Algebraic Companions

Robert M. Corless

June 2022

Editor-in-Chief, Maple Transactions
Western University, Canada

talk intended for ILAS 2022, Galway, Ireland
These slides available at rcorless.github.io

This talk is based on the following papers

1. Eunice Chan & RMC, A New Kind of Companion Matrix (ELA 2017)
2. Eunice Chan & RMC, Minimal Height Companion Matrices for Euclid Polynomials (Math. Comput. Sci. 2019)
3. Eunice Chan et al, Algebraic Linearizations (LAA 2019)
4. Eunice Chan, RMC, & Leili Rafiee Sevyeri, Generalized Standard Triples (ELA 2021)

Contributions of many: Neil Calkin, Lalo Gonzalez-Vega, Don Knuth, Piers Lawrence, Juana Sendra, Rafa Sendra, and Steven Thornton, are gratefully acknowledged. I also thank Froilán Dopico for exceptionally detailed and patient editorial work for that last paper!

Mandelbrot polynomials and Matrices

The talk is also related to *Mandelbrot polynomials and matrices*.

1. Piers Lawrence & RMC, The Largest Root of the Mandelbrot Polynomials (Jonfest proceedings, 2013)
2. Bini and Robol's MPSolve paper (JCAM 2014) (version 1 was 2000, Bini & Fiorentino)
3. Neil J Calkin, Eunice Chan, & RMC, Some Facts and Conjectures about Mandelbrot Polynomials (Maple Transactions 2021)
4. Eunice Chan et al, A Fractal Eigenvector (American Math Monthly 2022)

Piers Lawrence had the fundamental idea which opened the door to these results.

NB: There is also a strongly related paper from 2017 by Robol, Vandebril, and Van Dooren.

Bohemian matrices

Another related thread of work: Bohemian Matrices

1. cover image: London Mathematical Society Newsletter, November 2020, page 16 (RMC, NJ Higham, & SE Thornton)
2. Upper H–berg and Toeplitz Bohemians (Chan et al, 2020, LAA)
3. What can we learn from Bohemian matrices? (RMC, 2021)
4. Skew-symmetric tridiagonal Bohemian matrices (RMC 2021 Maple Transactions)
5. Computational Discovery on Jupyter (chapter 4) (an OER by Neil Calkin, Eunice Chan, and RMC 2022, and to be a SIAM book :)

Algebraic companions

Suppose we have local linearizations (A_a,B_a) for dimension n matrix polynomial $a(x)$, and (A_b, B_b) for $b(x)$, with

$$\begin{aligned} E_a(z)(zB_a - A_a)F_a(z) =& \mathrm{diag}(a(z), I_{N_a-n}) \\ E_b(z)(zB_b - A_b)F_b(z) =& \mathrm{diag}(b(z), I_{N_b-n}) \end{aligned} \quad (1)$$

and we wish to construct a local linearization (A_c,B_c) for $c(x) = xa(x)b(x) + d$.

Suppose that we do not wish to expand this out, because we are afraid of making the conditioning worse.

Theorem 1.7 in the GST paper

Let $\boldsymbol{E}_a(z)$ and $\boldsymbol{F}_a(z)$ be rational matrices such that if $z \in \Sigma_a$ (ie the region in which the local linearization of $\boldsymbol{a}$ is valid) then $\boldsymbol{E}_a(z)$ and $\boldsymbol{F}_a(z)$ are invertible and $\boldsymbol{E}_a(z)(z\boldsymbol{B}_a - \boldsymbol{A}_a)\boldsymbol{F}_a(z) = \mathrm{diag}(a(z), I_{N_a - n})$, and likewise let $\boldsymbol{E}_b(z)$ and $\boldsymbol{F}_b(z)$ be rational matrices such that if $z \in \Sigma_b$ then $\boldsymbol{E}_b(z)$ and $\boldsymbol{F}_b(z)$ are invertible and $\boldsymbol{E}_b(z)(z\boldsymbol{B}_b - \boldsymbol{A}_b)\boldsymbol{F}_b(z) = \mathrm{diag}(b(z), I_{N_b - n})$.

Then the pencil $z\boldsymbol{B}_c - \boldsymbol{A}_c$ is a local linearization of $\boldsymbol{c}(z) = z\boldsymbol{a}(z)\boldsymbol{b}(z) + d$ for $z \in \Sigma_a \cap \Sigma_b$, where the matrices $\boldsymbol{B}_c$ and $\boldsymbol{A}_c$ are given on the next slides:

The constructed (block upper Hessenberg) linearization

$$B_c = \begin{bmatrix} B_a & & \\ & I_n & \\ & & B_b \end{bmatrix} \tag{2}$$

and

$$A_c = \begin{bmatrix} A_a & 0_{N_a,n} & -Y_a d X_b \\ -X_a & 0_n & 0_{n,N_b} \\ 0_{N_b,N_a} & -Y_b & B_a \end{bmatrix}. \tag{3}$$

Here $X_a = [I_n, 0, \ldots, 0] F_a^{-1}(z)$, $Y_a = E_a^{-1}(z)[I_n, 0, \ldots, 0]^T$ and likewise $X_B = [I_n, 0, \ldots, 0] F_b^{-1}(z)$, and $Y_b = E_a^{-1}(z)[I_n, 0, \ldots, 0]^T$ give the elements of the (generalized) standard triples for $a(z)$ and $b(z)$.

