Determining the Determinant: learning in the footsteps of Cramer and Cauchy

Danny Otero

Xavier University

JMM 2018, San Diego, CA
January 11, 2018

TRIUMPHS

TRansforming Instruction in Undergraduate Mathematics via Primary Historical Sources (NSF grant no. 1524065)

Principal Investigators:

- Janet H. Barnett (Colorado State Univ.–Pueblo)
- Kathleen Clark (Florida State Univ.)
- Dominic Klyve (Central Washington State Univ.)
- Jerry Lodder (New Mexico State Univ.)
- Danny Otero (Xavier Univ.)
- Nicholas Scoville (Ursinus College)
- Diana White (Univ. of Colorado, Denver)

Goal

To motivate and provide context for the study of mathematics by undergraduates at all levels by closely reading primary source materials.

Goal

To motivate and provide context for the study of mathematics by undergraduates at all levels by closely reading primary source materials.

Benefits of Primary Sources for Learning

- Motivates abstract concepts
- Illustrates the human, creative, artistic, and dynamic nature of mathematics
- Promotes understanding of present-day paradigms through the reading of an historical source which requires no knowledge of that paradigm
- Draws attention to subtleties which modern texts often take for granted
- Engenders cognitive dissonance (*dépaysement*) when comparing a historical source with modern approaches

Goal of this Project

To present a meaningful definition of, and the fundamental properties of the matrix determinant.

Goal of this Project

To present a meaningful definition of, and the fundamental properties of the matrix determinant.

Challenges

- The determinant is a difficult concept to teach in a unified manner for its symbolic complexity. In Axler's famous 1995 article "Down with determinants!" he advocates avoiding it altogether in the undergraduate linear algebra course, as it is "non-intuitive, and often defined without motivation."
- The history of the determinant is long and complicated, and it took a long while for the idea to coalesce into its modern form.

The Plan

Present the determinant in a lesson of about 1.5 weeks of classroom time, through two important works in the history of the development of the determinant:

The Plan

Present the determinant in a lesson of about 1.5 weeks of classroom time, through two important works in the history of the development of the determinant:

* Gabriel Cramer's publication of his Rule in an appendix of *Introduction à l'Analyse des Lignes Courbes Algébriques* [*An Introduction to the Analysis of Algebraic Curves*] (1750)

The Plan

Present the determinant in a lesson of about 1.5 weeks of classroom time, through two important works in the history of the development of the determinant:

* Gabriel Cramer's publication of his Rule in an appendix of *Introduction à l'Analyse des Lignes Courbes Algébriques* [*An Introduction to the Analysis of Algebraic Curves*] (1750)
* Augustin-Louis Cauchy's monumental memoir *Functions qui ne peuvent obtenir que deux valeurs égales et de signes contraires par suit des transpositions opérées entre les variables qu'elles renferment* [*On functions which take only two values, equal but of opposite sign, by means of transpositions performed among the variables which are contained therein*] (1815), read to the Paris Academy in 1812 – at age 23!

Students Reading and Writing

Focus on student activity

- Students read excerpts from the primary texts (in English translation)

Students Reading and Writing

Focus on student activity

- Students read excerpts from the primary texts (in English translation)
- ...presented with commentary text to establish context, highlight important features, and make connections with modern modes of description (notation, terminology)

Students Reading and Writing

Focus on student activity

- Students read excerpts from the primary texts (in English translation)
- ...presented with commentary text to establish context, highlight important features, and make connections with modern modes of description (notation, terminology)
- ...and the careful crafting of student tasks to aid in the construction of robust conceptual understanding.

A Rule for Solving Linear Systems

- Cramer presents formulas for the solutions to a linear system of n equations in n variables for $n = 1, 2, 3$

A Rule for Solving Linear Systems

- Cramer presents formulas for the solutions to a linear system of n equations in n variables for $n = 1, 2, 3$
- Here are his formulas for $n = 3$: if the system is

$$\begin{array}{rcl} A^1 & = & Z^1 z + Y^1 y + X^1 x \\ A^2 & = & Z^2 z + Y^2 y + X^2 x \\ A^3 & = & Z^3 z + Y^3 y + X^3 x \end{array}$$

