Canonical Decompositions

Richard Hollister

Western Michigan University

20 Sep. 2018

What do we mean by "canonical"?

Believe it or not, the story begins with everyone's favorite axiom of set theory: the Axiom of Choice.

Axiom of Choice

Let $\mathcal{C}$ be a nonempty collection of nonempty sets. Then there is a function $f : \mathcal{C} \to \cup_{A \in \mathcal{C}} A$ with

$$f(A) \in A \quad \forall A \in \mathcal{C}.$$

This function is called a choice function, and the function value $f(A)$ is called the canonical representative of A.

Uniqueness of the canonical representative

Canonical representatives are unique!

There are some natural questions to ask:

- What is the collection $\mathcal{C}$?
- How do we find a choice function?
- How do we find a choice function so that the canonical representative has some nice properties?

What is the collection $\mathcal{C}$?

Let X be a set, and $\sim$ be an equivalence relation on X. Then

- $\sim$ partitions X into equivalence classes,
- $\mathcal{C}$ is the collection of equivalence classes, and
- the choice function is now invertible.

In most cases when canonical representatives are studied, the collection $\mathcal{C}$ is produced in this way.

We are now looking for unique representatives of equivalence classes.

Matrices under unitary equivalence

For our first example, let $X = M_{m,n}(\mathbb{C})$ with $m \leq n$, and let $\sim$ be defined by

$$A \sim B \text{ iff } \exists U \in O_m(\mathbb{C}) \text{ and } \exists V \in O_n(\mathbb{C}) \text{ s.t. } A = UBV.$$

Recall that $O_k(\mathbb{C})$ is the group of complex, orthogonal (i.e. unitary) matrices of size $k \times k$.

Matrices under unitary equivalence

For our first example, let $X = M_{m,n}(\mathbb{C})$ with $m \leq n$, and let $\sim$ be defined by

$$A \sim B \text{ iff } \exists U \in O_m(\mathbb{C}) \text{ and } \exists V \in O_n(\mathbb{C}) \text{ s.t. } A = UBV.$$

Recall that $O_k(\mathbb{C})$ is the group of complex, orthogonal (i.e. unitary) matrices of size $k \times k$.

This relation is called *unitary equivalence*, and the canonical representatives are real diagonal matrices $(m \times n)$

$$\Sigma = \operatorname{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_m\}$$

such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m \geq 0$.

Singular value decomposition

Matrix decompositions showcase canonical representatives.

Theorem

For every matrix $A \in M_{m,n}(\mathbb{C})$, there are unitary matrices $U \in O_m(\mathbb{C})$ and $V \in O_n(\mathbb{C})$ and a unique diagonal matrix Σ (as described on previous slide) such that

$$A = U\Sigma V.$$

Square matrices under similarity

Definition

Two square matrices A and B are similar if there is a nonsingular matrix S such that

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

This defines an equivalence relation.

Square matrices under similarity

Definition

Two square matrices A and B are similar if there is a nonsingular matrix S such that

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

This defines an equivalence relation.

Possible canonical representatives:

- diagonal matrices,
- uppertriangular matrices, or
- something else.

Diagonal matrices

Recall the eigenvalue decomposition:

$$A = S\Lambda S^{-1}$$

where $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ (Unique?).

Diagonal matrices

Recall the eigenvalue decomposition:

$$A = S\Lambda S^{-1}$$

where $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ (Unique?). However, this decomposition is not always possible. For example, the matrix

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

cannot be diagonalized.

Uppertriangular matrices

There is also Schur's triangularization theorem.

Theorem

Every square, complex matrix $A \in M_n(\mathbb{C})$ is unitarily similar to an uppertriangular matrix with its eigenvalues displayed on the diagonal.

Uppertriangular matrices

There is also Schur's triangularization theorem.

Theorem

Every square, complex matrix $A \in M_n(\mathbb{C})$ is unitarily similar to an uppertriangular matrix with its eigenvalues displayed on the diagonal.

