Numeric Computation with Matrix Polynomials

George Labahn

Symbolic Computation Group

University of Waterloo

Joint with Mark Giesbrecht and Joseph Haraldson

Sorbonne University, September 11, 2020

Problems

Given

$$A = [a_{ij}(t)]_{n\times n} = \sum_{0\le i\le d} A_i t^i \in \mathbb{R}[t]^{n\times n},$$

a square matrix polynomial with real coefficients, we want to:

- find $\widehat{A}$ such that $\widehat{A}$ is singular and $\|A - \widehat{A}\|$ is minimized,
- find nearest $\widehat{A}$ with *interesting* Smith Normal Form (SNF).

Problems

Given

$$A = [a_{ij}(t)]_{n\times n} = \sum_{0\le i\le d} A_i t^i \in \mathbb{R}[t]^{n\times n},$$

a square matrix polynomial with real coefficients, we want to:

- find $\widehat{A}$ such that $\widehat{A}$ is singular and $\|A - \widehat{A}\|$ is minimized,
  - if $\det(\widehat{A}) \equiv 0$
  - $\exists b \in \mathbb{R}[t]^{n\times 1}\setminus\{0\}$ such that $\widehat{A}b \equiv 0$
- find nearest $\widehat{A}$ with *interesting* Smith Normal Form (SNF).

Problems

Given

$$A = [a_{ij}(t)]_{n\times n} = \sum_{0\le i\le d} A_i t^i \in \mathbb{R}[t]^{n\times n},$$

a square matrix polynomial with real coefficients, we want to:

- find $\widehat{A}$ such that $\widehat{A}$ is singular and $\|A - \widehat{A}\|$ is minimized,
  - if $\det(\widehat{A}) \equiv 0$
  - $\exists b \in \mathbb{R}[t]^{n\times 1}\setminus\{0\}$ such that $\widehat{A}b \equiv 0$
- find nearest $\widehat{A}$ with *interesting* Smith Normal Form (SNF).
  - $\widehat{A} \equiv \mathrm{diag}(s_1, s_2, \ldots, s_n)$ with $s_1|s_2|\cdots|s_n$
  - 'SNF interesting' $\implies$ not $\mathrm{diag}(1, \ldots, 1, \det(\widehat{A}))$.

Problem 1: Nearest Singular Matrix Polynomial

- Find $\widehat{A}$ such that $\widehat{A}$ is singular and $\|A - \widehat{A}\|$ is minimized:
  - Equivalent to: solve the optimization problem

$$\min_{\widehat{A},b} \|A - \widehat{A}\| \text{ subject to } \begin{cases} \widehat{A}b = 0 \\ \|b\| = 1. \end{cases}$$

Problem 1: Nearest Singular Matrix Polynomial

- Find $\widehat{A}$ such that $\widehat{A}$ is singular and $\|A - \widehat{A}\|$ is minimized:
  - Equivalent to: solve the optimization problem

$$\min_{\widehat{A},b} \|A - \widehat{A}\| \text{ subject to } \begin{cases} \widehat{A}b = 0 \\ \|b\| = 1. \end{cases}$$

- Approximation Requires a Norm :
  - $f \in \mathbb{R}[t]$ as $\|f\|_2 = \|(f_0, f_1, \ldots, f_d, 0, \ldots, 0)\|_2$,
  - $A \in \mathbb{R}[t]^{n \times n}$ as $\|A\|_F = \sqrt{\sum_{1 \leq i,j \leq n} \|A_{ij}\|_2^2}$.

Applications

- Stability of solutions to linear time invariant systems
  - [Byers & Nicols '93] and [Byers, He & Mehrmann '98]
- Stability of polynomial eigenvalue problems
- Approximate GCRD of Ore polynomials
  - [Giesbrecht, Haraldson & Kaltofen '16]

Our Goals

For: $\min_{\Delta A, b} \|\Delta A\|_F$ subject to $\begin{cases} \|b\|_2 = 1 \\ (A + \Delta A)b = 0. \end{cases}$

1. Determine if *minimal solutions* exist
2. Determine if minimal solutions are isolated
   - if A_1 and A_2 are distinct minimal perturbations then $\inf \|\Delta A_1 - \Delta A_2\|_F > 0$ (when $\operatorname{rank} A = n$)

