The largest entry in the inverse of a Vandermonde matrix

The Largest Entry in the Inverse of a Vandermonde Matrix

Carlo Sanna*
Department of Mathematical Sciences
Politecnico di Torino
Corso Duca degli Abruzzi 24
10129 Torino
Italy
carlo.sanna.dev@gmail.com

Jeffrey Shallit and Shun Zhang
School of Computer Science
University of Waterloo
Waterloo, ON N2L 3G1
Canada
shallit@uwaterloo.ca
s385zhang@uwaterloo.ca

December 11, 2020

Abstract

We investigate the size of the largest entry (in absolute value) in the inverse of certain Vandermonde matrices. More precisely, for every real $b > 1$, let $M_b(n)$ be the maximum of the absolute values of the entries of the inverse of the $n \times n$ matrix $[b^{ij}]_{0 \le i,j < n}$. We prove that $\lim_{n \to +\infty} M_b(n)$ exists, and we provide some formulas for it.

1 Introduction

Let $\mathbf{a} = (a_0, a_1, \ldots, a_{n-1})$ be a list of n real numbers. The classical *Vandermonde matrix* $V(\mathbf{a})$ is defined as follows:

$$V(\mathbf{a}) := \begin{bmatrix} 1 & a_0 & a_0^2 & \cdots & a_0^{n-1} \\ 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_{n-1} & a_{n-1}^2 & \cdots & a_{n-1}^{n-1} \end{bmatrix}.$$

As is well-known, the Vandermonde matrix $V(\mathbf{a})$ is invertible if and only if the a_i are pairwise distinct [3]. Formulas for the inverse $V(\mathbf{a})$ (when it exists) have been known at least since 1958 [5].

*C. Sanna is a member GNSAGA of the INdAM and of CrypTO, the Group of Cryptography and Number Theory of Politecnico di Torino.

In what follows, n is a positive integer and $b > 1$ is a fixed real number. Let us define the entries $c_{i,j,n}$ by

$$[c_{i,j,n}]_{0\le i,j<n} = V(b^0, b^1, b^2, \dots, b^{n-1})^{-1}, \tag{1}$$

and let $M_b(n) = \max_{0\le i,j<n} |c_{i,j,n}|$, the maximum of the absolute values of the entries of $V(1, b, b^2, \dots, b^{n-1})^{-1}$. The size of the entries of inverses of Vandermonde matrices have been studied for a long time (e.g., [1]). Recently, in a paper by the first two authors and Daniel Kane [2], we needed to estimate $M_2(n)$, and we proved that $M_2(n) \le 34$. In fact, even more is true: the limit $\lim_{n\to\infty} M_2(n)$ exists and equals $3\prod_{i\ge 2}\left(1 + \frac{1}{2^i-1}\right) \doteq 5.19411992918\cdots$. In this paper, we generalize this result, replacing 2 with any real number greater than 1.

Our main results are as follows:

Theorem 1. *Let $b > 1$ and $n_0 = \lceil \log_b(1 + \frac{1}{b})\rceil$. Then $|c_{i,j,n}| \le |c_{n_0,n_0,n}|$ for $i, j \ge n_0$. Hence $M_b(n) \in \{|c_{i,j,n}| : 0 \le i, j \le n_0\}$.*

Theorem 2. *Let $b \ge \tau = (1 + \sqrt{5})/2$ and $n \ge 2$. Then $M_b(n) \in \{|c_{0,0,n}|, |c_{1,1,n}|\}$.*

Theorem 3. *For all real $b > 1$ the limit $\lim_{n\to\infty} M_b(n)$ exists.*

2 Preliminaries

For every real number x, and for all integers $0 \le i, j < n$, let us define the power sum

$$\sigma_{i,j,n}(x) := \sum_{\substack{0\le h_1<\cdots<h_i<n \\ h_1,\dots,h_i\ne j}} x^{h_1+\cdots+h_i}.$$

The following lemma will be useful in later arguments.

Lemma 4. *Let i, j, n be integers with $0 \le i < n$, $0 \le j < n - 1$, and let x be a positive real number.*

(a) If $x > 1$, then $\sigma_{i,j,n}(x) \ge \sigma_{i,j+1,n}(x)$.

(b) If $x < 1$, then $\sigma_{i,j,n}(x) \le \sigma_{i,j+1,n}(x)$.

