

Determinantal complexity

Markus Bläser, Saarland University

Draft

Chapter 1

Determinantal complexity

The question whether $\text{VP}_{ws} = \text{VNP}$ can be rephrased as the question whether \det is p -projection of per . Related questions have been studied. One of them is the so-called determinantal complexity.

1.0.1 Definition. *The determinantal complexity $\text{dc}(f)$ of a polynomial $f \in \mathbb{F}[X_1, \dots, X_n]$ is the smallest s such that there are affine linear forms $\alpha_{i,j} \in \mathbb{F}[X_1, \dots, X_n]$, $1 \leq i, j \leq s$, such that we can write $f = \det_s(\alpha_{i,j})$.*

1.0.2 Lemma. *$(f_n) \in \text{VP}_{ws}$ iff $\text{dc}(f_n)$ is p -bounded.*

Proof. If $(f_n) \in \text{VP}_{ws}$, then it is a p -projection of \det . Therefore, its determinantal complexity is p -bounded. For the other direction, note that the determinant has weakly skew circuits of polynomial size. We can compute the affine linear forms by weakly skew circuits of polynomial size. Therefore, (f_n) has weakly polynomial circuits of polynomial size. \square

1.1 Mignon-Ressayre bound

In this section, we prove the best lower bound for the determinantal complexity, due to Mignon and Ressayre [MR04].

1.1.1 Observation. *$\frac{\partial}{\partial X_{i,j}} \text{per}_n(X)$ is the permanent of the $(n-1) \times (n-1)$ matrix obtained from X by deleting the i th row and the j th column. The same is true for the determinant.*

For a matrix $A \in \mathbb{F}^{n \times n}$, let $H_{\text{per}}(A)$ denote the $n^2 \times n^2$ -matrix with the entry in row (i, j) and column (k, ℓ) being equal to

$$\frac{\partial^2}{\partial X_{i,j} \partial X_{k,\ell}} \text{per}_n(A).$$

That is, we take the permanent polynomial, differentiate it twice and then we plug in the values from the matrix C . In the same way, we define H_f and $H_f(A)$ for any polynomial f .

For the proof, we need to construct a matrix A such that $\text{per}(A) = 0$ but $H_{\text{per}}(A)$ has full rank n^2 . Let A be the matrix

$$A = \begin{pmatrix} 1-n & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix},$$

that is, $a_{1,1} = 1 - n$ and all other entries are 1.

1.1.2 Lemma. $\text{per}(A) = 0$.

Proof. Like for the determinant, we can do a Laplace expansion of the permanent. It is even easier, since there are no signs to keep track off. (If you prefer to think in terms of cycle covers, a Laplace expansion along the i th row just groups the cycle covers depending on which node is visited right after node i .)

So we do a Laplace expansion along the first row. Since all other rows are filled only with 1's, all the submatrices that we get are the same. We get it once multiplied by $1 - n$ and $n - 1$ times multiplied by 1. So the sum is 0. \square

1.1.3 Lemma. $H_{\text{per}}(A)$ has rank n^2 .

Proof. When $i = k$ or $j = \ell$, then

$$\frac{\partial^2}{\partial X_{i,j} \partial X_{k,\ell}} \text{per}_n(X) = 0.$$

This is due to the fact that every monomial contains only one variable from each row or column. (This property is called set-multilinear).

If $i \neq k$ and $j \neq \ell$, then

$$\frac{\partial^2}{\partial X_{i,j} \partial X_{k,\ell}} \text{per}_n(X)$$

is the permanent of the submatrix obtained by deleting rows i and k and columns j and ℓ from X . If $1 \in \{i, j, k, \ell\}$, then

$$\frac{\partial^2}{\partial X_{i,j} \partial X_{k,\ell}} \text{per}_n(A) = (n - 2)!,$$

since the matrix that we obtain from A after deleting the rows and columns is the all-ones-matrix of size $(n - 2) \times (n - 2)$, the permanent of which is $(n - 2)!$. If $1 \notin \{i, j, k, \ell\}$, then

