A then $\lambda = 0$, i.e. all eigenvalues are equal to 0.

Proof.

Let $k \geqslant 1$ be any number such that $A^k = 0$. Let $v \in \mathbb{R}^n$ be an eigenvector of A for the eigenvalue $\lambda \in \mathbb{R}$. Then

$$(A^k)v = \lambda^k v = 0 \Longrightarrow \lambda = 0,$$

since $v \neq 0$. $\square$

Nilpotent Matrix (continued)

Corollary

Matrix $A \in M(n \times n; \mathbb{R})$ *is nilpotent if and only if its all eigenvalues over complex numbers are equal to* 0 *(i.e. the characteristic polynomial* $w_A(\lambda) = (-1)^n \lambda^n$*).*

Companion Matrix

Proposition

For any $a_0, \dots, a_{n-1} \in \mathbb{R}$ *where* $n \geqslant 2$ *if*

$$A = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ 0 & 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix},$$

then

$$w_A(\lambda) = (-1)^n(\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0).$$

Companion Matrix

Proposition

For any $a_0, \ldots, a_{n-1} \in \mathbb{R}$ *where* $n \geqslant 2$ *if*

$$A = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ 0 & 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix},$$

then

$$w_A(\lambda) = (-1)^n(\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0).$$

Proof.

Induction on n. If $n = 2$ then

$$\begin{bmatrix} -\lambda & -a_0 \\ 1 & -a_1 - \lambda \end{bmatrix} = \lambda^2 + a_1\lambda + a_0.$$

Companion Matrix (continued)

Proof.

For $n \geqslant 3$, by the Laplace formula for the first column and the inductive assumption

$$\det \begin{bmatrix} -\lambda & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & -\lambda & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & -\lambda & \cdots & 0 & -a_2 \\ 0 & 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1}-\lambda \end{bmatrix} =$$

$$= -\lambda(-1)^{n-1}(\lambda^{n-1} + \ldots + a_2\lambda + a_1) -$$

$$-\det \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & -\lambda & \cdots & 0 & -a_2 \\ 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1}-\lambda \end{bmatrix} =$$

$$= (-1)^n(\lambda^n + \ldots + a_2\lambda^2 + a_1\lambda) - (-1)^n(-a_0).$$

Companion Matrix (continued)

Corollary

Up to a sign, each monic polynomial of degree n is a characteristic polynomial of some matrix $A \in M(n \times n; \mathbb{R})$.

Primitive and Irreducible Matrices

Definition

Let $A \in M(n \times n; \mathbb{R})$. Matrix A is **primitive** if $A \geqslant 0$ and there exists k such that $A^k > 0$.

Primitive and Irreducible Matrices

Definition

Let $A \in M(n \times n; \mathbb{R})$. Matrix A is **primitive** if $A \geqslant 0$ and there exists k such that $A^k > 0$. Matrix A is **irreducible** if $A \geqslant 0$ and for each $1 \leqslant i, j \leqslant n$ there exists k such that $(A^k)_{ij} > 0$.

Primitive and Irreducible Matrices

Definition

Let $A \in M(n \times n; \mathbb{R})$. Matrix A is **primitive** if $A \geqslant 0$ and there exists k such that $A^k > 0$. Matrix A is **irreducible** if $A \geqslant 0$ and for each $1 \leqslant i, j \leqslant n$ there exists k such that $(A^k)_{ij} > 0$.

Remark

If matrix A is primitive then it is irreducible. If matrix A is irreducible then matrix $A + I$ is primitive because $A^m \geqslant 0$ and

$$(A + I)^k = I + \binom{k}{1} A + \binom{k}{2} A^2 + \binom{k}{3} A^3 + \ldots + \binom{k}{k} A^k,$$

for $k = \max\{k_{ij}\}$.