Proof. We have

$$\sigma_{i,j+1,n}(x) - \sigma_{i,j,n}(x) = \sum_{(h_1,\dots,h_i)\in S_{i,j,n}} x^{h_1+\cdots+h_i} - \sum_{(h_1,\dots,h_i)\in T_{i,j,n}} x^{h_1+\cdots+h_i},$$

where

$$S_{i,j,n} := \{0 \le h_1 < \cdots < h_i < n : j \in \{h_1, \dots, h_i\}, j + 1 \notin \{h_1, \dots, h_i\}\}$$

and

$$T_{i,j,n} := \{0 \le h_1 < \cdots < h_i < n \,:\, j \notin \{h_1,\ldots,h_i\},\, j+1 \in \{h_1,\ldots,h_i\}\}.$$

Now there is a bijection $S_{i,j,n} \to T_{i,j,n}$ given by

$$(h_1,\ldots,h_i) \mapsto (h_1,\ldots,h_{i_0-1},h_{i_0}+1,h_{i_0+1},\ldots,h_i),$$

where i_0 is the unique integer such that $h_{i_0} = j$. Hence, it follows easily that $\sigma_{i,j,n}(x) \ge \sigma_{i,j+1,n}(x)$ for $x > 1$, and $\sigma_{i,j,n}(x) \le \sigma_{i,j+1,n}(x)$ for $x < 1$. □

Lemma 5. *Let i, j, n be integers with $0 \le i, j < n$, and let x be a nonzero real number. Then*

$$\frac{\sigma_{n-i-1,j,n}(x)}{x^{n(n-1)/2-j}} = \sigma_{i,j,n}(x^{-1}).$$

Proof. Note that the subsets $\mathcal{A}$ of $\{0,1,\ldots,n-1\} - \{j\}$ of cardinality $n-i-1$ are in one-to-one correspondence with the subsets $\mathcal{A}'$ of cardinality i. The correspondence is given by the complement $\mathcal{A} \mapsto \mathcal{A}' = \{0,1,\ldots,n-1\} - \{j\} - \mathcal{A}$. In particular, we have

$$\sum_{a\in\mathcal{A}} a = \sum_{k\in\{0,1,\ldots,n\}-\{j\}} k - \sum_{a\in\mathcal{A}'} a = \frac{n(n-1)}{2} - j - \sum_{a\in\mathcal{A}'} a.$$

As a consequence, we get that

$$\sigma_{n-i-1,j,n}(x) = \sum_{\mathcal{A}} x^{\sum_{a\in\mathcal{A}} a} = x^{n(n-1)/2-j} \sum_{\mathcal{A}'} x^{-\sum_{a\in\mathcal{A}'} a} = x^{n(n-1)/2-j}\sigma_{i,j,n}(x^{-1}),$$

as claimed. □

Recall the following formula for the entries of the inverse of a Vandermonde matrix (see, e.g., [4, §1.2.3, Exercise 40]).

Lemma 6. *Let $a_0,\ldots,a_{n-1}$ be pairwise distinct real numbers. If $V(a_0,a_1,\ldots,a_{n-1})^{-1} = [d_{i,j}]_{0\le i,j<n}$ then*

$$d_{n-1,j}X^{n-1} + d_{n-2,j}X^{n-2} + \cdots + d_{0,j}X^0 = \prod_{\substack{0\le i<n\\ i\ne j}} \frac{X-a_i}{a_j-a_i}.$$

For $0 \le i, j < n$ define

$$\pi_{j,n} := \prod_{\substack{0\le h<n\\ h\ne j}} |b^j - b^h|.$$

We now obtain a relationship between the entries of $V(b^0,b^1,\ldots,b^{n-1})^{-1}$ and $\sigma_{i,j,n}$ and $\pi_{j,n}$.

Lemma 7. *Let* $V(b^0, b^1, \ldots, b^{n-1})^{-1} = [c_{i,j,n}]_{0 \le i,j < n}$. *Then*

$$|c_{i,j,n}| = \frac{\sigma_{n-i-1,j,n}}{\pi_{j,n}} \tag{2}$$

for $0 \le i, j < n$.

Proof. By Lemma 6, we have

$$\prod_{\substack{0 \le h < n \\ h \ne j}} \frac{X - b^h}{b^j - b^h} = \sum_{0 \le i < n} c_{i,j,n} X^i.$$

which in turn, by Vieta's formulas, gives

$$c_{n-i-1,j,n} = (-1)^i \left(\prod_{\substack{0 \le h < n \\ h \ne j}} \frac{1}{b^j - b^h} \right) \sum_{\substack{0 \le h_1 < \cdots < h_i < n \\ h_1, \ldots, h_i \ne j}} b^{h_1 + \cdots + h_i} \tag{3}$$

for $0 \le i < n$. The result now follows by the definitions of σ and π. □

Next, we obtain some inequalities for π.