$$\frac{\partial^2}{\partial X_{i,j} \partial X_{k,\ell}} \text{per}_n(A) = -2(n - 3)!,$$

since the matrix that we obtain from A in this case has $n - 1$ in position $(1, 1)$ and 1's elsewhere. Using Laplace expansion, one can easily see that the permanent of this matrix is $-2(n - 3)!$. Therefore, we have that

$$H_{\text{per}}(A) = (n - 3)! \begin{pmatrix} 0 & B & B & \dots & B \\ B & 0 & C & \dots & C \\ B & C & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & C \\ B & C & \dots & C & 0 \end{pmatrix}$$

where

$$B = \begin{pmatrix} 0 & n - 2 & \dots & n - 2 \\ n - 2 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ n - 2 & \dots & n - 2 & 0 \end{pmatrix}$$

and

$$C = \begin{pmatrix} 0 & n - 2 & n - 2 & \dots & n - 2 \\ n - 2 & 0 & -2 & \dots & -2 \\ n - 2 & -2 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -2 \\ n - 2 & -2 & \dots & -2 & 0 \end{pmatrix}.$$

The matrix B has full rank: If we subtract the $(n-1)$ th row from the n th, then the $(n-2)$ th row from the $(n-1)$ th and so on until we subtract the first row from the second, we get the matrix

$$(n-2) \begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & -1 \end{pmatrix}.$$

From the structure of the rows 2 to n it follows that every nontrivial vector in the kernel of the matrix has to be a nonzero multiple of the all-ones-vector. The scalar product of this vector with the first row is however nonzero. Therefore, B has full rank. Doing the same transformation with C , but stopping one row earlier, we get the matrix

$$\begin{pmatrix} 0 & n-2 & n-2 & \dots & n-2 \\ n-2 & 0 & -2 & \dots & -2 \\ 0 & -2 & 2 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & -2 & 2 \end{pmatrix}.$$

From the structure of the third to n th row, we get that the entries at positions 2 to n of every nontrivial vector in the kernel have to be the same. From the first row, it follows that these entries have to be 0. And finally, the second row tells us that also the first entry has to be zero then. Therefore, C is also invertible.

We have

$$\begin{aligned} \begin{pmatrix} CB^{-1} & 0 & \dots & 0 \\ 0 & I & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I \end{pmatrix} H_{\text{per}}(A) \begin{pmatrix} B^{-1}C & 0 & \dots & 0 \\ 0 & I & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I \end{pmatrix} &= (n-3)! \begin{pmatrix} 0 & C & C & C \dots & C \\ C & 0 & C & \dots & C \\ C & C & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & C \\ C & C & \dots & C & 0 \end{pmatrix} \\ &= (n-3)! \begin{pmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix} \otimes C. \end{aligned}$$

Since the Kronecker product of two full rank matrices has itself full rank, $H_{\text{per}}(A)$ has full rank. \square

1.1.4 Lemma. *Let $Y \in \mathbb{F}^{s \times s}$. If $\det_s(Y) = 0$, then $\text{rk } H_{\det_s}(Y) \leq 2s$.*

Proof. Let S and T be invertible matrices such that

$$SYT = \begin{pmatrix} 0 & 0 \\ 0 & I_t \end{pmatrix}$$

for some $t < s$. Multiplying Y by

$$\begin{pmatrix} 1/\det S & 0 \\ 0 & I_{s-1} \end{pmatrix} S \quad \text{and} \quad T \begin{pmatrix} 1/\det T & 0 \\ 0 & I_{s-1} \end{pmatrix}$$

from the left and the right, respectively, does not change its determinant. As

$$\begin{pmatrix} 1/\det S & 0 \\ 0 & I_{s-1} \end{pmatrix} SYT \begin{pmatrix} 1/\det T & 0 \\ 0 & I_{s-1} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & I_t \end{pmatrix}$$

(recall that $t < s$), we can assume w.l.o.g. that $Y = (y_{i,j})$ is of this form.

