ALGEBRA
5. DIAGONALIZATION

Grau de Matemàtiques. FME
2018-2019

Marta Casanellas
Departament de Matemàtiques

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

OUTLINE

1. EIGENVALUES AND EIGENVECTORS
2. DIAGONALIZATION THEOREM
3. ANNIHILATOR POLYNOMIALS
4. APPLICATIONS
5. BIBLIOGRAPHY

Outline

1. Eigenvalues and Eigenvectors
2. Diagonalization theorem
3. Annihilator polynomials
4. Applications
5. Bibliography

Motivation

Throughout the chapter we work on a $\mathbb{K}$-e.v. E of finite dimension n.

- Goal: **compute powers of matrices**.
- If M is a diagonal matrix, then it is easy to compute M^m for any $m \in \mathbb{N}$.
- If M is not diagonal, is there a change of basis that converts it to a diagonal matrix?

Definition

We say that an endomorphism $f : E \to E$ is diagonalizable in $\mathbb{K}$ if there exists a basis $\mathbf{v}$ of E such that $M_{\mathbf{v}}(f)$ is a diagonal matrix $D \in \mathcal{M}_n(\mathbb{K})$.

In other words, f is diagonalizable in $\mathbb{K}$ if there exists an invertible matrix $P \in \mathcal{M}_n(\mathbb{K})$ such that

$$P^{-1} M_{\mathbf{e}}(f) P$$

is a diagonal matrix (P can be thought as a change of basis matrix).

- An $n \times n$ matrix diagonalizes if there exists an invertible matrix P such that $P^{-1}MP$ is a diagonal matrix.
- If M diagonalizes, then $M = PDP^{-1}$ for a certain diagonal matrix D. Hence, M^m can be easily computed:

$$M^m = PDP^{-1}PDP^{-1}\ldots PDP^{-1}PDP^{-1} = PD^mP^{-1}.$$

Remark: If $M_{\mathbf{v}}(f)$ is diagonal, then $f(v_i) = d_i v_i$ ($d_i = i$th value in the diagonal).

DEFINITION

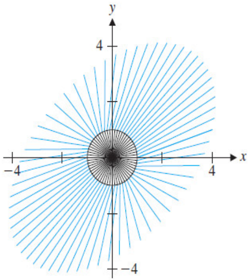
Let $f \in End(E)$. A vector $u \neq 0 \in E$ is an eigenvector (VEP) of f if $f(u) = \lambda\, u$ for some $\lambda \in \mathbb{K}$. In this case, we say that λ is an eigenvalue (VAP) of f and that u is an eigenvector with eigenvalue λ.

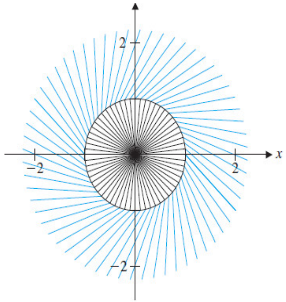
EXAMPLE

Consider the endomorphism of $\mathbb{R}^2$ given by $f(x, y) = (5x, 2y)$. Then, $e_1 = (1, 0)$ is an eigenvector of f with eigenvalue 1, and $e_2 = (0, 1)$ is an eigenvector of f with eigenvalue 2.
The standard matrix of f is

$$M_e(f) = \begin{pmatrix} 5 & 0 \\ 0 & 2 \end{pmatrix}.$$

In black v; in **blue** $f(v)$.





[D.Poole]

A map with eigenvectors $[(1,1)]$ and $[(1,-1)]$

A map with no eigenvectors.

Let f be an endomorphism of E.

LEMMA

- *$u \in E$ is a VEP of VAP $\lambda \in \mathbb{K} \Leftrightarrow u \in \mathrm{Nuc}(f - \lambda Id)$ and $u \neq 0$.*
- *$\lambda \in \mathbb{K}$ is an eigenvalue of $f \Leftrightarrow \det(f - \lambda Id) = 0$.*

DEFINITION

If λ is a VAP of f, we call E_λ the subspace $\mathrm{Nuc}(f - \lambda Id) \subseteq E$. This is the subspace formed by all VEP's of VAP λ plus $\mathbf{0}$.

Remark: A vector $u \neq 0$ is a VEP with VAP $0 \Leftrightarrow u \in \mathrm{Nuc}(f)$.
The spectrum of f is the set of all its eigenvalues and is denoted by $\sigma(M)$.

DEFINITION

The characteristic polynomial of $A \in \mathcal{M}_n(\mathbb{K})$ is

$$P_A(x) := det(A - xId) = \begin{vmatrix} a_{1,1} - x & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} - x & \dots & a_{2,n} \\ \dots & \dots & & \dots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} - x \end{vmatrix}.$$

If $f \in End(E)$, the characteristic polynomial of f is $p_A(x)$ where $A = M_{\mathbf{u}}(f)$ for some basis $\mathbf{u}$.

