Abstract

The notion of the tensor rank over a field F was introduced by Strassen for describing the minimal number (named multiplicative complexity) of multiplications/divisions necessary to compute a given set of bilinear forms over non-commutative indeterminates. The following results are presented in this paper. (I) An explicit formula for the rank of a pair of matrices (the most simple nontrivial case of the tensor rank) over an algebraically closed field F . (2) A new upper bound for the multiplicative complexity (over a finite field F) of polynomial multiplication is obtained. (3) An explicit description of the group of all the linear invertible transformations preserving the rank of tensors of a given dimension is suggested. (4) An existence (when F is algebraically closed) of the critical rank (for the space of tensors of a given dimension) is proved. This number is equal to bhe common rank of the tensors in some nonempty Zarisski-open set.As a corollary two "effective" methods for the constructing of the tensors with the rank no less than the critical one are exhibited. Also some bounds for the critical rank are formulated.

Let $V_1,\dots,V_K$ be the vector spaces over a field F and $\tau\in V_1\otimes_F\dots\otimes_F V_K$. We define a rank $rg_F(\tau)$ in the following manner ([I]):

$$rg_F(\tau)=min\{N:\tau=\sum_{i=1}^{N} v_1^i\otimes\dots\otimes v_K^i,\ v_j^i\in V_j\}$$

The rank of a set $\{A^{(1)},\dots,A^{(p)}\}$ of the matrices is defined as the minimal number of the rank I matrices by the linear combinations of which all the matrices $A^{(1)},\dots,A^{(p)}$ can be expressed. Evidently, $rg(A^{(1)},\dots,A^{(p)})$ is equal to the rank of the tensor $A=(a_{ij}^{\ell})$, where a_{ij}^{ℓ} is (i,j) -entry of $A^{(\ell)}$) The rank (or the multiplicative complexity) of a set of the bilinear forms is defined as the rank of the set of their coefficient matrices.

I. For two square matrices A,B we define a relation

$$B\preccurlyeq A\Longleftrightarrow rg(A,B)=rg(A)$$

LEMMA I.I. The relation $B\preccurlyeq A$ is equivalent to the existence of such a matrix C that

I) $B=AC$; 2) C has a basis of the eigen-vectors (such the matrices will be named diagonalizable); 3) Ker $C\supseteq$ Ker $B\supseteq$ Ker A .

LEMMA I.2. $$rg(A,B)=rg(A)+min_{C\preccurlyeq A}rg(B-C)$$

(The case A is a matrix unit was treated in [2]). Henceforth in the section I F is assumed to be algebraically closed.

PROPOSITION I.3. The rank of a pair of $m\times n\,(m\leqslant n)$ matrices is equal to $min\{n,2m\}$ everywhere in some nonempty Zariski-open set.

If the square matrices $C,\mathcal{D}$ are invertible, then $rg(A,B)=rg(CA\mathcal{D},CB\mathcal{D})$. So it is sufficient to find the rank of a pair of matrices in the Weierstrass-Kronecker canonical form ([3],ch.I2). By the Kronecker's theorem every pair of $m\times n$ matrices over an algebraically closed field can be reduced (with the help of the transformation $A,B\longrightarrow CA\mathcal{D},CB\mathcal{D}$ with invertible $C,\mathcal{D}$) to the following quasidiagonal form:

$$A=\left[\begin{array}{c|c} O\,L_{a_1}\cdots K_{b_1}\cdots & \\ \hline & E_E\cdots H_0\cdots \end{array}\right]\qquad B=\left[\begin{array}{c|c} O\,L'_{a_1}\cdots K'_{b_1}\cdots & \\ \hline & H_{\lambda_1}H_{\lambda_2}\cdots E\cdots \end{array}\right]$$

(all the corresponding blocks have the same dimensions). There are the following types of the corresponding pairs of the blocks:

singular blocks $a_i \times (a_i+1)$ of the type L :

$$L_{a_i} = \boxed{\begin{matrix} 1 & & & \\ & \ddots & & \\ & & 1 & 0 \end{matrix}}\ , \qquad L'_{a_i} = \boxed{\begin{matrix} 0 & 1 & & \\ & & \ddots & \\ & & & 1 \end{matrix}}\ ,$$

singular blocks $(b_j+1)\times b_j$ of the type K :

$$K_{b_j} = \boxed{\begin{matrix} 0 & & \\ 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}\ , \qquad K'_{b_j} = \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \\ & & 0 \end{matrix}}\ ,$$

regular square blocks of the type λ :

$$E = \boxed{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}\ , \qquad H_\lambda = \boxed{\begin{matrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{matrix}}$$

regular square blocks of the type $\infty : (H_0, E)$.