Lemma 8. *Define* $n_0 = \lceil \log_b(1 + \frac{1}{b}) \rceil$. *Then*

$$\pi_{j,n} \le \pi_{j+1,n} \quad \textit{for } n_0 \le j < n.$$

Proof. For $0 \le j < n - 1$, we have

$$\pi_{j+1,n} := \prod_{\substack{0 \le h < n \\ h \ne j+1}} |b^{j+1} - b^h| = b^{n-1} \prod_{\substack{0 \le h < n \\ h-1 \ne j}} |b^j - b^{h-1}| = \frac{b^{n+j-1} - b^{n-2}}{b^{n-1} - b^j} \pi_{j,n}.$$

A quick computation shows that the following inequalities are equivalent:

$$\frac{b^{n+j-1} - b^{n-2}}{b^{n-1} - b^j} \ge 1 \iff b^j \ge \frac{b^{n-1} + b^{n-2}}{b^{n-1} + 1}.$$

Let n_0 be the minimum positive integer such that $b^{n_0} \ge 1 + \frac{1}{b}$. Then $n_0 = \lceil \log_b(1 + \frac{1}{b}) \rceil$. Hence, for $n_0 \le j < n$, we have

$$b^j \ge 1 + \frac{1}{b} > \frac{b^{n-1} + b^{n-2}}{b^{n-1} + 1},$$

so that

$$\pi_{j,n} \le \pi_{j+1,n} \quad \text{for } n_0 \le j < n. \tag{4}$$

□

Finally, we have the easy

Lemma 9. *For* $0 \le i, j < n$ *we have* $c_{i,j,n} = c_{j,i,n}$.

Proof. $V(b^0, b^1, \ldots, b^{n-1})$ is a symmetric matrix, so its inverse is also. □

3 Proof of Theorem 1

Proof. Suppose $i, j \geq n_0$. Then

$$\begin{aligned} |c_{i,j,n}| &= \frac{\sigma_{n-i-1,j,n}}{\pi_{j,n}} \quad \text{(by (2))} \\ &\leq \frac{\sigma_{n-i-1,n_0,n}}{\pi_{j,n}} \quad \text{(by Lemma 4 (a))} \\ &\leq \frac{\sigma_{n-i-1,n_0,n}}{\pi_{n_0,n}} \quad \text{(by Lemma 8)} \\ &= |c_{i,n_0,n}| \quad \text{(by (2))}, \end{aligned}$$

and so we get

$$|c_{i,j,n}| \leq |c_{i,n_0,n}|. \tag{5}$$

But

$$c_{i,n_0,n} = c_{n_0,i,n} \tag{6}$$

by Lemma 9. Make the substitutions n_0 for i and i for j in (5) to get

$$|c_{n_0,i,n}| \leq |c_{n_0,n_0,n}|. \tag{7}$$

The result now follows by combining Eqs. (5), (6), and (7). □

4 Proof of Theorem 2

Proof. Since $b \geq \tau$, it follows that $b \geq 1 + 1/b$. Hence in Theorem 1 we can take $n_0 = 1$, and this gives $M_b(n) \in \{|c_{0,0,n}|, |c_{1,0,n}|, |c_{0,1,n}|, |c_{1,1,n}|\}$. However, by explicit calculation, we have

$$\begin{aligned} \sigma_{n-1,1,n} &= b^{n(n-1)/2-1} \\ \sigma_{n-2,1,n} &= b^{n(n-1)/2-1} + \sum_{(n-1)(n-2)/2-1 \leq i \leq n(n-1)/2-3} b^i, \end{aligned}$$

so that

$$\sigma_{n-1,1,n} \leq \sigma_{n-2,1,n}. \tag{8}$$

Hence

$$\begin{aligned} |c_{1,0,n}| &= |c_{0,1,n}| \quad \text{(by Lemma 9)} \\ &= \frac{\sigma_{n-1,1,n}}{\pi_{1,n}} \quad \text{(by (2))} \\ &\leq \frac{\sigma_{n-2,1,n}}{\pi_{1,n}} \quad \text{(by (8))} \\ &= |c_{1,1,n}| \quad \text{(by (2))}, \end{aligned}$$

and the result follows. □

5 Proof of Theorem 3

Proof. We have

$$\begin{aligned}
|c_{i,j,n}| &= \frac{\sigma_{n-i-1,j,n}}{\pi_{j,n}} \\
&= \frac{\sigma_{n-i-1,j,n}(b)}{\prod_{\substack{0\le h<n \\ h\neq j}} |b^j - b^h|} \\
&= \frac{\sigma_{n-i-1,j,n}(b)}{\prod_{\substack{0\le h<n \\ h\neq j}} (b^h \cdot |b^{j-h} - 1|)} \\
&= \frac{\sigma_{n-i-1,j,n}(b)}{b^{n(n-1)/2-j}} \cdot \frac{1}{\prod_{\substack{0\le h<n \\ h\neq j}} |b^{j-h} - 1|} \\
&= \sigma_{i,j,n}(b^{-1}) \frac{1}{\prod_{\substack{0\le h<n \\ h\neq j}} |b^{j-h} - 1|},
\end{aligned}$$

where we used Lemma 5.

