Linear Algebra and its Applications 438 (2013) 4500–4509

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

The rank of $n \times n$ matrix multiplication is at least $3n^2 - 2\sqrt{2}n^{\frac{3}{2}} - 3n$

Alex Massarenti[a], Emanuele Raviolo[b,*]

[a] SISSA, via Bonomea 265, 34136 Trieste, Italy
[b] Università di Pavia, via Ferrata 1, 27100 Pavia, Italy

ARTICLE INFO

Article history:
Received 2 December 2012
Accepted 23 January 2013
Available online 6 March 2013

Submitted by Shmuel Friedland

AMS classification:
Primary 14Q20
Secondary 13P99
68W30

Keywords:
Tensors
Matrix multiplication
Complexity theory

ABSTRACT

We prove that the rank of the $n \times n$ matrix multiplication is at least $3n^2 - 2\sqrt{2}n^{\frac{3}{2}} - 3n$. The previous bounds were $3n^2 - 4n^{\frac{3}{2}} - n$ due to Landsberg [2] and $\frac{5}{2}n^2 - 3n$ due to Bläser [1]. Our bound improves the previous bounds for any $n \geqslant 24$.

0. Introduction

The multiplication of two matrices is one of the most important operations in mathematics and applied sciences. To determine the complexity of matrix multiplication is a major open question in algebraic complexity theory.

Recall that the matrix multiplication $M_{n,l,m}$ is defined as the bilinear map

$$\begin{array}{rccc} M_{n,l,m}: & \mathrm{Mat}_{n\times l}(\mathbb{C}) \times \mathrm{Mat}_{l\times m}(\mathbb{C}) & \rightarrow & \mathrm{Mat}_{n\times m}(\mathbb{C}) \\ & (X, Y) & \mapsto & XY, \end{array}$$

[*] Corresponding author. Tel.: +39 3204924292.
E-mail addresses: alex.massarenti@sissa.it (A. Massarenti), emanuele.raviolo@unipv.it (E. Raviolo).

0024-3795/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.laa.2013.01.031

where $\mathrm{Mat}_{n\times l}(\mathbb{C})$ is the vector space of $n \times l$ complex matrices. A measure of the complexity of matrix multiplication, and of tensors in general, is the *rank*. For the bilinear map $M_{n,l,m}$ this is the smallest natural number r such that there exist $a_1, \dots, a_r \in \mathrm{Mat}_{n\times l}(\mathbb{C})^*$, $b_1, \dots, b_r \in \mathrm{Mat}_{l\times m}(\mathbb{C})^*$ and $c_1, \dots, c_r \in \mathrm{Mat}_{n\times m}(\mathbb{C})$ decomposing $M_{n,l,m}(X, Y)$ as

$$M_{n,l,m}(X, Y) = \sum_{i=1}^{r} a_i(X)b_i(X)c_i$$

for any $X \in \mathrm{Mat}_{n\times l}(\mathbb{C})$ and $Y \in \mathrm{Mat}_{l\times m}(\mathbb{C})$.

In the case of square matrices the standard algorithm gives an expression of the form $M_{n,n,n}(X, Y) = \sum_{i=1}^{n^3} a_i(X)b_i(X)c_i$. However *Strassen* showed that such algorithm is not optimal [5]. In this paper we are concerned with lower bounds on the rank of matrix multiplication. The first lower bound $\frac{3}{2}n^2$ was proved by *Strassen* [6] and then improved by *Bläser* [1], who found the lower bound $\frac{5}{2}n^2 - 3n$.

Recently *Landsberg* [2], building on work with *Ottaviani* [4], found the new lower bound $3n^2 - 4n^{\frac{3}{2}} - n$. The core of Landsberg's argument is the proof of the Key Lemma [2, Lemma 4.3]. In this paper we improve the Key Lemma and obtain new lower bounds for matrix multiplication.

Our main result is the following.

Theorem 0.1. *Let $p \leqslant \frac{n}{2}$ be a natural number. Then*

$$\mathrm{rk}(M_{n,n,m}) \geqslant \left(1 + \frac{p}{p+1}\right)nm + n^2 - (2p+3)n. \tag{0.1}$$

For example, when $\sqrt{\frac{n}{2}} \in \mathbb{Z}$, taking $p = \sqrt{\frac{n}{2}} - 1$, we get

$$\mathrm{rk}(M_{n,n,m}) \geqslant 2nm + n^2 - 2\sqrt{2}nm^{\frac{1}{2}} - n.$$

When $n = m$ we obtain

$$\mathrm{rk}(M_{n,n,n}) \geqslant \left(3 - \frac{1}{p+1}\right)n^2 - (2p+3)n. \tag{0.2}$$

This bound is maximized when $p = \left\lceil \sqrt{\frac{n}{2}} - 1 \right\rceil$ or $p = \left\lfloor \sqrt{\frac{n}{2}} - 1 \right\rfloor$, hence when $\sqrt{\frac{n}{2}} \in \mathbb{Z}$ we have

$$\mathrm{rk}(M_{n,n,n}) \geqslant 3n^2 - 2\sqrt{2}n^{\frac{3}{2}} - n.$$

In general we have the following bound

$$\mathrm{rk}(M_{n,n,n}) \geqslant 3n^2 - 2\sqrt{2}n^{\frac{3}{2}} - 3n. \tag{0.3}$$

The bound (0.3) improves Bläser's one, $\frac{5}{2}n^2 - 3n$, for $n \geqslant 32$. Nevertheless, when $p = 2$, the bound in (0.2) becomes $\frac{8}{3}n^2 - 7n$, which improves Bläser's one for every $n \geqslant 24$. Compared with Landsberg's bound $3n^2 - 4n^{\frac{3}{2}} - n$, our bound (0.3) is better for $n \geqslant 3$. More generally, our bound (0.2) improves Landsberg's one $\left(3 - \frac{1}{p+1}\right)n^2 - (2p+3)n$, for every $p \geqslant 1$.

