MULTIPLICATIVE COMPLEXITY OF A PAIR OF BILINEAR FORMS AND OF THE POLYNOMIAL MULTIPLICATION

Dima Grigoriev

MULTIPLICATIVE COMPLEXITY OF A PAIR OF BILINEAR FORMS AND OF THE POLYNOMIAL MULTIPLICATION

D.Yu.Grigoriev
Leningrad Branch of Mathematical Institute of Academy of Science, Fontanka 27, Leningrad, 191011, USSR

The report contains some new bounds on computational complexity of straight-line computations - a known model of computing ([1],[2]). This model well simulates usual computation procedures with branching and cycling instructions depending only on the size of the initial data. This model is also convenient for studying complexity properties of parallel computations - "width" is the minimal number of processors on which the given computation can be realized with the minimal time equal to the "depth". Many well known procedures for algebraic calculations (e.g. multiplication of polynomials, matrices) can be described as straight-line computations. From the technical viewpoint the model under consideration permits to apply for achieving of bounds of computational complexity with its help different algebraic apparatus ([3],[4]).

In the present report we'll consider a problem of computation of a set of bilinear forms over noncommutative indeterminates $\{x_i\},\{y_j\}$ considered earlier in literature ([1],[5],[6]). Straight-line computations will use 4 two-argument arithmetic operations and one-argument operations of multiplications by elements of some field F (further we'll mean it as the main field). We fix the following measure of complexity. By multiplicative complexity (or simply complexity) of a straight-line computation we'll mean a number of two-argument multiplications and divisions in it ([3],[4],[6]). The complexity of a given set S is defined as usually as the minimal complexity of straight-line computations which compute S. Using the results of [5],[6], we can bound ourselves (without increasing the bounds of complexity) only by straight-line computations of the following kind (bilinear chain): at the first stage - computation of some linear forms $\sum_i \alpha_{ij} x_i$, $\sum_i \beta_{ik} y_i$ at the second stage - execution of N two-argument multiplications of the kind $(\sum_i \alpha_{ik} x_i)\cdot(\sum_j \beta_{jk} y_j)$ $(1 \le k \le N)$; at the last stage - computation of some linear combinations of bilinear forms achieved at the second stage (N - is the complexity of the bilinear chain).

The complexity of a set of bilinear forms is equal to the rang of a set of its matrices of coefficients (the rang of a set of matrices is defined as the minimal number of matrices of rang 1 linear cover of which contains the given set of matrices). Analogously can be defined the rang of tensor [6], the rang of an algebra as the rang of its structure tensor and the rang of a group G over a field F as the rang of its group algebra $F(G)$ ([6]). The rang of one matrix in the above-mentioned definition is equal to its usual rang.

In this report the following results are presented: the explicit formula for the rang of a pair of matrices over an algebraically-closed field (theorem 1) and some its corollaries; the new upper bound on the multiplicative complexity over a finite field of the polynomial multiplication (theorem 2); the explicit form of the group of all rang-unchanging linear nonpeculiar transformations of the space of tensors of any given dimension (theorem 3); two effective methods of constructing of some tensors of rang non less that critical - such number that "almost every" tensor is of rang equal to this number (lemma 4.1 and theorem 4); some bounds on the critical rang (statement 4.2).

1. For any pair A,B of the square matrices we define the relation

$$B \lesssim A <=> rg(A,B) = rg(A)$$

Lemma 1.1. The relation $B \lesssim A$ is equivalent to the existence of such a matrix C that

1) $B=AC$;

2) C is of the simple spectrum;

3) $Ker\, C \supseteq Ker\, B \supseteq Ker\, A$

We define the relative rang of the matrix B relatively to the matrix A as follows:

$$rg(B/A) = \min_{C \lesssim A} rg(B-C)$$

Lemma 1.2. For every pair A,B of the square matrices

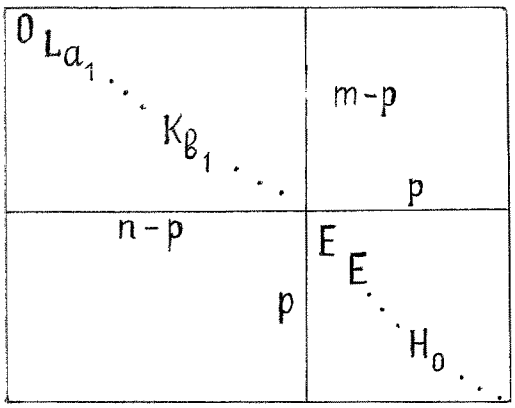
$$rg(A,B) = rg(A) + rg(B/A)$$

(The proof of these two lemmas in the particular case when the matrix A is a unit one can be found in [7]). We assume further in this item that the main field F is algebraically-closed.

