AN UPPER BOUND FOR THE PERMANENT VERSUS DETERMINANT PROBLEM

BRUNO GRENET

ABSTRACT. The Permanent versus Determinant problem is the following: Given an $n \times n$ matrix X of indeterminates over a field of characteristic different from two, find the smallest matrix M whose coefficients are linear functions in the indeterminates such that the permanent of X equals the determinant of M. We prove that the dimensions of M are at most $2^n - 1$.

The determinant and the permanent of an $(n \times n)$ matrix of indeterminates $X = (x_{ij})_{1 \leq i,j \leq n}$ over a field $\mathbb{K}$ are the two very similar polynomials defined respectively by

$$\det(X) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} x_{i\sigma(i)} \qquad \text{and} \qquad \operatorname{per}(X) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} x_{i\sigma(i)}$$

where S_n is the set of permutations of $\{1, \ldots, n\}$ and $\varepsilon(\sigma)$ is the signature of the permutation σ. The *determinantal complexity of the permanent* (over $\mathbb{K}$) is the function $\mathsf{dc}(n)$ defined as follows: Given an integer n, $\mathsf{dc}(n)$ is the dimension of the smallest possible matrix M whose coefficients are linear functions of the x_{ij}'s and such that $\det M = \operatorname{per} X$ (as polynomials over $\mathbb{K}$). The *Permanent versus Determinant problem* is to find the value of $\mathsf{dc}(n)$. This value depends on the field $\mathbb{K}$. In particular, if the characteristic of the field is two, permanent and determinant are the same polynomial and $\mathsf{dc}(n) = n$. In this following, we consider fields of characteristic different from two.

The Permanent versus Determinant problem goes back to a question by Pólya [8]. For 2×2 matrices, we have the equality

$$\operatorname{per} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det \begin{pmatrix} a & -b \\ c & d \end{pmatrix}.$$

The natural question of Pólya was to know whether this identity can be extended to higher dimensions. Szegő [10] proved that this is impossible, that is $\mathsf{dc}(n) \geq n + 1$ for $n \geq 3$. More recently, this question has been reinvestigated due to its relations to the algebraic version of the problem P versus NP [1]. It is conjectured that $\mathsf{dc}(n) = 2^{\Omega(n)}$ but the best known lower bound is $\Omega(n^2)$ [6, 2]. See the latter reference for a more extensive description on the history of the problem.

Date: May 25, 2012.

In this note, we are interested in an upper bound for this problem. Using Valiant's universality of the determinant for arithmetic formulas [12] (see also [4]) we can turn a formula for the permanent into a determinant. Thus using Ryser's [9] or Glynn's [3] formula, we obtain $\mathsf{dc}(n) \leq \mathcal{O}(n2^n)$. Actually, Cai, Chen and Li [2, Theorem 3] mention that using Ryser's formula, one can obtain the improved bound $\mathsf{dc}(n) \leq \mathcal{O}(2^n)$. We show in this note a direct construction to build a matrix of dimensions $(2^n - 1)$. The construction is based on an Arithmetic Branching Program (ABP) of size 2^n. Note that the existence of an ABP of this size can also be found as a consequence of a result of Nisan [7].

One can define the *strict* the determinantal complexity of the permanent $\mathsf{dc}_s(n)$ by allowing only constants or variables in the matrix M instead of general linear functions. Then $\mathsf{dc}_s(n) \geq \mathsf{dc}(n)$ and we can actually show that $\mathsf{dc}_s(n) \leq \mathsf{poly}(\mathsf{dc}(n))$. (For instance, $\mathsf{dc}(2) = 2$ and $\mathsf{dc}_s(2) = 3$.) In this note, we actually show that $\mathsf{dc}_s(n) \leq 2^n - 1$, whence $\mathsf{dc}(n) \leq 2^n - 1$. In the same way, Nisan's result [7] applies to ABP whose arcs are weighted by linear forms. The ABP we give is restricted to having only variables and constants as arc weights.

For small values of n, lower and upper bounds are not very far from each other. Mignon and Ressayre [6] proved that for fields of characteristic zero, $\mathsf{dc}(n) \geq n^2/2$. For $n = 3$, this proves that $\mathsf{dc}(3) \geq 5$, and with our upper bound bound, we get $5 \leq \mathsf{dc}(3) \leq \mathsf{dc}_s(3) \leq 7$. It would be interesting to investigate what the exact values of both $\mathsf{dc}(3)$ and $\mathsf{dc}_s(3)$ are.

