Polynomial eigenvalue problems: linearizations and global backward error analysis

Froilán M. Dopico

joint work with **Piers Lawrence** (KU Leuven, Belgium),
Javier Pérez (KU Leuven, Belgium), and
Paul Van Dooren (UC Louvain, Belgium)

Departamento de Matemáticas
Universidad Carlos III de Madrid, Spain

4th International Symposium on Riordan Arrays and Related Topics
Facultad de CC. Matemáticas. U. Complutense de Madrid
July 17-20, 2017

Outline

Outline

The simplest form of a Polynomial Eigenvalue Problem (PEP)

- Given a **regular** $n \times n$ **matrix polynomial**, that is,

$$P(\lambda) = P_d \lambda^d + \cdots + P_1 \lambda + P_0, \qquad P_i \in \mathbb{C}^{n \times n},$$

with $\det P(\lambda) \not\equiv 0$,

- a number $\lambda_0 \in \mathbb{C}$ **is called an eigenvalue** of $P(\lambda)$
- if **there exists a nonzero vector** $v \in \mathbb{C}^n$**, called eigenvector,** such that

$$P(\lambda_0)\, v = 0$$

- This problem generalizes in a highly nontrivial way the standard matrix eigenvalue problem (SMatEP)

$$Av = \lambda_0 v \iff (\lambda_0 I_n - A)\, v = 0, \quad A \in \mathbb{C}^{n \times n}.$$

- Given a **regular** $n \times n$ **matrix polynomial**, that is,

$$P(\lambda) = P_d \lambda^d + \cdots + P_1 \lambda + P_0 \quad , \qquad P_i \in \mathbb{C}^{n \times n},$$

with $\det P(\lambda) \not\equiv 0$,

- a number $\lambda_0 \in \mathbb{C}$ **is called an eigenvalue** of $P(\lambda)$
- if **there exists a nonzero vector** $v \in \mathbb{C}^n$**, called eigenvector,** such that

$$P(\lambda_0)\, v = 0$$

- This problem generalizes in a highly nontrivial way the standard matrix eigenvalue problem (SMatEP)

$$Av = \lambda_0 v \iff (\lambda_0 I_n - A)\, v = 0 \;, A \in \mathbb{C}^{n \times n}.$$

The simplest form of a Polynomial Eigenvalue Problem (PEP)

- Given a **regular** $n \times n$ **matrix polynomial**, that is,

$$P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0 , \qquad P_i \in \mathbb{C}^{n\times n},$$

  with $\det P(\lambda) \not\equiv 0$,

- a number $\lambda_0 \in \mathbb{C}$ **is called an eigenvalue** of $P(\lambda)$
- if **there exists a nonzero vector** $v \in \mathbb{C}^n$**, called eigenvector,** such that

$$P(\lambda_0)\, v = 0$$

- This problem generalizes in a highly nontrivial way the standard matrix eigenvalue problem (SMatEP)

$$Av = \lambda_0 v \iff (\lambda_0 I_n - A)\, v = 0, A \in \mathbb{C}^{n\times n}.$$

- Given a **regular** $n \times n$ **matrix polynomial**, that is,

$$P(\lambda) = P_d \lambda^d + \cdots + P_1 \lambda + P_0 \quad , \qquad P_i \in \mathbb{C}^{n \times n},$$

  with $\det P(\lambda) \not\equiv 0$,

- a number $\lambda_0 \in \mathbb{C}$ **is called an eigenvalue** of $P(\lambda)$
- if **there exists a nonzero vector** $v \in \mathbb{C}^n$**, called eigenvector,** such that

$$P(\lambda_0)\, v = 0$$

- This problem generalizes in a highly nontrivial way the standard matrix eigenvalue problem (SMatEP)

$$Av = \lambda_0 v \iff (\lambda_0 I_n - A)\, v = 0 \quad , A \in \mathbb{C}^{n \times n}.$$

Polynomial Eigenvalue Problems arise in many applications

- Some applications are:
  1. Vibration Analysis of Mechanical Structures,
  2. Vibro-Acoustics: fluid-structure interaction problems,
  3. Stability analysis in fluid mechanics,
  4. Signal Processing,
  5. Multivariable System Theory and Control Theory,
  6. Computer-aided geometric design,
  7. and, very recently, in Network (Graph) Analysis.
- The applications of PEPs are often related to systems of dth-Order Differential (Algebraic) Equations with constant coefficients:

$$P_d \frac{d^d y(t)}{dt^d} + \cdots + P_1 \frac{dy(t)}{dt} + P_0 y(t) = 0, \quad P_i \in \mathbb{C}^{n \times n},$$

  and to look for solutions of the form $y(t) = e^{\lambda t} v$ with $v \in \mathbb{C}^n$.
- As a consequence of the applications the numerical solution of PEPs has received considerable attention in the last 15 years.

Polynomial Eigenvalue Problems arise in many applications

- Some applications are:
  1. Vibration Analysis of Mechanical Structures,
  2. Vibro-Acoustics: fluid-structure interaction problems,
  3. Stability analysis in fluid mechanics,
  4. Signal Processing,
  5. Multivariable System Theory and Control Theory,
  6. Computer-aided geometric design,
  7. and, very recently, in Network (Graph) Analysis.
- The applications of PEPs are often related to systems of dth-Order Differential (Algebraic) Equations with constant coefficients:

$$P_d \frac{d^d y(t)}{dt^d} + \cdots + P_1 \frac{dy(t)}{dt} + P_0 y(t) = 0, \quad P_i \in \mathbb{C}^{n\times n},$$

  and to look for solutions of the form $y(t) = e^{\lambda t} v$ with $v \in \mathbb{C}^n$.
- As a consequence of the applications the numerical solution of PEPs has received considerable attention in the last 15 years.

Polynomial Eigenvalue Problems arise in many applications

- Some applications are:
  1. Vibration Analysis of Mechanical Structures,
  2. Vibro-Acoustics: fluid-structure interaction problems,
  3. Stability analysis in fluid mechanics,
  4. Signal Processing,
  5. Multivariable System Theory and Control Theory,
  6. Computer-aided geometric design,
  7. and, very recently, in Network (Graph) Analysis.
- The applications of PEPs are often related to systems of dth-Order Differential (Algebraic) Equations with constant coefficients:

$$P_d \frac{d^d y(t)}{dt^d} + \cdots + P_1 \frac{dy(t)}{dt} + P_0 y(t) = 0, \quad P_i \in \mathbb{C}^{n \times n},$$

  and to look for solutions of the form $y(t) = e^{\lambda t} v$ with $v \in \mathbb{C}^n$.
- As a consequence of the applications the numerical solution of PEPs has received considerable attention in the last 15 years.

How large is the degree of $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$ in practice?

- In most direct applications coming from vibrational problems in mechanics $\mathbf{d = 2}$: **the quadratic eigenvalue problem (QEP)**.
- Betcke, Higham, Mehrmann, Schröder and Tisseur report in *"NLEVP: A Collection of Nonlinear Eigenvalue Problems"* (ACMTMathSoft, 2013) on applications with
  - $\mathbf{d = 4}$: Hamiltonian control problems with T-even structure, homography-based method for calibrating a central cadioptric vision system, spatial stability analysis of the Orr-Sommerfeld equation, and finite element solution of the equation for the modes of a planar waveguide using piecewise linear basis functions.
  - $\mathbf{d = 3}$: modeling of drift instabilities in the plasma edge inside a Tokamak reactor, and the five point relative pose problem in computer vision.
- PEPs combined with interpolation are often used to solve approximately other nonlinear eigenvalue problems. Then d can be much larger. Kressner and Roman (Numer. Lin. Alg. Appl., 2014) report on $\mathbf{d = 30}$ (3D Laplace eigenvalue problem on the Fichera corner) and $\mathbf{d = 11}$ (nonlinear eigenvalue problem coming from a 3D fluid-structure interaction problem).

- In most direct applications coming from vibrational problems in mechanics $\mathbf{d = 2}$: **the quadratic eigenvalue problem (QEP)**.
- Betcke, Higham, Mehrmann, Schröder and Tisseur report in *"NLEVP: A Collection of Nonlinear Eigenvalue Problems"* (ACMTMathSoft, 2013) on applications with
  - $\mathbf{d = 4}$: Hamiltonian control problems with T-even structure, homography-based method for calibrating a central cadioptric vision system, spatial stability analysis of the Orr-Sommerfeld equation, and finite element solution of the equation for the modes of a planar waveguide using piecewise linear basis functions.
  - $\mathbf{d = 3}$: modeling of drift instabilities in the plasma edge inside a Tokamak reactor, and the five point relative pose problem in computer vision.
- PEPs combined with interpolation are often used to solve approximately other nonlinear eigenvalue problems. Then d can be much larger. Kressner and Roman (Numer. Lin. Alg. Appl., 2014) report on $\mathbf{d = 30}$ (3D Laplace eigenvalue problem on the Fichera corner) and $\mathbf{d = 11}$ (nonlinear eigenvalue problem coming from a 3D fluid-structure interaction problem).

- In most direct applications coming from vibrational problems in mechanics $\mathbf{d = 2}$: **the quadratic eigenvalue problem (QEP)**.
- Betcke, Higham, Mehrmann, Schröder and Tisseur report in *"NLEVP: A Collection of Nonlinear Eigenvalue Problems"* (ACMTMathSoft, 2013) on applications with
  - $\mathbf{d = 4}$: Hamiltonian control problems with T-even structure, homography-based method for calibrating a central cadioptric vision system, spatial stability analysis of the Orr-Sommerfeld equation, and finite element solution of the equation for the modes of a planar waveguide using piecewise linear basis functions.
  - $\mathbf{d = 3}$: modeling of drift instabilities in the plasma edge inside a Tokamak reactor, and the five point relative pose problem in computer vision.
- PEPs combined with interpolation are often used to solve approximately other nonlinear eigenvalue problems. Then d can be much larger. Kressner and Roman (Numer. Lin. Alg. Appl., 2014) report on $\mathbf{d = 30}$ (3D Laplace eigenvalue problem on the Fichera corner) and $\mathbf{d = 11}$ (nonlinear eigenvalue problem coming from a 3D fluid-structure interaction problem).

How large is the degree of $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$ in practice?

- In most direct applications coming from vibrational problems in mechanics $\mathbf{d = 2}$: **the quadratic eigenvalue problem (QEP)**.
- Betcke, Higham, Mehrmann, Schröder and Tisseur report in *"NLEVP: A Collection of Nonlinear Eigenvalue Problems"* (ACMTMathSoft, 2013) on applications with
  - $\mathbf{d = 4}$: Hamiltonian control problems with T-even structure, homography-based method for calibrating a central cadioptric vision system, spatial stability analysis of the Orr-Sommerfeld equation, and finite element solution of the equation for the modes of a planar waveguide using piecewise linear basis functions.
  - $\mathbf{d = 3}$: modeling of drift instabilities in the plasma edge inside a Tokamak reactor, and the five point relative pose problem in computer vision.
- PEPs combined with interpolation are often used to solve approximately other nonlinear eigenvalue problems. Then d can be much larger. Kressner and Roman (Numer. Lin. Alg. Appl., 2014) report on $\mathbf{d = 30}$ (3D Laplace eigenvalue problem on the Fichera corner) and $\mathbf{d = 11}$ (nonlinear eigenvalue problem coming from a 3D fluid-structure interaction problem).

How large is the degree of $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$ in practice?

- In most direct applications coming from vibrational problems in mechanics $\mathbf{d = 2}$: **the quadratic eigenvalue problem (QEP)**.
- Betcke, Higham, Mehrmann, Schröder and Tisseur report in *"NLEVP: A Collection of Nonlinear Eigenvalue Problems"* (ACMTMathSoft, 2013) on applications with
  - $\mathbf{d = 4}$: Hamiltonian control problems with T-even structure, homography-based method for calibrating a central cadioptric vision system, spatial stability analysis of the Orr-Sommerfeld equation, and finite element solution of the equation for the modes of a planar waveguide using piecewise linear basis functions.
  - $\mathbf{d = 3}$: modeling of drift instabilities in the plasma edge inside a Tokamak reactor, and the five point relative pose problem in computer vision.
- PEPs combined with interpolation are often used to solve approximately other nonlinear eigenvalue problems. Then d can be much larger. Kressner and Roman (Numer. Lin. Alg. Appl., 2014) report on $\mathbf{d = 30}$ (3D Laplace eigenvalue problem on the Fichera corner) and $\mathbf{d = 11}$ (nonlinear eigenvalue problem coming from a 3D fluid-structure interaction problem).

PEPs are "much" more difficult than Standard Matrix EPs (I)

- Given the **regular ($\det P(\lambda) \not\equiv 0$) matrix polynomial**

$$P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0, \qquad P_i \in \mathbb{C}^{n\times n},$$

then the related PEP $P(\lambda_0)\, v = 0, \quad 0 \neq v \in \mathbb{C}^n$

- has **at most dn finite eigenvalues** since

$$\det P(\lambda) = (\det P_d)\,\lambda^{dn} + \text{lower degree terms in } \lambda,$$

i.e., there may be much more eigenvalues in PEPs than in SMatEPs.

- If $\det P_d = 0$, then the number of finite eigenvalues of the PEP is degree $(\det P(\lambda))$ and it is said that
- the PEP has dn − degree $(\det P(\lambda))$ **infinite eigenvalues**.
- The **eigenvectors of a PEP** corresponding to different eigenvalues **are not necessarily linearly independent**, since, in fact, we can have more than n different eigenvalues. Stark contrast with Standard MatEP.

PEPs are "much" more difficult than Standard Matrix EPs (I)

- Given the **regular ($\det P(\lambda) \not\equiv 0$) matrix polynomial**

$$P(\lambda) = P_d \lambda^d + \cdots + P_1 \lambda + P_0, \qquad P_i \in \mathbb{C}^{n\times n},$$

then the related PEP $P(\lambda_0)\, v = 0, \quad 0 \neq v \in \mathbb{C}^n$

- has **at most dn finite eigenvalues** since

$$\det P(\lambda) = (\det P_d)\, \lambda^{dn} + \text{lower degree terms in } \lambda,$$

i.e., there may be much more eigenvalues in PEPs than in SMatEPs.

- If $\det P_d = 0$, then the number of finite eigenvalues of the PEP is degree $(\det P(\lambda))$ and it is said that
- the PEP has $dn -$ degree $(\det P(\lambda))$ **infinite eigenvalues**.
- The **eigenvectors of a PEP** corresponding to different eigenvalues **are not necessarily linearly independent**, since, in fact, we can have more than n different eigenvalues. Stark contrast with Standard MatEP.

- Given the **regular** ($\det P(\lambda) \not\equiv 0$) **matrix polynomial**

$$P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0 \quad , \qquad P_i \in \mathbb{C}^{n\times n},$$

then the related PEP $P(\lambda_0)\, v = 0, \quad 0 \neq v \in \mathbb{C}^n$

- has **at most dn finite eigenvalues** since

$$\det P(\lambda) = (\det P_d)\, \lambda^{dn} + \text{lower degree terms in } \lambda,$$

i.e., there may be much more eigenvalues in PEPs than in SMatEPs.

- If $\det P_d = 0$, then the number of finite eigenvalues of the PEP is degree $(\det P(\lambda))$ and it is said that
- the PEP has $dn -$ degree $(\det P(\lambda))$ **infinite eigenvalues**.
- The **eigenvectors of a PEP** corresponding to different eigenvalues **are not necessarily linearly independent**, since, in fact, we can have more than n different eigenvalues. Stark contrast with Standard MatEP.

PEPs are "much" more difficult than Standard Matrix EPs (I)

- Given the **regular ($\det P(\lambda) \not\equiv 0$) matrix polynomial**

$$P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0, \qquad P_i \in \mathbb{C}^{n\times n},$$

then the related PEP $P(\lambda_0)\,v = 0, \quad 0 \neq v \in \mathbb{C}^n$

- has **at most dn finite eigenvalues** since

$$\det P(\lambda) = (\det P_d)\,\lambda^{dn} + \text{lower degree terms in } \lambda,$$

i.e., there may be much more eigenvalues in PEPs than in SMatEPs.

- If $\det P_d = 0$, then the number of finite eigenvalues of the PEP is degree $(\det P(\lambda))$ and it is said that
- the PEP has $dn -$ degree $(\det P(\lambda))$ **infinite eigenvalues**.
- The **eigenvectors of a PEP** corresponding to different eigenvalues **are not necessarily linearly independent**, since, in fact, we can have more than n different eigenvalues. Stark contrast with Standard MatEP.

PEPs are "much" more difficult than Standard Matrix EPs (I)

- Given the **regular ($\det P(\lambda) \not\equiv 0$) matrix polynomial**

$$P(\lambda) = P_d \lambda^d + \cdots + P_1 \lambda + P_0 , \qquad P_i \in \mathbb{C}^{n \times n},$$

  then the related PEP $P(\lambda_0)\, v = 0, \quad 0 \neq v \in \mathbb{C}^n$

- has **at most dn finite eigenvalues** since

$$\det P(\lambda) = (\det P_d)\, \lambda^{dn} + \text{lower degree terms in } \lambda,$$

  i.e., there may be much more eigenvalues in PEPs than in SMatEPs.

- If $\det P_d = 0$, then the number of finite eigenvalues of the PEP is degree $(\det P(\lambda))$ and it is said that
- the PEP has dn – degree $(\det P(\lambda))$ **infinite eigenvalues**.
- The **eigenvectors of a PEP** corresponding to different eigenvalues **are not necessarily linearly independent**, since, in fact, we can have more than n different eigenvalues. Stark contrast with Standard MatEP.

- Another way to define the infinite eigenvalues of a PEP that can be generalized to non-regular matrix polynomials is through **the reversal polynomial**.
- Given $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$, its reversal is

$$\text{rev}P(\lambda) := \lambda^d P(\tfrac{1}{\lambda}) = P_0\lambda^d + \cdots + P_{d-1}\lambda + P_d \, .$$

- Then the **infinite eigenvalues** of $P(\lambda)$ correspond to the **zero eigenvalues** of rev$P(\lambda)$.
- Why the name **infinite eigenvalues**? A possible reason is that if a polynomial with infinite eigenvalues, i.e., with P_d singular, is perturbed a bit, then eigenvalues with very large absolute values often appears.
- Of course, numerically the challenge is to decide whether or not a very large eigenvalue should be considered as infinite, since exact singularity of P_d is almost always lost in computations.
- Infinite eigenvalues correspond to constraints in algebraic-differential equations.

The reversal polynomial and more on infinite eigenvalues

- Another way to define the infinite eigenvalues of a PEP that can be generalized to non-regular matrix polynomials is through **the reversal polynomial**.
- Given $P(\lambda) = P_d \lambda^d + \cdots + P_1 \lambda + P_0$, its reversal is

$$\text{rev}P(\lambda) := \lambda^d P(\tfrac{1}{\lambda}) = P_0 \lambda^d + \cdots + P_{d-1}\lambda + P_d \quad .$$

- Then the **infinite eigenvalues** of $P(\lambda)$ correspond to the **zero eigenvalues** of rev$P(\lambda)$.
- Why the name **infinite eigenvalues**? A possible reason is that if a polynomial with infinite eigenvalues, i.e., with P_d singular, is perturbed a bit, then eigenvalues with very large absolute values often appears.
- Of course, numerically the challenge is to decide whether or not a very large eigenvalue should be considered as infinite, since exact singularity of P_d is almost always lost in computations.
- Infinite eigenvalues correspond to constraints in algebraic-differential equations.

