The entry sum of the inverse Cauchy matrix

Darij Grinberg*

September 15, 2022

1. The Cauchy matrix

Let $x_1, x_2, \ldots, x_n$ be n numbers, and $y_1, y_2, \ldots, y_n$ be n further numbers chosen such that all n^2 pairwise sums $x_i + y_j$ are nonzero[1]. Consider the $n \times n$-matrix

$$C := \left(\frac{1}{x_i + y_j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n} = \begin{pmatrix} \frac{1}{x_1+y_1} & \frac{1}{x_1+y_2} & \cdots & \frac{1}{x_1+y_n} \\ \frac{1}{x_2+y_1} & \frac{1}{x_2+y_2} & \cdots & \frac{1}{x_2+y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{x_n+y_1} & \frac{1}{x_n+y_2} & \cdots & \frac{1}{x_n+y_n} \end{pmatrix}.$$

This matrix C is known as the *Cauchy matrix*, and has been studied for 180 years[2]. The first significant result was the formula for its determinant:

$$\det C = \frac{\prod\limits_{1 \leq i < j \leq n} \left(\left(x_i - x_j\right)\left(y_i - y_j\right)\right)}{\prod\limits_{(i,j) \in \{1,2,\ldots,n\}^2} \left(x_i + y_j\right)} \tag{1}$$

*Drexel University, Korman Center, 15 S 33rd Street, Philadelphia PA, 19104, USA

[1] Algebraists can replace the words "number" and "nonzero" by "element of a commutative ring" and "invertible", respectively. This generalization comes for free; we will not use anything specific to any kind of numbers in our proofs.

[2] Many authors define it to have entries $\frac{1}{x_i - y_j}$ instead of $\frac{1}{x_i + y_j}$. This boils down to replacing $y_1, y_2, \ldots, y_n$ by $-y_1, -y_2, \ldots, -y_n$.

found by Cauchy in 1841 [1] (see, e.g., [11, §1.3] or [2, Exercise 6.18 or Exercise 6.64] for modern proofs). Newer research focuses, e.g., on the LU decomposition [5], positivity properties [4], or generalizations [7]. See [6] for more on the history of the topic and for its connections to Lagrange interpolation (and for another proof of (1)). Applications range from the theoretical (an equivalent version [3, Lemma 5.15.3] of (1) is used in the classical representation theory of symmetric groups) to the practical (computing the inverse C^{-1} is a notoriously ill-conditioned problem that is used as a canary for numerical instability [13]).

2. The sum of the entries of the inverse

The following curious result appears to be known since at least the 1940s:

Theorem 2.1. Assume that the matrix C is invertible. Then, the sum of all entries of its inverse C^{-1} is $\sum\limits_{k=1}^{n} x_k + \sum\limits_{k=1}^{n} y_k$.

A natural, yet laborious approach to proving this theorem is to compute the entries of C^{-1} using (1), and then to add them up. The resulting sum can be seen (by a tricky induction) to simplify to $\sum\limits_{k=1}^{n} x_k + \sum\limits_{k=1}^{n} y_k$. Some details of this proof can be found in [8, §1.2.3, Exercise 44]. The proof given in [12, (13)] is simpler, avoiding the use of (1) but relying on Lagrange interpolation theory instead.

We propose a new proof of Theorem 2.1, which reflects the simplicity of the theorem. We let $A_{i,j}$ denote the (i,j)-th entry of any matrix A. The following simple lemma gets us half the way:

Lemma 2.2. Let A be an $n \times m$-matrix, and let B be an $m \times n$-matrix. Then,

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \left(x_i + y_j\right) A_{i,j} B_{j,i} = \sum_{i=1}^{n} x_i \left(AB\right)_{i,i} + \sum_{j=1}^{m} y_j \left(BA\right)_{j,j}.$$