Properties

Proposition

1. *$P_f(x)$ does not depend on the basis **u** chosen.*
2. *$P_f(x)$ is a polynomial of degree n,* $P_f(x) = c_n x^n + c_{n-1}x^{n-1} + \ldots + c_1 x + c_0$. *Moreover, if* $M_{\mathbf{u}}(f) = (a_{i,j})$, $\Rightarrow$
$$c_n = (-1)^n,$$
$$c_{n-1} = (-1)^{n-1}\,\mathrm{tr}(f) = (-1)^{n-1}(a_{1,1} + \ldots + a_{n,n}),$$
$$c_0 = \det(f).$$
3. *In particular,* det *and* tr *are the same for any matrix of the endomorphism f.*
4. *The roots of $P_f(x)$ are the eigenvalues of f, that is,* $\lambda \in \mathbb{K}$ *is an eigenvalue of f* $\Leftrightarrow P_f(\lambda) = 0$.

Algebraic and geometric multiplicity of an eigenvalue

Definition

If λ is an eigenvalue of f, the algebraic multiplicity of λ , denoted by a_λ, is the multiplicity as a root of $P_f(x)$.

Definition

The geometric multiplicity of λ, denoted by g_λ, is the dimension of the vector subspace $\text{Nuc}(f - \lambda Id)$, that is, $n - rk(A - \lambda I)$.

Proposition

For every eigenvalue λ, we have $1 \leq g_\lambda \leq a_\lambda$.

Outline

1. Eigenvalues and Eigenvectors
2. Diagonalization theorem
3. Annihilator polynomials
4. Applications
5. Bibliography

Let $f \in End(E)$. Then,

LEMMA

- *If u, v are VEP's of different VAP's $\Rightarrow u, v$ are l.i.*
- *If $\lambda_1, \ldots, \lambda_r$ are different VAP's $\Rightarrow$ the sum $E_{\lambda_1} + \ldots + E_{\lambda_r}$ is a direct sum,*

$$E_{\lambda_1} + \ldots + E_{\lambda_r} = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_r}$$

COROLLARY

If all the roots of $P_f(x)$ are in $\mathbb{K}$ and simple ($a_{\lambda_i} = 1$ for each λ_i), then f diagonalizes.

Diagonalization theorem

Theorem (Diagonalization)

An endomorphism f of E is diagonalizable in $\mathbb{K}$ if and only if

1. *$P_f(x)$ has all its roots $\lambda_1, \ldots, \lambda_k$ in $\mathbb{K}$ (P_f fully decomposes in $\mathbb{K}$) and*
2. *for every VAP λ_i, the algebraic multiplicity and geometric multiplicity are equal: $g_{\lambda_i} = a_{\lambda_i}$.*

If it diagonalizes, it does so in a basis of VEP's.

Procedure to diagonalize an endomorphism

Given an endomorphism f of $\mathbb{R}^n$, let A be its standard matrix.

1. Compute the characteristic polynomial $P_f(x) = det(A - x\,Id)$.
2. Compute the roots $\lambda_1, \ldots, \lambda_k$ of $P_f(x)$: if some $\lambda_i \notin \mathbb{R} \Rightarrow f$ does not diagonalize in $\mathbb{R}$. Otherwise,
3. For each eigenvalue λ_i, compute its algebraic multiplicity as a root of $P_f(x)$, a_{λ_i}.
4. For each λ_i, compute $\mathrm{Nuc}(A - \lambda_i Id)$: this is the set of all eigenvectors of f with eigenvalue λ_i. The dimension of this space is the geometric multiplicity g_{λ_i} of λ_i.
5. If $\lambda_i \in \mathbb{R}$ and $g_{\lambda_i} = a_{\lambda_i}$ for each eigenvalue λ_i, then f diagonalizes.