THEOREM I. Let a pair of matrices A, B over an algebraically closed field F in its Weierstrass-Kronecker canonical form contain:

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

b) K blocks of the type $K : (K_{b_1}, K'_{b_1}), \ldots, (K_{b_K}, K'_{b_K})$;

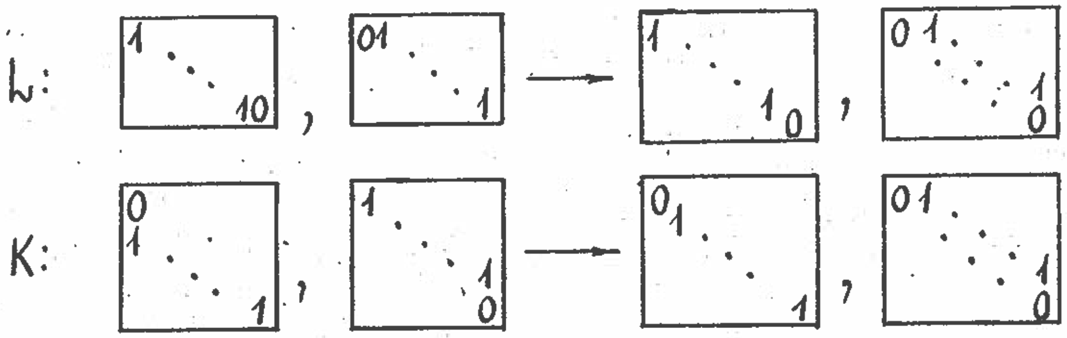
c) for every λ d_λ blocks of the type λ with dimensions no less than 2×2 (may be $\lambda=\infty$). Let $d = max_\lambda d_\lambda$ and all the regular blocks in both A and B form the square $p\times p$ matrices. Then

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

Let us prove the lower bound for $rg(A,B)$

Firstly we can replace the matrices A, B by the matrices $A' = \alpha A + \beta B,\ B' = \gamma A + \delta B$ ($\left|\begin{smallmatrix}\alpha & \beta \\ \gamma & \delta\end{smallmatrix}\right| \neq 0$ hence $rg(A,B) = rg(A',B')$ and the parametres $\{a_i\}, \{b_j\}, p, d$ remain unchanged), so that the canonical form of (A',B') doesn't contain blocks of the type ∞ and besides that $d = d_0$ (we shall denote A', B' again by A, B).

Then we add some null rows and columns to A and B simultaniously in order to make A, B $N\times N$ -square and A containing all its unities on the principal diagonal. We describe this transformation for the singular pairs of the blocks:



Assume the lower bound is wrong, then by lemmas I.I.,I.2. there exists such a diagonalizable (this property is implied by the structure of A) matrice $C \leqslant A$ that $rg(B-C)<K+\ell+d$.

We denote $r=rg\,C$ and choose the principal $r\times r$ minor C_1 (it is an undermatrix containing r entries of the principal diagonal) of C such that $rg(C_1)=r$.

Let S be the set of $(d+K+\ell)$ rows of B consisting of the first rows of (I) $(K+\ell)$ singular blocks; (2) $d=d_0$ regular blocks of the type O which have the dimensions no less than 2×2. Having done some permutation of the elements of the basis of the N-dimensional space, we place C_1 at the lower right corner of C. We demand also this permutation to give the rows from S the numbers from $(s-t+1)$ to $(s+d+K+\ell-t)$ for some $t(0\leq t\leq d+K+\ell)$, $s=N-r$ (we preserve the same notations for the permuted matrices).

Denote by $w_1,\dots,w_{r+t}$ the rows of B with the numbers from $(s-t+1)$ to N. Every $w_i\,(1\leq i\leq d+K+\ell)$ contains only one unity and these unities are the only non-zero entries in the rows and the columns containing them (we shall refer to this property as S-property). Moreover the columns of B with the numbers from $(s-t+1)$ to $(s+d+K+\ell-t)$ contain only zeroes (this property will be named B-property) owing to the choosing of S (together with the inequality $rg(B-C)<d+K+\ell$ it gives $t\geq 1$). S-property and B-property are deduced from the equality $d=d_0$ and the form of the transformations adding null rows and columns.