(variables in lowercase), then the solutions are given by

$$z = \frac{A^1Y^2X^3 - A^1Y^3X^2 - A^2Y^1X^3 + A^2Y^3X^1 + A^3Y^1X^2 - A^3Y^2X^1}{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1}$$

$$y = \frac{Z^1A^2X^3 - Z^1A^3X^2 - Z^2A^1X^3 + Z^2A^3X^1 + Z^3A^1X^2 - Z^3A^2X^1}{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1}$$

$$x = \frac{Z^1Y^2A^3 - Z^1Y^3A^2 - Z^2Y^1A^3 + Z^2Y^3A^1 + Z^3Y^1A^2 - Z^3Y^2A^1}{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1}$$

A Rule for Solving Linear Systems

- Cramer concludes by describing the solution formulas for the general case:

  we will find the value of each unknown by forming n fractions of which the common denominator has as many terms as there are diverse arrangements of n different things. Each term is composed of the letters $ZYXV$ &c., always written in the same order, but to which we distribute, as exponents, the first n numbers arranged in all possible ways.
- Signs are assigned to each such term according to the parity of the derangement of 1, 2, ..., n given by the sequence of "exponents"

A Rule for Solving Linear Systems

- *The common denominator being thus formed, we will have the value of z by giving to this denominator the numerator which is formed by changing, in all its terms, Z into A. ...And we find in a similar manner the value of the other unknowns.*

Alternating Symmetric Functions

Cauchy's memoir is in two parts:

I. He considers multivariable functions in a set of variables of the form $a_1, a_2, \ldots, a_n; b_1, b_2, \ldots, b_n; \ldots,$ in which transposing any one pair of indices changes only the sign of the function. These he calls *alternating symmetric functions*.

Alternating Symmetric Functions

Cauchy's memoir is in two parts:

I. He considers multivariable functions in a set of variables of the form $a_1, a_2, \ldots, a_n; b_1, b_2, \ldots, b_n; \ldots,$ in which transposing any one pair of indices changes only the sign of the function. These he calls *alternating symmetric functions*. He shows that arbitrary permutations performed upon these indices also produces either no change or a change of sign depending on the *parity* of the permutation (even or odd).

Alternating Symmetric Functions

Cauchy's memoir is in two parts:

I. He considers multivariable functions in a set of variables of the form $a_1, a_2, \ldots, a_n; b_1, b_2, \ldots, b_n; \ldots,$ in which transposing any one pair of indices changes only the sign of the function. These he calls *alternating symmetric functions*. He shows that arbitrary permutations performed upon these indices also produces either no change or a change of sign depending on the *parity* of the permutation (even or odd). He closes by saying that he will study

a certain kind of alternating symmetric function which present themselves in a great number of analytic investigations. It is by means of these functions that we express the general values of unknowns that many equations of the first degree contain. …Mr. Gauss has labeled these same functions with the name determinants.

Alternating Symmetric Functions

II. Cauchy calls a *symmetric system* an array in the n^2 quantities $a_{i,\,j}$, what we today call the $n \times n$ matrix (a_{ij}).

Alternating Symmetric Functions

II. Cauchy calls a *symmetric system* an array in the n^2 quantities $a_{i,j}$, what we today call the $n \times n$ matrix (a_{ij}).

- He uses the notation $\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n})$ to denote the determinant of the symmetric system ($\mathbf{S}(\pm\ \ )$ acts as an antisymmetrizing operator on the second set of indices of the *indicative* term $a_{1,1}a_{2,2}\ldots a_{n,n}$).

Alternating Symmetric Functions

II. Cauchy calls a *symmetric system* an array in the n^2 quantities $a_{i,j}$, what we today call the $n \times n$ matrix (a_{ij}).

- He uses the notation $\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n})$ to denote the determinant of the symmetric system ($\mathbf{S}(\pm\ \ )$ acts as an antisymmetrizing operator on the second set of indices of the *indicative* term $a_{1,1}a_{2,2}\ldots a_{n,n}$).
- Consideration of the action of the operator $\mathbf{S}(\pm\ \ )$ lets him conclude that $\det(a_{ij})^T = \det(a_{ij})$.

Alternating Symmetric Functions

II. Cauchy calls a *symmetric system* an array in the n^2 quantities $a_{i,j}$, what we today call the $n \times n$ matrix (a_{ij}).

- He uses the notation $\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n})$ to denote the determinant of the symmetric system ($\mathbf{S}(\pm\quad)$ acts as an antisymmetrizing operator on the second set of indices of the *indicative* term $a_{1,1}a_{2,2}\ldots a_{n,n}$).
- Consideration of the action of the operator $\mathbf{S}(\pm\quad)$ lets him conclude that $\det(a_{ij})^T = \det(a_{ij})$.
- $\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n}) = \mathbf{S}[\pm a_{n,n}\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n-1,n-1})]$

Alternating Symmetric Functions

II. Cauchy calls a *symmetric system* an array in the n^2 quantities $a_{i,j}$, what we today call the $n \times n$ matrix (a_{ij}).