Perron–Frobenius Theorem

Theorem

Let A be an irreducible matrix. Then there exist $\lambda_{max} \in \mathbb{R}, \lambda_{max} > 0$ a positive eigenvalue of A such that

i) *for any other eigenvalue $\lambda \in \mathbb{C}$ of matrix A*

$$|\lambda| \leqslant \lambda_{max},$$

Perron–Frobenius Theorem

Theorem

Let A be an irreducible matrix. Then there exist $\lambda_{max} \in \mathbb{R}, \lambda_{max} > 0$ a positive eigenvalue of A such that

i) *for any other eigenvalue $\lambda \in \mathbb{C}$ of matrix A*

$$|\lambda| \leqslant \lambda_{max},$$

ii) $V_{(\lambda_{max})} = \mathrm{lin}(v)$ *where $v \in \mathbb{R}^n$ and $v > 0$ (i.e., all entries of v are positive),*

Perron–Frobenius Theorem

Theorem

Let A be an irreducible matrix. Then there exist $\lambda_{max} \in \mathbb{R}, \lambda_{max} > 0$ a positive eigenvalue of A such that

i) *for any other eigenvalue $\lambda \in \mathbb{C}$ of matrix A*

$$|\lambda| \leqslant \lambda_{max},$$

ii) $V_{(\lambda_{max})} = \mathrm{lin}(v)$ *where $v \in \mathbb{R}^n$ and $v > 0$ (i.e., all entries of v are positive),*

iii) λ_{max} *is a simple root of $w_A(\lambda)$ (i.e $w_A(\lambda)$ is not divisible by $(\lambda - \lambda_{max})^2$),*

Perron–Frobenius Theorem

Theorem

Let A be an irreducible matrix. Then there exist $\lambda_{max} \in \mathbb{R}, \lambda_{max} > 0$ a positive eigenvalue of A such that

i) *for any other eigenvalue $\lambda \in \mathbb{C}$ of matrix A*

$$|\lambda| \leqslant \lambda_{max},$$

ii) $V_{(\lambda_{max})} = \mathrm{lin}(v)$ *where $v \in \mathbb{R}^n$ and $v > 0$ (i.e., all entries of v are positive),*

iii) λ_{max} *is a simple root of $w_A(\lambda)$ (i.e $w_A(\lambda)$ is not divisible by $(\lambda - \lambda_{max})^2$),*

iv) *if $w \in \mathbb{R}^n$, $w > 0$ and w is an eigenvalue of A then $w \in V_{(\lambda_{max})}$.*

Perron–Frobenius Theorem (continued)

Remark

If A is a primitive matrix then moreover

$$|\lambda| < \lambda_{max},$$

for any eigenvalue $\lambda \in \mathbb{C}$ of A.

Perron–Frobenius Theorem Proof

Let $k \in \mathbb{N}$ be a number such that

$$B = (I + A)^k > 0.$$

Obviously

$$\text{if } v \leqslant w, v \neq w \text{ then } Bv < Bw.$$

Let

$$Q = \{v \in \mathbb{R}^n \mid v \geqslant 0,\ \ v \neq 0\}, \quad C = Q \cap \{v \in \mathbb{R}^n \mid \|v\| = 1\}.$$

For any $v \geqslant 0$ such that $v \neq 0$ let

$$L(v) = \max\{\lambda \in \mathbb{R} \mid \lambda v \leqslant Av\} = \min_{\substack{1 \leqslant i \leqslant n \\ v_i \neq 0}} \frac{(Av)_i}{v_i}.$$

It is clear that $L(\mu v) = L(v)$ for $\mu > 0$, in particular $L\left(\frac{v}{\|v\|}\right) = L(v)$ for $v \geqslant 0, v \neq 0$.

Perron–Frobenius Theorem Proof (continued)

For any $v \geqslant 0$ such that $v \neq 0$

$$\text{if } \mu v \leqslant Av \text{ then } \mu Bv \leqslant BAv = ABv,$$

which implies that (maximum over a larger set)

$$L(v) \leqslant L(Bv).$$

Moreover, if $Av \neq L(v)v$ then, by the definition,

$$L(v)v \leqslant Av, \quad Av \neq L(v)v, \quad \text{hence} \quad B(L(v)v) < B(Av).$$

This is equivalent to

$$L(v)Bv < ABv, \text{ i.e., } L(v) < L(Bv),$$

(i-th components of $L(v)A$ and Av are equal for some i).