So there exist the rows $u_1,\dots,u_{d+K+\ell-t}$ of $(C-B)$ with the numbers greater than S whose projections on the columns with the numbers from $(s+1)$ to $(s+d+K+\ell-t)$ are linear independent (this proposition can be deduced from the B-property and invertibility of C_1). We supplement rows $u_1,\dots,u_{d+K+\ell-t}$

of $(C-B)$ with such rows $u_{d+K+\ell-t+1},\dots,u_h\,(h<d+K+\ell)$ that $u_1,\dots,u_h$ form the basis of the space of all the rows of $(C-B)$.

We denote by $v_1,\dots,v_{z+t}$ the rows of C with the numbers from $(s-t+1)$ to N. As $zg\,C=zg\,C_1=z$ then there are t linear equalities:

$$\begin{aligned} v_1 &= \sum_{i=1}^{z}\alpha_{i1}v_{i+t}\\ &\vdots\\ v_t &= \sum_{i=1}^{z}\alpha_{it}v_{i+t}\end{aligned}$$

We replace every $v_i\,(1\leqslant i\leqslant z+t)$ by the sum of row W_i of B with the number $(i+s-t)$ and row of $(C-B)$ with the number $(i+s-t)$, and we express the latter row as a linear combination of $u_1,\dots,u_h$. So we obtain the equalities

$$\begin{aligned} W_1 &= \sum_{i=1}^{z}\alpha_{i1}W_{i+t}+\sum_{j=1}^{h}\beta_{1j}u_j\\ &\vdots\\ W_t &= \sum_{i=1}^{z}\alpha_{it}W_{i+t}+\sum_{j=1}^{h}\beta_{tj}u_j\end{aligned}$$

As $h<d+K+\ell$ then for some nontrivial linear combination $\sum_{i=1}^{t}\gamma_iW_i$ we obtain

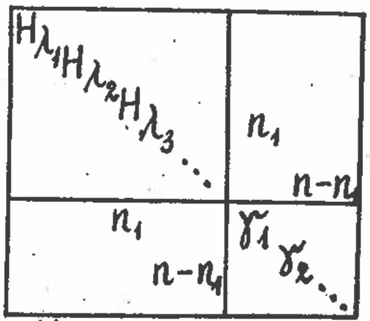
$$\sum_{i=1}^{t}\gamma_iW_i=\sum_{i=1}^{z}\alpha_iW_{i+t}+\sum_{j=1}^{d+K+\ell-t}\beta_ju_j$$

Projecting this equality on the columns with the numbers from $(s+1)$ to $(s+d+K+\ell-t)$ we infer that $\beta_1=\dots=\beta_{d+K+\ell-t}=0$, and so this equality leads to the contradiction with S-property. The lower bound is proved.

Let us prove the upper bound for $zg(A,B)$

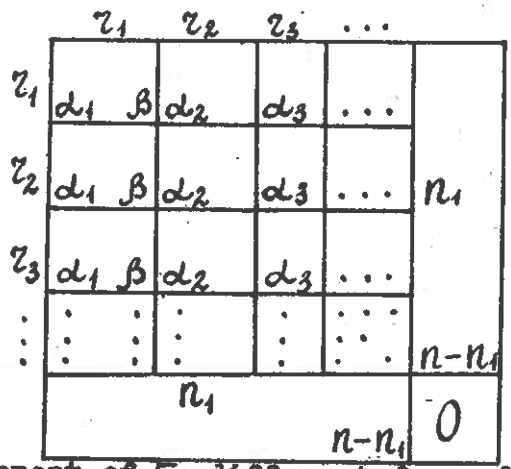
LEMMA I.4. Let $f_0,f_1,\dots,f_n\in F[x]$ and not all the derivatives $f_0',f_1',\dots,f_n'$ vanish identically. Assume for every $\alpha_1,\dots,\alpha_n\in F$ there exists a multiple root of $f_0+\sum_i\alpha_if_i$. Then $f_0,f_1,\dots,f_n$ have a nontrivial common divisor (here F is any infinite field).

