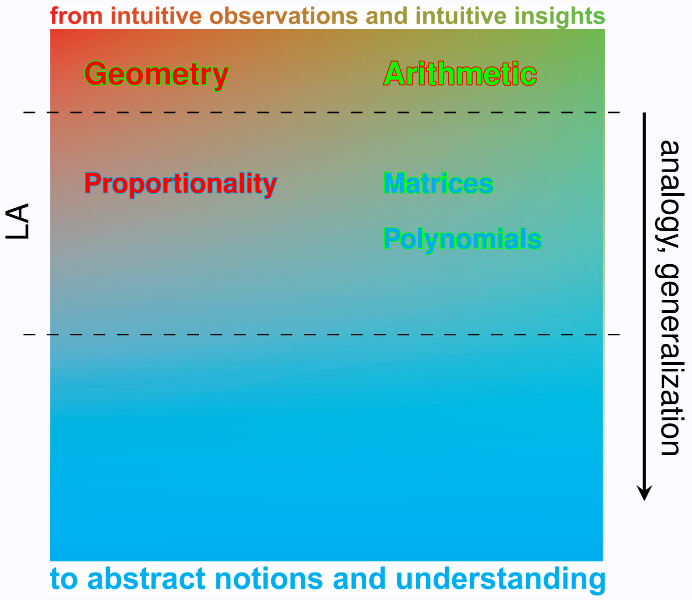
ILAS-ED Seminar

Zoom ID 929 2504 4372 Meeting, September 27, 2022

(Matrix) Zeros of polynomials

Damjan Kobal, University of Ljubljana

Mathematics learning



... from intuition to abstraction geometry-algebra sample ...

https://www.geogebra.org/m/vguhhxzj

Matrix zeros of polynomials

D. Kobal, *Matrix zeros of polynomials*, The Mathematical Gazette, Cambridge Univeristy Press, Vol. 104, March 2020, p. 27 - 35.

Presented also at the *Linear algebra education minisymposia*, ILAS 2022, Galway, Ireland, June 2022.

Introduction

- The concepts of polynomials and matrices generalize the elementary arithmetic of numbers.
- Our elementary exploration of polynomials and matrices offers an interesting matrix analogue to the concept of a zero of a polynomial.
- The discussion offers an opportunity for better comprehension of the fundamental concepts of polynomials and matrices.
- As an application we are lead to the meaningful questions of linear algebra and to an easy proof of otherwise advanced Cayley - Hamilton theorem.

The arithmetic of matrices and polynomials

- After mastering the basic polynomial and matrix calculations, including the concepts of square matrices (over a field, usually $\mathbb{C}$), identity and inverse matrices, we can discuss also 'polynomials of matrices' and 'matrices of polynomials.'
- In what follows, all our matrices will be square matrices.

Polynomials of matrices

- $p(x), M \longrightarrow p(M)$
- $p(x) = x^2 - 3x + 2, M = \begin{bmatrix} 3 & 2 & -2 \\ 1 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix} \longrightarrow p(M) = \begin{bmatrix} 0 & -2 & 0 \\ -1 & 0 & 1 \\ 0 & -2 & 0 \end{bmatrix}.$

Matrices of polynomials

- For example

$$\begin{bmatrix} x^2+x+1 & x \\ 1 & 2x^2 \end{bmatrix} = x^2 \cdot \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} + x \cdot \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}. \quad (1)$$

- Or

$$\begin{bmatrix} x^2+x+1 & x \\ 1 & 2x^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \cdot x^2 + \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \cdot x + \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

- Notation (*polynomial matrix*) $M(x) = \begin{bmatrix} x^2+x+1 & x \\ 1 & 2x^2 \end{bmatrix}$

Remark 1

For a polynomial matrix $M(x)$ and matrix A, it makes no sense to talk about $M(A)$.