The start of the proof

Form

$$E_1(z) = \begin{bmatrix} E_a(z) & & \\ & I_n & \\ & & E_b(z) \end{bmatrix} \tag{4}$$

and

$$F_1(z) = \begin{bmatrix} F_A(z) & & \\ & I_n & \\ & & F_B(z) \end{bmatrix}, \tag{5}$$

and form $E_1(xB_c - A_c)F_1$ to start. We will have to apply various block permutations, and the key fact, which follows.

Key point of the proof (for which we thank a referee)

$$\begin{bmatrix} I_n & -a(z) \\ & I_n \end{bmatrix} \begin{bmatrix} a(z) & d \\ I_n & -zb(z) \end{bmatrix} \begin{bmatrix} zb(z) & I_n \\ I_n & \end{bmatrix} = \begin{bmatrix} 0 & d + zab \\ I_n & -zb(z) \end{bmatrix} \begin{bmatrix} zb(z) & I_n \\ I_n & \end{bmatrix}$$
$$= \begin{bmatrix} za(z)b(z) + d & \\ & I_n \end{bmatrix} .$$

This gives a whole different class of possible linearizations. For instance, consider a variation of Newton's example polynomial, namely $p(x) = x^3 - Tx - 5 = x(x - \sqrt{T})(x + \sqrt{T}) - 5$. Algebraic linearization gives

$$\mathbf{A} = \begin{bmatrix} \sqrt{T} & 0 & 5 \\ -1 & 0 & 0 \\ 0 & -1 & -\sqrt{T} \end{bmatrix} \tag{6}$$

as a companion matrix. Computing the eigenvalues of *this* matrix, when $T = 2 \cdot 10^5$, results in a relative error of $1.4 \cdot 10^{-13}$ in the smallest eigenvalue, whereas using the Frobenius companion forces an error of about 10^{-9}.

Varying T

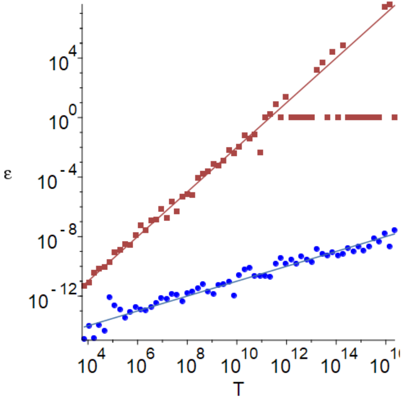

Figure 1: Relative error in smallest eigenvalue: Algebraic Linearization vs Frobenius Linearization, as the parameter T varies in $x^3 - Tx - 5$. Fits: $10^{-16} \cdot \sqrt{T}$ (blue, Algebraic), $10^{-17} \cdot T^{3/2}$ (red, Frobenius).

A bigger example

We created an $n = 3$, grade 5 example by choosing a grade 2 $\boldsymbol{A}$ and a grade 2 $\boldsymbol{B}$ and a $\boldsymbol{D}$ and forming $\boldsymbol{C} = z\boldsymbol{AB} + \boldsymbol{D}$. We perturbed it in two different ways, and compared the algebraic linearization (Frobenius for $\boldsymbol{A}$ and $\boldsymbol{B}$) to the ordinary (2nd) Frobenius linearization for the explicitly expanded $\boldsymbol{C}$.

Preliminary results

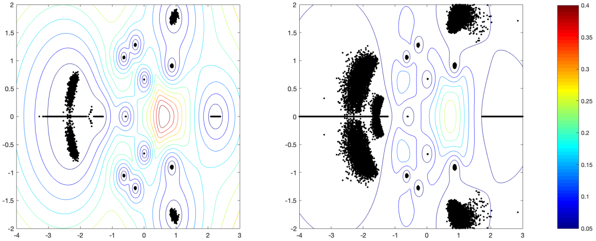

Figure 2: Pseudospectra of two different kinds of linearizations for our test equation which is expressed in the monomial basis. The linearization constructions used are algebraic linearization (left) and Frobenius linearization (right). [Graph courtesy Eunice Y. S. Chan.]

Unresolved questions

We *think* that the potentially improved numerical stability arises because the *height* of the new matrices can be lower.

$\mathrm{Height}(A) := \|\mathrm{vec}(A)\|_\infty$ is a matrix norm, but not a submultiplicative one. For instance, consider

$$\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} . \tag{7}$$

The height of $\boldsymbol{AB}$ is not necessarily less than the height of $\boldsymbol{A}$ times the height of $\boldsymbol{B}$.

Also, the height of a matrix can be forced to 1 by scaling, so we are really worrying about the smallest nonzero elements after such a scaling.

Minimal height companions/linearizations

If we are given a recursive construction, this idea makes sense. But if we are given a fully formed matrix polynomial $P(z)$, can we *construct* factors in a reasonable way? And how far can this be taken?

An alternative question: if the entries of the (matrix) polynomial coefficients are integers, what is the *minimal height* linearization? And how do we compute it? This looks like a discrete optimization problem. [I have asked some of my friends for advice but so far they have all looked rather helplessly at me.]

NB: As exemplified by the Mandelbrot matrices, the *minimal height* may be *exponentially smaller* than the size of the coefficients of the original polynomial.

Thank you!

Happy to take questions!

This work was partially supported by NSERC grant RGPIN-2020-06438, and partially supported by the grant PID2020-113192GB-I00 (Mathematical Visualization: Foundations, Algorithms and Applications) from the Spanish MICINN.