LEMMA I.5. Let $n\times n$ matrix A be equal to



where Jordan's block H_{λ_i} is of the dimension $r_i\times r_i (r_i\geq 2; 1\leq i\leq s)$ and $\lambda_i\neq\lambda_j$ $(1\leq i<j\leq s)$.Then $rg(E,A)\leq n+1$ (E is a matrix unit).

Let q be the characteristic of an algebraically closed field F .Consider the following matrix C of the rank no greater than I:



where β in any element of F different from all $(\lambda_j-\lambda_1)(1<j\leq s)$ if $q|K_i$ for all $1\leq i\leq s$, and $\beta=0$ otherwise. By lemma I.4. there exist such $\alpha_1,\dots,\alpha_s$ that the characteristic polynomial of the upper left $n_1\times n_1$ undermatrix of $(A+C)$ has no multiple roots, so $(A+C)$ is diagonalizable. Hence by lemmas I.I.,I.2. $rg(E,A)\leq n+1$.

LEMMA I.6. $rg(L_a,L'_a)\leq a+1$, $rg(K_b,K'_b)\leq b+1$.

Theorem I is concluded from the proved lower bound and by lemmas I.5.,I.6.

COROLLARY I.7. For $m\times n (m\leq n)$ matrices over an algebraically closed field

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

($[x]$ is an entire part x, $\lceil x\rceil=-[-x]$).

PROPOSITION I.8. Rank (over $\bar{\mathbb{Q}}$) of a pair of integer matrices can be computed in polynomial time.

Using [3], ch.I2,§ 3 we can find in polynomial time all the singular blocks of a pair of integer matrices (the solutions of the auxiliary systems of the linear equations with the rational coefficients can be found in polynomial time by [4]).

Now we restrict ourselves to considering a pair of integer matrices A, B having only regular blocks in its canonical form, and we have to find a parametre d (without loss of generality, we again assume the pair A, B has no blocks of the type ∞).We compute the elements of matrix $C=A^{-1}B$ again using [4]. Let $i_1(\lambda),\dots,i_K(\lambda)$ be the invariant polynomials of $C(i_{j+1}(\lambda)|i_j(\lambda)$ for all $1\leqslant j<K)$. Let $\rho_j(\lambda)=(i_j(\lambda),i_j'(\lambda))$ where $i_j'(\lambda)$ is a derivati e of $i_j(\lambda)$, then $d=max\{j: deg\,\rho_j(\lambda)>0\}$. To find $i_1(\lambda),\dots,i_K(\lambda)$ in polynomial time we can use Hu's method from supplement A in [5] extended from integer matrices to integer λ-matrices (M.A.Frumkin has drawn the author's attention to the possibility of such an extension).

The proposition is also valid for matrices over finite fields, the field of rational Gaussian numbers etc.

2. The multiplicative complexity (over a finite field F) of the polynomial multiplication is equal to the rank (over the field F) of the following set of bilinear forms: $\left\{z_K=\sum\limits_{0\leqslant i,K-i\leqslant n} x_i y_{K-i},\ 0\leqslant K\leqslant 2n\right\}$, we denote this number by $rg_F(P_n)$. Let F be a finite field of the characteristic q. We shall prove below that $rg_F(P_{n-1})\leqslant n\cdot g_q(n)$ where $g_q(n)$ is the inverse of some function which belongs to the class $\mathcal{E}^4$ of Grzegorczyk's hierarchy [6] but not to $\mathcal{E}^3$, and so $g_q(n)$ grows slower than $log\,log\cdots log\,n$ with any fixed number of iterations. This upper bound for the multiplicative complexity is better than one based on the fast Fourier transform [7],[8].

Strassen [I]defined the rank of an algebra as the rank of its structure tensor and defined the rank of a group G (over some field F) as the rank of group algebra $F(G)$ .Let $F(q^s)$ be the field consisting of q^s elements, Z_n be the cyclic group of order n.

LEMMA 2.I. For every integer n

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

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

In more general form these inequalities were formulated in the recently published [9].

LEMMA 2.2. Let $r=q^m-1$. Then $F(q)(Z_r)\simeq\sum_i\oplus F(q^{K_i})$ where $K_i|m$ for every i (certainly, $\sum_i K_i=r$).

