

An upper bound on the number of monomials in determinants of sparse matrices with symbolic entries

Michael Kalkbrener
Departement Mathematik
ETH-Zentrum
CH-8092 Zürich, Switzerland

1991 Math. Subj. Class.: 13P99, 15A15, 68Q40

Keywords: symbolic computation, determinants, Sylvester resultants, Toeplitz matrices, Hankel matrices, Sperner theory

Abstract

The objective of this paper is to gain some insight into how well sparsity is preserved under determinant computations. For a square matrix A whose elements are indeterminates x_1, \dots, x_n and zeros, the determinant $\det(A)$ is a polynomial in x_1, \dots, x_n with integer coefficients. We derive an upper bound on the number of monomials in $\det(A)$ for a class of determinants which includes bigradients, Sylvester resultants and determinants of Toeplitz and Hankel matrices. Our approach is based on a result by Stanley in the theory of partially ordered sets.

1 Introduction

Solving systems of linear and non-linear equations is a fundamental problem in computer algebra. Since many systems arising in practice are sparse it is important to develop algorithms which take advantage of the sparseness of the input (see, for instance, [Wie86], [KS91], [Stu91], [CE93]). In this context it is interesting to know how well sparsity is preserved under certain basic operations. The objective of this paper is to study this problem for determinant computations.

Let A be a square matrix whose elements are indeterminates x_1, \dots, x_n and zeros. Then the determinant $\det(A)$ of A is a polynomial in x_1, \dots, x_n with integer coefficients. In this paper we derive an upper bound on the number of monomials in $\det(A)$ for a class of determinants which includes bigradients (see [Hou70] for a definition), Sylvester resultants and determinants of Toeplitz and Hankel matrices. The bound depends on the order of the matrix A and on n , the number of distinct indeterminates in A .

Our approach is based on a result in the theory of partially ordered sets. We show that there exists a relation between the number of monomials in $\det(A)$ and the size s of the largest antichain in a specific partially ordered set P . Since P has the Sperner property [Sta80] we easily obtain a recursive formula for s . Our bound is based on this formula.

For Sylvester resultants of binomials and determinants of three-diagonal Toeplitz matrices we derive explicit formulas in terms of monomials and show that for these classes the upper bound is attained. In order to assess the quality of the bound we have computed several examples and we have compared the bound with the actual numbers. These comparisons are presented in the last section.

2 The Sperner property in partially ordered sets

We first recall a few definitions from the theory of partially ordered sets (see, for instance, [And87] and [Sta86]). Let P be a finite partially ordered set with zero element. A subset C of P is called a chain of P if any two elements of C are comparable. An antichain of P is a subset A of P such that any two distinct elements of A are incomparable. The length of a chain C is defined to be $|C| - 1$. If every maximal chain of P has the same length k then we say that P is graded of rank k . In this case all maximal chains between the same endpoints have the same length. We define the rank of $x \in P$ to be the length of a maximal chain between 0 and x , denoted by $rank(x)$. Let N_i denote the number of elements of rank i . The numbers N_0, N_1, \dots, N_k are called the rank numbers of P . The rank numbers form a unimodal sequence if there exists a j such that

$$N_0 \leq N_1 \leq \dots \leq N_j, \quad N_j \geq \dots \geq N_k.$$

A graded poset P is said to have the Sperner property if the size of the largest antichain in P is the largest rank number. A map h from P to a partially ordered set P' is an isomorphism if h is a bijection and $x \preceq y$ if and only if $h(x) \preceq h(y)$ for every $x, y \in P$.

Let $k \in \mathbb{N}$, $l \in \mathbb{N}_0$. We now turn our attention to the poset $L(k, l)$ defined by

$$L(k, l) := \{(x_1, \dots, x_k) \mid x_i \text{ integers, } 0 \leq x_1 \leq x_2 \leq \dots \leq x_k \leq l\}$$

with order relation \preceq defined by

$$x = (x_1, \dots, x_k) \preceq y = (y_1, \dots, y_k) \quad \text{if } x_i \leq y_i \text{ for each } i.$$

Obviously, the poset $L(k, l)$ is graded of rank kl and $rank(x) = x_1 + \dots + x_k$. For $i \in \mathbb{Z}$ let $N(k, l, i)$ denote the number of elements of $L(k, l)$ of rank i . It is well-known (see [And87, p.47]) that

- (a) the rank numbers of $L(k, l)$ form a unimodal sequence and
- (b) $N(k, l, i) = N(k, l, kl - i)$ for each i .