Our Goals

For: $\min_{\Delta A, b} \|\Delta A\|_F$ subject to $\begin{cases} \|b\|_2 = 1 \\ (A + \Delta A)b = 0. \end{cases}$

1. Determine if *minimal solutions* exist
2. Determine if minimal solutions are isolated
   - if A_1 and A_2 are distinct minimal perturbations then $\inf \|\Delta A_1 - \Delta A_2\|_F > 0$ (when $\operatorname{rank} A = n$)
3. Is the optimization problem is (locally) well-posed
   - Optimal value is isolated around minimizers
4. Derive and implement a quadratically convergent algorithm

Previous Work (Byers, Nichols (1993), Byers, He, Mehrmann (1998), etc)

Previous work mostly focused on *matrix pencils* : $\mathrm{P}(\mathrm{t}) = \mathrm{P}_1\mathrm{t} + \mathrm{P}_0$.

Previous Work (Byers, Nichols (1993), Byers, He, Mehrmann (1998), etc)

Previous work mostly focused on *matrix pencils* : $P(t) = P_1 t + P_0$.

- Convert $A(t) = A_d t^d + \cdots + A_1 t + A_0$ into matrix pencil

$$U(t)(P_1 t + P_0)V(t) = \mathrm{diag}(I_{d(n-1)}, A(t))$$

where

$$P_1 = \begin{bmatrix} 0 & & & A_d \\ -I & & & A_{d-1} \\ & \ddots & & \vdots \\ & & -I & A_1 \end{bmatrix}, \quad P_0 = \begin{bmatrix} I & & & \\ & \ddots & & \\ & & I & \\ & & & A_0 \end{bmatrix}$$

and $U(t), V(t)$ unimodular.

Previous Work (Byers, Nichols (1993), Byers, He, Mehrmann (1998), etc)

Previous work mostly focused on *matrix pencils* : $P(t) = P_1 t + P_0$.

- Convert $A(t) = A_d t^d + \cdots + A_1 t + A_0$ into matrix pencil

$$U(t)(P_1 t + P_0)V(t) = \mathrm{diag}(I_{d(n-1)}, A(t))$$

where

$$P_1 = \begin{bmatrix} 0 & & & A_d \\ -I & & & A_{d-1} \\ & \ddots & & \vdots \\ & & -I & A_1 \end{bmatrix}, \quad P_0 = \begin{bmatrix} I & & & \\ & \ddots & & \\ & & I & \\ & & & A_0 \end{bmatrix}$$

and $U(t), V(t)$ unimodular.

- Problem: conversion not distance preserving; not isomorphism

Possible Other Approaches Considered

- Restricted SVD [De Moor '93, '94]
  - At best linear convergence
- Structured Total Least Norm [Rosen, Park & Glick '96]
  - Super linear convergence (not quadratic) [Lemmerling '99]
- Variable Projection [Golub & Pereyra '73, '03]
  - Use Gauss-Newton (Pure Newton)
  - Converges super-linearly if normalized
  - Problem size is much larger
- Lift and Project methods [Spaenlehauer & Schost '16]

Goal 1: Properties of minimal solutions

Theorem

Let $\Delta A \in \mathbb{R}[t]^{n \times n}$ *have the same support as* A*, that is* $\deg \Delta A_{ij} \leq \deg A_{ij}$*. Then the optimization problem*

$$\min_{\Delta A, b} \|\Delta A\| \text{ subject to } \begin{cases} \|b\| = 1 \\ (A + \Delta A)b = 0 \end{cases}$$

has an attainable global minimum $(\Delta A^\star, b^\star)$*.*

- Nearest singular matrix polynomial always exists

  i.e. $\exists (\Delta A^\star, b^\star)$ s.t. $(A + \Delta A^\star)b^\star = 0$ and $\|\Delta A^\star\|$ is minimal