- Another way to define the infinite eigenvalues of a PEP that can be generalized to non-regular matrix polynomials is through **the reversal polynomial**.
- Given $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$, its reversal is

$$\text{rev}P(\lambda) := \lambda^d P(\tfrac{1}{\lambda}) = P_0\lambda^d + \cdots + P_{d-1}\lambda + P_d .$$

- Then the **infinite eigenvalues** of $P(\lambda)$ correspond to the **zero eigenvalues** of rev$P(\lambda)$.
- Why the name **infinite eigenvalues**? A possible reason is that if a polynomial with infinite eigenvalues, i.e., with P_d singular, is perturbed a bit, then eigenvalues with very large absolute values often appears.
- Of course, numerically the challenge is to decide whether or not a very large eigenvalue should be considered as infinite, since exact singularity of P_d is almost always lost in computations.
- Infinite eigenvalues correspond to constraints in algebraic-differential equations.

The reversal polynomial and more on infinite eigenvalues

- Another way to define the infinite eigenvalues of a PEP that can be generalized to non-regular matrix polynomials is through **the reversal polynomial**.
- Given $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$, its reversal is

$$\text{rev}P(\lambda) := \lambda^d P(\tfrac{1}{\lambda}) = P_0\lambda^d + \cdots + P_{d-1}\lambda + P_d \quad .$$

- Then the **infinite eigenvalues** of $P(\lambda)$ correspond to the **zero eigenvalues** of rev$P(\lambda)$.
- Why the name **infinite eigenvalues**? A possible reason is that if a polynomial with infinite eigenvalues, i.e., with P_d singular, is perturbed a bit, then eigenvalues with very large absolute values often appears.
- Of course, numerically the challenge is to decide whether or not a very large eigenvalue should be considered as infinite, since exact singularity of P_d is almost always lost in computations.
- Infinite eigenvalues correspond to constraints in algebraic-differential equations.

- Another way to define the infinite eigenvalues of a PEP that can be generalized to non-regular matrix polynomials is through **the reversal polynomial**.
- Given $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$, its reversal is

$$\text{rev}P(\lambda) := \lambda^d P(\tfrac{1}{\lambda}) = P_0\lambda^d + \cdots + P_{d-1}\lambda + P_d \quad .$$

- Then the **infinite eigenvalues** of $P(\lambda)$ correspond to the **zero eigenvalues** of rev$P(\lambda)$.
- Why the name **infinite eigenvalues**? A possible reason is that if a polynomial with infinite eigenvalues, i.e., with P_d singular, is perturbed a bit, then eigenvalues with very large absolute values often appears.
- Of course, numerically the challenge is to decide whether or not a very large eigenvalue should be considered as infinite, since exact singularity of P_d is almost always lost in computations.
- Infinite eigenvalues correspond to constraints in algebraic-differential equations.

The reversal polynomial and more on infinite eigenvalues

- Another way to define the infinite eigenvalues of a PEP that can be generalized to non-regular matrix polynomials is through **the reversal polynomial**.
- Given $P(\lambda) = P_d \lambda^d + \cdots + P_1 \lambda + P_0$, its reversal is

$$\text{rev}P(\lambda) := \lambda^d P(\tfrac{1}{\lambda}) = P_0 \lambda^d + \cdots + P_{d-1}\lambda + P_d .$$

- Then the **infinite eigenvalues** of $P(\lambda)$ correspond to the **zero eigenvalues** of rev$P(\lambda)$.
- Why the name **infinite eigenvalues**? A possible reason is that if a polynomial with infinite eigenvalues, i.e., with P_d singular, is perturbed a bit, then eigenvalues with very large absolute values often appears.
- Of course, numerically the challenge is to decide whether or not a very large eigenvalue should be considered as infinite, since exact singularity of P_d is almost always lost in computations.
- Infinite eigenvalues correspond to constraints in algebraic-differential equations.

Example 1

Let ϵ be a small parameter and consider the quadratic matrix polynomial

$$P(\lambda) = \begin{bmatrix} (\lambda-1)(\lambda-2) & 0 \\ 0 & \lambda(\epsilon\lambda - 1) \end{bmatrix}$$
$$= \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & \epsilon \end{bmatrix} + \lambda \begin{bmatrix} -3 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

- If $\epsilon \neq 0$, then the eigenvalues are $\{1, 2, 0, 1/\epsilon\}$, (very large if $|\epsilon| \ll 1$).
- If $\epsilon = 0$, then the eigenvalues are $\{1, 2, 0, \infty\}$.
- Eigenvector of $\lambda_0 = 1$: $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- Eigenvector of $\lambda_0 = 2$: $v_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- **The two eigenvectors are equal!!**

Example 1

Let ϵ be a small parameter and consider the quadratic matrix polynomial

$$P(\lambda) = \begin{bmatrix} (\lambda-1)(\lambda-2) & 0 \\ 0 & \lambda(\epsilon\lambda-1) \end{bmatrix}$$
$$= \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & \epsilon \end{bmatrix} + \lambda \begin{bmatrix} -3 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

- If $\epsilon \neq 0$, then the eigenvalues are $\{1, 2, 0, 1/\epsilon\}$, (very large if $|\epsilon| \ll 1$).
- If $\epsilon = 0$, then the eigenvalues are $\{1, 2, 0, \infty\}$.
- Eigenvector of $\lambda_0 = 1$: $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- Eigenvector of $\lambda_0 = 2$: $v_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- **The two eigenvectors are equal!!**

Example 1

Let ϵ be a small parameter and consider the quadratic matrix polynomial

$$P(\lambda) = \begin{bmatrix} (\lambda-1)(\lambda-2) & 0 \\ 0 & \lambda(\epsilon\lambda - 1) \end{bmatrix}$$
$$= \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & \epsilon \end{bmatrix} + \lambda \begin{bmatrix} -3 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

- If $\epsilon \neq 0$, then the eigenvalues are $\{1, 2, 0, 1/\epsilon\}$, (very large if $|\epsilon| \ll 1$).
- If $\epsilon = 0$, then the eigenvalues are $\{1, 2, 0, \infty\}$.
- Eigenvector of $\lambda_0 = 1$: $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- Eigenvector of $\lambda_0 = 2$: $v_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- **The two eigenvectors are equal!!**

Example 1

Let ϵ be a small parameter and consider the quadratic matrix polynomial

$$P(\lambda) = \begin{bmatrix} (\lambda-1)(\lambda-2) & 0 \\ 0 & \lambda(\epsilon\lambda - 1) \end{bmatrix}$$
$$= \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & \epsilon \end{bmatrix} + \lambda \begin{bmatrix} -3 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

- If $\epsilon \neq 0$, then the eigenvalues are $\{1, 2, 0, 1/\epsilon\}$, (very large if $|\epsilon| \ll 1$).
- If $\epsilon = 0$, then the eigenvalues are $\{1, 2, 0, \infty\}$.
- Eigenvector of $\lambda_0 = 1$: $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- Eigenvector of $\lambda_0 = 2$: $v_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- **The two eigenvectors are equal!!**

Example 1

Let ϵ be a small parameter and consider the quadratic matrix polynomial

$$P(\lambda) = \begin{bmatrix} (\lambda-1)(\lambda-2) & 0 \\ 0 & \lambda(\epsilon\lambda - 1) \end{bmatrix}$$
$$= \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & \epsilon \end{bmatrix} + \lambda \begin{bmatrix} -3 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

- If $\epsilon \neq 0$, then the eigenvalues are $\{1, 2, 0, 1/\epsilon\}$, (very large if $|\epsilon| \ll 1$).
- If $\epsilon = 0$, then the eigenvalues are $\{1, 2, 0, \infty\}$.
- Eigenvector of $\lambda_0 = 1$: $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- Eigenvector of $\lambda_0 = 2$: $v_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- **The two eigenvectors are equal!!**

Example 1

Let ϵ be a small parameter and consider the quadratic matrix polynomial

$$P(\lambda) = \begin{bmatrix} (\lambda-1)(\lambda-2) & 0 \\ 0 & \lambda(\epsilon\lambda - 1) \end{bmatrix}$$
$$= \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & \epsilon \end{bmatrix} + \lambda \begin{bmatrix} -3 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

- If $\epsilon \neq 0$, then the eigenvalues are $\{1, 2, 0, 1/\epsilon\}$, (very large if $|\epsilon| \ll 1$).
- If $\epsilon = 0$, then the eigenvalues are $\{1, 2, 0, \infty\}$.
- Eigenvector of $\lambda_0 = 1$: $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- Eigenvector of $\lambda_0 = 2$: $v_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- The two eigenvectors are equal!!

Example 1

Let ϵ be a small parameter and consider the quadratic matrix polynomial

$$P(\lambda) = \begin{bmatrix} (\lambda-1)(\lambda-2) & 0 \\ 0 & \lambda(\epsilon\lambda - 1) \end{bmatrix}$$
$$= \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & \epsilon \end{bmatrix} + \lambda \begin{bmatrix} -3 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

- If $\epsilon \neq 0$, then the eigenvalues are $\{1, 2, 0, 1/\epsilon\}$, (very large if $|\epsilon| \ll 1$).
- If $\epsilon = 0$, then the eigenvalues are $\{1, 2, 0, \infty\}$.
- Eigenvector of $\lambda_0 = 1$: $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- Eigenvector of $\lambda_0 = 2$: $v_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- **The two eigenvectors are equal!!**

- An additional important step of difficulty is that PEPs can be **singular**, which happens when

$$P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$$

is either **rectangular or square with** $\det P(\lambda) \equiv 0$, i.e., zero for all λ.

- **Singular PEPs appear in applications** (though not as often as regular). In particular in Multivariable System Theory and Control Theory.
- A key concept in singular PEPs is the **normal rank (nrank)**, defined as the size of the largest minor that is not identically zero.
- It allows us to define eigenvalues for singular PEPs: $\lambda_0 \in \mathbb{C}$ is said to be a **finite eigenvalue of a singular** $P(\lambda)$ if

$$\operatorname{rank} P(\lambda_0) < \text{nrank}\, P(\lambda)\,,$$

- and to say that $P(\lambda)$ has an eigenvalue at infinity if 0 is an eigenvalue of rev$P(\lambda)$.

PEPs are "much" more difficult than Standard Matrix EPs (II)

- An additional important step of difficulty is that PEPs can be **singular**, which happens when

$$P(\lambda) = P_d \lambda^d + \cdots + P_1 \lambda + P_0$$

is either **rectangular or square with** $\det P(\lambda) \equiv 0$, i.e., zero for all λ.

- **Singular PEPs appear in applications** (though not as often as regular). In particular in Multivariable System Theory and Control Theory.
- A key concept in singular PEPs is the **normal rank (nrank)**, defined as the size of the largest minor that is not identically zero.
- It allows us to define eigenvalues for singular PEPs: $\lambda_0 \in \mathbb{C}$ is said to be a **finite eigenvalue of a singular** $P(\lambda)$ if

$$\operatorname{rank} P(\lambda_0) < \text{nrank}\, P(\lambda)\,,$$

- and to say that $P(\lambda)$ has an eigenvalue at infinity if 0 is an eigenvalue of $\text{rev}P(\lambda)$.

PEPs are "much" more difficult than Standard Matrix EPs (II)

- An additional important step of difficulty is that PEPs can be **singular**, which happens when

$$P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$$

is either **rectangular or square with** $\det P(\lambda) \equiv 0$, i.e., zero for all λ.

- **Singular PEPs appear in applications** (though not as often as regular). In particular in Multivariable System Theory and Control Theory.
- A key concept in singular PEPs is the **normal rank (nrank)**, defined as the size of the largest minor that is not identically zero.
- It allows us to define eigenvalues for singular PEPs: $\lambda_0 \in \mathbb{C}$ is said to be a **finite eigenvalue of a singular** $P(\lambda)$ if

$$\operatorname{rank} P(\lambda_0) < \text{nrank}\, P(\lambda)\,,$$

- and to say that $P(\lambda)$ has an eigenvalue at infinity if 0 is an eigenvalue of rev$P(\lambda)$.

PEPs are "much" more difficult than Standard Matrix EPs (II)

- An additional important step of difficulty is that PEPs can be **singular**, which happens when

$$P(\lambda) = P_d \lambda^d + \cdots + P_1 \lambda + P_0$$

is either **rectangular or square with** $\det P(\lambda) \equiv 0$, i.e., zero for all λ.

- **Singular PEPs appear in applications** (though not as often as regular). In particular in Multivariable System Theory and Control Theory.
- A key concept in singular PEPs is the **normal rank (nrank)**, defined as the size of the largest minor that is not identically zero.
- It allows us to define eigenvalues for singular PEPs: $\lambda_0 \in \mathbb{C}$ is said to be a **finite eigenvalue of a singular** $P(\lambda)$ if

$$\operatorname{rank} P(\lambda_0) < \text{nrank}\, P(\lambda)\,,$$

- and to say that $P(\lambda)$ has an eigenvalue at infinity if 0 is an eigenvalue of rev$P(\lambda)$.

PEPs are "much" more difficult than Standard Matrix EPs (II)

- An additional important step of difficulty is that PEPs can be **singular**, which happens when

$$P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$$

is either **rectangular or square with** $\det P(\lambda) \equiv 0$, i.e., zero for all λ.

- **Singular PEPs appear in applications** (though not as often as regular). In particular in Multivariable System Theory and Control Theory.
- A key concept in singular PEPs is the **normal rank (nrank)**, defined as the size of the largest minor that is not identically zero.
- It allows us to define eigenvalues for singular PEPs: $\lambda_0 \in \mathbb{C}$ is said to be a **finite eigenvalue of a singular** $P(\lambda)$ if

$$\operatorname{rank} P(\lambda_0) < \text{nrank}\, P(\lambda)\,,$$

- and to say that $P(\lambda)$ has an eigenvalue at infinity if 0 is an eigenvalue of rev$P(\lambda)$.

Example II

Consider the matrix polynomial of degree 4

$$P(\lambda) = \begin{bmatrix} \lambda & -\lambda^4 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\lambda & 0 \\ 0 & 0 & 0 & 1 & -\lambda \end{bmatrix} \in \mathbb{C}[\lambda]^{3\times 5},$$

- which has nrank $P(\lambda) = 3$ (pay attention to columns $2, 3, 4$).
- Note that 0 is an eigenvalue since rank $P(0) = 2$, and
- that $P(\lambda)$ has an eigenvalue at infinity since

$$\text{rev}P(\lambda) = \begin{bmatrix} \lambda^3 & -1 & 0 & 0 & 0 \\ 0 & 0 & \lambda^4 & -\lambda^3 & 0 \\ 0 & 0 & 0 & \lambda^4 & -\lambda^3 \end{bmatrix} \in \mathbb{C}[\lambda]^{3\times 5}$$

has rank $(\text{rev}P(0)) = 1 < 3$, i.e., with "geometric multiplicity 2".

Example II

Consider the matrix polynomial of degree 4

$$P(\lambda) = \begin{bmatrix} \lambda & -\lambda^4 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\lambda & 0 \\ 0 & 0 & 0 & 1 & -\lambda \end{bmatrix} \in \mathbb{C}[\lambda]^{3\times 5},$$

- which has nrank $P(\lambda) = 3$ (pay attention to columns $2, 3, 4$).
- Note that 0 is an eigenvalue since rank $P(0) = 2$, and
- that $P(\lambda)$ has an eigenvalue at infinity since

$$\text{rev}P(\lambda) = \begin{bmatrix} \lambda^3 & -1 & 0 & 0 & 0 \\ 0 & 0 & \lambda^4 & -\lambda^3 & 0 \\ 0 & 0 & 0 & \lambda^4 & -\lambda^3 \end{bmatrix} \in \mathbb{C}[\lambda]^{3\times 5}$$

has rank $(\text{rev}P\,(0)) = 1 < 3$, i.e., with “geometric multiplicity 2”.

Example II

Consider the matrix polynomial of degree 4

$$P(\lambda) = \begin{bmatrix} \lambda & -\lambda^4 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\lambda & 0 \\ 0 & 0 & 0 & 1 & -\lambda \end{bmatrix} \in \mathbb{C}[\lambda]^{3\times 5},$$

- which has nrank $P(\lambda) = 3$ (pay attention to columns $2, 3, 4$).
- Note that 0 is an eigenvalue since rank $P(0) = 2$, and
- that $P(\lambda)$ has an eigenvalue at infinity since

$$\text{rev}P(\lambda) = \begin{bmatrix} \lambda^3 & -1 & 0 & 0 & 0 \\ 0 & 0 & \lambda^4 & -\lambda^3 & 0 \\ 0 & 0 & 0 & \lambda^4 & -\lambda^3 \end{bmatrix} \in \mathbb{C}[\lambda]^{3\times 5}$$

has rank $(\text{rev}P\,(0)) = 1 < 3$, i.e., with "geometric multiplicity 2".

Example II

Consider the matrix polynomial of degree 4

$$P(\lambda) = \begin{bmatrix} \lambda & -\lambda^4 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\lambda & 0 \\ 0 & 0 & 0 & 1 & -\lambda \end{bmatrix} \in \mathbb{C}[\lambda]^{3\times 5},$$

- which has nrank $P(\lambda) = 3$ (pay attention to columns $2, 3, 4$).
- Note that 0 is an eigenvalue since rank $P(0) = 2$, and
- that $P(\lambda)$ has an eigenvalue at infinity since

$$\text{rev}P(\lambda) = \begin{bmatrix} \lambda^3 & -1 & 0 & 0 & 0 \\ 0 & 0 & \lambda^4 & -\lambda^3 & 0 \\ 0 & 0 & 0 & \lambda^4 & -\lambda^3 \end{bmatrix} \in \mathbb{C}[\lambda]^{3\times 5}$$

has rank $(\text{rev}P\,(0)) = 1 < 3$, i.e., with "geometric multiplicity 2".

PEPs are "much" more difficult than Standard Matrix EPs (III)

- Apart from eigenvalues, **singular matrix polynomials have** other "interesting numbers" attached to them called **minimal indices**.
- They are related to the fact that a singular $m \times n$ matrix polynomial $P(\lambda)$ has non-trivial left and/or right null-spaces over the field $\mathbb{F}(\lambda)$ of rational functions:

$$\mathcal{N}_\ell(P) := \left\{ y(\lambda)^T \in \mathbb{F}(\lambda)^{1\times m} \, : \, y(\lambda)^T P(\lambda) \equiv 0^T \right\},$$
$$\mathcal{N}_r(P) := \left\{ x(\lambda) \in \mathbb{F}(\lambda)^{n\times 1} \, : \, P(\lambda) x(\lambda) \equiv 0 \right\},$$

- and $\mathcal{N}_\ell(P)$ and $\mathcal{N}_r(P)$ have bases consisting entirely of vector polynomials.

Definition (Minimal bases)

A right (resp. left) **minimal basis** of $P(\lambda)$ is a basis of $\mathcal{N}_r(P)$ (resp. $\mathcal{N}_\ell(P)$)

1. consisting of vector polynomials
2. whose sum of degrees is minimal among all bases of $\mathcal{N}_r(P)$ (resp. $\mathcal{N}_\ell(P)$) consisting of vector polynomials.

- Apart from eigenvalues, **singular matrix polynomials have** other “interesting numbers” attached to them called **minimal indices**.
- They are related to the fact that a singular $m \times n$ matrix polynomial $P(\lambda)$ has non-trivial left and/or right null-spaces over the field $\mathbb{F}(\lambda)$ of rational functions:

$$\begin{aligned} \mathcal{N}_\ell(P) &:= \left\{ y(\lambda)^T \in \mathbb{F}(\lambda)^{1\times m} \,:\, y(\lambda)^T P(\lambda) \equiv 0^T \right\}, \\ \mathcal{N}_r(P) &:= \left\{ x(\lambda) \in \mathbb{F}(\lambda)^{n\times 1} \,:\, P(\lambda) x(\lambda) \equiv 0 \right\}, \end{aligned}$$

- and $\mathcal{N}_\ell(P)$ and $\mathcal{N}_r(P)$ have bases consisting entirely of vector polynomials.

Definition (Minimal bases)

A right (resp. left) **minimal basis** of $P(\lambda)$ is a basis of $\mathcal{N}_r(P)$ (resp. $\mathcal{N}_\ell(P)$)

1. consisting of vector polynomials
2. whose sum of degrees is minimal among all bases of $\mathcal{N}_r(P)$ (resp. $\mathcal{N}_\ell(P)$) consisting of vector polynomials.