Using deep algebraic methods Stanley showed that $L(k, l)$ has the Sperner property [Sta80]. Hence, together with (a) and (b), we obtain the following theorem.

Theorem 1 *The largest antichain in $L(k, l)$ has cardinality $N(k, l, \lfloor kl/2 \rfloor)$.*

It follows from the definition of $L(k, l)$ that the $N(k, l, i)$ satisfy the following recurrence:

$$\begin{aligned} N(1, l, i) &= 1 && \text{if } 0 \leq i \leq l, \\ N(1, l, i) &= 0 && \text{otherwise,} \end{aligned}$$

$$N(k, l, i) = \sum_{j=0}^l N(k-1, j, i-j) \quad \text{if } k > 1.$$

3 A bound on generalized bigradients

We use the result in the previous section to obtain a bound on the number of monomials in a specific type of determinant which we will call generalized bigradient.

Throughout this paper let $m \in \mathbb{N}_0$, $n \in \mathbb{N}$, $x_1, \dots, x_r, y_1, \dots, y_s$ distinct indeterminates and $A = (a_{i,j})$ a square matrix of order $m+n$ such that for every $i \in \{1, \dots, m\}$, $j \in \{m+1, \dots, m+n\}$ and $k \in \{1, \dots, m+n\}$

$$a_{i,k} \in \{x_1, \dots, x_r, 0\} \quad \text{and} \quad a_{j,k} \in \{y_1, \dots, y_s, 0\}.$$

Then its determinant can be written as a polynomial in the x 's and y 's with integer coefficients, i.e.

$$\det(A) = \sum c_{pq} x^p y^q, \tag{1}$$

where $p = (p_1, \dots, p_r) \in \mathbb{N}_0^r$, $q = (q_1, \dots, q_s) \in \mathbb{N}_0^s$, $x^p = \prod_i x_i^{p_i}$, $y^q = \prod_j y_j^{q_j}$. Define $S(A) := \{(p, q) \in \mathbb{N}_0^{r+s} \mid c_{pq} \neq 0\}$. Note that $S(A) \subseteq S_m^r \times S_n^s$, where

$$S_i^j := \{(c_1, \dots, c_j) \in \mathbb{N}_0^j \mid \sum_{k=1}^j c_k = i\}.$$

We call $\det(A)$ a generalized bigradient if it has the following property:

For every $p \in S_m^r$ and every $q = (q_1, \dots, q_s)$, $q' = (q'_1, \dots, q'_s) \in C_p$ with $q \neq q'$ there exists a $j \in \{1, \dots, s\}$ with

$$\sum_{i=1}^j q_i > \sum_{i=1}^j q'_i,$$

where $C_p := \{q \in S_n^s \mid (p, q) \in S(A)\}$.

We now assume that $\det(A)$ is a generalized bigradient and derive a bound on the number of monomials in $\det(A)$, i.e. a bound on the cardinality of $S(A)$. Obviously,

$$|S(A)| = \sum_{p \in S_m^r} |C_p|. \tag{2}$$

We define a partial ordering \preceq on S_n^s . Let $q = (q_1, \dots, q_s)$ and $q' = (q'_1, \dots, q'_s)$ be elements of S_n^s . Then

$$q \preceq q' \quad \text{if} \quad \sum_{i=1}^j q_i \leq \sum_{i=1}^j q'_i \quad \text{for} \quad j = 1, \dots, s.$$

By definition of generalized bigradients,

$$C_p \text{ is an antichain for every } p \in S_m^r. \tag{3}$$

We define a map h from S_n^s to $L(s-1, n)$ by

$$(q_1, \dots, q_s) \mapsto (q_1, q_1 + q_2, \dots, q_1 + q_2 + \dots + q_{s-1}).$$