Tool : Converting Matrix polynomials $\Longrightarrow$ real Matrices

Given $a \in \mathbb{R}[t]$ at most degree d. Let $\mu = d + 1$

- Embed: $\mathbb{R}[t]$ in $\mathbb{R}^{(\mu+d)\times\mu}$:

$$a = a_d t^d + \cdots a_1 t + a_0 \Longrightarrow \phi(a) = \begin{bmatrix} a_0 & & \\ a_1 & \ddots & \\ \vdots & & a_0 \\ a_d & & \\ & \ddots & \vdots \\ & & a_d \end{bmatrix}$$

- Notice that $\phi(a) \cdot [\mathrm{coeffs}(b)]^t = [\mathrm{coeffs}(a \cdot b)]^t$

Embedding Matrix polynomials into Real Matrices

Given A, $n \times n$ with entries at most degree d. Let $\mu = nd + 1$

- Embed : $\mathbb{R}[t]^{n \times n}$ in $\mathbb{R}^{n(\mu+d) \times n\mu}$

$$A \implies \mathcal{A} = \begin{bmatrix} \phi(a_{11}) & \cdots & \phi(a_{1n}) \\ \vdots & & \vdots \\ \phi(a_{n1}) & \cdots & \phi(a_{nn}) \end{bmatrix}$$

- Similar embedding for vectors: $b \implies \beta$

Embedding Matrix polynomials into Real Matrices

Given A, $n \times n$ with entries at most degree d. Let $\mu = nd + 1$

- Embed : $\mathbb{R}[t]^{n \times n}$ in $\mathbb{R}^{n(\mu+d) \times n\mu}$

$$A \implies \mathcal{A} = \begin{bmatrix} \phi(a_{11}) & \cdots & \phi(a_{1n}) \\ \vdots & & \vdots \\ \phi(a_{n1}) & \cdots & \phi(a_{nn}) \end{bmatrix}$$

- Similar embedding for vectors: $b \implies \beta$
- Basically we capture $Ab = 0$ in terms of coefficients. That is:

$$Ab = 0 \implies \mathcal{A} \cdot \beta = 0.$$

Embedding Matrix polynomials into Real Matrices

Given A, $n \times n$ with entries at most degree d. Let $\mu = nd + 1$

- Embed : $\mathbb{R}[t]^{n \times n}$ in $\mathbb{R}^{n(\mu+d) \times n\mu}$

$$A \implies \mathcal{A} = \begin{bmatrix} \phi(a_{11}) & \cdots & \phi(a_{1n}) \\ \vdots & & \vdots \\ \phi(a_{n1}) & \cdots & \phi(a_{nn}) \end{bmatrix}$$

- Similar embedding for vectors: $b \implies \beta$
- Basically we capture $Ab = 0$ in terms of coefficients. That is:

$$Ab = 0 \implies \mathcal{A} \cdot \beta = 0.$$

- Embedding is quasi-distance preserving: $\|A\|_F^2 = \frac{\|\mathcal{A}\|_F^2}{\mu}$

Embedding Example

The embedding preserves kernel vectors...

$$A = \begin{bmatrix} t^2 - 1 & t + 1 \\ t^2 - 2t + 1 & t - 1 \end{bmatrix} \quad \text{and} \quad \ker A = \begin{bmatrix} -1 \\ t - 1 \end{bmatrix}$$

$$\mathcal{A} = \left[\begin{array}{ccccc|ccccc} -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & -2 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \qquad \ker \mathcal{A} = \left[\begin{array}{cccc} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ \hline -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

- The embedding does not preserve rank information

Useful property

- If $A \in \mathbb{R}[t]^{n \times n}$ is singular then there exists $b \in \ker A$ such that $\deg b \leq nd = \mu - 1$.
  - Embed A as $\mathcal{A} \in \mathbb{R}^{n(\mu+d) \times n\mu}$ and b as $\beta \in \mathbb{R}^{n\mu \times 1}$
  - $Ab = 0 \iff \mathcal{A}\beta = 0$

Useful property

- If $A \in \mathbb{R}[t]^{n \times n}$ is singular then there exists $b \in \ker A$ such that $\deg b \leq nd = \mu - 1$.
  - Embed A as $\mathcal{A} \in \mathbb{R}^{n(\mu+d) \times n\mu}$ and b as $\beta \in \mathbb{R}^{n\mu \times 1}$
  - $Ab = 0 \iff \mathcal{A}\beta = 0$
- Note that
  - SVD provides a cheap lower bound on the distance to a singular matrix. Hence also a singular matrix polynomial
  - If $A + \Delta A \in \mathbb{R}[t]^{n \times n}$ is singular, then $\|\Delta\mathcal{A}\|_F \geq \sigma_{n\mu}(\mathcal{A})$
- Set of rank deficient matrices in $\mathbb{R}[t]^{n \times n}$, deg $\leq d$, is closed.