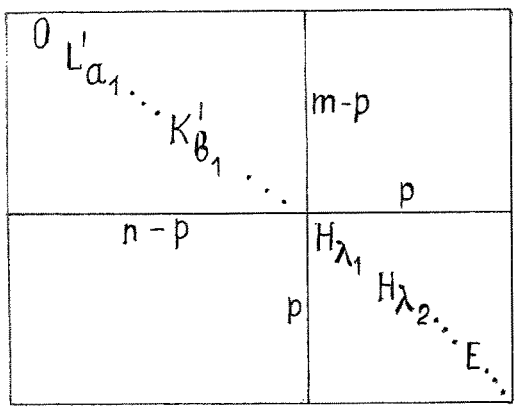
Corollary 1.3. For the fixed m,n $(m \le n)$ the rang of a pair of $m \times n$ matrices is equal to $\min\{n, 2m\}$ everywhere outside some Zarisski-closed set of the dimension less than $2mn$.

If the square matrices C, D are nonpeculiar then $rg(A,B)=rg(CAD,CBD)$. So it's sufficient to find the rang of a pair of the matrices in the canonical Weierstrass-Kronecker form ([8], ch.12). According to the Kronecker's theorem every pair A,B of $m \times n$ matrices (over an algebraically-closed field) by the mentioned transformation can be reduced to the following quasidiagonal form:

$A =$

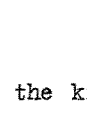
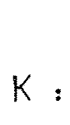


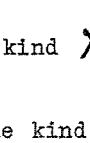
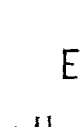
$B =$



At the table all the possible kinds of the blocks are presented (the matrices in any pair of the corresponding in A and B blocks are of the same dimensions).

Singular blocks $a \times (a+1)$ of the kind L : $L_a = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1\ 0 \end{bmatrix}$, $L'_a = \begin{bmatrix} 0\ 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}$

Singular blocks $(\beta+1)\times\beta$ of the kind K : $K_\beta=$  , $K'_\beta=$ 

Regular blocks $S\times S$ of the kind λ : $E=$  , $H_\lambda=$ 

Regular square blocks of the kind ∞ : H_0, E.

Theorem 1. Let a pair A,B of the matrices over an algebraically-closed field in its canonical Weierstrass-Kronecker form contains:

a) ℓ blocks of the kind L : $(L_{a_1}, L'_{a_1}),\dots,(L_{a_\ell}, L'_{a_\ell})$;

b) K blocks of the kind K : $(K_{\beta_1}, K'_{\beta_1}),\dots,(K_{\beta_K}, K'_{\beta_K})$;

c) for every λ d_λ blocks each of the kind λ and of the dimension non less than 2x2 (may be $\lambda=\infty$) and let $d=\max_\lambda d_\lambda$, and all the regular blocks in both A and B form the square $p\times p$ matrices. Then

$$rg(A,B)=\sum_{i=1}^{\ell}(a_i+1)+\sum_{j=1}^{K}(\beta_j+1)+p+d$$

The lemmas 1.1 and 1.2 are used in the proof of the theorem.

Corollary 1.4. For $m\times n (m\le n)$ matrices over an algebraically-closed field

$$\max_{A,B} rg(A,B)=\min\{m+[n/2], 2m\}$$

(here and further $[x]$ - is entier of x, $\lceil x\rceil=-[-x]$).

2. In the second item the new upper bound on the multiplicative complexity over the finite field F of the polynomial multiplication is proved.

The achieved upper bound has the form $n\cdot g_q(n)$ where q is the characteristic of the field F and the function $g_q(n)$ grows (about n) slowly than any fixed iteration of logarithm. It's better (in the sense of the multiplicative complexity) than earlier known upper bounds [9]-[12] (in [12] the bound $c\cdot n\cdot\ell g n\cdot\ell g\ell g n$ is presented).

The multiplicative complexity over the field F of the multiplication of two polynomials both of degree n (we denote this number by $rg_F(P_n)$) is equal to the rang (or multiplicative complexity) over the field F of the following set of bilinear forms:

$$\left\{z_K=\sum_{0\le i,K-i\le n} x_i y_{K-i}, 0\le K\le 2n\right\}$$ over the noncommutative

indeterminates $\{x_i\},\{y_j\}$.