Matrices and zeros of polynomials

- A number m is a zero of the polynomial $p(x)$ if and only if $p(m) = 0$.
- For matrix M and polynomial $p(x)$, we might get $p(M) = 0$
- For example, for $p(x) = x^2 - 3x + 2$ for

$$M = \begin{bmatrix} 3 & 2 & -2 \\ 1 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix} \text{ and for } N = \begin{bmatrix} 3 & 2 & -2 \\ 1 & 2 & -1 \\ 2 & 2 & -1 \end{bmatrix}$$

we get

$$p(M) = \begin{bmatrix} 0 & -2 & 0 \\ -1 & 0 & 1 \\ 0 & -2 & 0 \end{bmatrix} \text{ and } p(N) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

- *Definition :* We say that a matrix M is a *matrix zero* of the polynomial $p(x)$ if and only if $p(M)$ is the zero matrix.
- Could anything be said about matrix zeros of a given polynomial?

Matrix zero of a polynomial - exploration

- Could we find a matrix zero of a simple polynomial

$$p(x) = x^2 - 3x + 2 = (x-1)(x-2)?$$

  One such matrix zero, which looks quite arbitrary, was given above. Could we find another? Or maybe many others?
- Zeros 1 and 2 as 'one-dimensional' matrices $[1]$ and $[2]$. Zeros of the form $1 \cdot I$ or $2 \cdot I$.
- Diagonal matrices D

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

  seem mysterious, but anticipated, 'matrix zeros.'
- Calculating $(D-I)(D-2I)$ gives a nice intuitive understanding of the 'block matrix multiplication. (These observations will be valuable later, when we learn about eigenvalues.)

Matrix zero of a polynomial - further exploration

- $$\left.\begin{array}{r} p(D) = 0 \\ M = P \cdot D \cdot P^{-1} \end{array}\right\} \longrightarrow p(M) = 0$$
- At this point students should ... notice that
  - 'matrix zeros' of a polynomial are quite complex,
  - already a (simple) polynomial might have infinitely many 'matrix zeros,'
  - considering D and P as above, we have a pattern to find many 'matrix zeros,'
  - this pattern is well-hidden within obtained 'matrix zeros' (by free choice of the invertible matrix P, the obtained matrix M looks quite arbitrary).

Theorem 1

The number m is a zero of the polynomial $p(x)$ if and only if $p(x) = r(x) \cdot (x - m)$ for some polynomial $r(x)$.

Proof.

$(\Longleftarrow)$ $p(x) = r(x) \cdot (x - m) \longrightarrow p(m) = r(x) \cdot (m - m) = 0.$

$(\Longrightarrow)$ $p(m) = 0$

$$
\begin{aligned}
x^i - m^i &= (x^{i-1} + x^{i-2}m + \cdots + xm^{i-2} + m^{i-1})(x - m) \\
p(x) - p(m) &= \textstyle\sum_{i=1}^{n} a_i \cdot (x^{i-1} + x^{i-2}m + \cdots + xm^{i-2} + m^{i-1})(x - m) \\
&= r(x) \cdot (x - m) \\
p(x) &= p(m) + r(x) \cdot (x - m) = r(x) \cdot (x - m)
\end{aligned}
$$

□

What could be a sensible analogue of Theorem 1 for matrices?

$$p(x) = r(x) \cdot (x - m)$$

$$p(x) \cdot I = R(x) \cdot (x \cdot I - M)$$

Theorem 2

The matrix M is a matrix zero of the polynomial $p(x)$ if and only if

$$p(x) \cdot I = R(x) \cdot (xI - M)$$

for some polynomial matrix $R(x)$, where I is the identity matrix.

Example

For $p(x) = x^2 - 3x + 2$ and $N = \begin{bmatrix} 3 & 2 & -2 \\ 1 & 2 & -1 \\ 2 & 2 & -1 \end{bmatrix}$, we have $p(N) = 0$.

$$(x^2 - 3x + 2) \cdot I = \begin{bmatrix} x & 2 & -2 \\ 1 & x-1 & -1 \\ 2 & 2 & x-4 \end{bmatrix} \cdot (xI - N)$$

Remark 2

As in Theorem 1

$$p(M) = 0 \longrightarrow p(x) \cdot I = R(x) \cdot (xI - M)$$

Conversely (trivial in 'number-case')

$$p(x) \cdot I = R(x) \cdot (xI - M) \not\longrightarrow p(M) = 0$$

We can not simply calculate $p(M)I = R(M) \cdot (M \cdot I - M) = 0$, as $R(M)$ does not make sense.