Proof of Lemma 2.2. We have

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{m}(x_i+y_j)A_{i,j}B_{j,i} &= \sum_{i=1}^{n}\sum_{j=1}^{m}x_iA_{i,j}B_{j,i} + \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{m}}_{=\sum_{j=1}^{m}\sum_{i=1}^{n}} y_j \underbrace{A_{i,j}B_{j,i}}_{=B_{j,i}A_{i,j}} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m}x_iA_{i,j}B_{j,i} + \sum_{j=1}^{m}\sum_{i=1}^{n}y_jB_{j,i}A_{i,j} \\
&= \sum_{i=1}^{n}x_i \underbrace{\sum_{j=1}^{m}A_{i,j}B_{j,i}}_{\substack{=(AB)_{i,i} \\ \text{(by the definition of} \\ \text{the matrix product)}}} + \sum_{j=1}^{m}y_j \underbrace{\sum_{i=1}^{n}B_{j,i}A_{i,j}}_{\substack{=(BA)_{j,j} \\ \text{(by the definition of} \\ \text{the matrix product)}}} \\
&= \sum_{i=1}^{n}x_i(AB)_{i,i} + \sum_{j=1}^{m}y_j(BA)_{j,j}.
\end{aligned}$$

□

Proof of Theorem 2.1. Applying Lemma 2.2 to $m=n$, $A=C$ and $B=C^{-1}$, we obtain

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i+y_j)C_{i,j}\left(C^{-1}\right)_{j,i} &= \sum_{i=1}^{n}x_i \underbrace{\left(CC^{-1}\right)_{i,i}}_{\substack{=1 \\ \text{(since } CC^{-1} \text{ is the} \\ \text{identity matrix)}}} + \sum_{j=1}^{n}y_j \underbrace{\left(C^{-1}C\right)_{j,j}}_{\substack{=1 \\ \text{(since } C^{-1}C \text{ is the} \\ \text{identity matrix)}}} \\
&= \sum_{i=1}^{n}x_i + \sum_{j=1}^{n}y_j = \sum_{k=1}^{n}x_k + \sum_{k=1}^{n}y_k.
\end{aligned}$$

However, the factor $(x_i+y_j)C_{i,j}$ on the left hand side of this equality simplifies to 1 (since the definition of C yields $C_{i,j}=\dfrac{1}{x_i+y_j}$). Thus, the left hand side of this equality is $\sum_{i=1}^{n}\sum_{j=1}^{n}\underbrace{(x_i+y_j)C_{i,j}}_{=1}\left(C^{-1}\right)_{j,i} = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(C^{-1}\right)_{j,i}$, which is clearly the sum of all entries of C^{-1}. We have thus shown that the sum of all entries of C^{-1} is $\sum_{k=1}^{n}x_k + \sum_{k=1}^{n}y_k$. This proves Theorem 2.1. □

3. Variants

Theorem 2.1 was stated under the assumption that C be invertible. Using (1), it is easy to see that this assumption is equivalent to requiring that $x_1, x_2, \ldots, x_n$ be distinct and that $y_1, y_2, \ldots, y_n$ be distinct[3]. It is not hard to relieve Theorem 2.1 of this assumption: Just replace the inverse C^{-1} (which no longer exists) by the adjugate[4] $\operatorname{adj} C$ of the matrix C. The resulting theorem is as follows:

Theorem 3.1. The sum of all entries of the adjugate matrix $\operatorname{adj} C$ is $\left(\sum_{k=1}^{n} x_k + \sum_{k=1}^{n} y_k\right) \det C$.

Proof. Similar to our above proof of Theorem 2.1.[5] Use the classical result that $C \cdot \operatorname{adj} C = \operatorname{adj} C \cdot C = \det C \cdot I_n$ (where I_n denotes the $n \times n$ identity matrix). □

Theorem 3.1 can be transformed even further:

Theorem 3.2. Let D be the $(n+1) \times (n+1)$-matrix obtained from C by inserting a row full of 1's at the very bottom and a column full of 1's at the very right, and putting 0 in the bottom-right corner:

$$D = \begin{pmatrix} \frac{1}{x_1+y_1} & \frac{1}{x_1+y_2} & \cdots & \frac{1}{x_1+y_n} & 1 \\ \frac{1}{x_2+y_1} & \frac{1}{x_2+y_2} & \cdots & \frac{1}{x_2+y_n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{x_n+y_1} & \frac{1}{x_n+y_2} & \cdots & \frac{1}{x_n+y_n} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix}.$$