Perron–Frobenius Theorem Proof (continued)

By abuse of notation, let

$$B \colon \mathbb{R}^n \to \mathbb{R}^n,$$

denote the linear function given by the matrix B. Since the set C is compact then $B(C) \subset \mathbb{R}^n_{>0}$ is compact too and $v \geqslant 0, v \neq 0$ implies that $Bv > B0 = 0$. By the Weierstrass extreme value theorem, the function L (which is continuous as minimum of continuous functions and all components are non–zero) obtains its maximum on the set $B(C)$. Let

$$\lambda_{max} = \max_{v \in B(C)} L(v),$$

$$v = \arg\max_{v \in B(C)} L(v).$$

By the above $v > 0$ and

$$Av = L(v)v = \lambda_{max} v.$$

Perron–Frobenius Theorem Proof (continued)

Since $L(v) \leqslant L(Bv)$

$$\lambda_{max} = \max_{v \in B(C)} L(v) = \max_{v \in C} L(v).$$

Perron–Frobenius Theorem Proof (continued)

Since $L(v) \leqslant L(Bv)$

$$\lambda_{max} = \max_{v \in B(C)} L(v) = \max_{v \in C} L(v).$$

Beacuse

$$Av = \lambda_{max} v,$$

$$A \geqslant 0, v > 0 \Longrightarrow Av > 0,$$

it follows that

$$\lambda_{max} > 0.$$

Perron–Frobenius Theorem Proof (continued)

Let $w \in \mathbb{C}^n, \lambda \in \mathbb{C}$ be such that

$$\lambda w = Aw,$$

i.e., for $i = 1, \ldots, n$

$$\lambda w_i = \sum_{i=1}^{n} a_{ij} w_j,$$

$$|\lambda||w_i| \leqslant \sum_{i=1}^{n} a_{ij} |w_j|,$$

since $a_{ij} \geqslant 0$. This is equivalent to

$$|\lambda||w| \leqslant A|w|,$$

where

$$|w| = (|w_1|, \ldots, |w_n|) \in \mathbb{R}^n.$$

Perron–Frobenius Theorem Proof (continued)

Let $w \in \mathbb{C}^n, \lambda \in \mathbb{C}$ be such that

$$\lambda w = Aw,$$

i.e., for $i = 1, \ldots, n$

$$\lambda w_i = \sum_{i=1}^{n} a_{ij} w_j,$$

$$|\lambda||w_i| \leqslant \sum_{i=1}^{n} a_{ij}|w_j|,$$

since $a_{ij} \geqslant 0$. This is equivalent to

$$|\lambda||w| \leqslant A|w|,$$

where

$$|w| = (|w_1|, \ldots, |w_n|) \in \mathbb{R}^n.$$

By definition

$$|\lambda| \leqslant L(|w|) \leqslant \lambda_{max},$$

i.e. λ_{max} is a real eigenvalue with maximal modulus and positive eigenvector.

Perron–Frobenius Theorem Proof (continued)

It is now enough to prove that v is a unique eigenvector for the simple eigenvalue λ_{max} and all other positive eigenvectors are multiples or vector v.

Perron–Frobenius Theorem Proof (continued)

It is now enough to prove that v is a unique eigenvector for the simple eigenvalue λ_{max} and all other positive eigenvectors are multiples or vector v.

As $A^\intercal$ is irreducible too there exists left eigenvalue $\mu_{max} > 0$ and a positive eigenvector $w > 0$ such that $w^\intercal A = \mu_{max} w$. Then

$$\mu_{max} w^\intercal v = (w^\intercal A) v = w^\intercal (Av) = \lambda_{max} w^\intercal v,$$

hence $\mu_{max} = \lambda_{max}$ as $w^\intercal v > 0$.

Perron–Frobenius Theorem Proof (continued)

Suppose that there exist $\eta \in \mathbb{R}$ and $u \geqslant 0, u \neq 0$ such that $Au = \eta u$. Then

$$\eta w^\intercal u = (w^\intercal A)u = w^\intercal (Au) = \lambda_{max} w^\intercal u,$$

hence $\eta = \lambda_{max}$ as $w^\intercal u > 0$. If $v' \in \mathbb{R}$ is another eigenvector corresponding to λ_{max} linearly independent with v then there exist $\alpha, \beta \in \mathbb{R}$ such that vector $v'' = \alpha v + \beta v'$ has some component equal to 0 and $v' \geqslant 0, v' \neq 0$. Then

$$0 < Bv' = (I + A)^k v' = (1 + \lambda_{max})^k v',$$

which leads to a contradiction.
Therefore, in the Jordan decomposition of matrix A there exists a unique Jordan block corresponding to the eigenvalue λ_{max}.