Goal 2: Separation Bounds

Theorem (Conjecture)

Suppose ΔA and $\Delta A^\star$ are distinct (local) minimal solutions then

$$\|\Delta A - \Delta A^\star\|_F \geq \frac{\|\Delta \mathcal{A} - \Delta \mathcal{A}^\star\|_2}{nd+1} \geq \frac{\sigma_{\min}(\mathcal{A})}{nd+1} = \frac{\min_{\|x\|=1} \|\mathcal{A}x\|_2}{nd+1}.$$

Goal 2: Separation Bounds

Theorem (Conjecture)

Suppose ΔA and $\Delta A^\star$ are distinct (local) minimal solutions then

$$\|\Delta A - \Delta A^\star\|_F \geq \frac{\|\Delta \mathcal{A} - \Delta \mathcal{A}^\star\|_2}{nd+1} \geq \frac{\sigma_{\min}(\mathcal{A})}{nd+1} = \frac{\min_{\|x\|=1} \|\mathcal{A}x\|_2}{nd+1}.$$

Theorem (Can prove)

Suppose ΔA and $\Delta A^\star$ are distinct (local) minimal solutions then

$$\|\Delta A - \Delta A^\star\|_F > 0.$$

Goals 3: Solving Optimization problem

Look at optimization problem with objective function

$$\Psi = \|\Delta\mathcal{A}\|_F{}^2 + \|\beta\|^2 - 1$$

subject to constraints $(\mathcal{A} + \Delta\mathcal{A})\beta = 0$ and $\beta^t\beta - 1 = 0$.

- Variables are $x = ((\Delta\mathcal{A})_{11}, \ldots, (\Delta\mathcal{A})_{NM}, \beta_1, \ldots, \beta_M)$.

  For short write $x = vec(\Delta\mathcal{A}, \beta)$

- Lagrange Multiplier is

$$L = \Psi + \lambda^T \begin{bmatrix} \text{vec}((\mathcal{A} + \Delta\mathcal{A})\beta) \\ \beta^t\beta - 1 \end{bmatrix}.$$

Solving Optimization Problem

- First order necessary (Karush-Kuhn-Tucker (KKT)) conditions

$$\nabla L(\Delta A^\star, \beta^\star, \lambda^\star) = 0. \qquad (\nabla \text{ is gradient})$$

- Idea is to solve $\nabla L = 0$ by Newton's method, i.e. compute

$$\begin{bmatrix} x^{k+1} \\ \lambda^{k+1} \end{bmatrix} = \begin{bmatrix} x^k + \Delta x^k \\ \lambda^k + \Delta \lambda^k \end{bmatrix} \text{ such that } \nabla^2 L \begin{bmatrix} \Delta x \\ \Delta \lambda \end{bmatrix} = -\nabla L$$

- Closed form for Jacobian of the constraints is

$$J = \nabla \begin{pmatrix} \text{vec}((\mathcal{A} + \Delta\mathcal{A})\beta) \\ \beta^T\beta - 1 \end{pmatrix}^t = \begin{bmatrix} \psi(\beta) & \mathcal{A} + \Delta\mathcal{A} \\ 0 & 2\beta^T \end{bmatrix}.$$

  - Here $\psi(\beta)$ s.t. $\psi(\beta)\text{vec}(A + \Delta A) = 0 \iff (\mathcal{A} + \Delta\mathcal{A})\beta = 0$
  - J has full rank $\implies$ minimal solutions $(\Delta A^\star, b^\star)$ are isolated
  - Multiple kernel vectors $\implies$ J is rank deficient

- The Hessian matrix of L is $\nabla^2 L = \begin{bmatrix} \nabla^2_{xx} L & J \\ J^T & 0 \end{bmatrix}$.