Definition. An *arithmetic branching program* (ABP) over $\mathbb{K}[\bar{x}]$ is a directed acyclic graph with two distinguished vertices s (the source) and t (the sink). Each arc is weighted by a linear form. The *weight of a path* from s to t is the product of the weights of the arcs it uses. The *value* of the ABP is the sum of the weights of all possible paths from s to t.

An arithmetic branching program is said *layered* if its set of vertices V is the disjoint union of k sets $V_1, \ldots, V_k$ such that $V_1 = \{s\}$, $V_k = \{t\}$, and each arc goes from V_i to V_{i+1} for some $i < k$.

We shall construct an ABP of size 2^n that computes the permanent of X.

Lemma. *There exists an ABP of size 2^n with weights in $\{x_{ij} : 1 \leq i, j \leq n\}$ whose value equals* per X*.*

Proof. The set of vertices of the ABP is indexed by the subsets of $\{1, \ldots, n\}$. The distinguished vertices are $s = \emptyset$ and $t = \{1, \ldots, n\}$. The ABP is layered: Each layer consists of all the subsets of same cardinality. We denote by V_0, $\ldots$, V_n the $(n+1)$ layers. There is an arc from a set $S \in V_i$ of weight x_{ij} to the set $S \cup \{j\}$ whenever $j \notin S$. There is no other arc.

Clearly, the ABP has size 2^n. Consider a path from s to t and the variables that are weights on this path. The range of the first index of the variables in $\{1, \ldots, n\}$ since the path goes through each layer, and an arc falling in the layer V_i has weight x_{ij} for some j. Moreover, the second index j of this variable corresponds to the element that is new in the set. Since the last set in the path is $t = \{1, \ldots, n\}$, each integer between 1 and n appears as a second index of some variable. Since there are n variables on a path from s to t, this shows that each path corresponds to a permutation of $\{1, \ldots, n\}$. Conversely, to each permutation corresponds a path from s to t.

The value of the ABP is the sum over all possible paths from s to t of the weight of the path. Thus it equals the permanent of X. □

This ABP can be turned into a matrix whose determinant has the desired value.

Theorem. *There exists a* $(2^n - 1) \times (2^n - 1)$ *matrix* M *with entries in* $\{x_{ij} : 1 \leq i, j \leq n\} \cup \{-1, 0, 1\}$ *such that* $\det M = \operatorname{per} X$.

Proof. We consider the ABP obtained in Lemma , seen as a digraph. We build a new digraph G obtained by merging the two vertices s and t, and by adding a loop (of weight 1) to all other vertices of the digraph.

Since the vertex s in G is the only vertex with no loop on it, a cycle cover of the graph G must cover s by a cycle of length at least 2. The graph G was obtained by merging the source and the sink of a DAG, thus its only cycles of length at least 2 correspond to path from s to t in the ABP. This shows that a cycle cover of G is made of a large cycle corresponding to a path from s to t in the ABP and of loops. In other words, if we define the weight of a cycle cover by the product of the weights of the arcs it uses, the sum of the weight of all the cycle covers of G equals the value of the ABP.

Now it is a well-known fact that the permanent of the adjacency matrix M of G equals the sum of the weights of all the cycle covers of G. Since the cycle covers of G are made of one large cycle and loops, they correspond to permutations of $\{1, \ldots, n\}$ that have all the same signature. If n is odd, this signature is 1 and $\operatorname{per} M = \det M = \operatorname{per} X$ and we are done. Otherwise, the signature is -1. If we replace in the first layer (of size n) the weights of the loops by -1, then the weight of a cycle cover will be multiplied by $(-1)^{n-1} = -1$. Thus, we get a new matrix N such that $\det N = \operatorname{per} M = \operatorname{per} X$. This concludes the proof. □

When n is odd, we obtained a matrix which uses only 0 and 1 as constants. If n is even, we can also obtain such a matrix, but of dimensions $2^n \times 2^n$.

The determinant of a matrix can be computed by an arithmetic circuit of polynomial size. The circuit can even be made *skew*, meaning that one argument of each multiplication gate is an input [11, 5]. In particular, a skew circuit is non commutative in the sense of Nisan [7]. We prove here that we can obtain here a quite small skew circuit for the permanent based on our construction. Moreover, our construction does not use any negative constant and our construction therefore yields a *monotone* skew circuit.