Our strategy is the following. We prove Lemma 3.2, which is the improved version of [2, Lemma 4.3], using the classical identities for determinants of Lemma 1.1 and Lemma 1.2, to lower the degree of the equations that give the lower bound for border rank for matrix multiplication. Then we exploit this lower degree as Bläser and Landsberg did.

The paper is organized as follows. In Section 1 we give the basic definitions and explain the geometric meanings of the notions of rank and border rank in terms of secant varieties of Segre varieties. Section 2 is devoted to the Landsberg–Ottaviani equations [4]; we present them as rephrased in [2]. Finally in Section 3 we improve the Key Lemma [2, Lemma 4.3] and prove Theorem 0.1.

1. Preliminaries and notation

Let V, W be two complex vector spaces of dimension n and m. The contraction morphism

$$\begin{array}{rcc} V^*\otimes W & \rightarrow & \mathrm{Hom}(V,W) \\ T=\sum_{i,j} f_i\otimes w_j & \mapsto & L_T \end{array},$$

where $L_T(v)=\sum_{i,j} f_i(v)w_j$, defines an isomorphism between $V^*\otimes W$ and the space of linear maps from V to W.

Then, given three vector spaces A, B, C of dimension a, b and c, we can identify $A^*\otimes B$ with the space of linear maps $A\rightarrow B$, and $A^*\otimes B^*\otimes C$ with the space of bilinear maps $A\times B\rightarrow C$. Let $T : A^*\times B^*\rightarrow C$ be a bilinear map. Then T induces a linear map $A^*\otimes B^*\rightarrow C$ and may also be interpreted as:

- an element of $(A^*\otimes B^*)^*\otimes C = A\otimes B\otimes C$,
- a linear map $A^*\rightarrow B\otimes C$.

Segre varieties and their secant varieties. Let A, B and C be complex vector spaces. The three factor Segre map is defined as

$$\begin{array}{rcl} \sigma_{1,1,1} : \mathbb{P}(A)\times\mathbb{P}(B)\times\mathbb{P}(C) & \rightarrow & \mathbb{P}(A\otimes B\otimes C) \\ ([a],[b],[c]) & \mapsto & [a\otimes b\otimes c], \end{array}$$

where $[a]$ denotes the class in $\mathbb{P}(A)$ of the vector $a\in A$. The notation $\sigma_{1,1,1}$ is justified by the fact that the Segre map is induced by the line bundle $\mathcal{O}(1,1,1)$ on $\mathbb{P}(A)\times\mathbb{P}(B)\times\mathbb{P}(C)$. The two factor Segre map

$$\sigma_{1,1} : \mathbb{P}(B)\times\mathbb{P}(C)\rightarrow\mathbb{P}(B\otimes C)$$

is defined in a similar way. The Segre varieties are defined as the images of the Segre maps: $\Sigma_{1,1,1}=\sigma_{1,1,1}(\mathbb{P}(A)\times\mathbb{P}(B)\times\mathbb{P}(C))$, $\Sigma_{1,1}=\sigma_{1,1}(\mathbb{P}(B)\times\mathbb{P}(C))$. For each integer $r\geqslant 0$ we define the open secant variety and the secant variety of $\Sigma_{1,1,1}$ respectively as

$$\mathbb{S}ec_r(\Sigma_{1,1,1})^o=\bigcup_{x_1,\ldots,x_{r+1}\in\,\Sigma_{1,1,1}}\langle x_1,\ldots,x_{r+1}\rangle,\quad \mathbb{S}ec_r(\Sigma_{1,1,1})=\overline{\mathbb{S}ec_r(\Sigma_{1,1,1})^o}.$$

In the above formulas $\langle x_1,\ldots,x_{r+1}\rangle$ denotes the linear space generated by the points x_i and $\mathbb{S}ec_r(\Sigma_{1,1,1})$ is the closure of $\mathbb{S}ec_r(\Sigma_{1,1,1})^o$ with respect to the Zariski topology. Let us notice that with the above definition $\mathbb{S}ec_0(\Sigma_{1,1,1})=\Sigma_{1,1,1}$.

Rank and border rank of a bilinear map. The *rank* of a bilinear map $T : A^*\times B^*\rightarrow C$ is the smallest natural number $r := \mathrm{rk}(T)\in\mathbb{N}$ such that there exist $a_1,\ldots,a_r\in A$, $b_1,\ldots,b_r\in B$ and $c_1,\ldots,c_r\in C$ decomposing $T(\alpha,\beta)$ as

$$T(\alpha, \beta) = \sum_{i=1}^{r} a_i(\alpha)b_i(\beta)c_i$$

for any $\alpha \in A^*$ and $\beta \in B^*$. The number $\mathrm{rk}(T)$ has also two additional interpretations.