Lemma 1 *The map h is an isomorphism between the posets S_n^s and $L(s-1, n)$.*

Proof: Obviously, h is well-defined and injective. For a given $(u_1, \dots, u_{s-1}) \in L(s-1, n)$

$$(u_1, u_2 - u_1, \dots, u_{s-1} - u_{s-2}, n - u_{s-1}) \in S_n^s$$

is mapped to (u_1, \dots, u_{s-1}) . Hence, h is surjective. Let $q = (q_1, \dots, q_s)$ and $q' = (q'_1, \dots, q'_s)$ be elements of S_n^s . It follows from

$$q \preceq q' \quad \text{iff} \quad \sum_{i=1}^j q_i \leq \sum_{i=1}^j q'_i \quad \text{for } j = 1, \dots, s \quad \text{iff} \quad h(q) \preceq h(q')$$

that h is an isomorphism. \square

Theorem 2 *If $\det(A)$ is a generalized bigradient the number of monomials in $\det(A)$ is at most*

$$N(s-1, n, \lfloor (s-1)n/2 \rfloor) \cdot \binom{m+r-1}{m}, \quad (4)$$

where for every $k \in \mathbb{N}$, $l \in \mathbb{N}_0$ and $i \in \mathbb{Z}$

$$\begin{aligned} N(1, l, i) &:= 1 && \text{if } 0 \leq i \leq l, \\ N(1, l, i) &:= 0 && \text{otherwise,} \end{aligned}$$

$$N(k, l, i) := \sum_{j=0}^l N(k-1, j, i-j) \quad \text{if } k > 1.$$

Proof: Let $p \in S_m^r$. By Lemma 1 and (3), the sets C_p and $h(C_p)$ have the same cardinality and $h(C_p)$ is an antichain in $L(s-1, n)$. Hence, by Theorem 1, the number of elements in C_p is at most $N(s-1, n, \lfloor (s-1)n/2 \rfloor)$. Together with (2) and

$$|S_m^r| = \binom{m+r-1}{m},$$

this proves the bound. \square

Using induction and the recursion formula for binomial numbers (see, for instance, [Aig75, p.160]) we obtain for every $k \in \mathbb{N}$, $l \in \mathbb{N}_0$ and $i \in \mathbb{Z}$

$$N(k, l, i) \leq \binom{l+k-1}{l}$$

and therefore the simpler but weaker bound

$$|S(A)| \leq \binom{m+r-1}{m} \cdot \binom{n+s-2}{n}. \quad (5)$$

Hence, for fixed r and s the number of monomials in $\det(A)$ is bounded by a polynomial function in m and n of total degree $r+s-3$.

4 Which determinants are generalized bigradients?

Let $t \in \mathbb{N}$ with $m \leq t$ and $n \leq t$, $x_1, \dots, x_{2t-1}, x_{2t}, y_1, \dots, y_{2t-1}, y_{2t}$ distinct indeterminates and $B = (b_{i,j})$ the $2t \times 2t$ -matrix

$$\begin{pmatrix} x_1 & x_2 & \dots & x_{2t} \\ 0 & x_1 & \dots & x_{2t-1} \\ \dots & \dots & \dots & \dots \\ y_1 & y_2 & \dots & y_{2t} \\ 0 & y_1 & \dots & y_{2t-1} \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

with t rows of x 's and t rows of y 's.

Theorem 3 *Assume that $A = (a_{i,j})$ is a submatrix of B obtained by deleting $t - m$ arbitrary x -rows, $t - n$ arbitrary y -rows and $2t - m - n$ arbitrary columns. Then the determinant of A is a generalized bigradient.*

Proof: Let $v_1, \dots, v_{m+n}, w_1, \dots, w_{m+n} \in \{1, \dots, 2t\}$ such that $v_1 < v_2 < \dots < v_{m+n}$, $w_1 < w_2 < \dots < w_{m+n}$ and $a_{i,j} = b_{v_i, w_j}$ for every $i, j \in \{1, \dots, m+n\}$. By definition,

$$\begin{aligned} a_{i,j} &= x_{w_j - v_i + 1} && \text{if } i \leq m, w_j - v_i \geq 0, \\ a_{i,j} &= y_{w_j - v_i + 1 + t} && \text{if } i > m, w_j - v_i \geq -t, \\ a_{i,j} &= 0 && \text{otherwise.} \end{aligned}$$