The diagonal is unique up to the order of the eigenvalues, but the rest is not unique at all.

Uppertriangular matrices

There is also Schur's triangularization theorem.

Theorem

Every square, complex matrix $A \in M_n(\mathbb{C})$ is unitarily similar to an uppertriangular matrix with its eigenvalues displayed on the diagonal.

The diagonal is unique up to the order of the eigenvalues, but the rest is not unique at all.

We are 0 for 2.

Jordan matrices

Diagonal was asking too much, and upper triangular was asking too little. We are in a real Goldielocks situation.

Jordan matrices

Diagonal was asking too much, and upper triangular was asking too little. We are in a real Goldielocks situation.

The answer comes in the form of Jordan blocks:

$$J_k(\lambda) := \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}_{k \times k} .$$

Jordan matrices

Diagonal was asking too much, and upper triangular was asking too little. We are in a real Goldielocks situation.

The answer comes in the form of Jordan blocks:

$$J_k(\lambda) := \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}_{k\times k} .$$

This matrix has eigenvalue λ with algebraic multiplicity k and geometric multiplicity 1.

Jordan decomposition

Theorem

Every square, complex matrix is similar to a direct sum of Jordan blocks.

Jordan decomposition

Theorem

Every square, complex matrix is similar to a direct sum of Jordan blocks.

For $A \in M_n(\mathbb{C})$here is nonsingular S and

$$J = J_{k_1}(\lambda_1) \oplus J_{k_2}(\lambda_2) \oplus \cdots \oplus J_{k_\ell}(\lambda_\ell)$$

such that

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

Jordan decomposition

Theorem

Every square, complex matrix is similar to a direct sum of Jordan blocks.

For $A \in M_n(\mathbb{C})$here is nonsingular S and

$$J = J_{k_1}(\lambda_1) \oplus J_{k_2}(\lambda_2) \oplus \cdots \oplus J_{k_\ell}(\lambda_\ell)$$

such that

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

Uniqueness?

Matrix pencils

The traditional eigenvalue problem is

$$(A - \lambda I)x = 0.$$

Matrix pencils

The traditional eigenvalue problem is

$$(A - \lambda I)x = 0.$$

Replace the $-I$ with another matrix B:

$$(A + \lambda B)x = 0.$$

Matrix pencils

The traditional eigenvalue problem is

$$(A - \lambda I)x = 0.$$

Replace the $-I$ with another matrix B:

$$(A + \lambda B)x = 0.$$

This is the generalized eigenvalue problem.

Matrix pencils

The traditional eigenvalue problem is

$$(A - \lambda I)x = 0.$$

Replace the $-I$ with another matrix B:

$$(A + \lambda B)x = 0.$$

This is the generalized eigenvalue problem.

Expressions of the form $A + \lambda B$ are called *matrix pencils*.

Equivalence of matrix pencils

Definition

Two matrix pencils $A + \lambda B$ and $C + \lambda D$ are *equivalent* if there are nonsingular matrices R and S such that

$$R(A + \lambda B)S = C + \lambda D.$$

Equivalence of matrix pencils

Definition

Two matrix pencils $A + \lambda B$ and $C + \lambda D$ are *equivalent* if there are nonsingular matrices R and S such that

$$R(A + \lambda B)S = C + \lambda D.$$

This is

- an equivalence relation,
- denoted $A + \lambda B \approx C + \lambda D$,
- sometimes called "strict" equivalence.