Then,

$$\det D = -\left(\sum_{k=1}^{n} x_k + \sum_{k=1}^{n} y_k\right) \cdot \det C.$$

Proof sketch. This follows from Theorem 3.1 using the following more general fact: If A is any $n \times n$-matrix, and if B is the $(n+1) \times (n+1)$-matrix obtained from A in the same way as D was obtained from C (that is, by inserting a row full of 1's at the very bottom and a column full of 1's at the very right, and putting 0 in the bottom-right corner), then

$$\det B = -s,$$

[3] Algebraists working over an arbitrary commutative ring should read "distinct" as "strongly distinct" (where two elements a, b of a ring are said to be *strongly distinct* if their difference $a - b$ is invertible).

[4] The *adjugate* $\operatorname{adj} A$ of an $n \times n$-matrix A is the $n \times n$-matrix whose (i, j)-th entry is $(-1)^{i+j} \det\left(A_{\sim j, \sim i}\right)$, where $A_{\sim j, \sim i}$ is the result of removing the j-th row and the i-th column from A. Older texts often refer to the adjugate as the "classical adjoint" (or just as the "adjoint", which however has another meaning as well).

[5] I wrote up this proof in much more detail in [2, solution to Exercise 6.69 **(a)**].

where s is the sum of all entries of $\operatorname{adj} A$. This fact, in turn, can be proved by Laplace expansion of $\det B$ along the last row (followed by expanding each cofactor along the last column). We refer to [2, solution to Exercise 6.69 **(c)**] for all details. $\square$

Theorem 3.2 appears in [9, Chapter XI, Exercise 43]; we know nothing more about its origins.

4. Two little exercises

For all its aid in our proof, it appears that Lemma 2.2 is a one-trick pony: We are unaware of any other interesting results whose proofs it simplifies. The sum of all entries of a matrix is not generally a particularly well-behaved quantity (unlike the sum of its **diagonal** entries, which is known as the trace and has many good properties). However, some experimentation has led us to a surprising (if not very deep) twin to Theorem 2.1.

We assume that $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ are **real** numbers (and that $n \geq 1$). Consider the $n \times n$-matrix

$$F := \left(\min\left\{x_i, y_j\right\}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n} = \begin{pmatrix} \min\left\{x_1, y_1\right\} & \min\left\{x_1, y_2\right\} & \cdots & \min\left\{x_1, y_n\right\} \\ \min\left\{x_2, y_1\right\} & \min\left\{x_2, y_2\right\} & \cdots & \min\left\{x_2, y_n\right\} \\ \vdots & \vdots & \ddots & \vdots \\ \min\left\{x_n, y_1\right\} & \min\left\{x_n, y_2\right\} & \cdots & \min\left\{x_n, y_n\right\} \end{pmatrix}.$$

Thus, F is obtained from C by replacing the "inverted sums" $\dfrac{1}{x_i + y_j}$ by the minima $\min\left\{x_i, y_j\right\}$ [6]. It would almost be too much to ask for F^{-1} to have properties comparable to those of C^{-1}. But in fact, it behaves even better:

Proposition 4.1. Assume that F is invertible. Then:

- **(a)** The sum of all entries of F^{-1} is $\dfrac{1}{\min\left\{x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n\right\}}$.
- **(b)** Assume that $x_1 \leq x_2 \leq \cdots \leq x_n$ and $y_1 \leq y_2 \leq \cdots \leq y_n$ and $x_1 \leq y_1$. Then, for each $j \in \{1, 2, \ldots, n\}$, the sum of all entries in the j-th column of F^{-1} is $\dfrac{1}{x_1}$ if $j = 1$, and is 0 if $j > 1$.