Perron–Frobenius Theorem Proof (continued)

Without loss of generality one may replace A by $\frac{A}{\lambda_{max}}$ and assume that $\lambda_{max} = 1$. Recall that

$$\|A\|_\infty = \max\{\|r_1\|_1, \ldots, \|r_n\|\},$$

where $r_1, \ldots, r_n$ denote the rows of matrix A. Therefore (recall $A \geqslant 0, v > 0$)

$$\|v\|_\infty = \|A^m v\|_\infty = \max_{1 \leqslant i \leqslant n} \langle r_i^{(m)}, v_i \rangle \geqslant \max_{1 \leqslant i \leqslant n} \left\| r_i^{(m)} \right\|_1 \min_{1 \leqslant i \leqslant n} v_i =$$

$$= \|A^m\|_\infty \min_{1 \leqslant i \leqslant n} v_i,$$

where $r_i^{(m)}$ denote the rows of A^m. Therefore for any m

$$\|A^m\| \leqslant \frac{\|v\|_\infty}{\min_{1 \leqslant i \leqslant n} v_i}.$$

Perron–Frobenius Theorem Proof (continued)

Let J be the Jordan matrix of A and let

$$J = C^{-1}AC,$$

then

$$\|J^m\|_\infty \leqslant \|C^{-1}\|_\infty \|A^m\|_\infty \|C\|_\infty.$$

If the size of the Jordan block J_1 corresponding to $\lambda_{max} = 1$ is bigger or equal than 2 then

$$J_1^m = \begin{bmatrix} 1 & m & \dots \\ 0 & 1 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix},$$

which gives a contradiction as then

$$\|J^m\|_\infty \geqslant 1 + m \longrightarrow \infty,$$

when $m \longrightarrow \infty$.

Perron–Frobenius Theorem Proof (continued)

Finally, assume that A is primitive. Take λ an eigenvalue of A such that $|\lambda| = \lambda_{max}$. From the first part of the proof it follows that

$$|\lambda| = L(|w|) = \lambda_{max}.$$

The inequality

$$|\lambda||w_i| \leqslant \sum_{i=1}^{n} a_{ij}|w_j|,$$

becomes equality only if all arguments of w_j for non–zero a_{ij} are the same. Applying the same argument to A^k, and dividing w by a unit complex number we get a real, non–negative, non–zero eigenvector corresponding to the eigenvalue λ_{max}. Hence $\lambda = \lambda_{max}$.

Application – Discrete Markov Chains

Let

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

be the transition matrix of some Markov chain (see Lecture 5). The eigenvalues of Q are $\frac{1}{6}, 1$ hence $\lambda_{max} = 1$ (the vector $(1, \ldots, 1)$ is an eigenvector of any transition matrix). Moreover

$$V_{(1)} = \mathrm{lin}((1,1)),$$

$$V_{(\frac{1}{6})} = \mathrm{lin}((3,-2)).$$

$$Q^n = \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{6^n} \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix}^{-1}.$$

Application – Discrete Markov Chains (continued)

$$\lim_{n\to\infty} Q^n = \lim_{n\to\infty} \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{6^n} \end{bmatrix} \begin{bmatrix} \frac{2}{5} & \frac{3}{5} \\ \frac{1}{5} & -\frac{1}{5} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{2}{5} & \frac{3}{5} \\ \frac{1}{5} & -\frac{1}{5} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} \frac{2}{5} & \frac{3}{5} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{2}{5} & \frac{3}{5} \\ \frac{2}{5} & \frac{3}{5} \end{bmatrix}.$$

Therefore for any initial conditions $\mathrm{t} = (t_1, t_2) \in \mathbb{R}^2_{\geqslant 0},\ \ t_1 + t_2 = 1$

$$\lim_{n\to\infty} \mathrm{t}^{\mathsf{T}} Q^n = \left(\frac{2}{5}, \frac{3}{5}\right).$$