PEPs are "much" more difficult than Standard Matrix EPs (III)

- Apart from eigenvalues, **singular matrix polynomials have** other "interesting numbers" attached to them called **minimal indices**.
- They are related to the fact that a singular $m \times n$ matrix polynomial $P(\lambda)$ has non-trivial left and/or right null-spaces over the field $\mathbb{F}(\lambda)$ of rational functions:

$$\begin{aligned} \mathcal{N}_\ell(P) &:= \left\{ y(\lambda)^T \in \mathbb{F}(\lambda)^{1\times m} \, : \, y(\lambda)^T P(\lambda) \equiv 0^T \right\}, \\ \mathcal{N}_r(P) &:= \left\{ x(\lambda) \in \mathbb{F}(\lambda)^{n\times 1} \, : \, P(\lambda) x(\lambda) \equiv 0 \right\}, \end{aligned}$$

- and $\mathcal{N}_\ell(P)$ and $\mathcal{N}_r(P)$ have bases consisting entirely of vector polynomials.

Definition (Minimal bases)

A right (resp. left) **minimal basis** of $P(\lambda)$ is a basis of $\mathcal{N}_r(P)$ (resp. $\mathcal{N}_\ell(P)$)

1. consisting of vector polynomials
2. whose sum of degrees is minimal among all bases of $\mathcal{N}_r(P)$ (resp. $\mathcal{N}_\ell(P)$) consisting of vector polynomials.

- Apart from eigenvalues, **singular matrix polynomials have** other “interesting numbers” attached to them called **minimal indices**.
- They are related to the fact that a singular $m \times n$ matrix polynomial $P(\lambda)$ has non-trivial left and/or right null-spaces over the field $\mathbb{F}(\lambda)$ of rational functions:

$$\begin{aligned} \mathcal{N}_\ell(P) &:= \left\{ y(\lambda)^T \in \mathbb{F}(\lambda)^{1\times m} : y(\lambda)^T P(\lambda) \equiv 0^T \right\}, \\ \mathcal{N}_r(P) &:= \left\{ x(\lambda) \in \mathbb{F}(\lambda)^{n\times 1} : P(\lambda) x(\lambda) \equiv 0 \right\}, \end{aligned}$$

- and $\mathcal{N}_\ell(P)$ and $\mathcal{N}_r(P)$ have bases consisting entirely of vector polynomials.

Definition (Minimal bases)

A right (resp. left) **minimal basis** of $P(\lambda)$ is a basis of $\mathcal{N}_r(P)$ (resp. $\mathcal{N}_\ell(P)$)

1. consisting of vector polynomials
2. whose sum of degrees is minimal among all bases of $\mathcal{N}_r(P)$ (resp. $\mathcal{N}_\ell(P)$) consisting of vector polynomials.

Minimal indices

There are infinitely many right minimal bases of $P(\lambda)$ (if there is one), but...

Theorem (Forney, SIAM J. Control, 1975)

The ordered list of degrees of the vector polynomials in any minimal basis of $\mathcal{N}_r(P)$ is always the same.

Definition

These degrees are called the **right minimal indices** of $P(\lambda)$.

Analogous definition for **left minimal indices**.

There are infinitely many right minimal bases of $P(\lambda)$ (if there is one), but...

Theorem (Forney, SIAM J. Control, 1975)

The ordered list of degrees of the vector polynomials in any minimal basis of $\mathcal{N}_r(P)$ is always the same.

Definition

These degrees are called the **right minimal indices** of $P(\lambda)$.

Analogous definition for **left minimal indices**.

There are infinitely many right minimal bases of $P(\lambda)$ (if there is one), but...

Theorem (Forney, SIAM J. Control, 1975)

The ordered list of degrees of the vector polynomials in any minimal basis of $\mathcal{N}_r(P)$ is always the same.

Definition

These degrees are called the **right minimal indices** of $P(\lambda)$.

Analogous definition for **left minimal indices**.

There are infinitely many right minimal bases of $P(\lambda)$ (if there is one), but...

Theorem (Forney, SIAM J. Control, 1975)

The ordered list of degrees of the vector polynomials in any minimal basis of $\mathcal{N}_r(P)$ is always the same.

Definition

These degrees are called the **right minimal indices** of $P(\lambda)$.

Analogous definition for **left minimal indices**.

Example III: right minimal bases and minimal indices

$$P(\lambda) = \begin{bmatrix} \lambda & -\lambda^4 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\lambda & 0 \\ 0 & 0 & 0 & 1 & -\lambda \end{bmatrix} \in \mathbb{C}[\lambda]^{3\times 5}$$

$$\mathcal{N}_r(P) = \text{Span}\{ \underbrace{\begin{bmatrix} \lambda^3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{u_1}, \underbrace{\begin{bmatrix} 0 \\ 0 \\ \lambda^2 \\ \lambda \\ 1 \end{bmatrix}}_{u_2} \} = \text{Span}\{ \underbrace{\begin{bmatrix} \lambda^3 \\ 1 \\ \lambda^3 \\ \lambda^2 \\ \lambda \end{bmatrix}}_{w_1}, \underbrace{\begin{bmatrix} \lambda^5 \\ \lambda^2 \\ \lambda^2 \\ \lambda \\ 1 \end{bmatrix}}_{w_2} \}$$

Sum of degrees of $\{u_1, u_2\} = 3 + 2 = 5$ (right minimal bases of $P(\lambda)$)

Sum of degrees of $\{w_1, w_2\} = 3 + 5 = 8$

Right minimal indices of $P(\lambda)$ = $\{2, 3\}$

Example III: right minimal bases and minimal indices

$$P(\lambda) = \begin{bmatrix} \lambda & -\lambda^4 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\lambda & 0 \\ 0 & 0 & 0 & 1 & -\lambda \end{bmatrix} \in \mathbb{C}[\lambda]^{3\times 5}$$

$$\mathcal{N}_r(P) = \text{Span}\{ \underbrace{\begin{bmatrix} \lambda^3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{u_1}, \underbrace{\begin{bmatrix} 0 \\ 0 \\ \lambda^2 \\ \lambda \\ 1 \end{bmatrix}}_{u_2} \} = \text{Span}\{ \underbrace{\begin{bmatrix} \lambda^3 \\ 1 \\ \lambda^3 \\ \lambda^2 \\ \lambda \end{bmatrix}}_{w_1}, \underbrace{\begin{bmatrix} \lambda^5 \\ \lambda^2 \\ \lambda^2 \\ \lambda \\ 1 \end{bmatrix}}_{w_2} \}$$

Sum of degrees of $\{u_1, u_2\} = 3 + 2 = 5$ (right minimal bases of $P(\lambda)$)

Sum of degrees of $\{w_1, w_2\} = 3 + 5 = 8$

Right minimal indices of $P(\lambda)$ = $\{2, 3\}$

Example III: right minimal bases and minimal indices

$$P(\lambda) = \begin{bmatrix} \lambda & -\lambda^4 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\lambda & 0 \\ 0 & 0 & 0 & 1 & -\lambda \end{bmatrix} \in \mathbb{C}[\lambda]^{3\times 5}$$

$$\mathcal{N}_r(P) = \text{Span}\{ \underbrace{\begin{bmatrix} \lambda^3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{u_1}, \underbrace{\begin{bmatrix} 0 \\ 0 \\ \lambda^2 \\ \lambda \\ 1 \end{bmatrix}}_{u_2} \} = \text{Span}\{ \underbrace{\begin{bmatrix} \lambda^3 \\ 1 \\ \lambda^3 \\ \lambda^2 \\ \lambda \end{bmatrix}}_{w_1}, \underbrace{\begin{bmatrix} \lambda^5 \\ \lambda^2 \\ \lambda^2 \\ \lambda \\ 1 \end{bmatrix}}_{w_2} \}$$

Sum of degrees of $\{u_1, u_2\} = 3 + 2 = 5$ (right minimal bases of $P(\lambda)$)

Sum of degrees of $\{w_1, w_2\} = 3 + 5 = 8$

Right minimal indices of $P(\lambda)$ = $\{2, 3\}$

Example III: right minimal bases and minimal indices

$$P(\lambda) = \begin{bmatrix} \lambda & -\lambda^4 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\lambda & 0 \\ 0 & 0 & 0 & 1 & -\lambda \end{bmatrix} \in \mathbb{C}[\lambda]^{3\times 5}$$

$$\mathcal{N}_r(P) = \text{Span}\{ \underbrace{\begin{bmatrix} \lambda^3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{u_1}, \underbrace{\begin{bmatrix} 0 \\ 0 \\ \lambda^2 \\ \lambda \\ 1 \end{bmatrix}}_{u_2} \} = \text{Span}\{ \underbrace{\begin{bmatrix} \lambda^3 \\ 1 \\ \lambda^3 \\ \lambda^2 \\ \lambda \end{bmatrix}}_{w_1}, \underbrace{\begin{bmatrix} \lambda^5 \\ \lambda^2 \\ \lambda^2 \\ \lambda \\ 1 \end{bmatrix}}_{w_2} \}$$

Sum of degrees of $\{u_1, u_2\} = 3 + 2 = 5$ (right minimal bases of $P(\lambda)$)

Sum of degrees of $\{w_1, w_2\} = 3 + 5 = 8$

Right minimal indices of $P(\lambda) = \{2, 3\}$

Example III: right minimal bases and minimal indices

$$P(\lambda) = \begin{bmatrix} \lambda & -\lambda^4 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\lambda & 0 \\ 0 & 0 & 0 & 1 & -\lambda \end{bmatrix} \in \mathbb{C}[\lambda]^{3\times 5}$$

$$\mathcal{N}_r(P) = \text{Span}\{ \underbrace{\begin{bmatrix} \lambda^3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{u_1}, \underbrace{\begin{bmatrix} 0 \\ 0 \\ \lambda^2 \\ \lambda \\ 1 \end{bmatrix}}_{u_2} \} = \text{Span}\{ \underbrace{\begin{bmatrix} \lambda^3 \\ 1 \\ \lambda^3 \\ \lambda^2 \\ \lambda \end{bmatrix}}_{w_1}, \underbrace{\begin{bmatrix} \lambda^5 \\ \lambda^2 \\ \lambda^2 \\ \lambda \\ 1 \end{bmatrix}}_{w_2} \}$$

Sum of degrees of $\{u_1, u_2\} = 3 + 2 = 5$ (right minimal bases of $P(\lambda)$)

Sum of degrees of $\{w_1, w_2\} = 3 + 5 = 8$

Right minimal indices of $P(\lambda)$ = $\{2, 3\}$

The complete eigenstructure of a matrix polynomial

As a consequence of the previous discussion, we define:

Definition

The **complete eigenstructure** of an $m \times n$ matrix polynomial $P(\lambda)$ is comprised of:

- its finite eigenvalues, together with their partial multiplicities,
- its infinite eigenvalue, together with its partial multiplicities,
- $n - r$ right minimal indices $\varepsilon_1, \ldots, \varepsilon_{n-r}$, and
- $m - r$ left minimal indices $\eta_1, \ldots, \eta_{m-r}$,

where r is the normal rank of $P(\lambda)$.

Remarks

- Minimal indices only appear in singular polynomials.
- The partial multiplicities are rigorously defined through the Smith form of $P(\lambda)$ and for matrices they are just the sizes of the Jordan blocks associated to each eigenvalue.

The complete eigenstructure of a matrix polynomial

As a consequence of the previous discussion, we define:

Definition

The **complete eigenstructure** of an $m \times n$ matrix polynomial $P(\lambda)$ is comprised of:

- its finite eigenvalues, together with their partial multiplicities,
- its infinite eigenvalue, together with its partial multiplicities,
- $n - r$ right minimal indices $\varepsilon_1, \ldots, \varepsilon_{n-r}$, and
- $m - r$ left minimal indices $\eta_1, \ldots, \eta_{m-r}$,

where r is the normal rank of $P(\lambda)$.

Remarks

- Minimal indices only appear in singular polynomials.
- The partial multiplicities are rigorously defined through the Smith form of $P(\lambda)$ and for matrices they are just the sizes of the Jordan blocks associated to each eigenvalue.

Outline

The most reliable methods for solving numerically PEPs are based on the concept of linearization.

Definition (Linearizations of Matrix Polynomials)

A linearization $L(\lambda)$ of a matrix polynomial $P(\lambda)$ **is a linear matrix polynomial, or matrix pencil,** such that

(1) $L(\lambda)$ and $P(\lambda)$ **have the same number of right minimal indices**.

(2) $L(\lambda)$ and $P(\lambda)$ **have the same number of left minimal indices**.

(3) $L(\lambda)$ and $P(\lambda)$ **have the same finite eigenvalues** with the same partial multiplicities.

If, **in addition**,

(4) $L(\lambda)$ and $P(\lambda)$ **have the same infinite eigenvalues** with the same partial multiplicities,

then $L(\lambda)$ is called a **strong linearization** of $P(\lambda)$.

The most reliable methods for solving numerically PEPs are based on the concept of linearization.

Definition (Linearizations of Matrix Polynomials)

A linearization $L(\lambda)$ of a matrix polynomial $P(\lambda)$ **is a linear matrix polynomial, or matrix pencil,** such that

(1) $L(\lambda)$ and $P(\lambda)$ **have the same number of right minimal indices**.

(2) $L(\lambda)$ and $P(\lambda)$ **have the same number of left minimal indices**.

(3) $L(\lambda)$ and $P(\lambda)$ **have the same finite eigenvalues** with the same partial multiplicities.

If, **in addition**,

(4) $L(\lambda)$ and $P(\lambda)$ **have the same infinite eigenvalues** with the same partial multiplicities,

then $L(\lambda)$ is called a **strong linearization** of $P(\lambda)$.

The most reliable methods for solving numerically PEPs are based on the concept of linearization.

Definition (Linearizations of Matrix Polynomials)

A linearization $L(\lambda)$ of a matrix polynomial $P(\lambda)$ **is a linear matrix polynomial, or matrix pencil,** such that

(1) $L(\lambda)$ and $P(\lambda)$ **have the same number of right minimal indices**.

(2) $L(\lambda)$ and $P(\lambda)$ **have the same number of left minimal indices**.

(3) $L(\lambda)$ and $P(\lambda)$ **have the same finite eigenvalues** with the same partial multiplicities.

If, **in addition**,

(4) $L(\lambda)$ and $P(\lambda)$ **have the same infinite eigenvalues** with the same partial multiplicities,

then $L(\lambda)$ is called a **strong linearization** of $P(\lambda)$.

The classical **Frobenius companion form** of the $m \times n$ matrix polynomial $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$ is

$$C_1(\lambda) := \begin{bmatrix} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ -I_n & \lambda I_n & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda I_n & \\ & & & -I_n & \lambda I_n \end{bmatrix} \in \mathbb{C}[\lambda]^{(m+n(d-1))\times nd}$$

Theorem ($C_1(\lambda)$ is much more than a strong linearization!!)

(a) *If $0 \leq \varepsilon_1 \leq \cdots \leq \varepsilon_p$ are the right minimal indices of $P(\lambda)$, then the right minimal indices of $C_1(\lambda)$ are $\varepsilon_1 + d - 1 \leq \cdots \leq \varepsilon_p + d - 1$.*

(b) *If $0 \leq \eta_1 \leq \cdots \leq \eta_q$ are the left minimal indices of $P(\lambda)$, then the left minimal indices of $C_1(\lambda)$ are $\eta_1 \leq \cdots \leq \eta_q$.*

Example of strong linearization whose right (resp. left) minimal indices allow us to recover the ones of the polynomial via uniform shifts.

The most famous strong linearization (I)

The classical **Frobenius companion form** of the $m \times n$ matrix polynomial $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$ is

$$C_1(\lambda) := \begin{bmatrix} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ -I_n & \lambda I_n & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda I_n & \\ & & & -I_n & \lambda I_n \end{bmatrix} \in \mathbb{C}[\lambda]^{(m+n(d-1))\times nd}$$

Theorem ($C_1(\lambda)$ is much more than a strong linearization!!)

(a) *If $0 \leq \varepsilon_1 \leq \cdots \leq \varepsilon_p$ are the right minimal indices of $P(\lambda)$, then the right minimal indices of $C_1(\lambda)$ are $\varepsilon_1 + d - 1 \leq \cdots \leq \varepsilon_p + d - 1$.*

(b) *If $0 \leq \eta_1 \leq \cdots \leq \eta_q$ are the left minimal indices of $P(\lambda)$, then the left minimal indices of $C_1(\lambda)$ are $\eta_1 \leq \cdots \leq \eta_q$.*

Example of strong linearization whose right (resp. left) minimal indices allow us to recover the ones of the polynomial via uniform shifts.

The most famous strong linearization (I)

The classical **Frobenius companion form** of the $m \times n$ matrix polynomial $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$ is

$$C_1(\lambda) := \begin{bmatrix} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ -I_n & \lambda I_n & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda I_n & \\ & & & -I_n & \lambda I_n \end{bmatrix} \in \mathbb{C}[\lambda]^{(m+n(d-1))\times nd}$$

Theorem ($C_1(\lambda)$ is much more than a strong linearization!!)

(a) *If $0 \leq \varepsilon_1 \leq \cdots \leq \varepsilon_p$ are the right minimal indices of $P(\lambda)$, then the right minimal indices of $C_1(\lambda)$ are $\varepsilon_1 + d - 1 \leq \cdots \leq \varepsilon_p + d - 1$.*

(b) *If $0 \leq \eta_1 \leq \cdots \leq \eta_q$ are the left minimal indices of $P(\lambda)$, then the left minimal indices of $C_1(\lambda)$ are $\eta_1 \leq \cdots \leq \eta_q$.*

Example of strong linearization whose right (resp. left) minimal indices allow us to recover the ones of the polynomial via uniform shifts.

Theorem (recovery of eigenvectors from $C_1(\lambda)$)

Let $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$ *be a* **regular** *matrix polynomial,* $\lambda_0 \in \mathbb{C}$ **be a finite eigenvalue of** $P(\lambda)$*, and* $C_1(\lambda)$ *be the Frobenius companion form of* $P(\lambda)$*. Then,* **any eigenvector** z **of** $C_1(\lambda)$ **associated to** λ_0 **has the form**

$$z = \begin{bmatrix} \lambda_0^{d-1}\, x \\ \vdots \\ \lambda_0\, x \\ x \end{bmatrix} = \begin{bmatrix} \lambda_0^{d-1} \\ \vdots \\ \lambda_0 \\ 1 \end{bmatrix} \otimes x$$

with x **an eigenvector of** $P(\lambda)$ **associated to** λ_0**.**

- $C_1(\lambda)$ is one (among many others) strong linearization of $P(\lambda)$ that allows us to recover without computational cost the eigenvectors of the polynomial from those of the linearization,
- and, also, the minimal bases.

Theorem (recovery of eigenvectors from $C_1(\lambda)$)

Let $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$ be a **regular** *matrix polynomial,* $\lambda_0 \in \mathbb{C}$ **be a finite eigenvalue of** $P(\lambda)$, *and $C_1(\lambda)$ be the Frobenius companion form of $P(\lambda)$. Then,* **any eigenvector z of $C_1(\lambda)$ associated to λ_0 has the form**

$$z = \begin{bmatrix} \lambda_0^{d-1}\, x \\ \vdots \\ \lambda_0\, x \\ x \end{bmatrix} = \begin{bmatrix} \lambda_0^{d-1} \\ \vdots \\ \lambda_0 \\ 1 \end{bmatrix} \otimes x$$

with x an eigenvector of $P(\lambda)$ associated to λ_0.

- $C_1(\lambda)$ is one (among many others) strong linearization of $P(\lambda)$ that allows us to recover without computational cost the eigenvectors of the polynomial from those of the linearization,
- and, also, the minimal bases.