[6] This can be seen as an instance of tropicalization (see, e.g., [10]). More precisely, tropicalization (the sort that replaces $+$ and $\cdot$ by max and $+$) would replace $\dfrac{1}{x_i + y_j}$ by $-\max\left\{x_i, y_j\right\}$; but this turns into $\min\left\{x_i, y_j\right\}$ if we multiply all our numbers $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$ by -1.

The proof of this proposition is another neat exercise in working with inverse matrices – one we do not want to spoil for the reader. As with C, computing the determinant is not necessary. However, it is computable, and the result is another nice exercise:

Proposition 4.2. Assume that $x_1 \leq x_2 \leq \cdots \leq x_n$ and $y_1 \leq y_2 \leq \cdots \leq y_n$. For any $i, j \in \{1, 2, \ldots, n\}$, set $f_{i,j} := \min \{x_i, y_j\}$. Then,

$$\det F = f_{1,1} \cdot \prod_{k=2}^{n} \left(f_{k,k} - f_{k,k-1} - f_{k-1,k} + f_{k-1,k+1} \right). \tag{2}$$

Note that the product on the right hand side of (2) will often be 0 if the x_i and the y_j's are ordered in an "insufficiently balanced" way (e.g., if there are more than two y_j's between two consecutive x_i's). We leave it to the reader to establish more precise criteria for $\det F$ to be 0.

References

[1] Augustin-Louis Cauchy, *Memoire sur les fonctions alternées et sur les sommes alternées*, Exercices d'analyse et de phys. math., ii (1841). pp. 151–159. Reprinted in: Œuvres completes, 2e ser., tome xii.

[2] Darij Grinberg, *Notes on the combinatorial fundamentals of algebra*, 15 September 2022, arXiv:2008.09862v3.

[3] Pavel Etingof, Oleg Golberg, Sebastian Hensel, Tiankai Liu, Alex Schwendner, Dmitry Vaintrob, Elena Yudovina, *Introduction to Representation Theory*, with historical interludes by Slava Gerovitch, Student Mathematical Library **59**, AMS 2011.

[4] Miroslav Fiedler, *Notes on Hilbert and Cauchy matrices*, Linear Algebra and its Applications **432** (2010), pp. 351–356.

[5] Israel Gohberg, Israel Koltracht, *Triangular factors of Cauchy and Vandermonde matrices*, Integral Eq. Operator Theory **26** (1996), pp. 46–59.

[6] R. Gow, *Cauchy's matrix, the Vandermonde matrix and polynomial interpolation*, IMS Bulletin **28** (1992), pp. 45–52.

[7] Masao Ishikawa, Soichi Okada, Hiroyuki Tagawa, Jiang Zeng, *Generalizations of Cauchy's determinant and Schur's Pfaffian*, Advances in Applied Mathematics **36** (2006), pp. 251–287.

[8] Donald Ervin Knuth, *The Art of Computer Programming, volume 1: Fundamental Algorithms*, 3rd edition, Addison–Wesley 1997.
See `https://www-cs-faculty.stanford.edu/~knuth/taocp.html` for errata.

[9] Thomas Muir, *A Treatise on the Theory of Determinants*, revised and enlarged by William H. Metzler, Dover 1960.

[10] Masatoshi Noumi, Yasuhiko Yamada, *Tropical Robinson-Schensted-Knuth correspondence and birational Weyl group actions*, arXiv:math-ph/0203030v2.

[11] Viktor V. Prasolov, *Problems and Theorems in Linear Algebra*, Translations of Mathematical Monographs, vol. #134, AMS 1994.
See `https://staff.math.su.se/mleites/books/prasolov-1994-problems.pdf` for a preprint.
See also `https://drive.google.com/open?id=0B2UfTLwpN9okblBJbGxOZXc4Rm8` for a newer edition in Russian.

[12] Samuel Schechter, *On the Inversion of Certain Matrices*, Math. Comp. **13** (1959), pp. 73–77.

[13] John Todd, *Computational Problems Concerning the Hilbert Matrix*, Journal of Research of the National Bureau of Standards–B **65** (1961), no. 1, pp. 19–22.