Let us define $g_q(n)$ in the following manner. Set $g_q(2)=g_q(3)=g_q(4)-1=2$. If $m>4$ is equal to $[(q^{s-1})/2]+1$ for some integer s, then define $g_q(m)=2q\,g_q(s)$. If $m>4$ and $[(q^s)/2]<m\le[(q^{s+1})/2]$ for some integer s, then define $g_q(m)=g_q([(q^s)/2]+1)$.

THEOREM 2. For every integer n

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

The theorem will be proved by the induction on n. Let for $n\le s$ $(s>4)$ the inequality is proved. Set $r=q^s-1$ and using by turns the first inequality of lemma 2.I.,lemma 2.2. and the inequality $rg(\mathfrak{A}\oplus\mathfrak{B})\le rg(\mathfrak{A})+rg(\mathfrak{B})$ [I]which is valid for any algebras $\mathfrak{A},\mathfrak{B}$, the second inequality of lemma 2.I., the induction conjecture, again lemma 2.2 and the non-decreasing of the function $g_q(n)$ in n, we obtain a chain of inequalities:

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

$$\sum_i rg_{F(q)}(P_{K_i-1})\le\sum_i K_i\,g_q(K_i)\le r\cdot g_q(s).$$

Let $n>4$ so that $[(q^{s-1})/2]<n\le[(q^s)/2]$. Using by turns the non-decreasing of the function $rg_{F(q)}(P_n)$ in n, the inequality just proved, the definition of the function $g_q(n)$ and its non-decreasing in n, we obtain a chain of inequalities completing the proof of the theorem:

$$rg_{F(q)}(P_{n-1})\le rg_{F(q)}(P_{[(r-1)/2]})\le r\cdot g_q(s)\le$$

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

3. The following two kinds of transformations of a tensor-product space $V_1\otimes\cdots\otimes V_K$ of some vector spaces $V_1,\dots,V_K$ do not change the rank of tensors:

I) invertible linear transformation in every component V_i $(1\leqslant i\leqslant K)$

2) if for some i,j $f: V_i \to V_j$ is an isomorphism of the vector spaces, then the following transformation does not change the rank:

$$v_1\otimes\dots\otimes v_i\otimes\dots\otimes v_j\otimes\dots\otimes v_K \to v_1\otimes\dots\otimes f^{-1}(v_j)\otimes\dots\otimes f(v_i)\otimes\dots\otimes v_K$$

THEOREM 3. The group of all invertible linear transformations of $V_1\otimes\dots\otimes V_K$ mapping tensors with rank I in tensors with rank I coincides with the group generated by transformations of the kinds I), 2).

The proof is based on the following lemma:

LEMMA 3.I. Let $\tau_1,\tau_2\in V_1\otimes\dots\otimes V_K$ be such that $rg(\tau_1)=rg(\tau_2)=rg(\tau_1+\tau_2)=1$. Then there exist such $i (1\leqslant i\leqslant K)$, $v^{(1)}, v^{(2)}\in V_i$ and $v_j\in V_j\,(j\neq i)$ that

$$\tau_p = v_1\otimes\dots\otimes v_{i-1}\otimes v^{(p)}\otimes v_{i+1}\otimes\dots\otimes v_K\ (p=1,2).$$

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

Let $X^{(q)}(n_1,\dots,n_K)=\{x^{(q)}_{i_1,\dots,i_K}: 1\leqslant i_j\leqslant n_j,\ 1\leqslant j\leqslant K\}$ be an algebraically independent set over F_q and denote by $G_q=F_q(X^{(q)}(n_1,\dots,n_K))$ a field of the algebraic functions. We define a tensor $T_q(n_1,\dots,n_K)\in G_q^{n_1}\otimes\dots\otimes G_q^{n_K}$ whose $(i_1,\dots,i_K)$-entry is equal to $x^{(q)}_{i_1,\dots,i_K}$ and set $r_q(n_1,\dots,n_K)= rg_{G_q}(T_q(n_1,\dots,n_K))$.

LEMMA 4.I. There exist such primitive-recursive functions $d=d(n_1,\dots,n_K)$, $M=M(n_1,\dots,n_K)$ that the rank of a tensor from $F^{n_1}\otimes\dots\otimes F^{n_K}$ is equal to $r_q=r_q(n_1,\dots,n_K)$ everywhere in some nonempty Zariski-open set in $F^{n_1}\otimes\dots\otimes F^{n_K}$, and the entries of any tensor with rank less than r_q satisfy some algebraic equation having coefficients in F_q, the degree less than d and the sum of modules of the coefficients (in the case when $q=0$) less than M.