Let us first consider the case when $t = s - 1$. An entry of $H_{\det}(y_{i,j})$ is of the form

$$\frac{\partial^2}{\partial X_{e,f} \partial X_{k,\ell}} \det_s(y_{i,j}).$$

This entry can only be nonzero, if differentiating removes the first row and the first column and the h th row and h th column for any other h . This means that

1. $(e, f) = (1, 1)$ and $(k, \ell) = (h, h)$,
2. $(e, f) = (1, h)$ and $(k, \ell) = (h, 1)$,
3. $(e, f) = (h, 1)$ and $(k, \ell) = (1, h)$,
4. $(e, f) = (h, h)$ and $(k, \ell) = (1, 1)$.

In the first case, we get one row with $s - 1$ ones in it. In the fourth case, we get one column with $s - 1$ ones in it. In the second case, we get $s - 1$ different rows, each having a single one and zeros elsewhere. The third case is similar. Altogether, we get that the rank is at most $2 + 2(s - 1) = 2s$. (In fact, equality holds.) When $t = s - 2$, then we have to delete the first and second row and column, respectively, to get a nonzero entry. Therefore, the rank of the matrix can be at most four, which is less than $2s$.

When $t < s - 2$, then every entry of $H_{\det}(Y)$ will be zero. \square

1.1.5 Theorem. $\text{dc}(\text{per}_n) \geq n^2/2$

Proof. Let s be the determinantal complexity of per_n . We know that there are affine linear forms $\alpha_{i,j}(X)$, $1 \leq i, j \leq s$ such that

$$\text{per}_n(X) = \det_s(\alpha_{i,j}(X)).$$

Let $A = (a_{i,j})$ be the matrix from Lemma 1.1.3. Write $\alpha_{i,j}(X) = \lambda_{i,j}(X - A) + y_{i,j}$, where $\lambda_{i,j}$ is a homogeneous linear form and $y_{i,j}$ is a constant, i.e., we perform a translation on the coordinates. Thus,

$$\text{per}_n(X) = \det_s(\lambda_{i,j}(X - A) + y_{i,j}). \quad (1.1.6)$$

Since $\text{per}_n(A) = 0$, we have $\det_s(y_{i,j}) = 0$, so $Y := (y_{i,j})$ is not of full rank. Define H_{\det} in the same way as H_{per} . Now we differentiate both sides of (1.1.6). We get

$$H_{\text{per}}(X) = LH_{\det}(\alpha_{i,j}(X - A) + y_{i,j})L^T$$

for some matrix L with entries from \mathbb{F} by the chain rule (see below). Therefore,

$$H_{\text{per}}(A) = LH_{\det}(Y)L^T$$

and

$$\text{rk } H_{\text{per}}(A) \leq \text{rk } H_{\det}(Y).$$

By Lemma 1.1.3, we know that $\text{rank } H_{\text{per}}(A) = n^2$. Therefore, we are done when we show that $\text{rk}(H_{\det}(Y)) \leq 2s$. But this follows from Lemma 1.1.4, since $\det(Y) = 0$. \square

1.1.7 Observation. Let f be a polynomial in Y_1, \dots, Y_m variables and ℓ_1, \dots, ℓ_m be affine linear forms in X_1, \dots, X_n . Then by the chain rule

$$\frac{\partial^2}{\partial X_i \partial X_j} f(\ell_1, \dots, \ell_n) = \sum_{s=1}^m \sum_{t=1}^m \frac{\partial^2}{\partial Y_s \partial Y_t} f(\ell_1, \dots, \ell_n) \frac{\partial}{\partial X_i} \ell_s \frac{\partial}{\partial X_j} \ell_t$$

Note that $\frac{\partial}{\partial X_i} \ell_s$ and $\frac{\partial}{\partial X_j} \ell_t$ are just constants. Let $L = (\frac{\partial}{\partial X_i} \ell_s)_{1 \leq s \leq m, 1 \leq i \leq n}$. Then

$$\left(\frac{\partial^2}{\partial X_i \partial X_j} f(\ell_1, \dots, \ell_n) \right) = L \left(\frac{\partial^2}{\partial Y_s \partial Y_t} f(\ell_1, \dots, \ell_n) \right) L^T$$