- Considering T as an element of $A \otimes B \otimes C$ the rank r is the smallest number of rank one tensors in $A \otimes B \otimes C$ needed to span a linear space containing the point T. Equivalently, $\mathrm{rk}(T)$ is the smallest number of points $t_1, \dots, t_r \in \Sigma_{1,1,1}$ such that $[T] \in \langle t_1, \dots, t_r \rangle$. In the language of secant varieties this means that $[T] \in \mathbb{S}ec_{r-1}(\Sigma_{1,1,1})^o$ but $[T] \notin \mathbb{S}ec_{r-2}(\Sigma_{1,1,1})^o$.
- Similarly, if we consider T as a linear map $A^* \to B \otimes C$ then $\mathrm{rk}(T)$ is the smallest number of rank one tensors in $B \otimes C$ need to span a linear space containing the linear space $T(A^*)$. As before we have a geometric counterpart. In fact $\mathrm{rk}(T)$ is the smallest number of points $t_1, \dots, t_r \in \Sigma_{1,1}$ such that $\mathbb{P}(T(A^*)) \subseteq \langle t_1, \dots, t_r \rangle$.

The *border rank* of a bilinear map $T : A^* \times B^* \to C$ is the smallest natural number $r := \underline{\mathrm{rk}}(T)$ such that T is the limit of bilinear maps of rank r but is not a limit of tensors of rank s for any $s < r$. There is a geometric interpretation also for this notion: T has border rank r if $[T] \in \mathbb{S}ec_{r-1}(\Sigma_{1,1,1})$ but $[T] \notin \mathbb{S}ec_{r-2}(\Sigma_{1,1,1})$. Clearly $\mathrm{rk}(T) \geqslant \underline{\mathrm{rk}}(T)$.

Matrix multiplication. Now, let us consider a special tensor. Given three vector spaces $L = \mathbb{C}^l$, $M = \mathbb{C}^m$ and $N = \mathbb{C}^n$ we define $A = N \otimes L^*$, $B = L \otimes M^*$ and $C = N^* \otimes M$. We have a matrix multiplication map

$$M_{n,l,m} : A^* \times B^* \to C$$

As a tensor $M_{n,l,m} = Id_N \otimes Id_M \otimes Id_L \in (N^* \otimes L) \otimes (L \otimes M^*) \otimes (N^* \otimes M) = A \otimes B \otimes C$, where $Id_N \in N^* \otimes N$ is the identity map. If $n = l$ the choice of a linear map $\alpha^0 : N \to L$ of maximal rank allows us to identify $N \cong L$. Then the multiplication map $M_{n,n,m} \in (N \otimes N^*) \otimes (N \otimes M^*) \otimes (N^* \otimes M)$ induces a linear map $N^* \otimes N \to (N^* \otimes M) \otimes (N^* \otimes M)^*$ which is an inclusion of Lie algebras

$$M_A : \mathfrak{gl}(N) \to \mathfrak{gl}(B),$$

where $\mathfrak{gl}(N) \cong N^* \otimes N$ is the algebra of linear endomorphisms of N. In particular, the rank of the commutator $[M_A(\alpha^1), M_A(\alpha^2)]$ of $nm \times nm$ matrices is equal to m times the rank of the commutator $[\alpha^1, \alpha^2]$ of $n \times n$ matrices. This equality reflects a general philosophy, that is to translate expressions in commutators of $\mathfrak{gl}_{n^2}$ into expressions in commutators in $\mathfrak{gl}_n$.

Matrix equalities. The following lemmas are classical in linear algebra. However, for completeness, we give a proof.

Lemma 1.1. *The determinant of a* 2×2 *block matrix is given by*

$$\det \begin{pmatrix} X & Y \\ Z & W \end{pmatrix} = \det(X)\det(W - ZX^{-1}Y),$$

where X is an invertible $n \times n$ matrix, Y is a $n \times m$ matrix, Z is a $m \times n$ matrix, and W is a $m \times m$ matrix.

Proof. The statement follows from the equality

$$\begin{pmatrix} X & Y \\ Z & W \end{pmatrix} \begin{pmatrix} -X^{-1}Y & Id_n \\ Id_m & 0 \end{pmatrix} = \begin{pmatrix} 0 & X \\ W - ZX^{-1}Y & Z \end{pmatrix}. \quad \square$$

Lemma 1.2. *Let A be an $n \times n$ invertible matrix and U, V any $n \times m$ matrices. Then*

$$\det_{n\times n}(A + UV^t) = \det_{n\times n}(A) \det_{m\times m}(Id + V^t A^{-1} U),$$

where V^t is the transpose of V.

Proof. It follows from the equality

$$\begin{pmatrix} A & 0 \\ V^t & Id \end{pmatrix} \begin{pmatrix} Id & -A^{-1}U \\ 0 & Id + V^t A^{-1} U \end{pmatrix} \begin{pmatrix} Id & 0 \\ -V^t & Id \end{pmatrix} = \begin{pmatrix} A + UV^t & -U \\ 0 & Id \end{pmatrix}. \quad \square$$

2. Landsberg–Ottaviani equations

In [4] *Landsberg* and *Ottaviani* generalized Strassen's equations as introduced by *Strassen* in [6]. We follow the exposition of [2, Section 2].