Corollary. *There exists a monotone skew circuit of size* $(n-1)(2^n - 1)$ *to compute the permanent.*

Proof. This proof is also based on the ABP obtained in Lemma . Since the ABP is layered, we can consider the bipartite subgraphs induced by two consecutive layers. For $1 \leq i \leq n$, let M_i be the biadjacency matrix of the bipartite subgraph induced by the layers V_{i-1} and V_i. The layer V_i contains the subsets of $\{1, \dots, n\}$ of size i, therefore is of size $\binom{n}{i}$. This implies that M_i has $\binom{n}{i-1}$ rows and $\binom{n}{i}$ columns. A vertex $v \in V_i$ is indexed by a subset of size i, and it receives arrows from the i subsets obtained by removing one element in v. This means that each column of M_i contains exactly i nonzero entries.

The value of the ABP is given by the product $M_1 M_2 \cdots M_n$. We can parenthesize this product $(\cdots((M_1 M_2) M_3) \cdots M_n)$ to compute it. Then for $2 \leq i \leq n$, $M_1 \cdots M_i$ is a $(1 \times \binom{n}{i})$ matrix. Using the naive matrix vector multiplication algorithm, the computation of $(M_1 \cdots M_{i-1}) M_i$ requires $(2i-1)$ arithmetic operations per column since each column of M_i has i nonzero entries, that is $(2i-1)\binom{n}{i}$ to compute every entries. Thus the total number of arithmetic operations required to compute the complete product is

$$\begin{aligned} \sum_{i=2}^{n} (2i-1) \binom{n}{i} &= 2 \sum_{i=2}^{n} i \binom{n}{i} - \sum_{i=2}^{n} \binom{n}{i} \\ &= (n2^n - 2n) - (2^n - n - 1) \\ &= (n-1)(2^n - 1). \end{aligned}$$

Each multiplication has one of its argument being an entry of a matrix M_i, therefore the circuit is skew. Moreover, the circuit is monotone since it does not use any constant. □

References

[1] Agrawal, M. Determinant versus permanent. In *Proc. 25th International Congress of Mathematicians* (2006), vol. 3, pp. 985–997.

[2] Cai, J., Chen, X., and Li, D. A quadratic lower bound for the permanent and determinant problem over any characteristic≠ 2. In *Proc. STOC* (2008), ACM, pp. 491–498.

[3] Glynn, D. The permanent of a square matrix. *European J. Combin. 31*, 7 (2010), 1887–1891.

[4] Grenet, B., Kaltofen, E. L., Koiran, P., and Portier, N. Symmetric Determinantal Representation of Formulas and Weakly Skew Circuits. In *Randomization, Relaxation, and Complexity in Polynomial Equation Solving*, no. 556 in Contemp. Math. Amer. Math. Soc., 2011, pp. 61–96.

[5] Malod, G., and Portier, N. Characterizing valiant's algebraic complexity classes. *J. Complexity 24*, 1 (2008), 16–38.

[6] Mignon, T., and Ressayre, N. A quadratic bound for the determinant and permanent problem. *Int. Math. Res. Notices 2004*, 79 (2004), 4241.

[7] Nisan, N. Lower bounds for non-commutative computation. In *Proc. STOC* (1991), ACM, pp. 410–418.

[8] Pólya, G. Augfabe 424. *Arch. Math. Phys. 20* (1913), 271.

[9] Ryser, H. *Combinatorial Mathematics*, vol. 14 of *Carus Math. Monogr.* Mathematical Association of America, Washington, 1963.

[10] Szegő, G. Zu aufgabe 424. *Arch. Math. Phys. 21* (1913), 291–292.

[11] Toda, S. Classes of arithmetic circuits capturing the complexity of computing the determinant. *IEICE T. Inf. Syst. 75*, 1 (1992), 116–124.

[12] Valiant, L. Completeness classes in algebra. In *Proc. STOC* (1979), ACM, pp. 249–261.

LIP, UMR 5668, ÉNS de Lyon – CNRS – INRIA – UCBL, Université de Lyon

E-mail address: Bruno.Grenet@ens-lyon.fr