Remark 3

Somehow similar mistake: For characteristic polynomial

$$\chi_M(x) = \det(xI - M) \not\longrightarrow \chi_M(M) = \det(M \cdot I - M) = 0$$

Full understanding of Cayley-Hamilton theorem is hard ..., but every student to whom we consider 'worthy' to introduce the 'characteristic polynomial', should understand, that in this 'equation' we have a '$n \times n$' matrix on the left and a number '0' on the right.

Analogy between ‘number’ and ‘matrix’ case.

Lemma 1

If $p(x)$ is a polynomial such that $p(x) \cdot (x - a) = c$, where a and c are constants, then $c = 0$ and $p(x) \equiv 0$.

Proof.

$$p(x) = b_n x^n + b_{n-1} x^{n-1} + \cdots + b_1 x + b_0$$

$$c = p(x) \cdot (x - a) = b_n x^{n+1} + (b_{n-1} - ab_n) x^n + \cdots + (b_0 - b_1 a) x - b_0 a.$$

$$\longrightarrow b_n = 0 \longrightarrow b_{n-1} = 0 \longrightarrow \ldots \longrightarrow b_0 = 0 \longrightarrow p(x) \equiv 0 \longrightarrow c = 0$$

□

Lemma 2

If $N(x)$ is a polynomial matrix such that $N(x) \cdot (xI - A) = M$, where A and M are matrices (with constant entries), then M must be the zero matrix and $N(x) \equiv 0$.

Proof.

$N(x)$ can be written (see (1))

$$N(x) = N_n x^n + N_{n-1} x^{n-1} + \cdots + N_1 x + N_0$$

$$\longrightarrow M = N_n x^{n+1} + (N_{n-1} - N_n \cdot A) x^n + \cdots + (N_0 - N_1 \cdot A) x - N_0 \cdot A$$

$$\longrightarrow N_n = 0 \longrightarrow N_{n-1} = 0 \longrightarrow \ldots \longrightarrow N_0 = 0 \longrightarrow N(x) \equiv 0 \longrightarrow M = 0$$

□

Theorem 2 (again)

The matrix M is a matrix zero of the polynomial $p(x)$ if and only if

$$p(x) \cdot I = R(x) \cdot (xI - M)$$

for some polynomial matrix $R(x)$, where I is the identity matrix.

Proof.

$(\Longrightarrow)$ $p(M) = 0$

$$\begin{aligned} x^i I - M^i &= (x^{i-1} I + x^{i-2} M + \cdots + x M^{i-2} + M^{i-1}) \cdot (xI - M) \\ p(x)I - p(M) &= \textstyle\sum_{i=1}^{n} a_i (x^{i-1} I + x^{i-2} M + \cdots + M^{i-1}) \cdot (xI - M) \\ &= Q(x) \cdot (xI - M) \\ p(x)I &= p(M) + Q(x) \cdot (xI - M) \end{aligned}$$

$$p(x)I = p(M) + Q(x) \cdot (xI - M). \tag{2}$$

$$p(M) = 0 \longrightarrow p(x)I = Q(x) \cdot (xI - M).$$

$(\Longleftarrow)$ $p(x)I = R(x) \cdot (xI - M) \wedge (2)$:

$$\begin{aligned} R(x) \cdot (xI - M) &= p(M) + Q(x) \cdot (xI - M) \\ p(M) &= \big(R(x) - Q(x)\big) \cdot (xI - M) \end{aligned}$$

By Lemma 2 $R(x) \equiv Q(x)$ and $p(M) = 0$.