Let $T \in A \otimes B \otimes C$ be a tensor, and assume $b = c$. Let us consider T as a linear map $A^* \to B \otimes C$, and assume that there exists $\alpha \in A^*$ such that $T(\alpha) : B^* \to C$ is of maximal rank b. Via $T(\alpha)$ we can identify $B \cong C$, and consider $T(A^*) \subseteq B^* \otimes B$ as a subspace of the space of linear endomorphisms of B.

In [6] *Strassen* considered the case $a = 3$. Let $\alpha^0, \alpha^1, \alpha^2$ be a basis of A^*. Assume that $T(\alpha^0)$ has maximal rank and that $T(\alpha^1), T(\alpha^2)$ are diagonalizable, commuting endomorphisms. Then $T(\alpha^1), T(\alpha^2)$ are simultaneously diagonalizable and it is not difficult to prove that in this case $\mathrm{rk}(T) = b$. In general, $T(\alpha^1), T(\alpha^2)$ are not commuting. The idea of Strassen was to consider their commutator $[T(\alpha^1), T(\alpha^2)]$ to obtain results on the border rank of T. In fact, Strassen proved that, if $T(\alpha^0)$ is of maximal rank, then $\underline{\mathrm{rk}}(T) \geqslant b + \mathrm{rank}[T(\alpha^1), T(\alpha^2)]/2$ and $\underline{\mathrm{rk}}(T) = b$ if and only if $[T(\alpha^1), T(\alpha^2)] = 0$.

Now let us consider the case $a = 3, b = c$. Fix a basis a_0, a_1, a_2 of a A, and let a^0, a^1, a^2 be the dual basis of A^*. Choose bases of B and C, so that elements of $B \otimes C$ can be written as matrices. Then we can write $T = a_0 \otimes X_0 - a_1 \otimes X_1 + a_2 \otimes X_2$, where the X_i are $b \times b$ matrices. Consider $T \otimes Id_A \in A \otimes B \otimes C \otimes A^* \otimes A = A^* \otimes B \otimes A \otimes A \otimes C$,

$$T \otimes Id_A = (a_0 \otimes X_0 - a_1 \otimes X_1 + a_2 \otimes X_2) \otimes (a^0 \otimes a_0 + a^1 \otimes a_1 + a^2 \otimes a_2)$$

and its skew-symmetrization in the A factor $T_A^1 \in A^* \otimes B \otimes \bigwedge^2 A \otimes C$, given by

$$\begin{aligned} T_A^1 = {} & a^1 X_0 (a_0 \wedge a_1) + a^2 X_0 (a_0 \wedge a_2) - a^0 X_1 (a_1 \wedge a_0) - a^2 X_1 (a_1 \wedge a_2) \\ & + a^0 X_2 (a_2 \wedge a_0) + a^1 X_2 (a_2 \wedge a_1) \end{aligned}$$

where $a^i X_j (a_j \wedge a_i) := a^i \otimes X_j \otimes (a_j \wedge a_i)$. It can also be considered as a linear map

$$T_A^1 : A \otimes B^* \to \bigwedge^2 A \otimes C.$$

In the basis a_0, a_1, a_2 of A and $a_0 \wedge a_1, a_0 \wedge a_2, a_1 \wedge a_2$ of $\bigwedge^2 A$ the matrix of T_A^1 is the following

$$Mat(T_A^1) = \begin{pmatrix} X_1 & -X_2 & 0 \\ X_0 & 0 & -X_2 \\ 0 & X_0 & -X_1 \end{pmatrix}$$

Assume X_0 is invertible and change bases such that it is the identity matrix. By Lemma 1.1, on the matrix obtained by reversing the order of the rows of $Mat(T_A^1)$, with

$$X = \begin{pmatrix} 0 & X_0 \\ X_0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} -X_1 \\ -X_2 \end{pmatrix}, \quad Z = \begin{pmatrix} X_1 & -X_2 \end{pmatrix}, \quad W = 0$$

we get

$$\det(Mat(T_A^1)) = \det(X_1X_2 - X_2X_1) = \det([X_1, X_2]).$$

Now we want to generalize this construction as done in [4]. We consider the case $a = 2p + 1$, $T \otimes Id_{\bigwedge^p A} \in A \otimes B \otimes C \otimes \bigwedge^p A^* \otimes \bigwedge^p A = (\bigwedge^p A^* \otimes B) \otimes (\bigwedge^{p+1} A \otimes C)$, and its skew-symmetrization

$$T_A^p : \bigwedge^p A \otimes B^* \to \bigwedge^{p+1} A \otimes C.$$

Note that $\dim(\bigwedge^p A \otimes B^*) = \dim(\bigwedge^{p+1} A \otimes C) = \binom{2p+1}{p} b$. After choosing a basis $a_0, \dots, a_{2p}$ of A we can write $T = \sum_{i=0}^{2p} (-1)^i a_i \otimes X_i$. The matrix of T_A^p with respect the basis $a_0 \wedge \dots \wedge a_{p-1}, \dots, a_{p+1} \wedge \dots \wedge a_{2p}$ of $\bigwedge^p A$, and $a_0 \wedge \dots \wedge a_p, \dots, a_p \wedge \dots \wedge a_{2p}$ of $\bigwedge^{p+1} A$ is of the form

$$Mat(T_A^p) = \begin{pmatrix} Q & 0 \\ R & \overline{Q} \end{pmatrix} \tag{2.1}$$

where the matrix is blocked $(\binom{2p}{p+1} b, \binom{2p}{p} b) \times (\binom{2p}{p+1} b, \binom{2p}{p} b)$, the lower left block is given by

$$R = \begin{pmatrix} X_0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & X_0 \end{pmatrix}$$

and Q is a matrix having blocks $X_1, \dots, X_{2p}$ and zero, while $\overline{Q}$ is the block transpose of Q except that if an index is even, the block is multiplied by -1. We derive below the expression (2.1) in the case $p = 2$; the general case can be developed similarly, see [2, Section 3].