1.2 Grenet's construction

The best upper bound of the determinantal complexity is due to Grenet [Gre12]. It is (of course) exponential. Grenet's construction even writes the permanent as a projection of the determinant. It can be easily described in combinatorial terms (see [BES] for an alternative explanation). We construct a digraph G as follows: The nodes are all subsets of $\{1, \dots, n\}$. We identify \emptyset and $\{1, \dots, n\}$ with each other, so there are $2^n - 1$ nodes in total. Let S and T be two nodes of G . There will be an edge from S to T with weight $X_{i,j}$ if $|S| = i - 1$, $j \notin S$, and $T = S \cup \{j\}$. The node \emptyset will have outgoing edges with weight $X_{1,j}$ to the node $\{j\}$, $1 \leq j \leq n$ and incoming edges with weights $X_{n,j}$ from the node $\{1, \dots, n\} \setminus \{j\}$. Furthermore, every node except \emptyset gets a self loop of weight 1.

How does a cycle cover of G look like? Edges go only from nodes S to nodes T of larger cardinality. Therefore, the graph is "almost" acyclic, we only get cycles since we identified \emptyset with $\{1, \dots, n\}$. Every cycle has to go through \emptyset . So there can be only one cycle which is not a self loop and since \emptyset has no self loop, we have to use one such cycle and cover all other nodes with self loops. Therefore every cycle cover of G has the same number of cycles and therefore the same sign.

1.2.1 Observation. *Cycle covers of G stand in one-to-one correspondance with permutations in S_n .*

This is due to the fact that every cycle simulates the process of adding the numbers $1, \dots, n$ in some particular order to the empty set until we get $\{1, \dots, n\}$. Let π be this order. Then the weight of this cycle is $X_{1,\pi(1)} \cdots X_{n,\pi(n)}$. It follows that

$$\text{per}(G) = \text{per}(X)$$

Since all cycle covers of G have the same sign, $\text{per}(G) = \pm \det(G)$. Thus we haven proven the following theorem.

1.2.2 Theorem. $D(\text{per}_n) \leq 2^n - 1$.

Chapter 2

Extension to border complexity

The results of this chapter were first proven by [LMR10]. We follow the proof by Grochow [Gro15].

2.0.1 Lemma. *There is a set of polynomial equations $S_{e,d}$ in the coefficients of two polynomials $f \in \text{Sym}^e V^*$ and $g \in \text{Sym}^d V^*$ such that*

$$[\forall P \in S_{e,d} : P(c_f, c_g) = 0] \iff f|g,$$

where c_f and c_g are the coefficient vectors of f and g .

Proof. Consider the map $\text{Mult}_f : \text{Sym}^{d-e} V^* \rightarrow \text{Sym}^d V^*$ given by $h \mapsto f \cdot h$. Let M_f be the matrix of Mult_f , say in the monomial basis. The entries of M_f are either 0s or coefficients of f . M_f has dimensions $\binom{n+d}{n} \times \binom{n+d-e}{n}$. We have $g \in \text{im Mult}_f$ iff $c_g \in \text{col-span } M_f$ iff $\text{rk } M_f = \text{rk}(M_f|c_g)$. Here $(M_f|c_g)$ denotes the matrix M_f extended by the coefficient vector c_g of g .