If B is nonsingular

If B is nonsingular, then

$$(A + \lambda B) \approx (AB^{-1} + \lambda I),$$

If B is nonsingular

If B is nonsingular, then

$$(A + \lambda B) \approx (AB^{-1} + \lambda I),$$

but since AB^{-1} is similar to its Jordan form J,

$$(AB^{-1} + \lambda I) \approx (J + \lambda I).$$

If B is nonsingular

If B is nonsingular, then

$$(A + \lambda B) \approx (AB^{-1} + \lambda I),$$

but since AB^{-1} is similar to its Jordan form J,

$$(AB^{-1} + \lambda I) \approx (J + \lambda I).$$

The last pencil $J + \lambda I$ is a direct sum of blocks of the form

$$\begin{bmatrix} \lambda_i + \lambda & 1 & & \\ & \lambda_i + \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i + \lambda \end{bmatrix}.$$

These are just pencil versions of Jordan blocks.

Regular pencils

This is an example of a regular matrix pencil.

Definition

A matrix pencil is *regular* if it's determinant is not identically the zero function.

Regular pencils

This is an example of a regular matrix pencil.

Definition

A matrix pencil is *regular* if it's determinant is not identically the zero function.

Is every regular matrix pencils is equivalent to a direct sum of Jordan blocks (pencil version)?

Regular pencils

This is an example of a regular matrix pencil.

Definition

A matrix pencil is *regular* if it's determinant is not identically the zero function.

Is every regular matrix pencils is equivalent to a direct sum of Jordan blocks (pencil version)?

Unfortunately not.

Regular pencils

This is an example of a regular matrix pencil.

Definition

A matrix pencil is *regular* if it's determinant is not identically the zero function.

Is every regular matrix pencils is equivalent to a direct sum of Jordan blocks (pencil version)?

Unfortunately not. For example

$$L(\lambda) := \begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix}$$

has determinant 1, but the lead coefficient is singular.

Not enough eigenvalues!

This 2-by-2 example doesn't have any *finite* eigenvalues.

Not enough eigenvalues!

This 2-by-2 example doesn't have any *finite* eigenvalues. Let

$$L_n(\lambda) := \begin{bmatrix} 1 + \frac{\lambda}{n} & \lambda \\ 0 & 1 + \frac{\lambda}{n} \end{bmatrix},$$

and notice that $L(\lambda) = \lim_{n \to \infty} L_n(\lambda)$.

Not enough eigenvalues!

This 2-by-2 example doesn't have any *finite* eigenvalues. Let

$$L_n(\lambda) := \begin{bmatrix} 1 + \frac{\lambda}{n} & \lambda \\ 0 & 1 + \frac{\lambda}{n} \end{bmatrix},$$

and notice that $L(\lambda) = \lim_{n \to \infty} L_n(\lambda)$.

Each $L_n(\lambda)$ has eigenvalue n with multiplicity 2.

Not enough eigenvalues!

This 2-by-2 example doesn't have any *finite* eigenvalues. Let

$$L_n(\lambda) := \begin{bmatrix} 1 + \frac{\lambda}{n} & \lambda \\ 0 & 1 + \frac{\lambda}{n} \end{bmatrix},$$

and notice that $L(\lambda) = \lim_{n\to\infty} L_n(\lambda)$.

Each $L_n(\lambda)$ has eigenvalue n with multiplicity 2.

ANALYSIS

Not enough eigenvalues!

This 2-by-2 example doesn't have any *finite* eigenvalues. Let

$$L_n(\lambda) := \begin{bmatrix} 1 + \frac{\lambda}{n} & \lambda \\ 0 & 1 + \frac{\lambda}{n} \end{bmatrix},$$

and notice that $L(\lambda) = \lim_{n\to\infty} L_n(\lambda)$.

Each $L_n(\lambda)$ has eigenvalue n with multiplicity 2.

ANALYSIS

The pencil $L(\lambda)$ has an eigenvalue at ∞ with multiplicity 2.

Eigenvalues at ∞

Regular matrix pencils can have eigenvalues at ∞ if the lead coefficient is singular!

Eigenvalues at ∞

Regular matrix pencils can have eigenvalues at ∞ if the lead coefficient is singular!