Let $F(q^S)$ be the field consisting of q^S elements, Z_n be the cyclic group of the order n.

Lemma 2.1. For every n

$$rg_F(Z_n)\le rg_F(P_{n-1})\le rg_F(Z_{2n-1});$$

$$rg_{F(q)}(F(q^n))\le rg_{F(q)}(P_{n-1})\le rg_{F(q)}(F(q^{2n-1}))$$

In the more general form these inequalities were proved in the recently published [13].

Lemma 2.2. Let $\tau=q^m-1$. Then there is the following decomposition at the direct sum:

$$F(q)(Z_\tau)\simeq \sum_i \oplus F(q^{K_i})$$

where $K_i \mid m$ for every i (certainly, $\sum_i K_i=\tau$).

Let's define the function $g_q(n)$ in the following manner. We define $g_q(2)=g_q(3)=g_q(4)-1=2$. If $m>4$ is equal to $[(q^{S-1})/2]+1$ for some integer S, then we define $g_q(m)=2q\cdot g_q(S)$. If $m>4$ and for some integer S the following inequalities are fulfiled: $[(q^S)/2]<m\le[(q^{S+1})/2]$, then we define $g_q(m)=g_q([(q^S)/2]+1)$.

Theorem 2. For every n

$$rg_{F(q)}(P_{n-1})\le n\cdot g_q(n)$$

We use the induction on n. Let for $n\le S (S\ge 4)$ the inequality is true. We set $\tau=q^S-1$ and using in succession the first inequality from the lemma 2.1, the lemma 2.2 and the inequality $rg(\mathcal{A}\oplus\mathcal{L})\le rg(\mathcal{A})+rg(\mathcal{L})$ for any algebras $\mathcal{A}$, $\mathcal{L}$ ([6]), the second inequality from the lemma 2.1, the induction conjecture, again lemma 2.2 and the monotony about n of the function $g_q(n)$, we obtain a chain of inequalities:

$$rg_{F(q)}(P_{[(\tau-1)/2]})\le rg_{F(q)}(Z_\tau)\le \sum_i rg_{F(q)}(F(q^{K_i}))\le$$

$$\sum_i rg_{F(q)}(P_{K_i-1})\le \sum_i K_i g_q(K_i)\le \tau\cdot g_q(S)$$

Let $n>4$ and $[(q^{S-1})/2]<n\le[(q^S)/2]$. Using in succession the monotony about n of the function $rg_{F(q)}(P_n)$, the inequality proved, the definition of the function $g_q(n)$ and its monotony about n, we obtain a chain of inequalities completing the proof of the theorem:

$$\tau g_{F(q)}(p_{n-1}) \le \tau g_{F(q)}(p_{[(\tau-1)/2]}) \le \tau \cdot g_q(S) \le$$

$$([(q^{S-1})/2]+1)\cdot g_q([(q^{S-1})/2]+1) \le n \cdot g_q(n)$$

Let's remark that the function $g_q(n)$ is inverse to some function from the class $\mathscr{E}^4 \setminus \mathscr{E}^3$ of Grzegorczyk's hierarchy ([14]).

3. The following two kinds of transformations of the tensorproduct space $U_1 \otimes \cdots \otimes U_K$ of the vector spaces $U_1, \ldots, U_K$ doesn't change the rang of the tensors:

1) a nonpeculiar linear transformation in any component $U_i (1 \le i \le K)$;

2) if for some i, j the mapping $f: U_i \to U_j$ is an isomorphism of the vector spaces, then the rang is unchanged under the following transformation:

$$u_1 \otimes \cdots \otimes u_i \otimes \cdots \otimes u_j \otimes \cdots \otimes u_K \to u_1 \otimes \cdots \otimes f^{-1}(u_j) \otimes \cdots \otimes f(u_i) \otimes \cdots \otimes u_K$$

Theorem 3. The group of all nonpeculiar linear transformations of the space $U_1 \otimes \cdots \otimes U_K$, mapping the tensors of the rang 1 to the tensors of the rang 1, coincides with the group generated by the transformations of the kinds 1), 2).

4. Henceforth we assume that the main field F of the characteristic q is algebraically-closed, and let F_q be the primitive field of the characteristic q (q is prime or equal to zero).