Theorem (recovery of eigenvectors from $C_1(\lambda)$)

Let $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$ be a **regular** *matrix polynomial,* $\lambda_0 \in \mathbb{C}$ **be a finite eigenvalue of** $P(\lambda)$*, and $C_1(\lambda)$ be the Frobenius companion form of $P(\lambda)$. Then,* **any eigenvector** z **of** $C_1(\lambda)$ **associated to** λ_0 **has the form**

$$z = \begin{bmatrix} \lambda_0^{d-1}\, x \\ \vdots \\ \lambda_0\, x \\ x \end{bmatrix} = \begin{bmatrix} \lambda_0^{d-1} \\ \vdots \\ \lambda_0 \\ 1 \end{bmatrix} \otimes x$$

with x an eigenvector of $P(\lambda)$ associated to λ_0.

- $C_1(\lambda)$ is one (among many others) strong linearization of $P(\lambda)$ that allows us to recover without computational cost the eigenvectors of the polynomial from those of the linearization,
- and, also, the minimal bases.

Parenthesis: There are many other strong linearizations (I)

- Since 2006 (Mackey, Mackey, Mehl, Mehrmann, SIMAX), many “new” strong linearizations of matrix polynomials have been developed by many authors all around the world
- which also allow us to recover minimal indices via uniform shifts and eigenvectors of regular PEPs without any computational cost.
- One relevant motivation for developing new classes of linearizations is to preserve structures appearing in applications, which is important for saving operations in algorithms and for preserving properties of the eigenvalues in floating point arithmetic.
- For instance, if $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$ is Hermitian, i.e., it has Hermitian coefficients, **the Frobenius companion form is not!!**

$$C_1(\lambda) := \begin{bmatrix} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ -I_n & \lambda I_n & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda I_n & \\ & & & -I_n & \lambda I_n \end{bmatrix}$$

Parenthesis: There are many other strong linearizations (I)

- Since 2006 (Mackey, Mackey, Mehl, Mehrmann, SIMAX), many "new" strong linearizations of matrix polynomials have been developed by many authors all around the world
- which also allow us to recover minimal indices via uniform shifts and eigenvectors of regular PEPs without any computational cost.
- One relevant motivation for developing new classes of linearizations is to preserve structures appearing in applications, which is important for saving operations in algorithms and for preserving properties of the eigenvalues in floating point arithmetic.
- For instance, if $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$ is Hermitian, i.e., it has Hermitian coefficients, **the Frobenius companion form is not!!**

$$C_1(\lambda) := \begin{bmatrix} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ -I_n & \lambda I_n & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda I_n & \\ & & & -I_n & \lambda I_n \end{bmatrix}$$

- Since 2006 (Mackey, Mackey, Mehl, Mehrmann, SIMAX), many "new" strong linearizations of matrix polynomials have been developed by many authors all around the world
- which also allow us to recover minimal indices via uniform shifts and eigenvectors of regular PEPs without any computational cost.
- One relevant motivation for developing new classes of linearizations is to preserve structures appearing in applications, which is important for saving operations in algorithms and for preserving properties of the eigenvalues in floating point arithmetic.
- For instance, if $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$ is Hermitian, i.e., it has Hermitian coefficients, **the Frobenius companion form is not!!**

$$C_1(\lambda) := \begin{bmatrix} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ -I_n & \lambda I_n & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda I_n & \\ & & & -I_n & \lambda I_n \end{bmatrix}$$

- Since 2006 (Mackey, Mackey, Mehl, Mehrmann, SIMAX), many "new" strong linearizations of matrix polynomials have been developed by many authors all around the world
- which also allow us to recover minimal indices via uniform shifts and eigenvectors of regular PEPs without any computational cost.
- One relevant motivation for developing new classes of linearizations is to preserve structures appearing in applications, which is important for saving operations in algorithms and for preserving properties of the eigenvalues in floating point arithmetic.
- For instance, if $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0$ is Hermitian, i.e., it has Hermitian coefficients, **the Frobenius companion form is not!!**

$$C_1(\lambda) := \begin{bmatrix} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ -I_n & \lambda I_n & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda I_n & \\ & & & -I_n & \lambda I_n \end{bmatrix}$$

- but

$$\widetilde{L}(\lambda) = \begin{bmatrix} \lambda P_1 + P_0 & \lambda I_n & & & & & 0 \\ \lambda I_n & 0 & I_n & & & & \\ & I_n & \lambda P_3 + P_2 & \lambda I_n & & & \\ & & \lambda I_n & 0 & I_n & & \\ & & & I_n & \lambda P_5 + P_4 & \lambda I_n & \\ & & & & \lambda I_n & 0 & I_n \\ 0 & & & & & I_n & \lambda P_7 + P_6 \end{bmatrix},$$

is a **Hermitian strong linearization** of the $n \times n$ Hermitian matrix polynomial $P(\lambda) = P_7\lambda^7 + \cdots + P_1\lambda + P_0$ (Antoniou & Vologiannidis (ELA, 2004), Mackey & Mackey & Mehl & Mehrmann (LAA, 2010)).

There are "excellent algorithms" for solving linear PEPs!! (also called Generalized Eigenvalue Problems (GEP))

- In summary, "good" strong linearizations of a matrix polynomial $P(\lambda)$ are linear matrix polynomials that have the same eigenvalues as $P(\lambda)$ and that allow us to recover the eigenvectors when $P(\lambda)$ is regular, and the minimal indices when $P(\lambda)$ is singular.
- They are very important for solving numerically PEPs
- because there exist excellent algorithms for solving linear PEPs.
- The fundamental proposed approach

"linearization + linear eigenvalue algorithm on the linearization"

for solving numerically PEPs can be traced back at least to Van Dooren & De Wilde (LAA, 1983) and Van Dooren's PhD Thesis in 1979.

There are "excellent algorithms" for solving linear PEPs!!
(also called Generalized Eigenvalue Problems (GEP))

- In summary, "good" strong linearizations of a matrix polynomial $P(\lambda)$ are linear matrix polynomials that have the same eigenvalues as $P(\lambda)$ and that allow us to recover the eigenvectors when $P(\lambda)$ is regular, and the minimal indices when $P(\lambda)$ is singular.
- They are very important for solving numerically PEPs
- because there exist excellent algorithms for solving linear PEPs.
- The fundamental proposed approach

"linearization + linear eigenvalue algorithm on the linearization"

for solving numerically PEPs can be traced back at least to Van Dooren & De Wilde (LAA, 1983) and Van Dooren's PhD Thesis in 1979.

There are "excellent algorithms" for solving linear PEPs!!

(also called Generalized Eigenvalue Problems (GEP))

- In summary, "good" strong linearizations of a matrix polynomial $P(\lambda)$ are linear matrix polynomials that have the same eigenvalues as $P(\lambda)$ and that allow us to recover the eigenvectors when $P(\lambda)$ is regular, and the minimal indices when $P(\lambda)$ is singular.
- They are very important for solving numerically PEPs
- because there exist excellent algorithms for solving linear PEPs.
- The fundamental proposed approach

  "linearization + linear eigenvalue algorithm on the linearization"

  for solving numerically PEPs can be traced back at least to Van Dooren & De Wilde (LAA, 1983) and Van Dooren's PhD Thesis in 1979.

There are “excellent algorithms” for solving linear PEPs!!
(also called Generalized Eigenvalue Problems (GEP))

- In summary, “good” strong linearizations of a matrix polynomial $P(\lambda)$ are linear matrix polynomials that have the same eigenvalues as $P(\lambda)$ and that allow us to recover the eigenvectors when $P(\lambda)$ is regular, and the minimal indices when $P(\lambda)$ is singular.
- They are very important for solving numerically PEPs
- because there exist excellent algorithms for solving linear PEPs.
- The fundamental proposed approach

  “linearization + linear eigenvalue algorithm on the linearization”

  for solving numerically PEPs can be traced back at least to Van Dooren & De Wilde (LAA, 1983) and Van Dooren’s PhD Thesis in 1979.

- **QZ algorithm for regular GEPs** $(\det(A - \lambda B) \not\equiv 0)$ (Moler & Stewart, SINUM, 1973).
  - QZ is implemented in MATLAB command `eig(A,B)` and cost $O(n^3)$.
- **Staircase or GUPTRI algorithm for singular GEPs** (Van Dooren, LAA, 1979; Demmel & Kågström, ACMTMathSoft., 1993).
  - It computes eigenvalues and minimal indices and there are FORTRAN implementations.
- **QZ and GUPTRI are both backward stable** since only use orthogonal transformations.
- The command `polyeig` of MATLAB computes all the eigenvalues of a “not too large” regular PEP by applying QZ to the first companion form of the matrix polynomial defining the PEP.

Algorithms for solving "not too large" GEPs $A - \lambda B$

- **QZ algorithm for regular GEPs** $(\det(A - \lambda B) \not\equiv 0)$ (Moler & Stewart, SINUM, 1973).
  - QZ is implemented in MATLAB command `eig(A,B)` and cost $O(n^3)$.
- **Staircase or GUPTRI algorithm for singular GEPs** (Van Dooren, LAA, 1979; Demmel & Kågström, ACMTMathSoft., 1993).
  - It computes eigenvalues and minimal indices and there are FORTRAN implementations.
- **QZ and GUPTRI are both backward stable** since only use orthogonal transformations.
- The command `polyeig` of MATLAB computes all the eigenvalues of a "not too large" regular PEP by applying QZ to the first companion form of the matrix polynomial defining the PEP.

- **QZ algorithm for regular GEPs** $(\det(A - \lambda B) \not\equiv 0)$ (Moler & Stewart, SINUM, 1973).
  - QZ is implemented in MATLAB command `eig(A,B)` and cost $O(n^3)$.
- **Staircase or GUPTRI algorithm for singular GEPs** (Van Dooren, LAA, 1979; Demmel & Kågström, ACMTMathSoft., 1993).
  - It computes eigenvalues and minimal indices and there are FORTRAN implementations.
- **QZ and GUPTRI are both backward stable** since only use orthogonal transformations.
- The command `polyeig` of MATLAB computes all the eigenvalues of a “not too large” regular PEP by applying QZ to the first companion form of the matrix polynomial defining the PEP.

- **QZ algorithm for regular GEPs** $(\det(A - \lambda B) \not\equiv 0)$ (Moler & Stewart, SINUM, 1973).
  - QZ is implemented in MATLAB command `eig(A,B)` and cost $O(n^3)$.
- **Staircase or GUPTRI algorithm for singular GEPs** (Van Dooren, LAA, 1979; Demmel & Kågström, ACMTMathSoft., 1993).
  - It computes eigenvalues and minimal indices and there are FORTRAN implementations.
- **QZ and GUPTRI are both backward stable** since only use orthogonal transformations.
- The command `polyeig` of MATLAB computes all the eigenvalues of a "not too large" regular PEP by applying QZ to the first companion form of the matrix polynomial defining the PEP.

Algorithms for solving "not too large" GEPs $A - \lambda B$

- **QZ algorithm for regular GEPs** $(\det(A - \lambda B) \not\equiv 0)$ (Moler & Stewart, SINUM, 1973).
  - QZ is implemented in MATLAB command `eig(A,B)` and cost $O(n^3)$.
- **Staircase or GUPTRI algorithm for singular GEPs** (Van Dooren, LAA, 1979; Demmel & Kågström, ACMTMathSoft., 1993).
  - It computes eigenvalues and minimal indices and there are FORTRAN implementations.
- **QZ and GUPTRI are both backward stable** since only use orthogonal transformations.
- The command `polyeig` of MATLAB computes all the eigenvalues of a "not too large" regular PEP by applying QZ to the first companion form of the matrix polynomial defining the PEP.

- **QZ algorithm for regular GEPs** $(\det(A - \lambda B) \not\equiv 0)$ (Moler & Stewart, SINUM, 1973).
  - QZ is implemented in MATLAB command `eig(A,B)` and cost $O(n^3)$.
- **Staircase or GUPTRI algorithm for singular GEPs** (Van Dooren, LAA, 1979; Demmel & Kågström, ACMTMathSoft., 1993).
  - It computes eigenvalues and minimal indices and there are FORTRAN implementations.
- **QZ and GUPTRI are both backward stable** since only use orthogonal transformations.
- The command `polyeig` of MATLAB computes all the eigenvalues of a "not too large" regular PEP by applying QZ to the first companion form of the matrix polynomial defining the PEP.

- **They are based on linearizations** $A - \lambda B$ of the PEP
- and on **Krylov subspace methods on** $A - \lambda B$ (Arnoldi on $B^{-1}A$, Rational-Krylov with shifts on $(A - \theta_j B)^{-1}B$) for computing a few desired eigenvalues,
- **but the application of these Krylov methods is NOT direct**,
- **since this would be very expensive in terms of memory and orthogonalization costs**, because
- if $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0 \in \mathbb{C}[\lambda]^{n\times n}$ then its Frobenius companion form (and any other strong linearization) has size $nd \times nd$

$$C_1(\lambda) := \begin{bmatrix} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ -I_n & \lambda I_n & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda I_n & \\ & & & -I_n & \lambda I_n \end{bmatrix} \in \mathbb{C}[\lambda]^{nd\times nd}.$$

So, if n is very large, then nd is very very large.

- **They are based on linearizations** $A - \lambda B$ of the PEP
- and on **Krylov subspace methods on** $A - \lambda B$ (Arnoldi on $B^{-1}A$, Rational-Krylov with shifts on $(A - \theta_j B)^{-1}B$) for computing a few desired eigenvalues,
- **but the application of these Krylov methods is NOT direct**,
- **since this would be very expensive in terms of memory and orthogonalization costs**, because
- if $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0 \in \mathbb{C}[\lambda]^{n\times n}$ then its Frobenius companion form (and any other strong linearization) has size $nd \times nd$

$$C_1(\lambda) := \begin{bmatrix} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ -I_n & \lambda I_n & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda I_n & \\ & & & -I_n & \lambda I_n \end{bmatrix} \in \mathbb{C}[\lambda]^{nd\times nd}.$$

So, if n is very large, then nd is very very large.

- **They are based on linearizations** $A - \lambda B$ of the PEP
- and on **Krylov subspace methods on** $A - \lambda B$ (Arnoldi on $B^{-1}A$, Rational-Krylov with shifts on $(A - \theta_j B)^{-1}B$) for computing a few desired eigenvalues,
- **but the application of these Krylov methods is NOT direct**,
- **since this would be very expensive in terms of memory and orthogonalization costs**, because
- if $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0 \in \mathbb{C}[\lambda]^{n\times n}$ then its Frobenius companion form (and any other strong linearization) has size $nd \times nd$

$$C_1(\lambda) := \begin{bmatrix} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ -I_n & \lambda I_n & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda I_n & \\ & & & -I_n & \lambda I_n \end{bmatrix} \in \mathbb{C}[\lambda]^{nd\times nd}.$$

So, if n is very large, then nd is very very large.

- **They are based on linearizations** $A - \lambda B$ of the PEP
- and on **Krylov subspace methods on** $A - \lambda B$ (Arnoldi on $B^{-1}A$, Rational-Krylov with shifts on $(A - \theta_j B)^{-1}B$) for computing a few desired eigenvalues,
- **but the application of these Krylov methods is NOT direct**,
- **since this would be very expensive in terms of memory and orthogonalization costs**, because
- if $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0 \in \mathbb{C}[\lambda]^{n\times n}$ then its Frobenius companion form (and any other strong linearization) has size $nd \times nd$

$$C_1(\lambda) := \begin{bmatrix} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ -I_n & \lambda I_n & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda I_n & \\ & & & -I_n & \lambda I_n \end{bmatrix} \in \mathbb{C}[\lambda]^{nd\times nd}.$$

So, if n is very large, then nd is very very large.

- **They are based on linearizations** $A - \lambda B$ of the PEP
- and on **Krylov subspace methods on** $A - \lambda B$ (Arnoldi on $B^{-1}A$, Rational-Krylov with shifts on $(A - \theta_j B)^{-1} B$) for computing a few desired eigenvalues,
- **but the application of these Krylov methods is NOT direct**,
- **since this would be very expensive in terms of memory and orthogonalization costs**, because
- if $P(\lambda) = P_d \lambda^d + \cdots + P_1 \lambda + P_0 \in \mathbb{C}[\lambda]^{n \times n}$ then its Frobenius companion form (and any other strong linearization) has size $nd \times nd$

$$C_1(\lambda) := \begin{bmatrix} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ -I_n & \lambda I_n & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda I_n & \\ & & & -I_n & \lambda I_n \end{bmatrix} \in \mathbb{C}[\lambda]^{nd \times nd}.$$

So, if n is very large, then nd is very very large.

A few words on algorithms for solving large-scale regular PEPs (II)

- Therefore, Krylov subspace methods for PEPs take advantage of the structure of the linearization and of the bases of their Krylov subspaces
- to obtain **memory** and orthogonalization costs of the same order of those of an $n \times n$ standard matrix problem.
- The most stable and efficient methods in this family are
  1. TOAR (Two level Orthogonal ARnoldi) for QEPs (Su & Bai & Lu, 2008 and SIMAX 2016) based on $C_1(\lambda)$,
  2. CORK (COmpact Rational Krylov) for arbitrary PEPs (Van Beeumen & Meerbergen & Michiels, SIMAX, 2015) very general, it can use many linearizations and bases for expressing the PEP,
- which are the "final optimal product" of previous pioneering algorithms:
  1. SOAR (Second Order ARnoldi) for QEPs (Bai & Su, SIMAX, 2005),
  2. Q-Arnoldi for QEPs (Meerbergen, SIMAX, 2008),
  3. TOAR in Chebyshev basis and for arbitrary degree PEPs (Kressner & Roman, Num. Lin. Alg. Appl, 2014), etc....
- **Available HPC software:** parallel implementations of TOAR for any degree (including symm. versions) in SLEPc (Campos & Roman, 2016).

- Therefore, Krylov subspace methods for PEPs take advantage of the structure of the linearization and of the bases of their Krylov subspaces
- to obtain **memory** and orthogonalization costs of the same order of those of an $n \times n$ standard matrix problem.
- The most stable and efficient methods in this family are
  1. TOAR (Two level Orthogonal ARnoldi) for QEPs (Su & Bai & Lu, 2008 and SIMAX 2016) based on $C_1(\lambda)$,
  2. CORK (COmpact Rational Krylov) for arbitrary PEPs (Van Beeumen & Meerbergen & Michiels, SIMAX, 2015) very general, it can use many linearizations and bases for expressing the PEP,
- which are the "final optimal product" of previous pioneering algorithms:
  1. SOAR (Second Order ARnoldi) for QEPs (Bai & Su, SIMAX, 2005),
  2. Q-Arnoldi for QEPs (Meerbergen, SIMAX, 2008),
  3. TOAR in Chebyshev basis and for arbitrary degree PEPs (Kressner & Roman, Num. Lin. Alg. Appl, 2014), etc....
- **Available HPC software:** parallel implementations of TOAR for any degree (including symm. versions) in SLEPc (Campos & Roman, 2016).

- Therefore, Krylov subspace methods for PEPs take advantage of the structure of the linearization and of the bases of their Krylov subspaces
- to obtain **memory** and orthogonalization costs of the same order of those of an $n \times n$ standard matrix problem.
- The most stable and efficient methods in this family are
  1. TOAR (Two level Orthogonal ARnoldi) for QEPs (Su & Bai & Lu, 2008 and SIMAX 2016) based on $C_1(\lambda)$,
  2. CORK (COmpact Rational Krylov) for arbitrary PEPs (Van Beeumen & Meerbergen & Michiels, SIMAX, 2015) very general, it can use many linearizations and bases for expressing the PEP,
- which are the "final optimal product" of previous pioneering algorithms:
  1. SOAR (Second Order ARnoldi) for QEPs (Bai & Su, SIMAX, 2005),
  2. Q-Arnoldi for QEPs (Meerbergen, SIMAX, 2008),
  3. TOAR in Chebyshev basis and for arbitrary degree PEPs (Kressner & Roman, Num. Lin. Alg. Appl, 2014), etc....
- **Available HPC software:** parallel implementations of TOAR for any degree (including symm. versions) in SLEPc (Campos & Roman, 2016).