Example 2.1. Consider the case $p = 2$. The matrix of T_A^2 is

$$Mat(T_A^2) = \begin{pmatrix} X_2 & -X_3 & X_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X_1 & 0 & 0 & -X_3 & X_4 & 0 & 0 & 0 & 0 & 0 \\ 0 & X_1 & 0 & -X_2 & 0 & X_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & X_1 & 0 & -X_2 & X_3 & 0 & 0 & 0 & 0 \\ X_0 & 0 & 0 & 0 & 0 & 0 & -X_3 & X_4 & 0 & 0 \\ 0 & X_0 & 0 & 0 & 0 & 0 & -X_2 & 0 & X_4 & 0 \\ 0 & 0 & X_0 & 0 & 0 & 0 & 0 & -X_2 & X_3 & 0 \\ 0 & 0 & 0 & X_0 & 0 & 0 & -X_1 & 0 & 0 & X_4 \\ 0 & 0 & 0 & 0 & X_0 & 0 & 0 & -X_1 & 0 & X_3 \\ 0 & 0 & 0 & 0 & 0 & X_0 & 0 & 0 & -X_1 & X_2 \end{pmatrix}$$

If X_0 is the identity by Lemma 1.1 on $R = Id, Q$ and $\overline{Q}$ the determinant of $Mat(T_A^p)$ is equal to the determinant of

$$\begin{pmatrix} 0 & [X_1, X_2] & [X_1, X_3] & [X_1, X_4] \\ -[X_1, X_2] & 0 & [X_2, X_3] & [X_2, X_4] \\ -[X_1, X_3] & -[X_2, X_3] & 0 & [X_3, X_4] \\ -[X_1, X_4] & -[X_2, X_4] & -[X_3, X_4] & 0 \end{pmatrix}$$

In general the determinant of $Mat(T_A^p)$ is equal to the determinant of the $2pb \times 2pb$ matrix of commutators

$$\begin{pmatrix} 0 & X_{1,2} & X_{1,3} & X_{1,4} & \dots & X_{1,2p-1} & X_{1,2p} \\ -X_{1,2} & 0 & X_{2,3} & X_{2,4} & \dots & X_{2,2p-1} & X_{2,2p} \\ -X_{1,3} & -X_{2,3} & 0 & X_{3,4} & \dots & X_{3,2p-1} & X_{3,2p} \\ -X_{1,4} & -X_{2,4} & -X_{3,4} & 0 & \dots & X_{4,2p-1} & X_{4,2p} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -X_{1,2p-1} & -X_{2,2p-1} & -X_{3,2p-1} & -X_{4,2p-1} & \dots & 0 & X_{2p-1,2p} \\ -X_{1,2p} & -X_{2,2p} & -X_{3,2p} & -X_{4,2p} & \dots & -X_{2p-1,2p} & 0 \end{pmatrix}$$

where $X_{i,j}$ denotes the commutator matrix $[X_i, X_j] = X_iX_j - X_jX_i$.

3. Key Lemma

We use the same notation of [2] throughout the text.

Lemma 3.1 [3, Lemma 11.5.0.2]. *Let V be a n-dimensional vector space and let $P \in S^dV^* \setminus \{0\}$ be a polynomial of degree $d \leqslant n-1$ on V. For any basis $\{v_1, \dots, v_n\}$ of V there exists a subset $\{v_{i_1}, \dots, v_{i_s}\}$ of cardinality $s \leqslant d$ such that $P_{|\langle v_{i_1}, \dots, v_{i_s} \rangle}$ is not identically zero.*

Lemma 3.1 says, for instance, that a quadric surface in $\mathbb{P}^3$ can not contain six lines whose pairwise intersections span $\mathbb{P}^3$. Note that as stated Lemma 3.1 is sharp in the sense that under the same hypothesis the bound $s \leqslant d$ can not be improved. For example the polynomial $P(x, y, z, w) = xy$ vanishes on the four points $[1:0:0:0], \dots, [0:0:0:1] \in \mathbb{P}^3$.

Lemma 3.2. *Let $A = N^* \otimes L$, where $l = n$. Given any basis of A, there exists a subset of at least $n^2 - (2p+3)n$ basis vectors, and elements $\alpha^0, \alpha^1, \dots, \alpha^{2p}$ of A^*, such that*

- *α^0 is of maximal rank, and thus may be used to identify $L \simeq N$ and A as a space of endomorphisms. (I.e. in bases α^0 is the identity matrix.)*
- *Choosing a basis of L, so the α^j become $n \times n$ matrices, the size $2pn$ block matrix whose (i, j)-th block is $[\alpha^i, \alpha^j]$ has non-zero determinant.*
- *The subset of $n^2 - (2p+3)n$ basis vectors annihilate $\alpha^0, \alpha^1, \dots, \alpha^{2p}$.*

Proof. Let $\mathcal{B}$ be a basis of A, and consider the polynomial $P_0 = \det_n$. By Lemma 3.1 we get a subset S_0 of at most n elements of $\mathcal{B}$ and $\alpha^0 \in S_0$ with $\det_n(\alpha^0) \neq 0$. Now, via the isomorphism $\alpha^0 : L \to N$ we are allowed to identify $A = \mathfrak{gl}(L)$ as an algebra with identity element α^0. So, from now on, we work with $\mathfrak{sl}(L) = \mathfrak{gl}(L)/\left\langle \alpha^0 \right\rangle$ instead of $\mathfrak{gl}(L)$.