We have a new type of block:

$$\begin{bmatrix} 1 & \lambda & & \\ & 1 & \ddots & \\ & & \ddots & \lambda \\ & & & 1 \end{bmatrix}_{k \times k} .$$

Eigenvalues at ∞

Regular matrix pencils can have eigenvalues at ∞ if the lead coefficient is singular!

We have a new type of block:

$$\begin{bmatrix} 1 & \lambda & & \\ & 1 & \ddots & \\ & & \ddots & \lambda \\ & & & 1 \end{bmatrix}_{k \times k} .$$

This block has an eigenvalue at ∞ with algebraic multiplicity k and geometric multiplicity 1.

Weierstrass canonical form

Theorem

Every regular matrix pencil over $\mathbb{C}$ is equivalent to a direct sum of Jordan blocks (pencil version) and Weierstrass blocks (as on previous slide). The number and sizes of these blocks are uniquely determined by the matrix pencil.

Jordan blocks:

$$\begin{bmatrix} \lambda_i + \lambda & 1 & & \\ & \lambda_i + \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i + \lambda \end{bmatrix}.$$

Weierstrass blocks:

$$\begin{bmatrix} 1 & \lambda & & \\ & 1 & \ddots & \\ & & \ddots & \lambda \\ & & & 1 \end{bmatrix}.$$

Singular matrix pencils

A matrix pencil is *singular* if it has a nontrivial left or right nullsapce.

This happens if:

- its determinant is the zero function (for square pencils), or
- it's rectangular.

Minimal bases

The nullspace of a pencil is a subspace of *rational functions*.

This space has a polynomial basis.

1. Start with any rational basis.
2. Clear the denominators.
3. Now you have a polynomial basis.

Minimal bases

The nullspace of a pencil is a subspace of *rational functions*.

This space has a polynomial basis.

1. Start with any rational basis.
2. Clear the denominators.
3. Now you have a polynomial basis.

Definition

A *minimal basis* is a polynomial basis of *minimal order (sum of vector degrees)* amongst all polynomial bases.

Minimal indices

Definition

The *minimal indices* are the degrees of the basis vectors of a minimal basis.

Minimal indices

Definition

The *minimal indices* are the degrees of the basis vectors of a minimal basis.

This list of non-negative integers is uniquely determined by the rational subspace.

Minimal indices

Definition

The *minimal indices* are the degrees of the basis vectors of a minimal basis.

This list of non-negative integers is uniquely determined by the rational subspace.

The right(left) minimal indices of a pencil are the minimal indices of the right(left) nullspace.

Minimal indices

Consider the pencil

$$\begin{bmatrix} 1 & \lambda & & & \\ & 1 & \lambda & & \\ & & \ddots & \ddots & \\ & & & 1 & \lambda \end{bmatrix}_{k\times(k+1)} .$$

Minimal indices

Consider the pencil

$$\begin{bmatrix} 1 & \lambda & & & \\ & 1 & \lambda & & \\ & & \ddots & \ddots & \\ & & & 1 & \lambda \end{bmatrix}_{k\times(k+1)} .$$

Its right nullspace is spanned by the vector

$$\begin{bmatrix} 1 & -\lambda & \lambda^2 & \cdots & (-1)^k\lambda^k \end{bmatrix}^T .$$

Minimal indices

Consider the pencil

$$\begin{bmatrix} 1 & \lambda & & & \\ & 1 & \lambda & & \\ & & \ddots & \ddots & \\ & & & 1 & \lambda \end{bmatrix}_{k\times(k+1)} .$$

Its right nullspace is spanned by the vector

$$\begin{bmatrix} 1 & -\lambda & \lambda^2 & \cdots & (-1)^k\lambda^k \end{bmatrix}^T .$$

This is a minimal basis, so the pencil has a right minimal index of k.