Let $(p, q) = (p_1, \dots, p_{2t}, q_1, \dots, q_{2t}) \in S(A)$ and σ a permutation of $\{1, \dots, m+n\}$ with

$$\prod_{i=1}^{m+n} a_{i, \sigma(i)} = x^p y^q.$$

Hence, for $k \in \{1, \dots, 2t\}$

$$\begin{aligned} p_k &= |\{i \in \{1, \dots, m\} \mid w_{\sigma(i)} - v_i + 1 = k\}|, \\ q_k &= |\{i \in \{m+1, \dots, m+n\} \mid w_{\sigma(i)} - v_i + 1 + t = k\}|. \end{aligned}$$

Therefore,

$$\begin{aligned} \sum_{k=1}^{2t} k(p_k + q_k) &= \left(\sum_{i=1}^m w_{\sigma(i)} - v_i + 1 \right) + \left(\sum_{i=m+1}^{m+n} w_{\sigma(i)} - v_i + 1 + t \right) \\ &= m + n + tn + \sum_{i=1}^{m+n} w_i - \sum_{i=1}^{m+n} v_i. \end{aligned}$$

Let $p \in S_m^{2t}$ and $q = (q_1, \dots, q_{2t}), q' = (q'_1, \dots, q'_{2t}) \in C_p$ with $q \neq q'$. Then

$$\sum_{k=1}^{2t} k(q_k - q'_k) = 0.$$

Thus,

$$\sum_{k=1}^j q_k > \sum_{k=1}^j q'_k$$

for some $j \in \{1, \dots, 2t\}$. \square

Assume that the determinant of the matrix A is a generalized bigradient and that the matrix A' has been obtained from A by substituting zeros for some indeterminates. Then it is obvious that $\det(A')$ is a generalized bigradient as well. Therefore, we obtain the following two corollaries.

Corollary 1 *Each of the following conditions implies that $\det(A)$ is a generalized bigradient:*

- (a) $\det(A)$ is a bigradient.
- (b) $\det(A)$ is a Sylvester resultant.
- (c) $\det(A)$ is a determinant of a Toeplitz matrix.
- (d) $\det(A)$ is a determinant of a Hankel matrix.

Corollary 2 *Let $x_1, \dots, x_r, y_1, \dots, y_s$ be distinct indeterminates, $f = x_1 z^{d_1} + \dots + x_r z^{d_r}$ and $g = y_1 z^{e_1} + \dots + y_s z^{e_s}$ univariate polynomials in z and $S(j, f, g) = c_j z^j + \dots + c_0$ the j -th subresultant of f and g , where $0 \leq j < \min(m, n)$, $n := \deg(f)$ and $m := \deg(g)$. Then c_i is a generalized bigradient for every $i \in \{0, \dots, j\}$ and the number of monomials in $S(j, f, g)$ is at most*

$$(j+1) \cdot N(s-1, n-j, \lfloor (s-1)(n-j)/2 \rfloor) \cdot \binom{m+r-j-1}{m-j}.$$

5 Formulas for restricted classes of determinants

In this section we derive explicit formulas for two restricted classes of determinants and show that for both classes bound (4) is actually attained.

Theorem 4 *Let x_1, x_2, y_1, y_2 be distinct indeterminates and $f = x_1 z^n + x_2$ and $g = y_1 z^m + y_2$ univariate polynomials in z . Then the Sylvester resultant of f and g is of the form*

$$\text{res}(f, g) = (x_1^{m'} y_2^{n'} + (-1)^{m'n'} x_2^{m'} y_1^{n'})^d,$$

where $d := \gcd(m, n)$, $m' := m/d$ and $n' := n/d$.

Proof: Define $f' := x_1 z^{n'} + x_2$ and $g' := y_1 z^{m'} + y_2$ and let $B = (b_{i,j})$ be the Sylvester matrix of f' and g' and σ a permutation of $\{1, \dots, m' + n'\}$ such that $b_{i, \sigma(i)} \neq 0$ for every $i \in \{1, \dots, m' + n'\}$. Then

$$\begin{aligned} \sigma(i) &= i \text{ or } \sigma(i) = i + n' && \text{for } i = 1, \dots, m' \text{ and} \\ \sigma(j) &= j \text{ or } \sigma(j) = j - m' && \text{for } j = m' + 1, \dots, m' + n'. \end{aligned}$$