Following the proof of [2, Lemma 4.3], let $v_{1,0}, \dots, v_{2p,0} \in \mathfrak{sl}(L)$ be linearly independent and not equal to any of the given basis vectors, and let us work locally on an affine open neighbourhood $\mathbb{V} \subset G(2p, \mathfrak{sl}(L))$ of $E_0 = \langle v_{1,0}, \dots, v_{2p,0}\rangle$. We extend $v_{1,0}, \dots, v_{2p,0}$ to a basis $v_{1,0}, \dots, v_{2p,0}, w_1, \dots, w_{n^2-2p-1}$ of $\mathfrak{sl}(L)$, and take local coordinates (f_s^{μ}) with $1 \leqslant s \leqslant 2p$, $1 \leqslant \mu \leqslant n^2 - 2p - 1$, on V, so that $v_s = v_{s,0} + \sum_{\mu=1}^{n^2-2p-1} f_s^{\mu} w_{\mu}$.

We denote $v_{i,j} = [v_i, v_j]$ and let us define

$$A_{i,i+1} = \begin{pmatrix} 0 & v_{i,i+1} \\ -v_{i,i+1} & 0 \end{pmatrix}$$

for $i = 1, \dots, 2p$ and let A be the following diagonal block matrix

$$A = \mathrm{diag}(A_{1,2}, A_{3,4}, \dots, A_{2p-3,2p-2}, Id_{2n\times 2n})$$

which is a squared matrix of order $4pn$. Consider the $4pn \times 4pn$ matrix

$$M = \begin{pmatrix} 0 & v_{1,2} & v_{1,3} & v_{1,4} & \dots & v_{1,2p-1} & v_{1,2p} \\ -v_{1,2} & 0 & v_{2,3} & v_{2,4} & \dots & v_{2,2p-1} & v_{2,2p} \\ -v_{1,3} & -v_{2,3} & 0 & v_{3,4} & \dots & v_{3,2p-1} & v_{3,2p} \\ -v_{1,4} & -v_{2,4} & -v_{3,4} & 0 & \dots & v_{4,2p-1} & v_{4,2p} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -v_{1,2p-1} & -v_{2,2p-1} & -v_{3,2p-1} & -v_{4,2p-1} & \dots & 0 & v_{2p-1,2p} \\ -v_{1,2p} & -v_{2,2p} & -v_{3,2p} & -v_{4,2p} & \dots & -v_{2p-1,2p} & 0 \end{pmatrix}$$

The polynomial $\det_{4pn\times 4pn}(M)$ is not identically zero on $G(2p, \mathfrak{sl}(L))$, so it is not identically zero on $\mathbb{V}$. Furthermore we can write $M = A + UId_{4pn\times 4pn}$, where

$$U = \begin{pmatrix} 0 & 0 & v_{1,3} & v_{1,4} & \dots & v_{1,2p-1} & v_{1,2p} \\ 0 & 0 & v_{2,3} & v_{2,4} & \dots & v_{2,2p-1} & v_{2,2p} \\ -v_{1,3} & -v_{2,3} & 0 & 0 & \dots & v_{3,2p-1} & v_{3,2p} \\ -v_{1,4} & -v_{2,4} & 0 & 0 & \dots & v_{4,2p-1} & v_{4,2p} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -v_{1,2p-1} & -v_{2,2p-1} & -v_{3,2p-1} & -v_{4,2p-1} & \dots & -Id_{n\times n} & v_{2p-1,2p} \\ -v_{1,2p} & -v_{2,2p} & -v_{3,2p} & -v_{4,2p} & \dots & -v_{2p-1,2p} & -Id_{n\times n} \end{pmatrix}$$

By Lemma 1.2 we have

$$\det(M) = \det(A)\det(Id + A^{-1}U) = \det([v_1, v_2])^2 \dots \det([v_{2p-3}, v_{2p-2}])^2 \det(Id + A^{-1}U).$$

The entries of the $n \times n$ matrices $[v_k, v_{k+1}]$ are quadratic in the f_s^{μ}'s, so the polynomials $\det([v_k, v_{k+1}])$ have degree $2n$, and

$$P_1 = \det([v_1, v_2])^2 \dots \det([v_{2p-3}, v_{2p-2}])^2 = (\det([v_1, v_2]) \dots \det([v_{2p-3}, v_{2p-2}]))^2$$

is a polynomial of degree $4n(p-1)$. Since P_1 is a square, we can consider the polynomial $\widetilde{P}_1 = \det([v_1, v_2]) \dots \det([v_{2p-3}, v_{2p-2}])$ which has degree $2n(p-1)$. Applying Lemma 3.1 to $\widetilde{P}_1$ we find a subset S_1 of at most $2n(p-1)$ elements of our basis such that $\widetilde{P}_1$, and hence P_1, is not identically zero on $\langle S_1 \rangle$.