A few words on algorithms for solving large-scale regular PEPs (II)

- Therefore, Krylov subspace methods for PEPs take advantage of the structure of the linearization and of the bases of their Krylov subspaces
- to obtain **memory** and orthogonalization costs of the same order of those of an $n \times n$ standard matrix problem.
- The most stable and efficient methods in this family are
  1. TOAR (Two level Orthogonal ARnoldi) for QEPs (Su & Bai & Lu, 2008 and SIMAX 2016) based on $C_1(\lambda)$,
  2. CORK (COmpact Rational Krylov) for arbitrary PEPs (Van Beeumen & Meerbergen & Michiels, SIMAX, 2015) very general, it can use many linearizations and bases for expressing the PEP,
- which are the "final optimal product" of previous pioneering algorithms:
  1. SOAR (Second Order ARnoldi) for QEPs (Bai & Su, SIMAX, 2005),
  2. Q-Arnoldi for QEPs (Meerbergen, SIMAX, 2008),
  3. TOAR in Chebyshev basis and for arbitrary degree PEPs (Kressner & Roman, Num. Lin. Alg. Appl, 2014), etc....
- **Available HPC software:** parallel implementations of TOAR for any degree (including symm. versions) in SLEPc (Campos & Roman, 2016).

A few words on algorithms for solving large-scale regular PEPs (II)

- Therefore, Krylov subspace methods for PEPs take advantage of the structure of the linearization and of the bases of their Krylov subspaces
- to obtain **memory** and orthogonalization costs of the same order of those of an $n \times n$ standard matrix problem.
- The most stable and efficient methods in this family are
  1. TOAR (Two level Orthogonal ARnoldi) for QEPs (Su & Bai & Lu, 2008 and SIMAX 2016) based on $C_1(\lambda)$,
  2. CORK (COmpact Rational Krylov) for arbitrary PEPs (Van Beeumen & Meerbergen & Michiels, SIMAX, 2015) very general, it can use many linearizations and bases for expressing the PEP,
- which are the "final optimal product" of previous pioneering algorithms:
  1. SOAR (Second Order ARnoldi) for QEPs (Bai & Su, SIMAX, 2005),
  2. Q-Arnoldi for QEPs (Meerbergen, SIMAX, 2008),
  3. TOAR in Chebyshev basis and for arbitrary degree PEPs (Kressner & Roman, Num. Lin. Alg. Appl, 2014), etc....
- **Available HPC software:** parallel implementations of TOAR for any degree (including symm. versions) in SLEPc (Campos & Roman, 2016).

A few words on algorithms for solving large-scale regular PEPs (III)

The typical key result that is behind "memory-efficient" Krylov methods for PEPs is:

Theorem

Let $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0 \in \mathbb{C}[\lambda]^{n\times n}$ *with* P_d *nonsingular and* $C_1(\lambda) =: A - \lambda B \in \mathbb{C}^{nd\times nd}$ *be its first Frobenius companion form. Let the columns of*

$$V_j = \begin{bmatrix} V_j^{(1)} \\ V_j^{(2)} \\ \vdots \\ V_j^{(d)} \end{bmatrix} \in \mathbb{C}^{nd\times j}$$ *be orthonormal basis of span*$\{v, B^{-1}Av, \ldots, (B^{-1}A)^{j-1}v\}$.

where $V_j^{(\ell)} \in \mathbb{C}^{n\times j}$. *Then,* $\operatorname{rank}[V_j^{(1)}\ V_j^{(2)}\ \cdots\ V_j^{(d)}] < j + d$.

This implies that there exists $Q_j \in \mathbb{C}^{n\times r_j}$ with orthonormal columns and $r_j < d + j$ s.t.

$$V_j^{(\ell)} = Q_j R_j^{(\ell)}, \text{ for } \ell = 1, \ldots, d \quad \text{and} \quad R_j^{(\ell)} \in \mathbb{C}^{r_j\times j}$$

and that V_j **can be stored with** $< n(d+j) + d(d+j)j$ parameters instead of ndj.

Idea: to rewrite the Arnoldi iteration in terms of Q_j and $R_j^{(\ell)}$ without using V_j.

The typical key result that is behind “memory-efficient” Krylov methods for PEPs is:

Theorem

Let $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0 \in \mathbb{C}[\lambda]^{n\times n}$ *with* P_d *nonsingular and* $C_1(\lambda) =: A - \lambda B \in \mathbb{C}^{nd\times nd}$ *be its first Frobenius companion form. Let the columns of*

$$V_j = \begin{bmatrix} V_j^{(1)} \\ V_j^{(2)} \\ \vdots \\ V_j^{(d)} \end{bmatrix} \in \mathbb{C}^{nd\times j}$$ *be orthonormal basis of* span$\{v, B^{-1}Av, \ldots, (B^{-1}A)^{j-1}v\}$.

where $V_j^{(\ell)} \in \mathbb{C}^{n\times j}$. *Then,* $\operatorname{rank}\,[V_j^{(1)}\ V_j^{(2)}\ \cdots\ V_j^{(d)}] < j + d$.

This implies that there exists $Q_j \in \mathbb{C}^{n\times r_j}$ with orthonormal columns and $r_j < d + j$ s.t.

$$V_j^{(\ell)} = Q_j\, R_j^{(\ell)}, \text{ for } \ell = 1, \ldots, d \quad \text{and} \quad R_j^{(\ell)} \in \mathbb{C}^{r_j\times j}$$

and that V_j **can be stored with** $< n(d+j) + d(d+j)j$ parameters instead of ndj.

Idea: to rewrite the Arnoldi iteration in terms of Q_j and $R_j^{(\ell)}$ without using V_j.

A few words on algorithms for solving large-scale regular PEPs (III)

The typical key result that is behind "memory-efficient" Krylov methods for PEPs is:

Theorem

Let $P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0 \in \mathbb{C}[\lambda]^{n\times n}$ *with* P_d *nonsingular and* $C_1(\lambda) =: A - \lambda B \in \mathbb{C}^{nd\times nd}$ *be its first Frobenius companion form. Let the columns of*

$$V_j = \begin{bmatrix} V_j^{(1)} \\ V_j^{(2)} \\ \vdots \\ V_j^{(d)} \end{bmatrix} \in \mathbb{C}^{nd\times j}$$ *be orthonormal basis of span*$\{v, B^{-1}Av, \ldots, (B^{-1}A)^{j-1}v\}$.

where $V_j^{(\ell)} \in \mathbb{C}^{n\times j}$. *Then,* $\operatorname{rank}[V_j^{(1)}\ V_j^{(2)}\ \cdots\ V_j^{(d)}] < j + d$.

This implies that there exists $Q_j \in \mathbb{C}^{n\times r_j}$ with orthonormal columns and $r_j < d + j$ s.t.

$$V_j^{(\ell)} = Q_j\, R_j^{(\ell)}, \text{ for } \ell = 1, \ldots, d \quad \text{and} \quad R_j^{(\ell)} \in \mathbb{C}^{r_j\times j}$$

and that V_j **can be stored with** $< n(d+j) + d(d+j)j$ parameters instead of ndj.

Idea: to rewrite the Arnoldi iteration in terms of Q_j and $R_j^{(\ell)}$ without using V_j.

Outline

1. Basics on Polynomial Eigenvalue Problems (PEPs)
2. Numerical solution of PEPs through linearizations
3. **Other methods for solving PEPs without linearizations**
4. Global backward error problem for PEPs solved with linearizations
5. Block Kronecker pencils
6. The solution of the perturbation problem
7. The structured global backward error result
8. Conclusions

One can use for PEPs methods for general NLEPs

NLEP = "nonlinear eigenvalue problem". In large-scale setting and/or with refinement purposes, these methods can be applied to PEPs. Most of them require to evaluate $P(\lambda)$ and some very often.

- (Quasi)-Newton methods for systems nonlinear equations on $P(\lambda)v = 0$ (classical topic, recent survey-results by Jarlebring et al. 2017).
- Block Newton methods for approximating more than one eigenpair simultaneously (Kressner, 2009; Effenberger, 2013).
- Residual inverse iteration (Neumaier, 1985) (is a Quasi-Newton).
- Nonlinear Rayleigh-Ritz iterative method (Liao, Bai, Lee, Ko, 2010).
- Nonlinear Arnoldi, Nonlinear Jacobi-Davidson (Voss, Betcke, 2004, ...)
- Contour integral methods for finding all eigenvalues inside a certain region: good parallelism properties, if there are k eigenvalues the problem is reduced to $k \times k$ linear problem via evaluation of two contour integrals. (Sakurai et al., 2003, 2009, 2013; Beyn, 2012)
- **Available HPC software:** parallel implementations of (block)-Newton, polynomial Jacobi-Davidson in SLEPc (Campos & Roman, 2016).

NLEP = "nonlinear eigenvalue problem". In large-scale setting and/or with refinement purposes, these methods can be applied to PEPs. Most of them require to evaluate $P(\lambda)$ and some very often.

- (Quasi)-Newton methods for systems nonlinear equations on $P(\lambda)v = 0$ (classical topic, recent survey-results by Jarlebring et al. 2017).
- Block Newton methods for approximating more than one eigenpair simultaneously (Kressner, 2009; Effenberger, 2013).
- Residual inverse iteration (Neumaier, 1985) (is a Quasi-Newton).
- Nonlinear Rayleigh-Ritz iterative method (Liao, Bai, Lee, Ko, 2010).
- Nonlinear Arnoldi, Nonlinear Jacobi-Davidson (Voss, Betcke, 2004, ...)
- Contour integral methods for finding all eigenvalues inside a certain region: good parallelism properties, if there are k eigenvalues the problem is reduced to $k \times k$ linear problem via evaluation of two contour integrals. (Sakurai et al., 2003, 2009, 2013; Beyn, 2012)
- **Available HPC software:** parallel implementations of (block)-Newton, polynomial Jacobi-Davidson in SLEPc (Campos & Roman, 2016).

Outline

- We consider a **general** $m \times n$ **matrix polynomial**, square or rectangular, regular or singular,

$$P(\lambda) = P_d \lambda^d + \cdots + P_1 \lambda + P_0, \qquad P_i \in \mathbb{C}^{m \times n},$$

- and we assume that its **complete eigenstructure**
- has been computed by applying a **backward stable algorithm** (QZ for regular, Staircase for singular)
- to a **strong linearization** $\mathcal{L}(\lambda)$ of $P(\lambda)$
- that allows us to **recover the minimal indices** of $P(\lambda)$ from those of $\mathcal{L}(\lambda)$ via **uniform shifts**.
- In this talk, we restrict most of the results to the **new wide class of block Kronecker linearizations**.

The Setting

- We consider a **general** $m \times n$ **matrix polynomial**, square or rectangular, regular or singular,

$$P(\lambda) = P_d \lambda^d + \cdots + P_1 \lambda + P_0, \qquad P_i \in \mathbb{C}^{m \times n},$$

- and we assume that its **complete eigenstructure**
- has been computed by applying a **backward stable algorithm** (QZ for regular, Staircase for singular)
- to a **strong linearization** $\mathcal{L}(\lambda)$ of $P(\lambda)$
- that allows us to **recover the minimal indices** of $P(\lambda)$ from those of $\mathcal{L}(\lambda)$ via **uniform shifts**.
- In this talk, we restrict most of the results to the **new wide class of block Kronecker linearizations**.

- We consider a **general** $m \times n$ **matrix polynomial**, square or rectangular, regular or singular,

$$P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0 \ , \qquad P_i \in \mathbb{C}^{m\times n},$$

- and we assume that its **complete eigenstructure**
- has been computed by applying a **backward stable algorithm** (QZ for regular, Staircase for singular)
- to a **strong linearization** $\mathcal{L}(\lambda)$ of $P(\lambda)$
- that allows us to **recover the minimal indices** of $P(\lambda)$ from those of $\mathcal{L}(\lambda)$ via **uniform shifts**.
- In this talk, we restrict most of the results to the **new wide class of block Kronecker linearizations**.

- We consider a **general** $m \times n$ **matrix polynomial**, square or rectangular, regular or singular,

$$P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0, \qquad P_i \in \mathbb{C}^{m\times n},$$

- and we assume that its **complete eigenstructure**
- has been computed by applying a **backward stable algorithm** (QZ for regular, Staircase for singular)
- to a **strong linearization** $\mathcal{L}(\lambda)$ of $P(\lambda)$
- that allows us to **recover the minimal indices** of $P(\lambda)$ from those of $\mathcal{L}(\lambda)$ via **uniform shifts**.
- In this talk, we restrict most of the results to the **new wide class of block Kronecker linearizations**.

- We consider a **general** $m \times n$ **matrix polynomial**, square or rectangular, regular or singular,

$$P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0 , \qquad P_i \in \mathbb{C}^{m\times n},$$

- and we assume that its **complete eigenstructure**
- has been computed by applying a **backward stable algorithm** (QZ for regular, Staircase for singular)
- to a **strong linearization** $\mathcal{L}(\lambda)$ of $P(\lambda)$
- that allows us to **recover the minimal indices** of $P(\lambda)$ from those of $\mathcal{L}(\lambda)$ via **uniform shifts**.
- In this talk, we restrict most of the results to the **new wide class of block Kronecker linearizations**.

The Setting

- We consider a **general** $m \times n$ **matrix polynomial**, square or rectangular, regular or singular,

$$P(\lambda) = P_d\lambda^d + \cdots + P_1\lambda + P_0 , \qquad P_i \in \mathbb{C}^{m\times n},$$

- and we assume that its **complete eigenstructure**
- has been computed by applying a **backward stable algorithm** (QZ for regular, Staircase for singular)
- to a **strong linearization** $\mathcal{L}(\lambda)$ of $P(\lambda)$
- that allows us to **recover the minimal indices** of $P(\lambda)$ from those of $\mathcal{L}(\lambda)$ via **uniform shifts**.
- In this talk, we restrict most of the results to the **new wide class of block Kronecker linearizations**.

- *The computed complete eigenstructure of $\mathcal{L}(\lambda)$ is the exact complete eigenstructure of a matrix pencil $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ such that*

$$\frac{\|\Delta\mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F} = O(\mathbf{u}),$$

  where $\mathbf{u} \approx 10^{-16}$ is the unit roundoff and

- $\|\cdot\|_F$ is the Frobenius norm, i.e., for any matrix polynomial

$$\|Q_k\lambda^k + \cdots + Q_1\lambda + Q_0\|_F = \sqrt{\|Q_k\|_F^2 + \cdots + \|Q_1\|_F^2 + \|Q_0\|_F^2}\,.$$

- But, does this imply that (after shifting properly the minimal indices) the computed complete eigenstructure of $P(\lambda)$ is the exact complete eigenstructure of a matrix polynomial of the same degree $P(\lambda) + \Delta P(\lambda)$ such that

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} = O(\mathbf{u})\ ??$$

- For solving this question, we pose the following **theoretical problems** of matrix perturbation theory.

- *The computed complete eigenstructure of $\mathcal{L}(\lambda)$ is the exact complete eigenstructure of a matrix pencil $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ such that*

$$\frac{\|\Delta\mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F} = O(\mathbf{u}),$$

  where $\mathbf{u} \approx 10^{-16}$ is the unit roundoff and

- $\|\cdot\|_F$ is the Frobenius norm, i.e., for any matrix polynomial

$$\|Q_k\lambda^k + \cdots + Q_1\lambda + Q_0\|_F = \sqrt{\|Q_k\|_F^2 + \cdots + \|Q_1\|_F^2 + \|Q_0\|_F^2}\ .$$

- But, does this imply that (after shifting properly the minimal indices) the computed complete eigenstructure of $P(\lambda)$ is the exact complete eigenstructure of a matrix polynomial of the same degree $P(\lambda) + \Delta P(\lambda)$ such that

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} = O(\mathbf{u}) \text{ ??}$$

- For solving this question, we pose the following **theoretical problems** of matrix perturbation theory.

- *The computed complete eigenstructure of $\mathcal{L}(\lambda)$ is the exact complete eigenstructure of a matrix pencil $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ such that*

$$\frac{\|\Delta\mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F} = O(\mathbf{u}),$$

  where $\mathbf{u} \approx 10^{-16}$ is the unit roundoff and

- $\|\cdot\|_F$ is the Frobenius norm, i.e., for any matrix polynomial

$$\|Q_k\lambda^k + \cdots + Q_1\lambda + Q_0\|_F = \sqrt{\|Q_k\|_F^2 + \cdots + \|Q_1\|_F^2 + \|Q_0\|_F^2}\ .$$

- But, does this imply that (after shifting properly the minimal indices) the computed complete eigenstructure of $P(\lambda)$ is the exact complete eigenstructure of a matrix polynomial of the same degree $P(\lambda) + \Delta P(\lambda)$ such that

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} = O(\mathbf{u})\ ??$$

- For solving this question, we pose the following **theoretical problems** of matrix perturbation theory.

- *The computed complete eigenstructure of $\mathcal{L}(\lambda)$ is the exact complete eigenstructure of a matrix pencil $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ such that*

$$\frac{\|\Delta\mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F} = O(\mathbf{u}),$$

where $\mathbf{u} \approx 10^{-16}$ is the unit roundoff and

- $\| \cdot \|_F$ is the Frobenius norm, i.e., for any matrix polynomial

$$\|Q_k\lambda^k + \cdots + Q_1\lambda + Q_0\|_F = \sqrt{\|Q_k\|_F^2 + \cdots + \|Q_1\|_F^2 + \|Q_0\|_F^2}\,.$$

- But, does this imply that (after shifting properly the minimal indices) the computed complete eigenstructure of $P(\lambda)$ is the exact complete eigenstructure of a matrix polynomial of the same degree $P(\lambda) + \Delta P(\lambda)$ such that

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} = O(\mathbf{u}) \text{ ??}$$

- For solving this question, we pose the following **theoretical problems** of matrix perturbation theory.

- **Data:**
  1. Matrix polynomial $P(\lambda)$ of degree d.
  2. Strong linearization $\mathcal{L}(\lambda)$ of $P(\lambda)$ enjoying uniform shift-relations for the minimal indices.
  3. Perturbation pencil $\Delta\mathcal{L}(\lambda)$.
- **Problem 1:** To establish conditions on $\|\Delta\mathcal{L}(\lambda)\|_F$ such that $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ is a strong linearization for some matrix polynomial $P(\lambda) + \Delta P(\lambda)$ of degree d, and such that
- **Problem 2:** the shift-relations between minimal indices of $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ and $P(\lambda) + \Delta P(\lambda)$ are equal to those between $\mathcal{L}(\lambda)$ and $P(\lambda)$.
- **Problem 3:** To prove a perturbation bound

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \leq C_{P,\mathcal{L}} \frac{\|\Delta\mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F},$$

  with $C_{P,\mathcal{L}}$ a number depending on $P(\lambda)$ and $\mathcal{L}(\lambda)$.
- For those $P(\lambda)$ and $\mathcal{L}(\lambda)$ s.t. $C_{P,\mathcal{L}}$ is moderate, to use global backward stable algorithms on $\mathcal{L}(\lambda)$ gives global backward stability for $P(\lambda)$, i.e., from the polynomial point of view.

- **Data:**
  1. Matrix polynomial $P(\lambda)$ of degree d.
  2. Strong linearization $\mathcal{L}(\lambda)$ of $P(\lambda)$ enjoying uniform shift-relations for the minimal indices.
  3. Perturbation pencil $\Delta\mathcal{L}(\lambda)$.
- **Problem 1:** To establish conditions on $\|\Delta\mathcal{L}(\lambda)\|_F$ such that $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ is a strong linearization for some matrix polynomial $P(\lambda) + \Delta P(\lambda)$ of degree d, and such that
- **Problem 2:** the shift-relations between minimal indices of $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ and $P(\lambda) + \Delta P(\lambda)$ are equal to those between $\mathcal{L}(\lambda)$ and $P(\lambda)$.
- **Problem 3:** To prove a perturbation bound

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \leq C_{P,\mathcal{L}} \frac{\|\Delta\mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F},$$

  with $C_{P,\mathcal{L}}$ a number depending on $P(\lambda)$ and $\mathcal{L}(\lambda)$.
- For those $P(\lambda)$ and $\mathcal{L}(\lambda)$ s.t. $C_{P,\mathcal{L}}$ is moderate, to use global backward stable algorithms on $\mathcal{L}(\lambda)$ gives global backward stability for $P(\lambda)$, i.e., from the polynomial point of view.