- He uses the notation $\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n})$ to denote the determinant of the symmetric system ($\mathbf{S}(\pm\quad)$ acts as an antisymmetrizing operator on the second set of indices of the *indicative* term $a_{1,1}a_{2,2}\ldots a_{n,n}$).
- Consideration of the action of the operator $\mathbf{S}(\pm\quad)$ lets him conclude that $\det(a_{ij})^T = \det(a_{ij})$.
- $\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n}) = \mathbf{S}[\pm a_{n,n}\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n-1,n-1})]$

  The expression on the right is interpreted as an expansion of the form $a_{1,\mu}b_{1,\mu} + a_{2,\mu}b_{2,\mu} + \cdots + a_{n,\mu}b_{n,\mu}$, where "*$b_{\nu,\mu}$ is the adjoint to $a_{\nu,\mu}$.*"

Alternating Symmetric Functions

II. Cauchy calls a *symmetric system* an array in the n^2 quantities $a_{i,j}$, what we today call the $n \times n$ matrix (a_{ij}).

- He uses the notation $\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n})$ to denote the determinant of the symmetric system ($\mathbf{S}(\pm\ \ )$ acts as an antisymmetrizing operator on the second set of indices of the *indicative* term $a_{1,1}a_{2,2}\ldots a_{n,n}$).
- Consideration of the action of the operator $\mathbf{S}(\pm\ \ )$ lets him conclude that $\det(a_{ij})^T = \det(a_{ij})$.
- $\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n}) = \mathbf{S}[\pm a_{n,n}\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n-1,n-1})]$

  The expression on the right is interpreted as an expansion of the form $a_{1,\mu}b_{1,\mu} + a_{2,\mu}b_{2,\mu} + \cdots + a_{n,\mu}b_{n,\mu}$, where "*$b_{\nu,\mu}$ is the adjoint to $a_{\nu,\mu}$.*"
- From this he derives what in modern terms is the adjoint formula: $A \cdot \operatorname{adj} A = (\det A) \cdot I_n$.

Alternating Symmetric Functions

- From the adjoint formula, he then derives Cramer's Rule.

Alternating Symmetric Functions

- From the adjoint formula, he then derives Cramer's Rule.
- He develops the *composition* of symmetric systems, i.e., the multiplication of matrices: $m_{i,j}$ is a system whose *component systems* are $a_{i,j}$ and $\alpha_{i,j}$ (in modern form, $(a_{i,j})(\alpha_{j,k}) = (m_{i,k})$)

Alternating Symmetric Functions

- From the adjoint formula, he then derives Cramer's Rule.
- He develops the *composition* of symmetric systems, i.e., the multiplication of matrices: $m_{i,j}$ is a system whose *component systems* are $a_{i,j}$ and $\alpha_{i,j}$ (in modern form, $(a_{i,j})(\alpha_{j,k}) = (m_{i,k})$)
- Finally, a modification of the adjoint formula leads him to the product formula for determinants: where M_n, D_n, δ_n are the respective determinants of the systems $m_{i,j}, a_{i,j}$ and $\alpha_{i,j}$, he shows that $M_n = D_n\delta_n$.

Sample Tasks

Task 7 (Cramer's Rule)

(a) Write down the expression that is "the common denominator" of the formulas that "find the value of each of the unknowns" in the system of three linear equations in three unknowns. It is found both in each of the formulas in the source text.

(b) Each term in this expression is "composed of the letters ZYX, which receive successive exponents 123, 132, 213, 231, 312, 321." For each of the terms, beginning with $Z^1Y^2X^3$, count the number of derangements.

(c) According to Cramer, how does the number of derangements in each term identify the sign ($+$ or ?) it is supposed to receive in the "common denominator" expression? Verify this for each of the six terms.

Sample Tasks

Task 13 (Determinant of the Transpose)

Verify that the alternating symmetric sum $\mathbf{S}(\pm a_{1,1}\, a_{2,2}\, a_{3,3})$ in which the permutations are applied to the second indices in each term is identical to the the alternating symmetric sum $\mathbf{S}(\pm a_{1,1}\, a_{2,2}\, a_{3,3})$ in which the permutations are applied to the first indices. This shows directly that for any 3×3 matrix A,

$$detA^T = detA.$$

Thanks for your attention!

Questions?