Assume that σ has a k -cycle (l_1, l_2, \dots, l_k) with $1 < k < m' + n'$. Since $l_{i+1} = l_i + n'$ or $l_{i+1} = l_i - m'$ it follows that there exist $0 < i < n'$ and $0 < j < m'$ with $im' = jn'$. This is a contradiction because m' and n' are relatively prime. Therefore,

$$\begin{aligned} \sigma(i) &= i && \text{for } i = 1, \dots, m' + n' && \text{or} \\ \sigma(i) &= i + n', \sigma(j) = j - m' && \text{for } i = 1, \dots, m', j = m' + 1, \dots, m' + n' \end{aligned}$$

and

$$\text{res}(f', g') = x_1^{m'} y_2^{n'} + (-1)^{m'n'} x_2^{m'} y_1^{n'}. \quad (6)$$

Since $\text{res}(f', g')$ can be represented as symmetric function of the roots of f' and g' (see [vdW71, p.107]) we obtain

$$\text{res}(f, g) = \text{res}(f'(z^d), g'(z^d)) = \text{res}(f', g')^d$$

and together with (6) the desired formula. \square

Hence, the number of monomials in $\text{res}(f, g)$ is $\gcd(m, n) + 1$. If $m = n$ then this number is $m + 1$ and bound (4) as well as bound (5) is attained. Note that

$$\text{res}(z^k f, g) = y_2^k \text{res}(f, g).$$

Therefore, Theorem 4 can be used for computing Sylvester resultants of arbitrary binomials.

Let $l, n \in \mathbb{N}$ with $1 \leq l \leq n - 1$ and let y_1, \dots, y_{2l+1} be distinct indeterminates. We now turn to $2l + 1$ -diagonal Toeplitz matrices $T_{l,n} = (t_{i,j})$ of order n defined by

$$t_{i,j} := y_{j-i+l+1} \text{ if } -l \leq j - i \leq l \quad \text{and} \quad t_{i,j} := 0 \text{ otherwise.}$$

Marr and Vineyard [MV88] give a closed expression for determinants of five-diagonal Toeplitz matrices $T_{2,n}$ which involves Chebyshev polynomials of the second kind of order $n + 1$. In the following theorem we present a decomposition of determinants of three-diagonal Toeplitz matrices into a sum of monomials.

Theorem 5 *The determinant of $T_{1,n}$ has the form*

$$\det(T_{1,n}) = \sum_{i=0}^{\lfloor n/2 \rfloor} \binom{n-i}{i} (-1)^i y_1^i y_2^{n-2i} y_3^i.$$

Proof: The proof of this formula easily follows from the recurrence relation

$$\det(T_{1,n}) = y_2 \cdot \det(T_{1,n-1}) - y_1 y_3 \cdot \det(T_{1,n-2})$$

and the recursion formula for binomial numbers. \square

For $n \times n$ Toeplitz matrices with 3 indeterminates bound (4) becomes

$$N(2, n, n) = \sum_{j=0}^n N(1, j, n-j) = \sum_{j=\lfloor n/2 \rfloor}^n N(1, j, n-j) = \lfloor n/2 \rfloor + 1.$$

By Theorem 5, this upper bound is exact for three-diagonal Toeplitz matrices.

6 Computational experiments and open problems

Motivated by Theorem 4 we consider Sylvester resultants of the following form: for natural numbers i, j let $R_{i,j}$ be the Sylvester resultant of

$$x_i z^{ij} + x_{i-1} z^{(i-1)j} + \dots + x_0 \quad \text{and} \quad y_i z^{ij} + y_{i-1} z^{(i-1)j} + \dots + y_0,$$

where $x_i, \dots, x_0, y_i, \dots, y_0$ are distinct indeterminates. We know from Theorem 4 that bound (4) gives the exact number for $i = 1$. For small $i > 1$ we have increased j as much as possible. For those $R_{i,j}$ which we could compute bound (4) is approximately twice the exact number of monomials in $R_{i,j}$. The results are listed below. The first row gives the exact number of monomials in the resultant, the second row gives the bound.