- Note : $\nabla^2 L$ has full rank $\iff$ J has full rank

Minimal Kernel Embedding

A kernel vector β corresponding to $\mathrm{b} \in \ker(A + \Delta A)$ is minimally degree embedded in $\mathcal{A} + \Delta\mathcal{A}$ if

1. $\ker(\mathcal{A} + \Delta\mathcal{A}) = \mathrm{span}(\beta)$ and b is primitive
2. β comes from a column echelon reduced basis.

Note:

- J full rank at $(\Delta A^\star, \mathrm{b}^\star)$ if $\mathrm{b}^\star$ is minimally degree embedded.
  - Minimally embed by deleting rows/columns of $\mathcal{A}$ and $\mathcal{B}$
  - Compute via orthogonal eliminations
- Newton's method converges quadratically to $(\Delta A^\star, \mathrm{b}^\star)$ with a suitable initial guess

Recall : Embedding Example

$$A = \begin{bmatrix} t^2 - 1 & t + 1 \\ t^2 - 2t + 1 & t - 1 \end{bmatrix} \quad \text{and} \quad \ker A = \begin{bmatrix} -1 \\ t - 1 \end{bmatrix}$$

$$\mathcal{A} = \left[\begin{array}{ccccc|ccccc} -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & -2 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \qquad \ker \mathcal{A} = \left[\begin{array}{cccc} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ \hline -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

Minimal Kernel Example

$$A = \begin{bmatrix} t^2 - 1 & t + 1 \\ t^2 - 2t + 1 & t - 1 \end{bmatrix} \quad \text{and} \quad \ker A = \begin{bmatrix} -1 \\ t - 1 \end{bmatrix}$$

The minimal embedding gives us

$$\mathcal{A}_{\min} = \left[\begin{array}{c|cc} -1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ \hline 1 & -1 & 0 \\ -2 & 1 & -1 \\ 1 & 0 & 1 \end{array}\right] \quad \text{and} \quad \mathcal{B}_{\min} = \left[\begin{array}{c} -1 \\ \hline -1 \\ 1 \end{array}\right]$$

Implementation Information

Our Implementation:

- Initialize ΔA and b via SVD (or other method)
- Minimally degree embed system over $\Delta \mathcal{A}$ and β
- Compute Lagrange multipliers via linear least squares

In general:

- Use any method to approximate $\nabla L = 0$
- Minimally degree embed and switch to Newton's method later

Cost Per Iteration:

- Maximum cost per iteration is $O(n^9 d^6)$
- Exploiting structure and sparsity reduces costs considerably

Algorithm

Input: Full rank $A \in \mathbb{R}[t]^{n \times n}$ with structure ΔA and
$C \in \mathbb{R}[t]^{n \times n}$ with (approx) kernel vector b

Output: $A + \Delta A^{\star}$ or an indication of failure

1. Embed input over $\mathbb{R}$
2. Compute λ^0 by solving $\nabla L|_{x^0} = 0$ via linear least squares
3. Compute $\begin{bmatrix} x + \Delta x \\ \lambda + \Delta\lambda \end{bmatrix}$ until $\left\| \begin{bmatrix} \Delta x \\ \Delta\lambda \end{bmatrix} \right\|_2$ is sufficiently small
4. Return the locally optimal solution or an indication of failure

Problem 2: Nearest Interesting The Smith Normal Form

Recall: Any $A \in \mathbb{R}[t]^{n\times n}$ there exists $U, V \in \mathbb{R}[t]^{n\times n}$

$$UAV = S = \mathrm{diag}(s_1, s_2, \ldots, s_n) \quad \text{where} \quad s_j | s_{j+1} \quad \text{and} \quad s_j \in \mathbb{R}[t].$$

Here U, V unimodular, that is, $\det(U), \det(V) \in \mathbb{R}\backslash\{0\}$

- S is the Smith Normal Form (SNF) of A
- The $\{s_j\}_{j=1}^n$ are the invariant factors
- Computing S is well understood in exact-arithmetic