- **Data:**
  1. Matrix polynomial $P(\lambda)$ of degree d.
  2. Strong linearization $\mathcal{L}(\lambda)$ of $P(\lambda)$ enjoying uniform shift-relations for the minimal indices.
  3. Perturbation pencil $\Delta\mathcal{L}(\lambda)$.
- **Problem 1:** To establish conditions on $\|\Delta\mathcal{L}(\lambda)\|_F$ such that $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ is a strong linearization for some matrix polynomial $P(\lambda) + \Delta P(\lambda)$ of degree d, and such that
- **Problem 2:** the shift-relations between minimal indices of $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ and $P(\lambda) + \Delta P(\lambda)$ are equal to those between $\mathcal{L}(\lambda)$ and $P(\lambda)$.
- **Problem 3:** To prove a perturbation bound

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \leq C_{P,\mathcal{L}} \frac{\|\Delta\mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F},$$

  with $C_{P,\mathcal{L}}$ a number depending on $P(\lambda)$ and $\mathcal{L}(\lambda)$.
- For those $P(\lambda)$ and $\mathcal{L}(\lambda)$ s.t. $C_{P,\mathcal{L}}$ is moderate, to use global backward stable algorithms on $\mathcal{L}(\lambda)$ gives global backward stability for $P(\lambda)$, i.e., from the polynomial point of view.

- **Data:**
  1. Matrix polynomial $P(\lambda)$ of degree d.
  2. Strong linearization $\mathcal{L}(\lambda)$ of $P(\lambda)$ enjoying uniform shift-relations for the minimal indices.
  3. Perturbation pencil $\Delta\mathcal{L}(\lambda)$.
- **Problem 1:** To establish conditions on $\|\Delta\mathcal{L}(\lambda)\|_F$ such that $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ is a strong linearization for some matrix polynomial $P(\lambda) + \Delta P(\lambda)$ of degree d, and such that
- **Problem 2:** the shift-relations between minimal indices of $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ and $P(\lambda) + \Delta P(\lambda)$ are equal to those between $\mathcal{L}(\lambda)$ and $P(\lambda)$.
- **Problem 3:** To prove a perturbation bound

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \leq C_{P,\mathcal{L}} \frac{\|\Delta\mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F},$$

  with $C_{P,\mathcal{L}}$ a number depending on $P(\lambda)$ and $\mathcal{L}(\lambda)$.
- For those $P(\lambda)$ and $\mathcal{L}(\lambda)$ s.t. $C_{P,\mathcal{L}}$ is moderate, to use global backward stable algorithms on $\mathcal{L}(\lambda)$ gives global backward stability for $P(\lambda)$, i.e., from the polynomial point of view.

- **Data:**
  1. Matrix polynomial $P(\lambda)$ of degree d.
  2. Strong linearization $\mathcal{L}(\lambda)$ of $P(\lambda)$ enjoying uniform shift-relations for the minimal indices.
  3. Perturbation pencil $\Delta\mathcal{L}(\lambda)$.
- **Problem 1:** To establish conditions on $\|\Delta\mathcal{L}(\lambda)\|_F$ such that $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ is a strong linearization for some matrix polynomial $P(\lambda) + \Delta P(\lambda)$ of degree d, and such that
- **Problem 2:** the shift-relations between minimal indices of $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ and $P(\lambda) + \Delta P(\lambda)$ are equal to those between $\mathcal{L}(\lambda)$ and $P(\lambda)$.
- **Problem 3:** To prove a perturbation bound

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \leq C_{P,\mathcal{L}} \frac{\|\Delta\mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F},$$

  with $C_{P,\mathcal{L}}$ a number depending on $P(\lambda)$ and $\mathcal{L}(\lambda)$.
- For those $P(\lambda)$ and $\mathcal{L}(\lambda)$ s.t. $C_{P,\mathcal{L}}$ is moderate, to use global backward stable algorithms on $\mathcal{L}(\lambda)$ gives global backward stability for $P(\lambda)$, i.e., from the polynomial point of view.

The perturbation analysis we present for these problems...

...has a number of key features which are not present in any other analyses published so far:

1. for the first time, it is NOT a first order analysis, since it is a rigorous analysis valid for perturbations $\Delta\mathcal{L}(\lambda)$ of finite norm,
2. it provides very detailed bounds, and not just vague big-O bounds as other analyses do,
3. it is valid simultaneously for all the linearizations in the very large new class of block Kronecker pencils, which includes Frobenius and many other "famous linearizations" for which this type of backward error analyses were not available before our work,
4. it establishes a framework that probably can be generalized to other classes of linearizations.

The perturbation analysis we present for these problems...

...has a number of key features which are not present in any other analyses published so far:

1. for the first time, it is NOT a first order analysis, since it is a rigorous analysis valid for perturbations $\Delta\mathcal{L}(\lambda)$ of finite norm,
2. it provides very detailed bounds, and not just vague big-O bounds as other analyses do,
3. it is valid simultaneously for all the linearizations in the very large new class of block Kronecker pencils, which includes Frobenius and many other "famous linearizations" for which this type of backward error analyses were not available before our work,
4. it establishes a framework that probably can be generalized to other classes of linearizations.

The perturbation analysis we present for these problems...

...has a number of key features which are not present in any other analyses published so far:

1. for the first time, it is NOT a first order analysis, since it is a rigorous analysis valid for perturbations $\Delta\mathcal{L}(\lambda)$ of finite norm,
2. it provides very detailed bounds, and not just vague big-O bounds as other analyses do,
3. it is valid simultaneously for all the linearizations in the very large new class of block Kronecker pencils, which includes Frobenius and many other "famous linearizations" for which this type of backward error analyses were not available before our work,
4. it establishes a framework that probably can be generalized to other classes of linearizations.

The perturbation analysis we present for these problems...

...has a number of key features which are not present in any other analyses published so far:

1. for the first time, it is NOT a first order analysis, since it is a rigorous analysis valid for perturbations $\Delta\mathcal{L}(\lambda)$ of finite norm,
2. it provides very detailed bounds, and not just vague big-O bounds as other analyses do,
3. it is valid simultaneously for all the linearizations in the very large new class of block Kronecker pencils, which includes Frobenius and many other “famous linearizations” for which this type of backward error analyses were not available before our work,
4. it establishes a framework that probably can be generalized to other classes of linearizations.

...has a number of key features which are not present in any other analyses published so far:

1. for the first time, it is NOT a first order analysis, since it is a rigorous analysis valid for perturbations $\Delta\mathcal{L}(\lambda)$ of finite norm,
2. it provides very detailed bounds, and not just vague big-O bounds as other analyses do,
3. it is valid simultaneously for all the linearizations in the very large new class of block Kronecker pencils, which includes Frobenius and many other "famous linearizations" for which this type of backward error analyses were not available before our work,
4. it establishes a framework that probably can be generalized to other classes of linearizations.

Previous works on this type of “global” backward error analyses

There are just a few: only first order results, only for Frobenius linearizations or their counterparts in other bases, often only valid for regular polynomials, or do not pay attention to minimal indices...

- Van Dooren & De Wilde (LAA 1983).
- Edelman & Murakami (Math. Comp. 1995).
- Lawrence & Corless (SIMAX 2015).
- Lawrence & Van Barel & Van Dooren (SIMAX 2016).
- Noferini & Pérez (Math. Comp., 2017).

A few comments on "local" backward error analyses of PEPs (I)

- The QZ algorithm for regular GEPs $A - \lambda B$ gives a stronger backward error result than mentioned before
- since computes the complete set of eigenvalues of $(A + \Delta A) - \lambda(B + \Delta B)$ with

$$\frac{\|\Delta A\|_F}{\|A\|_F} = O(\mathbf{u}) \quad \text{and} \quad \frac{\|\Delta B\|_F}{\|B\|_F} = O(\mathbf{u}),$$

- i.e., with "relative coefficientwise" backward stability.
- Therefore, it might seem natural to ask for the same type of "relative coefficientwise" backward stability in the numerical solution of higher degrees PEPs,
- but it has been proved that **it is impossible to guarantee this stability** (Mastronardi & Van Dooren, ETNA, 2015):

"There does not exist any algorithm that computes in floating point arithmetic the two roots of a quadratic scalar polynomial and that guarantees a priori that the computed two roots are the exact two roots of a nearby polynomial with a coefficientwise backward error of $O(\mathbf{u})$."

- The QZ algorithm for regular GEPs $A - \lambda B$ gives a stronger backward error result than mentioned before
- since computes the complete set of eigenvalues of $(A + \Delta A) - \lambda(B + \Delta B)$ with

$$\frac{\|\Delta A\|_F}{\|A\|_F} = O(\mathbf{u}) \quad \text{and} \quad \frac{\|\Delta B\|_F}{\|B\|_F} = O(\mathbf{u}),$$

- i.e., with "relative coefficientwise" backward stability.
- Therefore, it might seem natural to ask for the same type of "relative coefficientwise" backward stability in the numerical solution of higher degrees PEPs,
- but it has been proved that **it is impossible to guarantee this stability** (Mastronardi & Van Dooren, ETNA, 2015):

"There does not exist any algorithm that computes in floating point arithmetic the two roots of a quadratic scalar polynomial and that guarantees a priori that the computed two roots are the exact two roots of a nearby polynomial with a coefficientwise backward error of $O(\mathbf{u})$."

A few comments on "local" backward error analyses of PEPs (I)

- The QZ algorithm for regular GEPs $A - \lambda B$ gives a stronger backward error result than mentioned before
- since computes the complete set of eigenvalues of $(A + \Delta A) - \lambda(B + \Delta B)$ with

$$\frac{\|\Delta A\|_F}{\|A\|_F} = O(\mathbf{u}) \quad \text{and} \quad \frac{\|\Delta B\|_F}{\|B\|_F} = O(\mathbf{u}),$$

- i.e., with "relative coefficientwise" backward stability.
- Therefore, it might seem natural to ask for the same type of "relative coefficientwise" backward stability in the numerical solution of higher degrees PEPs,
- but it has been proved that **it is impossible to guarantee this stability** (Mastronardi & Van Dooren, ETNA, 2015):

"There does not exist any algorithm that computes in floating point arithmetic the two roots of a quadratic scalar polynomial and that guarantees a priori that the computed two roots are the exact two roots of a nearby polynomial with a coefficientwise backward error of $O(\mathbf{u})$."

A few comments on “local” backward error analyses of PEPs (I)

- The QZ algorithm for regular GEPs $A - \lambda B$ gives a stronger backward error result than mentioned before
- since computes the complete set of eigenvalues of $(A + \Delta A) - \lambda(B + \Delta B)$ with

$$\frac{\|\Delta A\|_F}{\|A\|_F} = O(\mathbf{u}) \quad \text{and} \quad \frac{\|\Delta B\|_F}{\|B\|_F} = O(\mathbf{u}),$$

- i.e., with “relative coefficientwise” backward stability.
- Therefore, it might seem natural to ask for the same type of “relative coefficientwise” backward stability in the numerical solution of higher degrees PEPs,
- but it has been proved that **it is impossible to guarantee this stability** (Mastronardi & Van Dooren, ETNA, 2015):

  “There does not exist any algorithm that computes in floating point arithmetic the two roots of a quadratic scalar polynomial and that guarantees a priori that the computed two roots are the exact two roots of a nearby polynomial with a coefficientwise backward error of $O(\mathbf{u})$.”

A few comments on "local" backward error analyses of PEPs (I)

- The QZ algorithm for regular GEPs $A - \lambda B$ gives a stronger backward error result than mentioned before
- since computes the complete set of eigenvalues of $(A + \Delta A) - \lambda(B + \Delta B)$ with

$$\frac{\|\Delta A\|_F}{\|A\|_F} = O(\mathbf{u}) \quad \text{and} \quad \frac{\|\Delta B\|_F}{\|B\|_F} = O(\mathbf{u}),$$

- i.e., with "relative coefficientwise" backward stability.
- Therefore, it might seem natural to ask for the same type of "relative coefficientwise" backward stability in the numerical solution of higher degrees PEPs,
- but it has been proved that **it is impossible to guarantee this stability** (Mastronardi & Van Dooren, ETNA, 2015):

  "There does not exist any algorithm that computes in floating point arithmetic the two roots of a quadratic scalar polynomial and that guarantees a priori that the computed two roots are the exact two roots of a nearby polynomial with a coefficientwise backward error of $O(\mathbf{u})$."

- The QZ algorithm for regular GEPs $A - \lambda B$ gives a stronger backward error result than mentioned before
- since computes the complete set of eigenvalues of $(A + \Delta A) - \lambda(B + \Delta B)$ with

$$\frac{\|\Delta A\|_F}{\|A\|_F} = O(\mathbf{u}) \quad \text{and} \quad \frac{\|\Delta B\|_F}{\|B\|_F} = O(\mathbf{u}),$$

- i.e., with "relative coefficientwise" backward stability.
- Therefore, it might seem natural to ask for the same type of "relative coefficientwise" backward stability in the numerical solution of higher degrees PEPs,
- but it has been proved that **it is impossible to guarantee this stability** (Mastronardi & Van Dooren, ETNA, 2015):
  "There does not exist any algorithm that computes in floating point arithmetic the two roots of a quadratic scalar polynomial and that guarantees a priori that the computed two roots are the exact two roots of a nearby polynomial with a coefficientwise backward error of $O(\mathbf{u})$."

- This is disappointing because there are applied regular PEPs (in particular QEPs) with coefficients of very different magnitudes and “relative coefficientwise” backward stability is desirable in such cases.
- An option to circumvent this problem is to try to guarantee a priori only **tiny “local” “relative coefficientwise” backward errors**, i.e.,
- that **each particular computed eigenpair** is the exact eigenpair of a nearby matrix polynomial with a coefficientwise backward error of $O(\mathbf{u})$
- **with a different nearby polynomial for each eigenpair**.
- Several authors have worked on this approach: Tisseur (LAA 2000), Higham & Li & Tisseur (SIMAX 2007), Li & Lin & Wang (Numer. Math. 2010), Hammarling & Munro & Tisseur (ACMTMatSoftw 2013), Zeng & Su (SIMAX, 2014).
- With additional assumptions and scalings, solving twice the PEP with QZ on two different linearizations, these algorithms and analyses may guarantee a priori **coefficient-wise “local” backward stability only for regular QEPs**.
- The problem remains open in higher degrees.

A few comments on "local" backward error analyses of PEPs (II)

- This is disappointing because there are applied regular PEPs (in particular QEPs) with coefficients of very different magnitudes and "relative coefficientwise" backward stability is desirable in such cases.
- An option to circumvent this problem is to try to guarantee a priori only **tiny "local" "relative coefficientwise" backward errors**, i.e.,
- that **each particular computed eigenpair** is the exact eigenpair of a nearby matrix polynomial with a coefficientwise backward error of $O(\mathbf{u})$
- **with a different nearby polynomial for each eigenpair**.
- Several authors have worked on this approach: Tisseur (LAA 2000), Higham & Li & Tisseur (SIMAX 2007), Li & Lin & Wang (Numer. Math. 2010), Hammarling & Munro & Tisseur (ACMTMatSoftw 2013), Zeng & Su (SIMAX, 2014).
- With additional assumptions and scalings, solving twice the PEP with QZ on two different linearizations, these algorithms and analyses may guarantee a priori **coefficient-wise "local" backward stability only for regular QEPs**.
- The problem remains open in higher degrees.

A few comments on "local" backward error analyses of PEPs (II)

- This is disappointing because there are applied regular PEPs (in particular QEPs) with coefficients of very different magnitudes and "relative coefficientwise" backward stability is desirable in such cases.
- An option to circumvent this problem is to try to guarantee a priori only **tiny "local" "relative coefficientwise" backward errors**, i.e.,
- that **each particular computed eigenpair** is the exact eigenpair of a nearby matrix polynomial with a coefficientwise backward error of $O(\mathbf{u})$
- **with a different nearby polynomial for each eigenpair**.
- Several authors have worked on this approach: Tisseur (LAA 2000), Higham & Li & Tisseur (SIMAX 2007), Li & Lin & Wang (Numer. Math. 2010), Hammarling & Munro & Tisseur (ACMTMatSoftw 2013), Zeng & Su (SIMAX, 2014).
- With additional assumptions and scalings, solving twice the PEP with QZ on two different linearizations, these algorithms and analyses may guarantee a priori **coefficient-wise "local" backward stability only for regular QEPs**.
- The problem remains open in higher degrees.

A few comments on "local" backward error analyses of PEPs (II)

- This is disappointing because there are applied regular PEPs (in particular QEPs) with coefficients of very different magnitudes and "relative coefficientwise" backward stability is desirable in such cases.
- An option to circumvent this problem is to try to guarantee a priori only **tiny "local" "relative coefficientwise" backward errors**, i.e.,
- that **each particular computed eigenpair** is the exact eigenpair of a nearby matrix polynomial with a coefficientwise backward error of $O(\mathbf{u})$
- **with a different nearby polynomial for each eigenpair**.
- Several authors have worked on this approach: Tisseur (LAA 2000), Higham & Li & Tisseur (SIMAX 2007), Li & Lin & Wang (Numer. Math. 2010), Hammarling & Munro & Tisseur (ACMTMatSoftw 2013), Zeng & Su (SIMAX, 2014).
- With additional assumptions and scalings, solving twice the PEP with QZ on two different linearizations, these algorithms and analyses may guarantee a priori **coefficient-wise "local" backward stability only for regular QEPs**.
- The problem remains open in higher degrees.

- This is disappointing because there are applied regular PEPs (in particular QEPs) with coefficients of very different magnitudes and "relative coefficientwise" backward stability is desirable in such cases.
- An option to circumvent this problem is to try to guarantee a priori only **tiny "local" "relative coefficientwise" backward errors**, i.e.,
- that **each particular computed eigenpair** is the exact eigenpair of a nearby matrix polynomial with a coefficientwise backward error of $O(\mathbf{u})$
- **with a different nearby polynomial for each eigenpair**.
- Several authors have worked on this approach: Tisseur (LAA 2000), Higham & Li & Tisseur (SIMAX 2007), Li & Lin & Wang (Numer. Math. 2010), Hammarling & Munro & Tisseur (ACMTMatSoftw 2013), Zeng & Su (SIMAX, 2014).
- With additional assumptions and scalings, solving twice the PEP with QZ on two different linearizations, these algorithms and analyses may guarantee a priori **coefficient-wise "local" backward stability only for regular QEPs**.
- The problem remains open in higher degrees.

A few comments on "local" backward error analyses of PEPs (II)

- This is disappointing because there are applied regular PEPs (in particular QEPs) with coefficients of very different magnitudes and "relative coefficientwise" backward stability is desirable in such cases.
- An option to circumvent this problem is to try to guarantee a priori only **tiny "local" "relative coefficientwise" backward errors**, i.e.,
- that **each particular computed eigenpair** is the exact eigenpair of a nearby matrix polynomial with a coefficientwise backward error of $O(\mathbf{u})$
- **with a different nearby polynomial for each eigenpair**.
- Several authors have worked on this approach: Tisseur (LAA 2000), Higham & Li & Tisseur (SIMAX 2007), Li & Lin & Wang (Numer. Math. 2010), Hammarling & Munro & Tisseur (ACMTMatSoftw 2013), Zeng & Su (SIMAX, 2014).
- With additional assumptions and scalings, solving twice the PEP with QZ on two different linearizations, these algorithms and analyses may guarantee a priori **coefficient-wise "local" backward stability only for regular QEPs**.
- The problem remains open in higher degrees.