Now, let us fix some particular value of the coordinates f_s^{μ} such that on the corresponding matrices $\overline{v}_1, \dots, \overline{v}_{2p-2}$ the matrix A is invertible. For these values the expression $\det(Id + A^{-1}U)$ makes sense. Let us consider the matrix

$$Id + A^{-1}U = \begin{pmatrix} Id & 0 & -v_{1,2}^{-1}v_{2,3} & -v_{1,2}^{-1}v_{2,4} & \dots & -v_{1,2}^{-1}v_{2,2p-1} & -v_{1,2}^{-1}v_{2,2p} \\ 0 & Id & v_{1,2}^{-1}v_{1,3} & v_{1,2}^{-1}v_{1,4} & \dots & v_{1,2}^{-1}v_{1,2p-1} & v_{1,2}^{-1}v_{1,2p} \\ v_{3,4}^{-1}v_{1,4} & v_{3,4}^{-1}v_{2,4} & Id & 0 & \dots & -v_{3,4}^{-1}v_{4,2p-1} & -v_{3,4}^{-1}v_{4,2p} \\ -v_{3,4}^{-1}v_{1,3} & -v_{3,4}^{-1}v_{2,3} & 0 & Id & \dots & v_{3,4}^{-1}v_{3,2p-1} & v_{3,4}^{-1}v_{3,2p} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -v_{1,2p-1} & -v_{2,2p-1} & -v_{3,2p-1} & -v_{4,2p-1} & \dots & 0 & v_{2p-1,2p} \\ -v_{1,2p} & -v_{2,2p} & -v_{3,2p} & -v_{4,2p} & \dots & -v_{2p-1,2p} & 0 \end{pmatrix}$$

By Lemma 1.1 on $Id + A^{-1}U$ with

$$X = \begin{pmatrix} Id & 0 \\ 0 & Id \end{pmatrix}, \quad Y = \begin{pmatrix} -v_{1,2}^{-1}v_{2,3} & -v_{1,2}^{-1}v_{2,4} & \dots & -v_{1,2}^{-1}v_{2,2p-1} & -v_{1,2}^{-1}v_{2,2p} \\ v_{1,2}^{-1}v_{1,3} & v_{1,2}^{-1}v_{1,4} & \dots & v_{1,2}^{-1}v_{1,2p-1} & v_{1,2}^{-1}v_{1,2p} \end{pmatrix},$$

$$Z = \begin{pmatrix} v_{3,4}^{-1}v_{1,4} & v_{3,4}^{-1}v_{2,4} \\ -v_{3,4}^{-1}v_{1,3} & -v_{3,4}^{-1}v_{2,3} \\ \vdots & \vdots \\ -v_{1,2p-1} & -v_{2,2p-1} \\ -v_{1,2p} & -v_{2,2p} \end{pmatrix}, \quad W = \begin{pmatrix} Id & 0 & \dots & -v_{3,4}^{-1}v_{4,2p-1} & -v_{3,4}^{-1}v_{4,2p} \\ 0 & Id & \dots & v_{3,4}^{-1}v_{3,2p-1} & v_{3,4}^{-1}v_{3,2p} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -v_{3,2p-1} & -v_{4,2p-1} & \dots & 0 & v_{2p-1,2p} \\ -v_{3,2p} & -v_{4,2p} & \dots & -v_{2p-1,2p} & 0 \end{pmatrix}$$

we get $\det(Id + A^{-1}U) = \det(W - ZY)$. Note that the coordinates f_s^{μ} appear in the terms indexed by $2p-1$ and $2p$, while all the other terms are constant once we fixed $\overline{v}_1, \dots, \overline{v}_{2p-2}$. Then $P_2 = \det(W - ZY)$ is a polynomial of degree $4n$. By Lemma 3.1 we find a subset S_2 of at most $4n$ elements of the basis $\mathcal{B}$ such that P_2 is not identically zero on $\langle S_2 \rangle$.

Summing up we found a subset S of at most $n + 2n(p-1) + 4n = (2p+3)n$ elements of $\mathcal{B}$ such that $\det(M)$ is not identically zero on $\langle S \rangle$. □

Remark 3.3. In [2, Lemma 4.3] the author proved the analogous statement for $n^2 - (4p+1)n$.

Proof of Theorem 0.1. The proof of the bound (0.1) is the same of [2, Theorem 1.2]; the only difference is that one uses Lemma 3.2 instead of [2, Lemma 4.3].

To prove the other assertions, let us consider the function $f : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$ defined by $f(p) = (3 - \frac{1}{p+1})n^2 - (2p+3)n$. The first derivative is $f'(p) = \frac{1}{(p+1)^2}n^2 - 2n$, which vanishes in $p = \sqrt{\frac{n}{2}} - 1$. Moreover $f''(p) = -\frac{2}{(p+1)^3}n^2 < 0$, hence $p = \sqrt{\frac{n}{2}} - 1$ is the maximum of f.

Then the bound (0.2) is maximized for $p = \left\lceil \sqrt{\frac{n}{2}} - 1 \right\rceil$ or $p = \left\lfloor \sqrt{\frac{n}{2}} - 1 \right\rfloor$, depending on the value of n.