In this case, for each eigenvalue λ_i, let $\{v_1^i, \dots, v_{a_{\lambda_i}}^i\}$ be a basis for $\mathrm{Nuc}(A - \lambda_i\, Id)$. Then,

1. $\mathbf{v} = \bigcup_{i=1}^{k} \{v_1^i, \dots, v_{a_{\lambda_i}}^i\}$ is a basis of $\mathbb{R}^n$.
2. $M_{\mathbf{v}}(f)$ is a diagonal matrix:

$$M_{\mathbf{v}}(f) = D = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_k \end{pmatrix}.$$

Recall that $M_{\mathbf{v}}(f)$ can be computed by doing a change of basis: if $\mathbf{e}$ is the standard basis of $\mathbb{R}^n$ and , then

$$A_{\mathbf{e}\to\mathbf{v}}\, A\, A_{\mathbf{v}\to\mathbf{e}} = D.$$

$$\text{(Equivalently, } A_{\mathbf{v}\to\mathbf{e}}\, D\, A_{\mathbf{e}\to\mathbf{v}} = A\text{).}$$

Real matrices with complex eigenvalues

If we allow to work in $\mathbb{K} = \mathbb{C}$ then,

Theorem

An endomorfism f is diagonalizable ***in*** $\mathbb{C}$ *if and only if*

- *for every VAP* λ_i*, the algebraic multiplicity and geometric multiplicity are equal:* $g_{\lambda_i} = a_{\lambda_i}$.

If A is a **real** matrix and we allow diagonalization in $\mathbb{K} = \mathbb{C}$, then VAP's and VEP's go "in pairs":

- $p_A(x) \in \mathbb{R}[x] \Rightarrow \lambda$ is a VAP of A if and only if $\bar{\lambda}$ is also a VAP.
- If $\lambda \in \mathbb{C} \setminus \mathbb{R}$, and $w \in \mathbb{C}^n$ is a VEP with VAP λ, then we can write $w = u + iv$ where $u, v \in \mathbb{R}^n$.
- Then, $\bar{w} := u - iv$ is an eigenvector with eigenvalue $\bar{\lambda}$.

If we want to work only in $\mathbb{R}$, the one can rearrange complex VAP's and VEP's in conjugate pairs to obtain a "diagonalization" of A in 2×2 blocks.

Outline

1 Eigenvalues and Eigenvectors

2 Diagonalization theorem

3 Annihilator polynomials

4 Applications

5 Bibliography

Lemma

Let $f \in End(E)$. Then, there exists a basis $\mathbf{u}$ in which $M_{\mathbf{u}}$ is triangular if and only if $P_f(x)$ fully decomposes in $\mathbb{K}$.

In this case the triangular matrix has the VAP's $\lambda_1, \ldots, \lambda_n$ (repeated if necessary) of f in the diagonal and

$$\det(A) = \lambda_1 \ldots \lambda_n,$$

$$tr(A) = \lambda_1 + \ldots + \lambda_n.$$

Theorem (First decomposition)

Let f in $End(E)$ such that

$$P_f(x) = (-1)^n (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\dots(x-\lambda_r)^{m_r}.$$

Then $\dim \mathrm{Nuc}(f-\lambda_i Id)^{m_i} = m_i$ *for each i and*

$$E = \mathrm{Nuc}(f-\lambda_1 Id)^{m_1} \oplus \dots \oplus \mathrm{Nuc}(f-\lambda_r Id)^{m_r}.$$

Lemma

Let $f \in End(E)$, $\lambda \neq \mu \in \mathbb{K}$ and let $k \in \mathbb{N}_{>0}$. Then,

- *the endomorphisms $f-\lambda Id$ and $f-\mu Id$ commute,*
- $V = \mathrm{Nuc}(f-\lambda Id)^k$ *is invariant by $f-\mu Id$*
- *$f-\mu Id$ restricted to V is bijective.*

Cayley-Hamilton theorem

Definition

A polynomial $q(x)$ is an annihilator polynomial for f if the endomorphism $q(f)$ is 0 (equivalently, $q(A) = 0$ if A is the matrix of f in any basis).

Theorem (Cayley-Hamilton)

The characteristic polynomial $P_f(x)$ of f is an annihilator polynomial for f, $p_f(f) = 0$.

Consequences: if A is the standard matrix of $f \in \mathbb{R}^n$ and $P_f(x) = a_0 + a_1 x + \ldots + a_n x^n$, then

- $a_0 Id + a_1 A + \ldots + a_n A^n = 0$
- A^n can be computed as a linear combination of $Id, A, A^2, \ldots, A^{n-1}$
- If A is invertible $\Rightarrow A^{-1}$ can be computed as a linear combination of $Id, A, \ldots, A^{n-1}$

DEFINITION

There exists a monic annihilator polynomial $m_f(x)$ such that any other annihilator polynomial for f is multiple of it. This polynomial is called the minimal annihilator polynomial.