High Powers of a Primitive Matrix

Corollary

Let $A \in M(n \times n; \mathbb{R})$ *be a primitive matrix (i.e.* $A > 0$*). Let* $v \in \mathbb{R}^n, v > 0$ *be the (right) eigenvector of* A *for the eigenvalue* λ_{max} *and let* $w \in \mathbb{R}^n, w > 0$ *be the (left) eigenvector of* A *for the eigenvalue* λ_{max} *such that* $w^\intercal v = 1$*. Then*

$$\lim_{n \to \infty} \left(\frac{Q}{\lambda_{max}} \right)^n = vw^\intercal.$$

High Powers of a Primitive Matrix (continued)

Example

For $Q = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$ we have $\lambda_{max} = 1$ and $v = (1, 1)$, $w = \frac{1}{5}(2, 3)$,

i.e.

$$\frac{1}{5}\begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 2 & 3 \end{bmatrix}, \qquad \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

therefore

$$\lim_{n \to \infty} Q^n = vw^{\mathsf{T}} = \frac{1}{5}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \end{bmatrix} = \begin{bmatrix} \frac{2}{5} & \frac{3}{5} \\ \frac{2}{5} & \frac{3}{5} \end{bmatrix}.$$

Graph of a Non–Negative Matrix

Definition

Let $A = [a_{ij}] \in M(n \times n; \mathbb{R})$ be a matrix such that $A \geqslant 0$. The directed graph given by A is a graph $G_A = G = (V, E)$, where

$$V = \{1, 2, \ldots, n\},$$

and for any $i, j \in V$,

$$(i, j) \in E \text{ if and only if } a_{ij} > 0.$$

Graph of a Non–Negative Matrix

Definition

Let $A = [a_{ij}] \in M(n \times n; \mathbb{R})$ be a matrix such that $A \geqslant 0$. The directed graph given by A is a graph $G_A = G = (V, E)$, where

$$V = \{1, 2, \dots, n\},$$

and for any $i, j \in V$,

$$(i, j) \in E \text{ if and only if } a_{ij} > 0.$$

Remark

Note that self–loops are allowed. The matrix G_A is closely related to the adjacency matrix of graph G.

Graph of a Non–Negative Matrix

Definition

Let $A = [a_{ij}] \in M(n \times n; \mathbb{R})$ be a matrix such that $A \geqslant 0$. The directed graph given by A is a graph $G_A = G = (V, E)$, where

$$V = \{1, 2, \ldots, n\},$$

and for any $i, j \in V$,

$$(i, j) \in E \text{ if and only if } a_{ij} > 0.$$

Remark

Note that self–loops are allowed. The matrix G_A is closely related to the adjacency matrix of graph G.

Definition

A directed graph $G = (V, E)$ is **strongly connected** if for each $i, j \in V$ there exists a path joining i with j.

Graph of a Non–Negative Matrix (continued)

Proposition

Let $A = [a_{ij}] \in M(n \times n; \mathbb{R})$ *be a matrix such that* $A \geqslant 0$. *The following conditions are equivalent*

i) *the matrix* A *is irreducible,*

ii) *the graph* G_A *is strongly connected.*

Graph of a Non–Negative Matrix (continued)

Proposition

Let $A = [a_{ij}] \in M(n \times n; \mathbb{R})$ be a matrix such that $A \geqslant 0$. The following conditions are equivalent

i) *the matrix A is irreducible,*

ii) *the graph G_A is strongly connected.*

Proof.

Follows directly from definitions. □

Graph of a Non–Negative Matrix (continued)

Proposition

Let $A = [a_{ij}] \in M(n \times n; \mathbb{R})$ *be a matrix such that* $A \geqslant 0$. *The following conditions are equivalent*

i) *the matrix* A *is primitive,*

ii) *the graph* G_A *is strongly connected and contains two cycles of relatively prime lengths.*

Graph of a Non–Negative Matrix (continued)

Proposition

Let $A = [a_{ij}] \in M(n \times n; \mathbb{R})$ be a matrix such that $A \geqslant 0$. The following conditions are equivalent

i) *the matrix A is primitive,*

ii) *the graph G_A is strongly connected and contains two cycles of relatively prime lengths.*

Proof.