Uninteresting SNF

Example (Uninteresting SNF over $\mathbb{R}[t]^{3\times 3}$)

$$A = \begin{bmatrix} t^3+3t+1 & 1 & t+1 \\ 0 & t^2+2t+2 & 0 \\ t+1 & t+1 & t^3+5t+1 \end{bmatrix} \quad \mathrm{SNF}(A) = \begin{bmatrix} 1 & & \\ & 1 & \\ & & \det(A) \end{bmatrix}$$

$$\det(A) = t^8 + 2t^7 + 10t^6 + 18t^5 + 34t^4 + 38t^3 + 40t^2 + 12t.$$

Interesting SNF

Example (Interesting SNF over $\mathbb{R}[t]^{3\times 3}$)

$$A = \begin{bmatrix} t+1 & t+1 & t-1 \\ 0 & t+1 & t^3 \\ 0 & 0 & t^2-1 \end{bmatrix} \quad \mathrm{SNF}(A) = \begin{bmatrix} 1 & & \\ & t+1 & \\ & & (t+1)(t^2-1) \end{bmatrix}$$

Questions for numeric SNF computation

When does $\mathcal{A}$ have a non-trivial Smith Normal Form?

- Small perturbations to $\mathcal{A}$ generically produce a trivial SNF
- How far is $\mathcal{A}$ from a matrix polynom. $\widehat{\mathcal{A}}$ with non-trivial SNF?
- Is there a radius of triviality?
  - i.e., if $\mathcal{A}$ is perturbed by a small amount is the SNF still trivial?

Questions for numeric SNF computation (cont.)

When is Computing the SNF Well-Posed?

Is there a nearest matrix polynomial $\widehat{A}$ with an interesting SNF?

- Is $\widehat{A}$ locally unique?
- How do we compute $\widehat{A}$?
- How do perturbations to $\mathcal{A}$ affect $\widehat{A}$?

Nearby SNF via Optimization

The McCoy Rank - Number of 1's in the SNF

Formally: McCoy rank of $\mathcal{A} \in \mathbb{R}[t]^{n\times n}$ is $\min_{\omega\in\mathbb{C}} \operatorname{rank}(\mathcal{A}(\omega))$.

Main Problem: Nearby Interesting SNF

Given $\mathcal{A} \in \mathbb{R}[t]^{n\times n}$ with McCoy rank $< n$: find $\widehat{\mathcal{A}} \in \mathbb{R}[t]^{n\times n}$ that (locally) solves the optimization problem

$$\min \|\mathcal{A} - \widehat{\mathcal{A}}\| \text{ such that } \begin{cases} \operatorname{SNF}(\widehat{\mathcal{A}}) = \operatorname{diag}(\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_{n-1}, \hat{s}_n), \\ \\ \deg(s_n) \geq \deg(\hat{s}_{n-1}) \geq 1. \end{cases}$$

Main results in Haraldson PhD thesis for SNF problem

1. Tight lower bounds on the radius of triviality
2. Polynomial-time procedure to decide ill-posed
3. Stability analysis on SNF via Optimization
4. Iterative algorithms with local quadratic convergence
   - Nearest matrix with reduced McCoy rank
   - Nearest matrix with McCoy rank at most $n - r$
   - Reasonable initial guess heuristics for both algorithms
   - Polynomial per-iteration cost

Previous Work on Floating Point SNF Computations

Reduction to Degree One

Every matrix polynomial $\mathcal{A} \in \mathbb{R}[t]^{n \times n}$ can be linearized to

$$\mathcal{P} = \mathcal{P}_0 + t\mathcal{P}_1 \quad \text{for some} \quad \mathcal{P}_0, \mathcal{P}_1 \in \mathbb{R}^{nd \times nd}.$$

- Extract the SNF from Kronecker's Canonical Form ($\mathcal{K}$) of $\mathcal{P}$

  via $\mathrm{SNF}(\mathcal{K}) = \mathrm{diag}(1, 1, \ldots, 1, \mathrm{SNF}(\mathcal{A}))$

Algorithms in Matrix Pencil case

Backward Stable: Finds the SNF of a nearby matrix.