$(M_f|c_g)$ has full rank $R := \binom{n+d-e}{n}$. We have $\text{rk}(M_f|c_g) = R$ iff all $(R+1) \times (R+1)$ -minors of $(M_f|c_g)$ vanish. These minors are polynomials in the coefficients of f and g and form the set $S_{e,d}$. \square

Recall that we define $\text{dc}(f)$ as the minimum s such that there are affine linear forms $a_{i,j}$ in the variables of f , $1 \leq i, j \leq s$ such that $f = \det(a_{i,j})$. Instead of looking at affine forms, we will now only consider homogeneous linear forms and padded polynomials. Let T be some variable that does not appear in f . It is easy to see that $\text{dc}(f) = s$ iff there are homogeneous linear forms $(\ell_{i,j})$, $1 \leq i, j \leq s$, such that $T^{s-\text{deg}(f)} f = \det(\ell_{i,j})$, we simply need to replace the constant term $\alpha_{i,j}$ of $a_{i,j}$ by $\alpha_{i,j}T$. Therefore, we will now look at $\text{per}_{m,n} = X_{m,m}^{m-n} \text{per}_n(X|_n)$ and compare it to $\det_m(X)$. Here X is an $m \times m$ -matrix with indeterminate entries and $X|_n$ denotes the upper left $n \times n$ -submatrix. We say that $\underline{\text{dc}}(f) \leq s$ if f is contained in the closure of all h with $\text{dc}(h) \leq s$.

Recall that for a polynomial $f(X)$, $H_f(X) = (\frac{\partial^2}{\partial X_i \partial X_j} f(X))$ denotes the Hessian matrix of f .

2.0.2 Lemma. $\det_s X$ divides all $(2s+1) \times (2s+1)$ -minors of $H_{\det_s}(X)$.

Proof. Since \det_s is irreducible, it is sufficient to show that for all $Y \in \mathbb{F}^{s \times s}$, if $\det_s Y = 0$, then also all $(2s+1) \times (2s+1)$ -minors of $H_{\det_s}(Y)$ vanish. This is the same as showing that for all Y , $\text{rk}(Y) \leq s-1$ implies $\text{rk}(H_{\det_s}(Y)) \leq 2s$. However, this has been shown in Lemma 1.1.4. \square

Let $f(X)$ be a polynomial in variables $X_{1,1}, \dots, X_{s,s}$ with coefficient vector c_f . E_s denotes the set of all equations in the entries of c_f that express the fact that f divides all $(2s+1) \times (2s+1)$ -minors of H_f . (The set E_s is the union over all minors of the equations of Lemma 2.0.1.)

2.0.3 Lemma. *Let $f(X) = \det_s L(X)$, where $L(X) = (\ell_{i,j}(X))$ is a matrix with homogeneous linear forms in X as entries. Then c_f fulfills E_s .*

Proof. By chain rule (see the previous chapter), we have

$$H_f(X) = BH_{\det_s}(L(X))B^T$$

where $B = (\frac{\partial}{\partial X_{i,j}} \ell_{s,t})$ is the Jacobian matrix. B contains only constants, since the $\ell_{s,t}$ are linear forms. Any submatrix S of $H_f(S)$ is for the form

$$B_1 H_{\det_s}(L(X)) B_2$$

where B_1 consists of some rows of B and B_2 of columns of B^T . Now assume that S is a square matrix. By the Binet-Cauchy formula (see below), we have

$$\begin{aligned} \det S &= \sum_I \det(B_1^I) \cdot \det[H_{\det_s}(L(X)) \cdot B_2]_I \\ &= \sum_I \det(B_1^I) \cdot \sum_J [I \times J\text{-minor of } H_{\det_s}(L(X))] \cdot \det B_{2,J} \\ &= \sum_{I,J} \det(B_1^I) [I \times J\text{-minor of } H_{\det_s}(L(X))] \cdot \det B_{2,J}, \end{aligned}$$

where I and J are sets of indices as given by the Binet-Cauchy formula. Since the coefficient vectors c_{\det_s} fulfill E_s by Lemma 2.0.2, $f = \det(L(X))$ also divides all $(2s+1) \times (2s+1)$ -minors of $H_{\det_s}(L(X))$. By the considerations above, every $(2s+1) \times (2s+1)$ -minors of H_f is a linear combination of such minors of $H_{\det_s}(L(X))$. This f divides all $(2s+1) \times (2s+1)$ -minors of H_f . Therefore c_f fulfills E_s . \square

2.0.4 Lemma (Binet-Cauchy formula, without proof). *Let M be an $n \times m$ and N be an $m \times n$ -matrix, $m \geq n$. Then*

$$\det(MN) = \sum_{\substack{I \subseteq \{1, \dots, m\} \\ |I|=n}} \det(M^I N_I),$$

where M^I is the $n \times n$ -submatrix with columns chosen according to I and N_I is the $n \times n$ -submatrix with columns chosen according to I .