- This is disappointing because there are applied regular PEPs (in particular QEPs) with coefficients of very different magnitudes and "relative coefficientwise" backward stability is desirable in such cases.
- An option to circumvent this problem is to try to guarantee a priori only **tiny "local" "relative coefficientwise" backward errors**, i.e.,
- that **each particular computed eigenpair** is the exact eigenpair of a nearby matrix polynomial with a coefficientwise backward error of $O(\mathbf{u})$
- **with a different nearby polynomial for each eigenpair**.
- Several authors have worked on this approach: Tisseur (LAA 2000), Higham & Li & Tisseur (SIMAX 2007), Li & Lin & Wang (Numer. Math. 2010), Hammarling & Munro & Tisseur (ACMTMatSoftw 2013), Zeng & Su (SIMAX, 2014).
- With additional assumptions and scalings, solving twice the PEP with QZ on two different linearizations, these algorithms and analyses may guarantee a priori **coefficient-wise "local" backward stability only for regular QEPs**.
- The problem remains open in higher degrees.

Outline

Two fundamental auxiliary matrix polynomials in the rest of the talk

$$L_k(\lambda) := \begin{bmatrix} -1 & \lambda & & & \\ & -1 & \lambda & & \\ & & \ddots & \ddots & \\ & & & -1 & \lambda \end{bmatrix} \in \mathbb{C}[\lambda]^{k\times(k+1)},$$

$$\Lambda_k(\lambda)^T := \begin{bmatrix} \lambda^k & \lambda^{k-1} & \cdots & \lambda & 1 \end{bmatrix} \in \mathbb{C}[\lambda]^{1\times(k+1)},$$

and their Kronecker products by identities

$$L_k(\lambda) \otimes I_n := \begin{bmatrix} -I_n & \lambda I_n & & & \\ & -I_n & \lambda I_n & & \\ & & \ddots & \ddots & \\ & & & -I_n & \lambda I_n \end{bmatrix} \in \mathbb{C}[\lambda]^{nk\times n(k+1)},$$

$$\Lambda_k(\lambda)^T \otimes I_n := \begin{bmatrix} \lambda^k I_n & \lambda^{k-1} I_n & \cdots & \lambda I_n & I_n \end{bmatrix} \in \mathbb{C}[\lambda]^{n\times n(k+1)}.$$

Two fundamental auxiliary matrix polynomials in the rest of the talk

$$L_k(\lambda) := \begin{bmatrix} -1 & \lambda & & & \\ & -1 & \lambda & & \\ & & \ddots & \ddots & \\ & & & -1 & \lambda \end{bmatrix} \in \mathbb{C}[\lambda]^{k\times(k+1)},$$

$$\Lambda_k(\lambda)^T := \begin{bmatrix} \lambda^k & \lambda^{k-1} & \cdots & \lambda & 1 \end{bmatrix} \in \mathbb{C}[\lambda]^{1\times(k+1)},$$

and their Kronecker products by identities

$$L_k(\lambda) \otimes I_n := \begin{bmatrix} -I_n & \lambda I_n & & & \\ & -I_n & \lambda I_n & & \\ & & \ddots & \ddots & \\ & & & -I_n & \lambda I_n \end{bmatrix} \in \mathbb{C}[\lambda]^{nk\times n(k+1)},$$

$$\Lambda_k(\lambda)^T \otimes I_n := \begin{bmatrix} \lambda^k I_n & \lambda^{k-1} I_n & \cdots & \lambda I_n & I_n \end{bmatrix} \in \mathbb{C}[\lambda]^{n\times n(k+1)}.$$

Revisiting the Frobenius companion form with these matrices at hand

The Frobenius companion form of the $m \times n$ matrix polynomial $P(\lambda) = P_d \lambda^d + \cdots + P_1 \lambda + P_0$ is

$$C_1(\lambda) := \begin{bmatrix} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ -I_n & \lambda I_n & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda I_n & \\ & & & -I_n & \lambda I_n \end{bmatrix},$$

and can be written as

$$C_1(\lambda) := \left[\begin{array}{ccccc} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ \hline & & L_{d-1}(\lambda) \otimes I_n & & \end{array} \right].$$

The Frobenius companion form of the $m \times n$ matrix polynomial $P(\lambda) = P_d \lambda^d + \cdots + P_1 \lambda + P_0$ is

$$C_1(\lambda) := \begin{bmatrix} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ -I_n & \lambda I_n & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda I_n & \\ & & & -I_n & \lambda I_n \end{bmatrix},$$

and can be written as

$$C_1(\lambda) := \left[\begin{array}{ccccc} \lambda P_d + P_{d-1} & P_{d-2} & \cdots & P_1 & P_0 \\ \hline \multicolumn{5}{c}{L_{d-1}(\lambda) \otimes I_n} \end{array} \right].$$

Definition

Let $\lambda M_1 + M_0$ be an arbitrary pencil. Then any pencil of the form

$$\mathcal{L}(\lambda) = \underbrace{\left[\begin{array}{c|c} \lambda M_1 + M_0 & L_\eta(\lambda)^T \otimes I_m \\ \hline L_\varepsilon(\lambda) \otimes I_n & 0 \end{array}\right]}_{(\varepsilon+1)n \quad\quad \eta m} \begin{array}{l} \}\,(\eta+1)m \\ \}\,\varepsilon n \end{array},$$

is called a block Kronecker pencil (one-block row and column cases included).

Theorem (key theorem of block Kronecker pencils)

Any block Kronecker pencil $\mathcal{L}(\lambda)$ is a strong linearization of the matrix polynomial

$$Q(\lambda) := (\Lambda_\eta(\lambda)^T \otimes I_m)(\lambda M_1 + M_0)(\Lambda_\varepsilon(\lambda) \otimes I_n) \in \mathbb{C}[\lambda]^{m\times n},$$

the right minimal indices of $\mathcal{L}(\lambda)$ are those of $Q(\lambda)$ shifted by ε, and the left minimal indices of $\mathcal{L}(\lambda)$ are those of $Q(\lambda)$ shifted by η.

Definition and key properties of Block Kronecker Pencils

Definition

Let $\lambda M_1 + M_0$ be an arbitrary pencil. Then any pencil of the form

$$\mathcal{L}(\lambda) = \underbrace{\left[\begin{array}{c|c} \lambda M_1 + M_0 & L_\eta(\lambda)^T \otimes I_m \\ \hline L_\varepsilon(\lambda) \otimes I_n & 0 \end{array}\right]}_{(\varepsilon+1)n \quad\quad \eta m} \begin{array}{l} \}\, (\eta+1)m \\ \}\, \varepsilon n \end{array} ,$$

is called a block Kronecker pencil (one-block row and column cases included).

Theorem (key theorem of block Kronecker pencils)

Any block Kronecker pencil $\mathcal{L}(\lambda)$ is a strong linearization of the matrix polynomial

$$Q(\lambda) := (\Lambda_\eta(\lambda)^T \otimes I_m)(\lambda M_1 + M_0)(\Lambda_\varepsilon(\lambda) \otimes I_n) \in \mathbb{C}[\lambda]^{m\times n},$$

the right minimal indices of $\mathcal{L}(\lambda)$ are those of $Q(\lambda)$ shifted by ε, and the left minimal indices of $\mathcal{L}(\lambda)$ are those of $Q(\lambda)$ shifted by η.

Theorem

- *Let* $P(\lambda) = \sum_{k=0}^{d} P_k \lambda^k \in \mathbb{C}[\lambda]^{m\times n}$,
- *let* $\mathcal{L}(\lambda)$ *be a block Kronecker pencil with* $\varepsilon + \eta + 1 = d$, *and*
- *let us consider* M_0 *and* M_1 *partitioned into* $(\eta + 1) \times (\varepsilon + 1)$ *blocks each of size* $m \times n$.

If the sum of the blocks on the $(d-k)$**th block antidiagonal of** M_0 **plus the sum of the blocks on the** $(d-k+1)$**th block antidiagonal of** M_1 **is equal to** P_k, *for* $k = 0, \dots, d$,

then $\mathcal{L}(\lambda)$ **is a strong linearizations of** $P(\lambda)$ **with uniform shift relations (**ε **and** η**) for the minimal indices.**

Examples of block Kronecker pencils (I)

$$P(\lambda) = \lambda^5 P_5 + \lambda^4 P_4 + \lambda^3 P_3 + \lambda^2 P_2 + \lambda P_1 + P_0 \in \mathbb{C}[\lambda]^{m\times n}$$

$$\left[\begin{array}{ccc|cc} \lambda P_5 + P_4 & 0 & 0 & -I_m & 0 \\ 0 & \lambda P_3 + P_2 & 0 & \lambda I_m & -I_m \\ 0 & 0 & \lambda P_1 + P_0 & 0 & \lambda I_m \\ \hline -I_n & \lambda I_n & 0 & 0 & 0 \\ 0 & -I_n & \lambda I_n & 0 & 0 \end{array}\right]$$

$$P(\lambda) = \lambda^5 P_5 + \lambda^4 P_4 + \lambda^3 P_3 + \lambda^2 P_2 + \lambda P_1 + P_0 \in \mathbb{C}[\lambda]^{m\times n}$$

$$\left[\begin{array}{ccc|cc} \lambda P_5 & \lambda P_4 & \lambda P_3 & -I_m & 0 \\ 0 & 0 & \lambda P_2 & \lambda I_m & -I_m \\ 0 & 0 & \lambda P_1 + P_0 & 0 & \lambda I_m \\ \hline -I_n & \lambda I_n & 0 & 0 & 0 \\ 0 & -I_n & \lambda I_n & 0 & 0 \end{array}\right]$$

Examples of block Kronecker pencils (III)

$$P(\lambda) = \lambda^5 P_5 + \lambda^4 P_4 + \lambda^3 P_3 + \lambda^2 P_2 + \lambda P_1 + P_0 \in \mathbb{C}[\lambda]^{m\times n}$$

$$\left[\begin{array}{ccc|cc} \lambda P_5 & A & P_2 & -I_m & 0 \\ \lambda P_4 & -\lambda A & \lambda B + P_1 & \lambda I_m & -I_m \\ \lambda P_3 & -\lambda B & P_0 & 0 & \lambda I_m \\ \hline -I_n & \lambda I_n & 0 & 0 & 0 \\ 0 & -I_n & \lambda I_n & 0 & 0 \end{array}\right]$$

for any matrices A and B.

Outline

The main perturbation theorem

Theorem

Let $\mathcal{L}(\lambda)$ be a block Kronecker pencil for $P(\lambda) = \sum_{i=0}^{d} P_i \lambda^i \in \mathbb{C}[\lambda]^{m \times n}$, i.e.,

$$\mathcal{L}(\lambda) = \left[\begin{array}{c|c} \lambda M_1 + M_0 & L_\eta(\lambda)^T \otimes I_m \\ \hline L_\varepsilon(\lambda) \otimes I_n & 0 \end{array} \right].$$

If $\Delta\mathcal{L}(\lambda)$ is any pencil with the same size as $\mathcal{L}(\lambda)$ and such that

$$\|\Delta\mathcal{L}(\lambda)\|_F < \frac{(\sqrt{2}-1)^2}{d^{5/2}} \frac{1}{1 + \|\lambda M_1 + M_0\|_F},$$

then $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ is a strong linearization of a matrix poly $P(\lambda) + \Delta P(\lambda)$ with grade d and such that

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \leq 14\, d^{5/2} \frac{\|\mathcal{L}(\lambda)\|_F}{\|P(\lambda)\|_F} \left(1 + \|\lambda M_1 + M_0\|_F + \|\lambda M_1 + M_0\|_F^2\right) \frac{\|\Delta\mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F}.$$

In addition, the right (resp. left) minimal indices of $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ are those of $P(\lambda) + \Delta P(\lambda)$ shifted by ε (resp. η), i.e., the shift relations are preserved.

How is this theorem proved? STEP 1. Restoring the zero block.

The perturbation destroys the $(2,2)$-zero block and the block Kronecker structure

$$\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda) = \left[\begin{array}{c|c} \lambda M_1 + M_0 + \Delta\mathcal{L}_{11}(\lambda) & L_\eta(\lambda)^T \otimes I_m + \Delta\mathcal{L}_{12}(\lambda) \\ \hline L_\varepsilon(\lambda) \otimes I_n + \Delta\mathcal{L}_{21}(\lambda) & \Delta\mathcal{L}_{22}(\lambda) \end{array}\right].$$

Our first step restores the $(2,2)$-zero block via a strict equivalence close to the identity

$$\begin{bmatrix} I_{(\eta+1)m} & 0 \\ C & I_{\varepsilon n} \end{bmatrix} (\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)) \begin{bmatrix} I_{(\varepsilon+1)n} & D \\ 0 & I_{\eta m} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda M_1 + M_0 + \Delta\mathcal{L}_{11}(\lambda) & L_\eta(\lambda)^T \otimes I_m + \Delta\widetilde{\mathcal{L}}_{12}(\lambda) \\ L_\varepsilon(\lambda) \otimes I_n + \Delta\widetilde{\mathcal{L}}_{21}(\lambda) & 0 \end{bmatrix} =: \mathcal{L}(\lambda) + \Delta\widetilde{\mathcal{L}}(\lambda).$$

- C and D are solutions of an **underdetermined quadratic system of two matrix equations** whose existence is proved and whose norms are properly bounded.
- Strict equivalences preserve complete eigenstructures of matrix pencils.

The perturbation destroys the $(2,2)$-zero block and the block Kronecker structure

$$\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda) = \left[\begin{array}{c|c} \lambda M_1 + M_0 + \Delta\mathcal{L}_{11}(\lambda) & L_\eta(\lambda)^T \otimes I_m + \Delta\mathcal{L}_{12}(\lambda) \\ \hline L_\varepsilon(\lambda) \otimes I_n + \Delta\mathcal{L}_{21}(\lambda) & \Delta\mathcal{L}_{22}(\lambda) \end{array}\right].$$

Our first step restores the $(2,2)$-zero block via a strict equivalence close to the identity

$$\begin{bmatrix} I_{(\eta+1)m} & 0 \\ C & I_{\varepsilon n} \end{bmatrix} (\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)) \begin{bmatrix} I_{(\varepsilon+1)n} & D \\ 0 & I_{\eta m} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda M_1 + M_0 + \Delta\mathcal{L}_{11}(\lambda) & L_\eta(\lambda)^T \otimes I_m + \Delta\widetilde{\mathcal{L}}_{12}(\lambda) \\ L_\varepsilon(\lambda) \otimes I_n + \Delta\widetilde{\mathcal{L}}_{21}(\lambda) & 0 \end{bmatrix} =: \mathcal{L}(\lambda) + \Delta\widetilde{\mathcal{L}}(\lambda).$$

- C and D are solutions of an **underdetermined quadratic system of two matrix equations** whose existence is proved and whose norms are properly bounded.
- Strict equivalences preserve complete eigenstructures of matrix pencils.

The perturbation destroys the $(2,2)$-zero block and the block Kronecker structure

$$\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda) = \left[\begin{array}{c|c} \lambda M_1 + M_0 + \Delta\mathcal{L}_{11}(\lambda) & L_\eta(\lambda)^T \otimes I_m + \Delta\mathcal{L}_{12}(\lambda) \\ \hline L_\varepsilon(\lambda) \otimes I_n + \Delta\mathcal{L}_{21}(\lambda) & \Delta\mathcal{L}_{22}(\lambda) \end{array}\right].$$

Our first step restores the $(2,2)$-zero block via a strict equivalence close to the identity

$$\begin{bmatrix} I_{(\eta+1)m} & 0 \\ C & I_{\varepsilon n} \end{bmatrix} (\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)) \begin{bmatrix} I_{(\varepsilon+1)n} & D \\ 0 & I_{\eta m} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda M_1 + M_0 + \Delta\mathcal{L}_{11}(\lambda) & L_\eta(\lambda)^T \otimes I_m + \Delta\widetilde{\mathcal{L}}_{12}(\lambda) \\ L_\varepsilon(\lambda) \otimes I_n + \Delta\widetilde{\mathcal{L}}_{21}(\lambda) & 0 \end{bmatrix} =: \mathcal{L}(\lambda) + \Delta\widetilde{\mathcal{L}}(\lambda).$$

- C and D are solutions of an **underdetermined quadratic system of two matrix equations** whose existence is proved and whose norms are properly bounded.
- Strict equivalences preserve complete eigenstructures of matrix pencils.

The perturbation destroys the $(2,2)$-zero block and the block Kronecker structure

$$\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda) = \left[\begin{array}{c|c} \lambda M_1 + M_0 + \Delta\mathcal{L}_{11}(\lambda) & L_\eta(\lambda)^T \otimes I_m + \Delta\mathcal{L}_{12}(\lambda) \\ \hline L_\varepsilon(\lambda) \otimes I_n + \Delta\mathcal{L}_{21}(\lambda) & \Delta\mathcal{L}_{22}(\lambda) \end{array}\right].$$

Our first step restores the $(2,2)$-zero block via a strict equivalence close to the identity

$$\begin{bmatrix} I_{(\eta+1)m} & 0 \\ C & I_{\varepsilon n} \end{bmatrix} (\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)) \begin{bmatrix} I_{(\varepsilon+1)n} & D \\ 0 & I_{\eta m} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda M_1 + M_0 + \Delta\mathcal{L}_{11}(\lambda) & L_\eta(\lambda)^T \otimes I_m + \Delta\widetilde{\mathcal{L}}_{12}(\lambda) \\ L_\varepsilon(\lambda) \otimes I_n + \Delta\widetilde{\mathcal{L}}_{21}(\lambda) & 0 \end{bmatrix} =: \mathcal{L}(\lambda) + \Delta\widetilde{\mathcal{L}}(\lambda).$$

- C and D are solutions of an **underdetermined quadratic system of two matrix equations** whose existence is proved and whose norms are properly bounded.
- Strict equivalences preserve complete eigenstructures of matrix pencils.

Step 2. We prove that

$$\mathcal{L}(\lambda) + \Delta\widetilde{\mathcal{L}}(\lambda) := \begin{bmatrix} \lambda M_1 + M_0 + \Delta\mathcal{L}_{11}(\lambda) & L_\eta(\lambda)^T \otimes I_m + \Delta\widetilde{\mathcal{L}}_{12}(\lambda) \\ L_\varepsilon(\lambda) \otimes I_n + \Delta\widetilde{\mathcal{L}}_{21}(\lambda) & 0 \end{bmatrix}$$

is a **strong block minimal bases linearization** of a matrix polynomial

$$P(\lambda) + \Delta P(\lambda)$$
$$:= \left(\Lambda_\eta(\lambda)^T \otimes I_m + \Delta R_\eta(\lambda)^T\right)\left(\lambda M_1 + M_0 + \Delta\mathcal{L}_{11}(\lambda)\right)\left(\Lambda_\varepsilon(\lambda) \otimes I_n + \Delta R_\varepsilon(\lambda)\right),$$

where $\|\Delta R_\eta(\lambda)\|_F$ and $\|\Delta R_\varepsilon(\lambda)\|_F$ are carefully bounded.

Step 3. $\|\Delta P(\lambda)\|_F$ is finally bounded.