Properties of $m_f(X)$:

- $m_f(X)$ is (up to a scalar) the annihilator polynomial of smallest defree for f.
- $m_f(X)|P_f(X)$
- $\lambda \in \mathbb{K}$ is a zero of $P_f(x)$ if and only if it is a zero of $m_f(X)$.
- If $P_f(x) = (-1)^n(x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\dots(x-\lambda_r)^{m_r} \Rightarrow$

$$m_f(x) = (-1)^n(x-\lambda_1)^{n_1}(x-\lambda_2)^{n_2}\dots(x-\lambda_r)^{n_r} \quad \text{with} \quad n_i \leq m_i.$$

- How do we find n_i? n_i is the smallest exponent k such that

$$\{\mathbf{0}\} \subsetneq \mathrm{Nuc}(f-\lambda_i) \subsetneq \cdots \subsetneq \mathrm{Nuc}(f-\lambda_i)^k = \mathrm{Nuc}(f-\lambda_i)^{k+1} = \dots.$$

Outline

1 Eigenvalues and Eigenvectors

2 Diagonalization theorem

3 Annihilator polynomials

4 Applications

5 Bibliography

If a matrix does not diagonalize in $\mathbb{C}$, one can do a change of basis to obtain a **triangular** matrix and one can reduce it even more to obtain its **Jordan canonical form**; that is, a block-diagonal matrix formed by blocks of type:

$$\begin{pmatrix} \lambda & 0 & \dots & 0 \\ 1 & \lambda & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \vdots & \\ 0 & \dots & 1 & \lambda \end{pmatrix}$$

In the exercise list there are some examples of how to obtain a Jordan form.

Let $A \in \mathcal{M}_n(\mathbb{R})$, with $A = PDP^{-1}$, $D = diag(\lambda_1, \dots, \lambda_n)$, $P =$ change-of-basis $= A_{\mathbf{v} \to \mathbf{e}}$, $\mathbf{v} = \{v_1, \dots, v_n\}$. Then,

- $A^k = PD^k P^{-1}$.
- If $x = c_1 v_1 + \cdots + c_n v_n \Rightarrow A^k x = c_1 \lambda_1^k v_1 + \cdots + c_n \lambda_n^k v_n$
- If $|\lambda_1| > |\lambda_i|$, then λ_1^k grows faster than λ_i^k so, if $c_1 \neq 0$,

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

- This is the technical basis to efficiently compute VAPs (and VEPs) $\rightsquigarrow$ Algebra Lineal Numèrica (Q2).

Definition

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

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

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

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

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

Lemma

The solutions to $\mathbf{x}(k+1) = A\,\mathbf{x}(k)$ are $\{\mathbf{x}(k) = A^k\mathbf{x}(0)\}_k$.

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.
If A is a stochastic matrix we have:

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

We say that a matrix A is positive if all its entries are > 0.

Theorem (Perron-Frobenius theorem)

Let A be a positive $n \times n$ matrix then A has a real eigenvalue λ_1 with the following properties:

- $\lambda_1 > 0$
- $\lambda_1 > |\lambda|$ *for any other eigenvalue* λ *(*λ_1 *is dominant).*
- λ_1 *has algebraic multiplicity 1.*
- λ_1 *has a corresponding positive eigenvector,* v*,* $Null(A - \lambda_1 Id) = [v]$*.*
- *There are no other positive eigenvectors for other eigenvalues.*

This theorem also holds for certain types of non-negative matrices.

Perron-Frobenius for stochastic matrices

Theorem

If A is a stochastic matrix, then 1 is VAP and, if A is positive,

- *$1 > |\lambda|$ for any other VAP λ*
- *1 has a positive VEP v.*
- *no other VAP has positive eigenvectors.*
- *If we take v to **sum to 1**, then we have*

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

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

*for **any** positive vector $\boldsymbol{x}$ that sums to 1.*

$\rightsquigarrow$ *This is the theory behind Google Pagerank algorithm.*

If f is a diagonalizable endomorphism with standard matrix $A \in M_n(\mathbb{R})$, so that $A = S\ D\ S^{-1}$, where D diagonal: $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, $S = A_{\mathbf{v}\to\mathbf{e}}$, and $\mathbf{v} = \{v_1, \dots, v_n\}$ is the corresponding basis of eigenvectors. Then, we define the exponential of the matrix A:

$$e^A = S\ e^D\ S^{-1} = S \begin{pmatrix} e^{\lambda_1} & 0 & \dots & 0 \\ 0 & e^{\lambda_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda_n} \end{pmatrix} S^{-1}$$

and this coincides with $\sum_{n\geq 0} \frac{A^n}{n!}$.

Outline

1. Eigenvalues and Eigenvectors
2. Diagonalization theorem
3. Annihilator polynomials
4. Applications
5. Bibliography

Bibliography

- Poole D., Linear Algebra A Modern Introduction, Brooks Cole, 2011
- S. Friedberg, A. Insel, L. Spence, Linear Algebra, Pearson Education 2014
- Castellet, LLerena, Algebra lineal i geometria, Edicions UAB, 2000.