2.0.5 Lemma. $\text{per}_{s,n}(X)$ does not divide all $(2s+1) \times (2s+1)$ -minors of $H_{\text{per}_{s,n}}(X)$ for $s < \frac{n^2}{2}$.

Proof. Recall that

$$\text{per}_{s,n}(X) = X_{s,s}^{s-n} \text{per}_n(X|_n).$$

Let $i, j, i', j' \leq n$. Then

$$\frac{\partial^2}{\partial X_{i,j} \partial X_{i',j'}} \text{per}_{s,n}(X) = X_{s,s}^{s-n} \frac{\partial^2}{\partial X_{i,j} \partial X_{i',j'}} \text{per}_n(X|_n).$$

Therefore, the $r \times r$ -minors of $H_{\text{per}_{s,n}}$ with rows and column indices $\leq n$ are of the form

$$X_{s,s}^{r(s-n)} \cdot [\text{corresponding minor } M(X|_n) \text{ of } H_{\text{per}_n}(X|_n)].$$

$\text{per}(X|_n)$ and $M(X|_n)$ are independent of $X_{s,s}$. Therefore $X_{s,s}^{s-n} \text{per}_n(X|_n)$ divides $X_{s,s}^{r(s-n)} M(X|_n)$ iff $\text{per}_n(X|_n)$ divides $M(X|_n)$. Since $\text{per}_n(X|_n)$ is irreducible, it is enough to find a matrix $A \in \mathbb{F}^{n \times n}$ with $\text{per}_n(A) = 0$ but $M(A) \neq 0$ in order to prove that $\text{per}_n(X|_n) \nmid M(X|_n)$.

Mignon and Ressayre (see the previous chapter) construct a matrix A with $\text{per}_n(A) = 0$ but $H_{\text{per}_n(A)}$ has full rank. Therefore, per_n does not divide an $n^2 \times n^2$ -minor of H_{per_n} and therefore, $\text{per}_{s,n}$ does not divide an $n^2 \times n^2$ -minor of $H_{\text{per}_{s,n}}$. If $s = n^2/2 - 1$, then $2s+1 = n^2 - 1$ and the claim follows. \square

2.0.6 Theorem. $\underline{\text{dc}}(\text{per}_n) \geq n^2/2$.

Proof. Assume that $\underline{\text{dc}}(\text{per}_n) = s < n^2/2$. Then there is a sequence of polynomials p_ϵ with $\text{dc}(p_\epsilon) \leq s$ and $p_\epsilon \rightarrow \text{per}_{s,n}$. By Lemma 2.0.3, the coefficient vectors of p_ϵ fulfill the system E_s . But then $c_{\text{per}_{s,n}}$ also fulfills E_s . This contradicts Lemma 2.0.5. \square

Bibliography

- [BES] Markus Bläser, David Eisenbud, and Frank-Olaf Schreyer. Ulrich complexity. *Differential Geometry and Applications*. to appear.
- [Gre12] Bruno Grenet. An upper bound for the permanent versus determinant problem, 2012.
- [Gro15] Joshua A. Grochow. Unifying known lower bounds via geometric complexity theory. *Computational Complexity*, 24(2):393–475, 2015.
- [LMR10] J. M. Landsberg, Laurent Manivel, and Nicolas Ressayre. Hypersurfaces with degenerate duals and the geometric complexity theory program. *CoRR*, abs/1004.4802, 2010.
- [MR04] Thierry Mignon and Nicolas Ressayre. A quadratic bound for the determinant and permanent problem. *International Mathematics Research Notices*, 2004(79):4241–4253, 2004.