Lemma 4.1. There exist such primitive-recursive functions

$$\tau_q = \tau_q(n_1, \ldots, n_K),\ d = d(n_1, \ldots, n_K),\ M = M(n_1, \ldots, n_K)$$

that the rang of any tensor from the space $F^{n_1} \otimes \ldots \otimes F^{n_K}$ is equal to $\tau_q(n_1, \ldots, n_K)$ (the critical rang) everywhere on some nonempty Zarisski-open set, and the coefficients of any tensor which rang is less than τ_q satisfy some algebraic equation with the coefficients from F_q, of degree less than d and with the sum of the modules of the coefficients (in the case when $q = 0$) less than M.

The functions M, d can be found in the class $\mathscr{E}^3$ of Grzegorczyk's hierarchy.

Theorem 4. 1) Let $\mu_1, \ldots, \mu_S \in \bar{F}_q$ $(S = n_1 \cdot \ldots \cdot n_K)$ be some elements of the degrees $d^{2^1}, \ldots, d^{2^S}$ over F_q, and let $\mu_1, \ldots, \mu_S$ be the coefficients (in any order) of some tensor $\tau \in \bar{F}_q^{n_1} \otimes \cdots \otimes \bar{F}_q^{n_K}$. Then $\tau g_{\bar{F}_q}(\tau) \ge \tau_q$;

2) Let $\mu'_1 = 1, \ldots, \mu'_{\ell+1} = M(\mu'_\ell (S+1))^d + 1, \ldots, \mu'_S$ be the coefficients (in any order) of the integer tensor $\tau' \in \mathbb{Q}^{n_1} \otimes \ldots \otimes \mathbb{Q}^{n_K}$. Then $\tau g_{\bar{\mathbb{Q}}}(\tau') \ge \tau_0$ (the numbers τ_q, M, d are taken from the lemma 4.1).

The idea of constructing tensors of the rang non less the critical in the theorem 4 is like the idea of Strassen [4] for constructing the polynomials which are hard to compute, but our idea (in applying to the problem under consideration) gives some more strong lower bound (the critical rang), using unfortunately very fast-growing functions M, d.

In conclusion we bound the value of the critical rang.

Statement 4.2. For every $n_1, \dots, n_K$

$$n_1 \cdot \dots \cdot n_K / (n_1 + \dots + n_K - (K-1)) \le z_q(n_1, \dots, n_K) \le [n_1/2] \max\{n_2, n_3\} \cdot n_4 \cdot \dots \cdot n_K$$

Acknowledgment. I would like to thank A.O.Slisenko for his constant attention to my work.

References

1. A.Borodin, M.Munro, The computational complexity of algebraic and numeric problems, Ser.Theory of Computations, Amer.Elsevier, N.Y., 1975.
2. V.Strassen, Berechnung und Program 1, Acta Inform., 1, 1972, p.320-335.
3. V.Strassen, Die Berechnungskomplexität von elementarsymmetrischen Funktionen und von Interpolationskoeffizienten, Numer.Math., 20, 1973, p.238-251.
4. V.Strassen, Polynomials which are hard to compute, SIAM J.Comput., 3,2, 1974, p.128-149.
5. S.Winograd, On the number of multiplications necessary to compute certain functions, Communs Pure and Appl.Math., 23, 1970, p.165-179.
6. V.Strassen, Vermeidung von Divisionen, J.reine und angew Math., 264, 1973, p.184-202.
7. Д.Ю.Григорьев, Об алгебраической сложности вычисления пары билинейных форм, Зап.научн.семинаров Ленингр.отд.Матем.ин-та АН СССР, 47, с.159-163, 1974.
8. Ф.Р.Гантмахер, Теория матриц, М., 1954.
9. A.A.Karatsuba, Yu.P.Ofman, Multiplication of multidigit numbers on automata, Soviet Physics Dokl., 7, 1963, p.595-596.
10. A.L.Toom, The complexity of a scheme of functional elements realizing the multiplication of integers, Soviet Math.Dokl., 4, 1963, p.714-716.
11. A.Schönhage, V.Strassen, Schnelle Multiplikation großer Zahlen, Computing, Arch.für elektr. Rechn., 7, 3-4, 1971, p.281-292.
12. A.Schönhage, Schnelle Multiplikation von Polynomen über Körpern der Charakteristik 2, Acta Inform., 7, 4, 1977, p.395-398.
13. C.M.Fiduccia, Y.Zalcstein, Algebras having linear multiplicative complexity, J.Assoc.Comput.Mach., 24, 2, 1977, p.311-331.
14. A.Grzegorczyk, Some classes of recursive functions, Rozprawy, Matematiczne IV, Warszawa, 1953, p.1-46.