- Full Rank Case: QZ Algorithm
  - Wilkinson (1979)
- Singular Case: Fast Staircase/Deflation Algorithms
  - Beelen and Van Dooren (1984,1988)
- Current: GUPTRI
  - Demmel and Edelman (1995)

Reducing Approximate SNF to Approximate GCD

- We can define the SNF in terms of the minors

$$s_j = \frac{\delta_j}{\delta_{j+1}} \text{ where } \delta_j = \text{GCD}\{\text{all } j \times j \text{ minors of } \mathcal{A}\}$$

- Requiring $\delta_j \neq 1 \implies \mathcal{A}$ has McCoy rank at most $n - j - 1$
- Use Sylvester matrices and approximate GCD techniques
  - δ_j's are approximate GCD's of several polynomials
  - Coefficient structure is multi-linear in the entries of $\mathcal{A}$
- $\mathcal{A}$ has McCoy rank at most $n - 2$ iff entries of the adjoint matrix have a non-trivial GCD.
- We can compute the adjoint matrix quickly and robustly

Issue with unattainable infimum

Recall from approximate gcd:

- There are co-prime polynomials with nearby polynomials with a non-trivial GCD at distances arbitrarily approaching an infimum, while at the infimum itself the GCD is trivial .

- Issue carries over to approximate Smith Normal Form

Example: Find Nearest 2×2 matrix with a non-trivial SNF

$$\mathcal{A} = \begin{bmatrix} t^2 - 2t + 1 & 0 \\ 0 & t^2 + 2t + 2 \end{bmatrix}$$ find a lower McCoy rank $\widetilde{\mathcal{A}}$.

Reduction to Approximate GCD

For this $\mathcal{A} = \begin{bmatrix} t^2 - 2t + 1 & 0 \\ 0 & t^2 + 2t + 2 \end{bmatrix}$

- Approximate GCD of a_{11} and a_{22} (Karmarkar and Lakshman)

  $$\inf \left\{ \|a_{11} - \widetilde{a}_{11}\|_2^2 + \|a_{22} - \widetilde{a}_{22}\|_2^2 \right\} \quad \text{s.t.} \quad \gcd(\widetilde{a}_{11}, \widetilde{a}_{22}) \neq 1.$$

  If: $\widetilde{a}_{11} = (c_{11}t + c_{10})(ht + 1)$ and $\widetilde{a}_{22} = (c_{21}t + c_{20})(ht + 1)$.

- The distance to a matrix with a non-trivial SNF is

  $$\inf_{h \in \mathbb{R}} \frac{5h^4 - 4h^3 + 14h + 2}{h^4 + h^2 + 1} = 2 \text{ when } h = 0.$$

- Thus $\gcd(\widetilde{a}_{11}, \widetilde{a}_{22}) = 1$ at the infima.

Generalized Sylvester matrices

Let $a \in \mathbb{R}[t]$ with $\deg a \leq d$.

$$\phi_r(a) = \begin{bmatrix} a_0 & \cdots & a_d & & \\ & \ddots & & \ddots & \\ & & a_0 & \cdots & a_d \end{bmatrix} \in \mathbb{R}^{r \times (r+d)}.$$

- Let $\mathbf{f} = (f_1, \ldots, f_k) \in \mathbb{R}[t]^k$ be ordered by decreasing degree
- $\mathbf{d} = (\deg(f_1), \ldots, \deg(f_k))$, $r = \deg f_1$ and $d = \max\{\deg_{f_j}\}_{j=2}^k$

$$\underbrace{\mathrm{Syl}(\mathbf{f}) = \mathrm{Syl}_{\mathbf{d}}(\mathbf{f})}_{\text{Generalized Sylvester Matrix}} = \begin{bmatrix} \phi_r(f_1) \\ \phi_d(f_2) \\ \vdots \\ \phi_d(f_k) \end{bmatrix} \in \mathbb{R}^{(r+(k-1)d) \times (r+d)}.$$

Generalized Sylvester Matrices

Theorem
$\gcd(f) = 1 \iff \mathrm{Syl}(f)$ *has full rank.*

Problem: What if the degrees of f can increase?