Step 2. We prove that

$$\mathcal{L}(\lambda) + \Delta\widetilde{\mathcal{L}}(\lambda) := \begin{bmatrix} \lambda M_1 + M_0 + \Delta\mathcal{L}_{11}(\lambda) & L_\eta(\lambda)^T \otimes I_m + \Delta\widetilde{\mathcal{L}}_{12}(\lambda) \\ L_\varepsilon(\lambda) \otimes I_n + \Delta\widetilde{\mathcal{L}}_{21}(\lambda) & 0 \end{bmatrix}$$

is a **strong block minimal bases linearization** of a matrix polynomial

$$P(\lambda) + \Delta P(\lambda)$$
$$:= \left(\Lambda_\eta(\lambda)^T \otimes I_m + \Delta R_\eta(\lambda)^T\right)\left(\lambda M_1 + M_0 + \Delta\mathcal{L}_{11}(\lambda)\right)\left(\Lambda_\varepsilon(\lambda) \otimes I_n + \Delta R_\varepsilon(\lambda)\right),$$

where $\|\Delta R_\eta(\lambda)\|_F$ and $\|\Delta R_\varepsilon(\lambda)\|_F$ are carefully bounded.

Step 3. $\|\Delta P(\lambda)\|_F$ is finally bounded.

Discussion of the perturbation bounds for block Kronecker pencils

$$\mathcal{L}(\lambda) = \left[\begin{array}{c|c} \lambda M_1 + M_0 & L_\eta(\lambda)^T \otimes I_m \\ \hline L_\varepsilon(\lambda) \otimes I_n & 0 \end{array} \right].$$

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \leq \underbrace{14\, d^{5/2} \frac{\|\mathcal{L}(\lambda)\|_F}{\|P(\lambda)\|_F} \left(1 + \|\lambda M_1 + M_0\|_F + \|\lambda M_1 + M_0\|_F^2\right)}_{C_{P,\mathcal{L}}} \frac{\|\Delta \mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F}.$$

- It can be proved that if $\|P(\lambda)\|_F \ll 1$ or $\|P(\lambda)\|_F \gg 1$, then $C_{P,\mathcal{L}} \gg 1$,
- and that, if $\|\lambda M_1 + M_0\|_F \gg 1$, then $C_{P,\mathcal{L}} \gg 1$.
- Therefore, for getting "backward stability" from Block Kronecker linearizations, one needs to normalize the matrix poly $\|P(\lambda)\|_F = 1$ and to use pencils such that $\|\lambda M_1 + M_0\|_F \approx \|P(\lambda)\|_F$, then

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \lesssim d^3 \sqrt{m+n}\, \frac{\|\Delta \mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F}.$$

For Fiedler and Frobenius linearizations $\|\lambda M_1 + M_0\|_F = \|P(\lambda)\|_F$

Discussion of the perturbation bounds for block Kronecker pencils

$$\mathcal{L}(\lambda) = \left[\begin{array}{c|c} \lambda M_1 + M_0 & L_\eta(\lambda)^T \otimes I_m \\ \hline L_\varepsilon(\lambda) \otimes I_n & 0 \end{array} \right].$$

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \leq \underbrace{14\, d^{5/2} \frac{\|\mathcal{L}(\lambda)\|_F}{\|P(\lambda)\|_F} \left(1 + \|\lambda M_1 + M_0\|_F + \|\lambda M_1 + M_0\|_F^2\right)}_{C_{P,\mathcal{L}}} \frac{\|\Delta \mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F}.$$

- It can be proved that if $\|P(\lambda)\|_F \ll 1$ or $\|P(\lambda)\|_F \gg 1$, then $C_{P,\mathcal{L}} \gg 1$,
- and that, if $\|\lambda M_1 + M_0\|_F \gg 1$, then $C_{P,\mathcal{L}} \gg 1$.
- Therefore, for getting "backward stability" from Block Kronecker linearizations, one needs to normalize the matrix poly $\|P(\lambda)\|_F = 1$ and to use pencils such that $\|\lambda M_1 + M_0\|_F \approx \|P(\lambda)\|_F$, then

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \lesssim d^3 \sqrt{m+n}\, \frac{\|\Delta \mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F}.$$

For Fiedler and Frobenius linearizations $\|\lambda M_1 + M_0\|_F = \|P(\lambda)\|_F$

Discussion of the perturbation bounds for block Kronecker pencils

$$\mathcal{L}(\lambda) = \left[\begin{array}{c|c} \lambda M_1 + M_0 & L_\eta(\lambda)^T \otimes I_m \\ \hline L_\varepsilon(\lambda) \otimes I_n & 0 \end{array} \right].$$

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \leq \underbrace{14\, d^{5/2} \frac{\|\mathcal{L}(\lambda)\|_F}{\|P(\lambda)\|_F} \left(1 + \|\lambda M_1 + M_0\|_F + \|\lambda M_1 + M_0\|_F^2\right)}_{C_{P,\mathcal{L}}} \frac{\|\Delta \mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F}.$$

- It can be proved that if $\|P(\lambda)\|_F \ll 1$ or $\|P(\lambda)\|_F \gg 1$, then $C_{P,\mathcal{L}} \gg 1$,
- and that, if $\|\lambda M_1 + M_0\|_F \gg 1$, then $C_{P,\mathcal{L}} \gg 1$.
- Therefore, for getting "backward stability" from Block Kronecker linearizations, one needs to normalize the matrix poly $\|P(\lambda)\|_F = 1$ and to use pencils such that $\|\lambda M_1 + M_0\|_F \approx \|P(\lambda)\|_F$, then

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \lesssim d^3 \sqrt{m+n}\, \frac{\|\Delta \mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F}.$$

For Fiedler and Frobenius linearizations $\|\lambda M_1 + M_0\|_F = \|P(\lambda)\|_F$

Discussion of the perturbation bounds for block Kronecker pencils

$$\mathcal{L}(\lambda) = \left[\begin{array}{c|c} \lambda M_1 + M_0 & L_\eta(\lambda)^T \otimes I_m \\ \hline L_\varepsilon(\lambda) \otimes I_n & 0 \end{array} \right].$$

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \leq \underbrace{14\, d^{5/2} \frac{\|\mathcal{L}(\lambda)\|_F}{\|P(\lambda)\|_F} \left(1 + \|\lambda M_1 + M_0\|_F + \|\lambda M_1 + M_0\|_F^2\right)}_{C_{P,\mathcal{L}}} \frac{\|\Delta \mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F}.$$

- It can be proved that if $\|P(\lambda)\|_F \ll 1$ or $\|P(\lambda)\|_F \gg 1$, then $C_{P,\mathcal{L}} \gg 1$,
- and that, if $\|\lambda M_1 + M_0\|_F \gg 1$, then $C_{P,\mathcal{L}} \gg 1$.
- Therefore, for getting "backward stability" from Block Kronecker linearizations, one needs to normalize the matrix poly $\|P(\lambda)\|_F = 1$ and to use pencils such that $\|\lambda M_1 + M_0\|_F \approx \|P(\lambda)\|_F$, then

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \lesssim d^3 \sqrt{m+n}\, \frac{\|\Delta \mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F}.$$

For Fiedler and Frobenius linearizations $\|\lambda M_1 + M_0\|_F = \|P(\lambda)\|_F$

$$\mathcal{L}(\lambda) = \left[\begin{array}{c|c} \lambda M_1 + M_0 & L_\eta(\lambda)^T \otimes I_m \\ \hline L_\varepsilon(\lambda) \otimes I_n & 0 \end{array} \right].$$

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \leq \underbrace{14\, d^{5/2} \frac{\|\mathcal{L}(\lambda)\|_F}{\|P(\lambda)\|_F} \left(1 + \|\lambda M_1 + M_0\|_F + \|\lambda M_1 + M_0\|_F^2\right)}_{C_{P,\mathcal{L}}} \frac{\|\Delta \mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F}.$$

- It can be proved that if $\|P(\lambda)\|_F \ll 1$ or $\|P(\lambda)\|_F \gg 1$, then $C_{P,\mathcal{L}} \gg 1$,
- and that, if $\|\lambda M_1 + M_0\|_F \gg 1$, then $C_{P,\mathcal{L}} \gg 1$.
- Therefore, for getting "backward stability" from Block Kronecker linearizations, one needs to normalize the matrix poly $\|P(\lambda)\|_F = 1$ and to use pencils such that $\|\lambda M_1 + M_0\|_F \approx \|P(\lambda)\|_F$, then

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \lesssim d^3 \sqrt{m+n}\, \frac{\|\Delta \mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F}.$$

For Fiedler and Frobenius linearizations $\|\lambda M_1 + M_0\|_F = \|P(\lambda)\|_F$

Discussion of the perturbation bounds for block Kronecker pencils

$$\mathcal{L}(\lambda) = \left[\begin{array}{c|c} \lambda M_1 + M_0 & L_\eta(\lambda)^T \otimes I_m \\ \hline L_\varepsilon(\lambda) \otimes I_n & 0 \end{array} \right].$$

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \leq \underbrace{14\, d^{5/2} \frac{\|\mathcal{L}(\lambda)\|_F}{\|P(\lambda)\|_F} \left(1 + \|\lambda M_1 + M_0\|_F + \|\lambda M_1 + M_0\|_F^2\right)}_{C_{P,\mathcal{L}}} \frac{\|\Delta \mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F}.$$

- It can be proved that if $\|P(\lambda)\|_F \ll 1$ or $\|P(\lambda)\|_F \gg 1$, then $C_{P,\mathcal{L}} \gg 1$,
- and that, if $\|\lambda M_1 + M_0\|_F \gg 1$, then $C_{P,\mathcal{L}} \gg 1$.
- Therefore, for getting "backward stability" from Block Kronecker linearizations, one needs to normalize the matrix poly $\|P(\lambda)\|_F = 1$ and to use pencils such that $\|\lambda M_1 + M_0\|_F \approx \|P(\lambda)\|_F$, then

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \lesssim d^3 \sqrt{m+n}\, \frac{\|\Delta \mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F}.$$

For Fiedler and Frobenius linearizations $\|\lambda M_1 + M_0\|_F = \|P(\lambda)\|_F$.

Outline

- The **PEPs appearing in applications have particular structures**,
- among them: Hermitian and skew-Hermitian, symmetric and skew-symmetric, palindromic and anti-palindromic, and alternating.
- **For any structured matrix polynomial of odd degree $d = 2k + 1$ in any of these classes, there exist (quasi) block Kronecker pencils with the same structure (called structured block Kronecker pencils)**,
- which can be defined in an elegant unified way through Möbius transformations,
- and whose left and right minimal indices are those of the matrix polynomial shifted by k.
- In addition, such **structured block Kronecker pencils can be constructed easily** from the coefficients of the matrix polynomial,
- and **satisfy an structured perturbation result**, which is not easy to prove.

Structured block Kronecker pencils

- The **PEPs appearing in applications have particular structures**,
- among them: Hermitian and skew-Hermitian, symmetric and skew-symmetric, palindromic and anti-palindromic, and alternating.
- **For any structured matrix polynomial of odd degree $d = 2k + 1$ in any of these classes, there exist (quasi) block Kronecker pencils with the same structure (called structured block Kronecker pencils)**,
- which can be defined in an elegant unified way through Möbius transformations,
- and whose left and right minimal indices are those of the matrix polynomial shifted by k.
- In addition, such **structured block Kronecker pencils can be constructed easily** from the coefficients of the matrix polynomial,
- and **satisfy an structured perturbation result**, which is not easy to prove.

Structured block Kronecker pencils

- The **PEPs appearing in applications have particular structures**,
- among them: Hermitian and skew-Hermitian, symmetric and skew-symmetric, palindromic and anti-palindromic, and alternating.
- **For any structured matrix polynomial of odd degree $d = 2k + 1$ in any of these classes, there exist (quasi) block Kronecker pencils with the same structure (called structured block Kronecker pencils)**,
- which can be defined in an elegant unified way through Möbius transformations,
- and whose left and right minimal indices are those of the matrix polynomial shifted by k.
- In addition, such **structured block Kronecker pencils can be constructed easily** from the coefficients of the matrix polynomial,
- and **satisfy an structured perturbation result**, which is not easy to prove.

Structured block Kronecker pencils

- The **PEPs appearing in applications have particular structures**,
- among them: Hermitian and skew-Hermitian, symmetric and skew-symmetric, palindromic and anti-palindromic, and alternating.
- **For any structured matrix polynomial of odd degree $d = 2k + 1$ in any of these classes, there exist (quasi) block Kronecker pencils with the same structure (called structured block Kronecker pencils)**,
- which can be defined in an elegant unified way through Möbius transformations,
- and whose left and right minimal indices are those of the matrix polynomial shifted by k.
- In addition, such **structured block Kronecker pencils can be constructed easily** from the coefficients of the matrix polynomial,
- and **satisfy an structured perturbation result**, which is not easy to prove.

- The **PEPs appearing in applications have particular structures**,
- among them: Hermitian and skew-Hermitian, symmetric and skew-symmetric, palindromic and anti-palindromic, and alternating.
- **For any structured matrix polynomial of odd degree $d = 2k + 1$ in any of these classes, there exist (quasi) block Kronecker pencils with the same structure (called structured block Kronecker pencils)**,
- which can be defined in an elegant unified way through Möbius transformations,
- and whose left and right minimal indices are those of the matrix polynomial shifted by k.
- In addition, such **structured block Kronecker pencils can be constructed easily** from the coefficients of the matrix polynomial,
- and **satisfy an structured perturbation result**, which is not easy to prove.

Structured block Kronecker pencils

- The **PEPs appearing in applications have particular structures**,
- among them: Hermitian and skew-Hermitian, symmetric and skew-symmetric, palindromic and anti-palindromic, and alternating.
- **For any structured matrix polynomial of odd degree $d = 2k + 1$ in any of these classes, there exist (quasi) block Kronecker pencils with the same structure (called structured block Kronecker pencils)**,
- which can be defined in an elegant unified way through Möbius transformations,
- and whose left and right minimal indices are those of the matrix polynomial shifted by k.
- In addition, such **structured block Kronecker pencils can be constructed easily** from the coefficients of the matrix polynomial,
- and **satisfy an structured perturbation result**, which is not easy to prove.

Structured block Kronecker pencils

- The **PEPs appearing in applications have particular structures**,
- among them: Hermitian and skew-Hermitian, symmetric and skew-symmetric, palindromic and anti-palindromic, and alternating.
- **For any structured matrix polynomial of odd degree $d = 2k + 1$ in any of these classes, there exist (quasi) block Kronecker pencils with the same structure (called structured block Kronecker pencils)**,
- which can be defined in an elegant unified way through Möbius transformations,
- and whose left and right minimal indices are those of the matrix polynomial shifted by k.
- In addition, such **structured block Kronecker pencils can be constructed easily** from the coefficients of the matrix polynomial,
- and **satisfy an structured perturbation result**, which is not easy to prove.

The structured perturbation theorem

Theorem

Let $\mathcal{L}(\lambda)$ be a structured block Kronecker pencil for any structured $n \times n$ matrix polynomial $P(\lambda)$ in the previous classes of odd degree $d = 2k+1$

$$\mathcal{L}(\lambda) = \left[\begin{array}{c|c} \lambda M_1 + M_0 & \widetilde{L}_k(\lambda)^T \otimes I_n \\ \hline L_k(\lambda) \otimes I_n & 0 \end{array}\right].$$

If $\Delta\mathcal{L}(\lambda)$ is any pencil with the same size and structure as $\mathcal{L}(\lambda)$ and such that

$$\|\Delta\mathcal{L}(\lambda)\|_F < \frac{(\sqrt{2}-1)^2}{d^{5/2}} \frac{1}{1 + \|\lambda M_1 + M_0\|_F},$$

then $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ is a strong linearization of a matrix poly $P(\lambda) + \Delta P(\lambda)$ with grade d, with the same structure as $P(\lambda)$, and such that

$$\frac{\|\Delta P(\lambda)\|_F}{\|P(\lambda)\|_F} \leq 14\, d^{5/2} \frac{\|\mathcal{L}(\lambda)\|_F}{\|P(\lambda)\|_F} \left(1 + \|\lambda M_1 + M_0\|_F + \|\lambda M_1 + M_0\|_F^2\right) \frac{\|\Delta\mathcal{L}(\lambda)\|_F}{\|\mathcal{L}(\lambda)\|_F}.$$

In addition, the right (resp. left) minimal indices of $\mathcal{L}(\lambda) + \Delta\mathcal{L}(\lambda)$ are those of $P(\lambda) + \Delta P(\lambda)$ shifted by k (resp. k).

Outline

- We have proved that the computation of the complete eigenstructure of a matrix polynomial $P(\lambda)$ (regular or singular, square or rectangular)
- applying a global backward stable algorithm to any block Kronecker pencil of $P(\lambda)$ **is globally backward stable from the polynomial point of view**
- **if** $\|P(\lambda)\|_F = 1$ **and** $\|\lambda M_1 + M_0\|_F \approx \|P(\lambda)\|_F$.
- **These results can be extended "in a structured way"** to matrix polynomials of odd degree in any of the following structured classes: Hermitian and skew-Hermitian, symmetric and skew-symmetric, palindromic and anti-palindromic, and alternating.
- The new perturbation analysis presents a number of novel features and establishes a framework that can be probably generalized to other linearizations, including linearizations of matrix polynomials expressed in other bases.

Conclusions

- We have proved that the computation of the complete eigenstructure of a matrix polynomial $P(\lambda)$ (regular or singular, square or rectangular)
- applying a global backward stable algorithm to any block Kronecker pencil of $P(\lambda)$ **is globally backward stable from the polynomial point of view**
- **if** $\|P(\lambda)\|_F = 1$ **and** $\|\lambda M_1 + M_0\|_F \approx \|P(\lambda)\|_F$.
- **These results can be extended "in a structured way"** to matrix polynomials of odd degree in any of the following structured classes: Hermitian and skew-Hermitian, symmetric and skew-symmetric, palindromic and anti-palindromic, and alternating.
- The new perturbation analysis presents a number of novel features and establishes a framework that can be probably generalized to other linearizations, including linearizations of matrix polynomials expressed in other bases.

Conclusions

- We have proved that the computation of the complete eigenstructure of a matrix polynomial $P(\lambda)$ (regular or singular, square or rectangular)
- applying a global backward stable algorithm to any block Kronecker pencil of $P(\lambda)$ **is globally backward stable from the polynomial point of view**
- **if** $\|P(\lambda)\|_F = 1$ **and** $\|\lambda M_1 + M_0\|_F \approx \|P(\lambda)\|_F$.
- **These results can be extended "in a structured way"** to matrix polynomials of odd degree in any of the following structured classes: Hermitian and skew-Hermitian, symmetric and skew-symmetric, palindromic and anti-palindromic, and alternating.
- The new perturbation analysis presents a number of novel features and establishes a framework that can be probably generalized to other linearizations, including linearizations of matrix polynomials expressed in other bases.

- We have proved that the computation of the complete eigenstructure of a matrix polynomial $P(\lambda)$ (regular or singular, square or rectangular)
- applying a global backward stable algorithm to any block Kronecker pencil of $P(\lambda)$ **is globally backward stable from the polynomial point of view**
- **if** $\|P(\lambda)\|_F = 1$ **and** $\|\lambda M_1 + M_0\|_F \approx \|P(\lambda)\|_F$.
- **These results can be extended "in a structured way"** to matrix polynomials of odd degree in any of the following structured classes: Hermitian and skew-Hermitian, symmetric and skew-symmetric, palindromic and anti-palindromic, and alternating.
- The new perturbation analysis presents a number of novel features and establishes a framework that can be probably generalized to other linearizations, including linearizations of matrix polynomials expressed in other bases.

Conclusions

- We have proved that the computation of the complete eigenstructure of a matrix polynomial $P(\lambda)$ (regular or singular, square or rectangular)
- applying a global backward stable algorithm to any block Kronecker pencil of $P(\lambda)$ **is globally backward stable from the polynomial point of view**
- **if** $\|P(\lambda)\|_F = 1$ **and** $\|\lambda M_1 + M_0\|_F \approx \|P(\lambda)\|_F$.
- **These results can be extended "in a structured way"** to matrix polynomials of odd degree in any of the following structured classes: Hermitian and skew-Hermitian, symmetric and skew-symmetric, palindromic and anti-palindromic, and alternating.
- The new perturbation analysis presents a number of novel features and establishes a framework that can be probably generalized to other linearizations, including linearizations of matrix polynomials expressed in other bases.