$R_{2,1}$	$R_{2,5}$	$R_{2,10}$	$R_{2,15}$	$R_{2,20}$	$R_{2,25}$	$R_{3,1}$	$R_{3,3}$	$R_{3,5}$	$R_{4,1}$	$R_{4,2}$
7	201	1252	3976	9051	17207	34	1530	13382	219	6054
12	396	2541	7936	18081	34476	60	3300	29376	560	16335

We know from Theorem 5 that bound (4) gives the exact number of monomials in the determinant $\det(T_{1,n})$ of a three-diagonal Toeplitz matrix for every order $n \in \mathbb{N}$. The results below show that also for small $i > 1$ bound (4) is close to the exact number of monomials in $\det(T_{i,n})$. Again the first row gives the exact number, the second row gives the bound.

$T_{2,3}$	$T_{2,6}$	$T_{2,9}$	$T_{2,12}$	$T_{2,15}$	$T_{2,18}$	$T_{2,21}$	$T_{2,24}$	$T_{2,27}$	$T_{2,30}$
5	14	43	85	148	225	360	517	713	921
5	18	43	86	150	241	362	519	715	956
$T_{3,4}$	$T_{3,5}$	$T_{3,7}$	$T_{3,8}$	$T_{3,10}$	$T_{3,11}$	$T_{3,13}$	$T_{3,14}$	$T_{3,16}$	$T_{3,17}$
16	31	86	141	314	462	901	1202	2123	2717
18	32	94	151	338	480	920	1242	2137	2739
$T_{4,5}$	$T_{4,6}$	$T_{4,7}$	$T_{4,8}$	$T_{4,9}$	$T_{4,10}$	$T_{4,11}$			
58	131	270	478	830	1339	2270			
73	151	289	526	910	1514	2430			

In this paper we have presented an upper bound on the number of monomials in $\det(A)$. We also consider it interesting to derive lower bounds on the number of monomials and formulas respectively bounds for the coefficients c_{pq} in the decomposition (1) of $\det(A)$ (see [GKZ90]).

The convex hull of $S(A)$ in \mathbb{R}^{r+s} is called the Newton polytope of $\det(A)$. An interesting problem is the characterization of the Newton polytopes of specific classes of determinants. This has been done for Sylvester resultants and multivariate resultants in [GKZ90] respectively [Stu94].

Acknowledgement: I am grateful to Rafael Sendra and Bernd Sturmfels for many interesting discussions.

References

- [Aig75] M. Aigner. *Kombinatorik I (in German)*. Springer, Berlin Heidelberg New York, 1975.
- [And87] I. Anderson. *Combinatorics of Finite Sets*. Clarendon Press, Oxford, 1987.
- [CE93] J. Canny and I. Emiris. An efficient algorithm for the sparse mixed resultant. In *Proc. AAECC-10*, pages 89–104, 1993.
- [GKZ90] I.M. Gelfand, M.M. Kapranov, and A.V. Zelevinsky. Newton polytopes of the classical resultant and discriminant. *Adv. in Math.*, 84(2):237–254, 1990.
- [Hou70] A.S. Householder. *The Numerical Treatment of a Single Nonlinear Equation*. McGraw Hill, 1970.
- [KS91] E. Kaltofen and B.D. Saunders. On Wiedemann’s method of solving sparse linear systems. In *Proc. AAECC-9*, pages 29–38, New Orleans, USA, 1991.
- [MV88] R.B. Marr and G.H. Vineyard. Five-diagonal Toeplitz determinants and their relation to Chebyshev polynomials. *SIAM J. Matrix Anal. Appl.*, 9:579–586, 1988.
- [Sta80] R.P. Stanley. Weyl groups, the hard Lefschetz theorem, and the Sperner property. *SIAM J. Alg. Disc. Meth.*, 1(2):168–184, 1980.
- [Sta86] R.P. Stanley. *Enumerative Combinatorics I*. Wadsworth & Brooks/Cole, Monterey, California, 1986.
- [Stu91] B. Sturmfels. Sparse elimination theory. In *Proc. Comput. Algebraic Geom. and Commut. Algebra*, pages 264–298, Cortona, Italy, 1991.
- [Stu94] B. Sturmfels. On the Newton polytope of the resultant. *Journal of Algebraic Combinatorics*, 3:207–236, 1994.
- [vdW71] B.L. van der Waerden. *Algebra I (in German)*. Springer, Berlin Heidelberg New York, 8. edition, 1971.
- [Wie86] D.H. Wiedemann. Solving sparse linear equations over finite fields. *IEEE Trans. Inf. Theory*, 32:54–62, 1986.