

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 Neil Calkin 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, Vandebriël, and Van Dooren.

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 $(\mathbf{A}_a, \mathbf{B}_a)$ for dimension n matrix polynomial $a(x)$, and $(\mathbf{A}_b, \mathbf{B}_b)$ for $b(x)$, with

$$\begin{aligned} E_a(z)(z\mathbf{B}_a - \mathbf{A}_a)F_a(z) &= \text{diag}(a(z), I_{N_a-n}) \\ E_b(z)(z\mathbf{B}_b - \mathbf{A}_b)F_b(z) &= \text{diag}(b(z), I_{N_b-n}) \end{aligned} \tag{1}$$

and we wish to construct a local linearization $(\mathbf{A}_c, \mathbf{B}_c)$ 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 $E_a(z)$ and $F_a(z)$ be rational matrices such that if $z \in \Sigma_a$ (ie the region in which the local linearization of \mathbf{a} is valid) then $E_a(z)$ and $F_a(z)$ are invertible and $E_a(z)(z\mathbf{B}_a - \mathbf{A}_a)F_a(z) = \text{diag}(a(z), I_{N_a-n})$, and likewise let $E_b(z)$ and $F_b(z)$ be rational matrices such that if $z \in \Sigma_b$ then $E_b(z)$ and $F_b(z)$ are invertible and $E_b(z)(z\mathbf{B}_b - \mathbf{A}_b)F_b(z) = \text{diag}(b(z), I_{N_b-n})$.

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

The constructed (block upper Hessenberg) linearization

$$B_c = \begin{bmatrix} B_a & & \\ & I_n & \\ & & B_b \end{bmatrix} \quad (2)$$

and

$$A_c = \begin{bmatrix} A_a & \mathbf{0}_{N_a, n} & -Y_a dX_b \\ -X_a & \mathbf{0}_n & \mathbf{0}_{n, N_b} \\ \mathbf{0}_{N_b, N_a} & -Y_b & A_b \end{bmatrix}. \quad (3)$$

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

The start of the proof

Form

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

and

$$F_1(z) = \begin{bmatrix} F_a(z) & & \\ & I_n & \\ & & F_b(z) \end{bmatrix}, \quad (5)$$

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)

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

Why this might be interesting

This gives a whole different class of possible linearizations¹. 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

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

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} .

¹The theory is actually in Gohberg, Lancaster, and Rodman, though!

Varying T

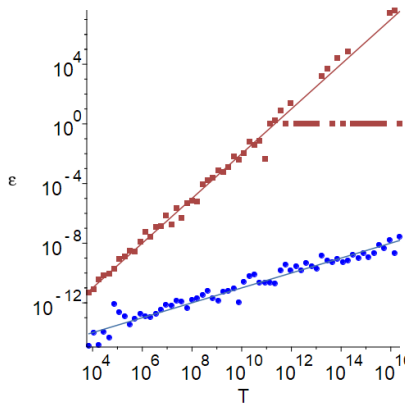


Figure 1: Relative error in smallest eigenvalue: Algebraic Linearization vs Frobenius Linearization, as the parameter T varies in $x^3 - Tx - 5$. Fits: $10^{-16} \cdot \sqrt{T}$ (blue, Algebraic), $10^{-17} \cdot T^{3/2}$ (red, Frobenius).

A bigger example

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

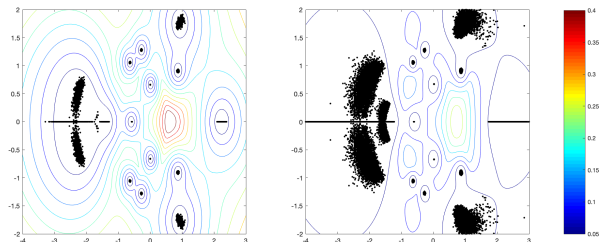


Figure 2: Pseudospectra of two different kinds of linearizations for our test equation which is expressed in the monomial basis. The linearization constructions used are algebraic linearization (left) and Frobenius linearization (right). [Graph courtesy Eunice Y. S. Chan.]

Unresolved questions

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

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

$$\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}. \quad (7)$$

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