$i) \Rightarrow ii)$ let k be a number such that $A^k > 0$. Then $A^{k+1} > 0$ so there are cycles of lengths k and $k+1$,

Graph of a Non–Negative Matrix (continued)

Proposition

Let $A = [a_{ij}] \in M(n \times n; \mathbb{R})$ be a matrix such that $A \geqslant 0$. The following conditions are equivalent

i) *the matrix A is primitive,*

ii) *the graph G_A is strongly connected and contains two cycles of relatively prime lengths.*

Proof.

$i) \Rightarrow ii)$ let k be a number such that $A^k > 0$. Then $A^{k+1} > 0$ so there are cycles of lengths k and $k+1$,
$ii) \Rightarrow i)$ see S. Sternberg *Dynamical Systems*, Section 9.2, the problem reduces to a statement from arithmetic: if $\mathrm{GCD}(a,b) = 1$ then there exists a $N \in \mathbb{N}$ such that

$$(\mathbb{N}a + \mathbb{N}b) \cap [N, +\infty) = \{N, N+1, N+2, \ldots\}.$$

$\square$

Example – Irreducible Not Primitive Matrix

$$
\begin{tikzcd}[column sep=small]
 & & 1 \arrow[dll] & & \\
2 \arrow[dr] & & & & 5 \arrow[ull] \\
 & 3 \arrow[rr] & & 4 \arrow[ur] &
\end{tikzcd}
$$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \ldots$$

$$A^5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = I_5, A^6 = A.$$

Example – Irreducible Not Primitive Matrix

In particular, if

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

then

$$w_A(\lambda) = \lambda^5 - 1,$$

hence $\lambda_{max} = 1$, any other eigenvalue λ of matrix A is a $5-$th root of unity and

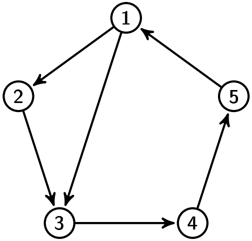
$$|\lambda| \leqslant \lambda_{max},$$

moreover λ_{max} has algebraic multiplicity 1 and $Av = v$, where

$$v = (1, 1, 1, 1, 1) > 0.$$

Incidentally, A is a particular case of a 5×5 circulant matrix with $c_{n-1} = 1,\ n = 5$ and all other c_i's equal to 0.

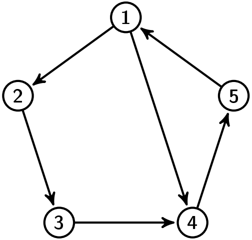
Example – Primitive Matrix



$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}, \ldots$$

$$A^{17} = \begin{bmatrix} 4 & 1 & 2 & 4 & 6 \\ 3 & 1 & 1 & 1 & 3 \\ 3 & 3 & 4 & 1 & 1 \\ 1 & 3 & 6 & 4 & 1 \\ 1 & 1 & 4 & 6 & 4 \end{bmatrix}.$$

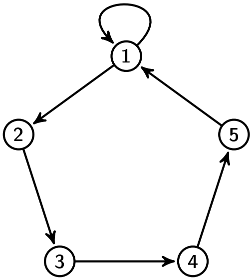
Example – Another Primitive Matrix



$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}, \ldots$$

$$A^{14} = \begin{bmatrix} 4 & 3 & 1 & 6 & 2 \\ 1 & 1 & 2 & 1 & 3 \\ 3 & 1 & 1 & 3 & 1 \\ 1 & 3 & 1 & 4 & 3 \\ 3 & 1 & 3 & 2 & 4 \end{bmatrix}.$$

Example – And Another Primitive Matrix



$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}, \ldots.$$

$$A^8 = \begin{bmatrix} 5 & 4 & 3 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 1 & 1 \\ 4 & 3 & 2 & 1 & 1 \end{bmatrix}.$$

Example – And Another Primitive Matrix (continued)

If

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

it can be checked that $w_A(\lambda) = \lambda^5 - \lambda^4 - 1 = (\lambda^2 - \lambda + 1)(\lambda^3 - \lambda - 1)$, with $\lambda_{max} \approx 1.3247$, and $v \approx (0.6765, 0.2197, 0.2910, 0.3855, 0.5107)$. Other eigenvalues have magnitudes smaller than λ_{max}