Minimal indices

Consider the pencil

$$\begin{bmatrix} 1 & \lambda & & & \\ & 1 & \lambda & & \\ & & \ddots & \ddots & \\ & & & 1 & \lambda \end{bmatrix}_{k\times(k+1)} .$$

Its right nullspace is spanned by the vector

$$\begin{bmatrix} 1 & -\lambda & \lambda^2 & \cdots & (-1)^k\lambda^k \end{bmatrix}^T .$$

This is a minimal basis, so the pencil has a right minimal index of k.

We will call these Kronecker blocks.

Kronecker canonical form

Theorem

Every matrix pencil over $\mathbb{C}$ is equivalent to a direct sum of

1. *Jordan blocks–one for each finite elementary divisor,*
2. *Weierstrass blocks–one for each infinite elementary divisor,*
3. *Kronecker blocks–one for each right minimal index, and*
4. *transpose Kronecker blocks–one for each left minimal index.*

Moreover, the number and sizes of these blocks are all uniquely determined by the matrix pencil.

Matrix polynomials

Definition

A *matrix polynomial* is an expression of the form

$$A_0 + A_1\lambda + \cdots + A_d\lambda^d.$$

Matrix polynomials

Definition

A *matrix polynomial* is an expression of the form

$$A_0 + A_1\lambda + \cdots + A_d\lambda^d.$$

Like matrix pencils, matrix polynomials can have

1. finite eigenvalues,
2. infinite eigenvalues,
3. right minimal indices, and
4. left minimal indices.

Unimodular equivalence

Definition

A square matrix polynomial is *unimodular* if its determinant is a nonzero constant.

Unimodular polynomial matrices have polynomial inverse.

Unimodular equivalence

Definition

A square matrix polynomial is *unimodular* if its determinant is a nonzero constant.

Unimodular polynomial matrices have polynomial inverse.

Definition

Matrix polynomials $P(\lambda)$ and $Q(\lambda)$ are *unimodularly equivalent* if there are unimodular $E(\lambda)$ and $F(\lambda)$ such that

$$P(\lambda) = E(\lambda)Q(\lambda)F(\lambda).$$

Smith canonical form

Theorem

Every polynomial matrix is unimodularly equivalent to a matrix of the form

$$\left[\begin{array}{ccc|c} d_1(\lambda) & & & \\ & \ddots & & 0 \\ & & d_r(\lambda) & \\ \hline & 0 & & 0 \end{array}\right]$$

with $0 \neq d_r(\lambda)|d_{r-1}(\lambda)|\cdots|d_1(\lambda)$*. Moreover, the number* r *and the polynomials* $p_1(\lambda), \ldots, p_r(\lambda)$ *are uniquely determined by polynomial matrix.*

Smith canonical form

Theorem

Every polynomial matrix is unimodularly equivalent to a matrix of the form

$$\left[\begin{array}{ccc|c} d_1(\lambda) & & & \\ & \ddots & & 0 \\ & & d_r(\lambda) & \\ \hline & 0 & & 0 \end{array}\right]$$

with $0 \neq d_r(\lambda) | d_{r-1}(\lambda) | \cdots | d_1(\lambda)$*. Moreover, the number* r *and the polynomials* $p_1(\lambda), \ldots, p_r(\lambda)$ *are uniquely determined by polynomial matrix.*

Pros and cons

Can we do better?

Is there a "Kronecker-like" canonical form for matrix polynomials?

Can we do better?

Is there a "Kronecker-like" canonical form for matrix polynomials?

For $d = 2$, there is, but there are over 20 different block types.

Can we do better?

Is there a "Kronecker-like" canonical form for matrix polynomials?

For $d = 2$, there is, but there are over 20 different block types.

For $d = 3$, no one knows, and no one is willing to try.

Can we do better?

Is there a "Kronecker-like" canonical form for matrix polynomials?

For $d = 2$, there is, but there are over 20 different block types.

For $d = 3$, no one knows, and no one is willing to try.

We can do better by abandoning the direct-sum-of-blocks approach. A new approach is the subject of my thesis and will be introduced at my proposal talk.

Thank you!