For $x < 1$ define

$$\sigma_{i,j,\infty}(x) = \sum_{\substack{0\le h_1<\cdots<h_i<\infty \\ h_1,\ldots,h_i\neq j}} \frac{1}{x^{h_1+\cdots+h_i}},$$

with the convention $\sigma_{0,j,\infty}(x) := 1$.

Hence the limits

$$\begin{aligned}
\ell_{i,j} &:= \lim_{n\to+\infty} |c_{i,j,n}| \\
&= \lim_{n\to+\infty} \sigma_{i,j,n}(b^{-1}) \frac{1}{\prod_{\substack{0\le h<n \\ h\neq j}} |b^{j-h} - 1|} \\
&= \lim_{n\to+\infty} \sigma_{i,j,n}(b^{-1}) \prod_{0\le h<j} \frac{1}{b^{j-h} - 1} \prod_{j<h<n} \frac{1}{1 - b^{j-h}} \\
&= \lim_{n\to+\infty} \sigma_{i,j,n}(b^{-1}) \prod_{1\le s\le j} \frac{1}{b^s - 1} \prod_{1\le t<n-j} \frac{1}{1 - b^{-t}} \\
&= \sigma_{i,j,\infty}(b^{-1}) \left(\prod_{1\le s\le j} \frac{1}{b^s - 1} \right) \left(\prod_{t\ge 1} \frac{1}{1 - b^{-t}} \right) \qquad (9)
\end{aligned}$$

exist and are finite.

From Theorem 1 we see that

$$\lim_{n\to+\infty} M_b(n) = \max_{0\le i\le j<n_0} \lim_{n\to+\infty} |c_{i,j,n}| = \max_{0\le i\le j\le n_0} \ell_{i,j},$$

and the proof is complete. □

From this theorem we can explicitly compute $\lim_{n\to+\infty} M_b(n)$ for $b \geq \tau$.

Corollary 10. *Let* $\alpha \doteq 2.324717957$ *be the real zero of the polynomial* $X^3 - 3X^2 + 2X - 1$.

(a) If $b \geq \alpha$, *then* $\lim_{n\to\infty} M_b(n) = \prod_{t\geq 1}(1-b^{-t})^{-1}$.

(b) If $\tau \leq b \leq \alpha$, *then* $\lim_{n\to\infty} M_b(n) = \frac{b^2-b+1}{b(b-1)^2}\prod_{t\geq 1}(1-b^{-t})^{-1}$.

Proof. From Theorem 2 we know that for $b \geq \tau$ we have $\lim_{n\to\infty} M_b(n) \in \{\ell_{0,0}, \ell_{1,1}\}$. Now an easy calculation based on (9) shows that

$$\ell_{0,0} = \prod_{t\geq 1}(1-b^{-t})^{-1}$$

$$\ell_{1,1} = \frac{b^2-b+1}{b(b-1)^2} \prod_{t\geq 1}(1-b^{-t})^{-1}.$$

By solving the equation $\frac{b^2-b+1}{b(b-1)^2} = 1$, we see that for $b \geq \alpha$ we have $\ell_{0,0} \geq \ell_{1,1}$, while if $\tau \leq b \leq \alpha$ we have $\ell_{1,1} \geq \ell_{0,0}$. This proves both parts of the claim. □

Remark 11. The quantity $M_b(n)$ converges rather slowly to its limit when b is close to 1. The following table gives some numerical estimates for $M_b(n)$.

b	$\lim_{n\to\infty} M_b(n)$
3	1.785312341998534190367486
$\alpha \doteq 2.3247$	2.4862447382651613433
2	5.194119929182595417
$\tau \doteq 1.61803$	26.788216012030303413
1.5	67.3672156
1.4	282.398
1.3	3069.44
1.2	422349.8

6 Final remarks

We close with a conjecture we have been unable to prove.

Conjecture 12. Let $b > 1$ and $n_0 = \lceil \log_b(1 + \frac{1}{b}) \rceil$. Then, for all sufficiently large n, we have $M_b(n) = |c_{i,i,n}|$ for some i, $0 \leq i \leq n_0$.

Acknowledgment

We thank the referee for a careful reading of the paper and several helpful suggestions.

References

[1] W. Gautschi. On inverses of Vandermonde and confluent Vandermonde matrices. *Numer. Mathematik* **4** (1962), 117–123.

[2] D. M. Kane, C. Sanna, and J. Shallit. Waring's theorem for binary powers. *Combinatorica* **39** (2019), 1335–1350.

[3] A. Klinger. The Vandermonde matrix. *Amer. Math. Monthly* **74** (1967), 571–574.

[4] D. E. Knuth. *The Art of Computer Programming, Vol. 1, Fundamental Algorithms.* Addison-Wesley, third edition, 1997.

[5] N. Macon and A. Spitzbart. Inverses of Vandermonde matrices. *Amer. Math. Monthly* **65** (1958), 95–100.