Conclusion

- Our 'matrix theorem' is a nice analogue of 'number theorem'.
- Still do not know how to get such a $p(x)$ that $p(M) = 0$ for a given M.
- No simple answer $\longrightarrow$ complexity of the matrix algebra.
- The unanswered question how to find $p(x)\dots$ could be nicely addressed later:
  - Matrix M is an element of the vector space of $n \times n$ matrices, which has dimension n^2.
  - $I = M^0, M^1, M^2, \dots, M^{n^2}$ must be a set of linearly dependent vectors...
  - So, there must be constants $a_0, a_1, a_2, \dots, a_{n^2}$ such that

    $a_0 I + a_1 M + a_2 M^2 + \dots + a_{n^2} M^{n^2} = 0.$
  - Therefore, for $p(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_{n^2} x^{n^2}$, we have $p(M) = 0$.
  - With Cayley-Hamilton theorem the statement is refined in the form of proving the existence of such a polynomial of much smaller degree.

Conclusion - Elementary proof of Cayley-Hamilton theorem

- Cayley-Hamilton theorem: For *characteristic polynomial* $\chi_M(x) = \det(xI - M)$, we have $\chi_M(M) = 0$.
- To understand characteristic polynomial, one must understand *determinant*.
- To understand *determinant*, one must understand 'row and column Laplace expansion', which is also an algorithm for obtaining the matrix inverse, if it exists:

$$\det(M) \cdot I = \operatorname{adj}(M) \cdot M. \tag{3}$$

- If M is any matrix, then $xI - M$ is a rather simple polynomial matrix. Putting $xI - M$ into the equation 3, we get

$$\det(xI - M) \cdot I = \operatorname{adj}(xI - M) \cdot (xI - M).$$

- This equality requires an understanding of the Laplace expansion, which is usually done for matrices with constant entries. It is a straightforward argument that analogous statements hold if elements of the matrix are polynomials.
- Cayley-Hamilton theorem is a trivial consequence of our theorem.

Appendix 1

- Standard Cayley-Hamilton theorem proofs are commonly 'computational' (for example, Lemma 1.9, p. 426, J. Hefferon, *Linear Algebra*, Orthogonal Publishing L3C, 2017) with essentially elementary reasoning but using advanced and unnecessary specific notations.
- The proof of the Cayley-Hamilton theorem is sometimes given within more abstract setting (for example p. 194 and p. 237, K. Hoffman and R. Kunze, *Linear Algebra*, Prentice Hall, 1971). Also our Theorem 2 and consequentially the Cayley-Hamilton theorem proof could be stated for polynomials and matrices over a ring R and considering associated R-modules instead of vector spaces.
- Cayley-Hamilton theorem proof is often avoided and only illustrated by examples. Sometimes the proof is given only for 'diagonalizable matrices'.
- At an advanced level 'the proof for only diagonalizable matrices' might be connected to *metric space* concepts, where one could use the fact (for example for matrices over $\mathbb{C}$, which suffice for most purposes) that diagonalizable matrices are dense within the space of all matrices. With good understanding of continuity, this suffices to conclude that the Cayley-Hamilton theorem holds for all matrices.

Appendix 2

- Arthur Cayley (1821 - 1895) in his original paper (A Memoir on the Theory of Matrices, *Philosophical Transactions of the Royal Society of London*, **148**, 1858, p. 17–37) only proved his theorem for an arbitrary 2×2 matrix

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

in the form of explicitly calculating the expression $M^2 - (a+d)M^1 + (ad - bc)M^0$ to be the zero matrix. In modern terminology, this corresponds to the statement that $\chi_M(M) = 0$ for $\chi_M(x) = \det(xI - M)$. Cayley indicated a similar proof for an arbitrary 3×3 matrix. For matrices of higher dimensions he wrote "*I have not thought it necessary to undertake the labour of a formal proof of the theorem in the general case of a matrix of any degree*".
- The theorem was first proved in its general form by Ferdinand Georg Frobenius (1849 - 1917) in his paper Über vertauschbare Matrizen, *Sitz. Preuß. Akad. Wiss. Berlin*, 1896, p. 601–614. Our proof somehow resembles Frobenious proof.