- Degrees of f can be at-most $\mathbf{d}' = (d'_1, \ldots, d'_k)$
- Spurious solutions can occur due to over-padding of zeros
- Define $\mathrm{rev}_{d'_j}(f_j) = t^{d_j} f(t^{-1})$
- Define $\mathrm{rev}_{\mathbf{d}'}(f)$ in the obvious way

Theorem
If $\mathrm{Syl}_{\mathbf{d}'}(f)$ *rank deficient then* $\gcd(f) = 1$ *iff* $\mathrm{Syl}(\mathrm{rev}_{\mathbf{d}'}(f))$ *full rank.*

Approximate SNF via Sylvester Matrices

Theorem

A nearest rank at most e Sylvester matrix always exists.

Theorem

Suppose that $\mathbf{d}' = (\gamma, \gamma \ldots, \gamma)$ *and* $\mathrm{Syl}_{\mathbf{d}'}(\mathrm{Adj}(\mathcal{A}))$ *has rank* e.

$$\frac{\sigma_e(\mathrm{Syl}_{\mathrm{dd}'}(\mathrm{Adj}(\mathcal{A})))}{\gamma n^3(d+1)^{3/2}\|\mathcal{A}\|_\infty^n n^{n/2}} \leq \|\mathcal{A}-\widehat{\mathcal{A}}\|_F, \text{ where } \mathrm{SNF}(\widehat{\mathcal{A}}) \text{ is non-trivial.}$$

- $\sigma_e(\mathrm{Syl}_{\mathbf{d}'}(\mathrm{Adj}(\mathcal{A})))$ *is the distance to a nearest singular matrix*

Example (Same $\mathcal{A}$ as the First Example)

A lower bound on the distance to non-triviality is $4.3556e-4$.

Nearest Matrix Polynomial with an Interesting SNF

Constrained Optimization Approach

$$\min \| \underbrace{\mathcal{A} - \widehat{\mathcal{A}}}_{\Delta\mathcal{A}} \|_F^2 \quad \text{such that} \quad \begin{cases} \operatorname{Adj}(\widehat{\mathcal{A}}) = \mathcal{F}\mathrm{h}, \\ \qquad \mathcal{F} \in \mathbb{R}[\mathrm{t}]^{n \times n}, \\ \qquad \mathrm{h} = \mathrm{h}_0 + \mathrm{h}_1 \mathrm{t} + \mathrm{h}_2 \mathrm{t}^2, \\ \qquad \mathrm{h}_2^2 + \mathrm{h}_1^2 - 1 = 0. \end{cases}$$

- Assume the adjoint has a finite approximate GCD
  - Otherwise the reversal has a non-trivial GCD
- Generically, the approximate GCD has degree 1 or 2
- $\mathrm{h}_2^2 + \mathrm{h}_1^2 - 1 = 0 \implies \mathrm{h}$ has degree at least 1
- Solve with Lagrange Multipliers and Levenberg-Marquardt

Levenberg-Marquardt Iteration

Theorem

The Levenberg-Marquardt iteration converges quadratically to the minimum value with a suitable initial guess.

Corollary

Under small perturbations:

- *Well-posed approximate SNF instances remain well-posed.*
- *Ill-posed instances cannot be regularized to be well-posed.*
- Theory applies by induction to arbitrary McCoy rank
- Applies to infinite eigenvalues: consider $\mathrm{t}^{\mathrm{d}}\mathcal{A}(\mathrm{t}^{-1})$
- This is why existing algorithms fail and cannot be saved

Algorithm and Implementation in `Maple 2017`

- Compute derivatives quickly
  - Partial two variable ansatz and evaluation
- $\mathrm{Adj}(\mathcal{A} + \Delta\mathcal{A})$ has exponentially many coefficients
- Compute derivatives of $\mathrm{Adj}(\cdot)$ quickly
  - Details in an upcoming paper!
- LM iteration cost is polynomial $\mathrm{O}(\mathrm{n}^9\mathrm{d}^3)$ flops for $\mathrm{r} = 2$
  - Grows exponentially in r, the specified McCoy Rank deficiency

Initial Guess

- Compute approximate GCD of the adjoint matrix
- Approximate Lagrange multipliers with linear least squares