This lemma can be proved by the method of quantifier-elimination and the author conjectures that such M,d can be found in $\mathcal{E}^3$ of Grzegorczyk's hierarchy.

THEOREM 4. I) Let $\mu_1,\dots,\mu_s\in F_q$ $(s=n_1\cdot\ldots\cdot n_K)$ be some elements of degrees $d^{2^1},\dots,d^{2^s}$ over F_q and let $\mu_1,\dots,\mu_s$ be entries (in any order) of some tensor $\tau\in\overline{F}_q^{n_1}\otimes\dots\overline{F}_q^{n_k}$. Then $rg_{\overline{F}_q}(\tau)\geqslant r_q$.

2) Let $\mu'_1=1,\dots,\mu'_{\ell+1}=M\cdot(\mu'_\ell(s+1))^d+1,\dots,\mu'_s$ be entries (in any order) of an integer tensor $\tau'\in\mathbb{Q}^{n_1}\otimes\dots\otimes\mathbb{Q}^{n_k}$. Then $rg_{\overline{\mathbb{Q}}}(\tau')\geqslant r_0$ (the numbers r_q, M, d are taken from lemma 4.I).

The idea of constructing τ,τ' is similar to one used by Strassen [IO] to construct polynomials which are "hard to compute". But our method gives some stronger (for the problem under consideration) lower bound (namely, the critical rank) using, unfortunately, functions M,d which grow very fast.

PROPOSITION 4.2. For every integer $n_1,\dots,n_k$

$$n_1\cdot\dots\cdot n_k/(n_1+\dots+n_k-(k-1))\leqslant r_q(n_1,\dots,n_k)\leqslant\lceil n_1/2\rceil\max\{n_2,n_3\}n_4\cdot\dots\cdot n_k$$

and for every prime q greater than some number (depending on $n_1,\dots,n_k$) $r_q(n_1,\dots,n_k)=r_0(n_1,\dots,n_k)$.

REMARK. According to proposition I.3. $r_q(2,m,n)=\min\{2m,n\}$, where $m\leqslant n$.

Acknowledgment. I would like to thank A.O.Slisenko for his permanent attention to my work and also to thank Yu.V.Matijasevich and G.E.Mints for the valuable remarks.

References

I. V. Strassen. Vermeidung von Divisionen, J.reine und angew. Math.,264,I973,p.I84-202.

2. Д.Ю. Григорьев. Об алгебраической сложности вычисления пары билинейных форм, Зап.научн.семинаров Ленингр.отд.Матем.ин-та АН СССР,47,I974,с.I59-I63.

3. Ф.Р. Гантмахер. Теория матриц, М.,I954.

4. I.Borosh, A.S. Frankel. Exact solutions of linear equations with rational coefficients by congruence techniques, Math.of Comp.,20,93,I966,p.I07-II7.

5. T.C. Hu. Integer programming and network flows, A-W,I970.

6. A. Grzegorczyk. Some classes of recursive functions, Rozprawy Matematiczne,IV,Warszawa,I953,p.I-46.

7. A. Schönhage, V. Strassen. Schnelle Multiplikation großer Zahlen,Computing,Archiv für electronisches Rech-

nen, 7,3-4,I97I,p.28I-292.

8. A. Schönhage. Schnelle Multiplikation von Polynomen über Körpern der Charakteristik 2,Acta Inform.,7,4,I977,p.395-398.

9. C.M. Fiduccia, Y. Zalcstein. Algebras having linear multiplicative complexity, J.Assoc.Comput.Math.,24,2, I977,p.3II-33I.

I0. V.Strassen. Polynomials with rational coefficients which are hard to compute,SIAM J.Comput.,3,2,I974,p.I28-I49.

Подписано к печати I7/Ш-78 г. М-08370
Формат 60х90 I/I6. Бумага офсетная. Печать офсетная.
Печ.л. 3/4=0.75 усл.печ.л. Уч.-изд.л. 0 72
Тираж 250 экз.Зак.276. Цена I0 коп.

Издательство „Наука", Ленинградское отделение
I99I64, Ленинград, В-I64, Менделеевская лин., I

I-я типография издательства „Наука"
I99034, Ленинград, В-34, 9 линия, I2