If $\left(\sqrt{\frac{n}{2}}-1\right)-\left\lfloor\sqrt{\frac{n}{2}}-1\right\rfloor \geqslant \frac{1}{2}$ we may consider $p=\left\lceil\sqrt{\frac{n}{2}}-1\right\rceil$. In this case $\sqrt{\frac{n}{2}}-1 \leqslant p \leqslant \sqrt{\frac{n}{2}}-\frac{1}{2}$, and we get $f(\left\lceil\sqrt{\frac{n}{2}}-1\right\rceil) \geqslant \lceil f\rceil(n) := 3n^2-2\sqrt{2}n^{\frac{3}{2}}-2n$.

If $\left(\sqrt{\frac{n}{2}}-1\right)-\left\lfloor\sqrt{\frac{n}{2}}-1\right\rfloor < \frac{1}{2}$ we consider $p=\left\lfloor\sqrt{\frac{n}{2}}-1\right\rfloor$. Then $\sqrt{\frac{n}{2}}-\frac{3}{2} \leqslant p \leqslant \sqrt{\frac{n}{2}}-1$, and we have $f(\left\lfloor\sqrt{\frac{n}{2}}-1\right\rfloor) \geqslant \lfloor f\rfloor(n) := (3-\frac{2\sqrt{2}}{2n-\sqrt{2}})n^2-\sqrt{2}n^{\frac{3}{2}}-n$.

Finally to prove (0.3) it is enough to observe that both $\lceil f\rceil(n)$ and $\lfloor f\rfloor(n)$ are greater than $3n^2-2\sqrt{2}n^{\frac{3}{2}}-3n$. □

Acknowledgements

Both the authors were introduced to this topic by *J.M. Landsberg* during the Summer School "*Tensors: Waring problems and Geometric Complexity Theory*" held in Cortona in July 2012. We would like to thank *J.M. Landsberg* and *M. Mella* for their beautiful lectures and all the participants for the stimulating atmosphere. We thank primarily *J.M. Landsberg* for suggesting us the problem, for his interest and his generous suggestions.

References

[1] M. Bläser, A $\frac{5}{2}n^2$ lower bound for the rank of $n \times n$-matrix multiplication over arbitrary fields, in: 440th Annual Symposium on Foundations of Computer Science (New York, 1999), IEEE Computer Soc., Los Alamitos, CA, 1999, pp. 45–50 (MR MR1916183).
[2] J.M. Landsberg, New lower bounds for the rank of matrix multiplication, arXiv:1206.1530.
[3] J.M. Landsberg, Tensors: Geometry and Applications, Graduate Studies in Mathematics, vol. 12, American Mathematical Society, Providence, RI, 2012 (MR 2865915).
[4] J.M. Landsberg, G. Ottaviani, New lower bounds for the border rank of matrix multiplication, arXiv:1112.6007.
[5] V. Strassen, Gaussian elimination is not optimal, Numer. Math. 13 (1969) 354–356.
[6] V. Strassen, Rank and optimal computation of generic tensors, Linear Algebra Appl. 52/53 (1983) 645–685 (MR 85b:15039).

Update

Linear Algebra and Its Applications

Volume 445, Issue , 15 March 2014, Page 369–371

DOI: https://doi.org/10.1016/j.laa.2013.12.009

Linear Algebra and its Applications 445 (2014) 369–371

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Corrigendum

Corrigendum to “The rank of $n \times n$ matrix multiplication is at least $3n^2 - 2\sqrt{2}n^{\frac{3}{2}} - 3n$” [Linear Algebra Appl. 438 (11) (2013) 4500–4509]

Alex Massarenti[a], Emanuele Raviolo[b,*]

[a] *SISSA, via Bonomea 265, 34136 Trieste, Italy*
[b] *Università di Pavia, via Ferrata 1, 27100 Pavia, Italy*

ARTICLE INFO

Article history:
Received 6 November 2013
Accepted 5 December 2013
Available online 23 December 2013
Submitted by S. Friedland

MSC:
primary 14Q20
secondary 13P99, 68W30

Keywords:
Tensors
Matrix multiplication
Complexity theory

The main result in [3] is incorrect, because of a mistake in the first version of [2], which was propagated in our paper. For the full and correct version of our paper see [4]. The correct statement is the following.

DOI of original article: http://dx.doi.org/10.1016/j.laa.2013.01.031.

[*] Corresponding author.
E-mail addresses: alex.massarenti@sissa.it (A. Massarenti), emanuele.raviolo@unipv.it (E. Raviolo).

0024-3795/$ – see front matter
http://dx.doi.org/10.1016/j.laa.2013.12.009

Theorem 0.1. *(See [4, Theorem 0.1].) Let $p \leqslant n$ be a positive natural number. Then*

$$\operatorname{rk}(M_{n,n,m}) \geqslant \left(1+\frac{p}{p+1}\right)nm+n^2-\left(2\binom{2p}{p+1}-\binom{2p-2}{p-1}+2\right)n. \tag{0.1}$$

When $n = m$ we obtain

$$\operatorname{rk}(M_{n,n,n}) \geqslant \left(3-\frac{1}{p+1}\right)n^2-\left(2\binom{2p}{p+1}-\binom{2p-2}{p-1}+2\right)n. \tag{0.2}$$

When $p = 2$ we can improve the bound to

$$\operatorname{rk}(M_{n,n,n}) \geqslant \frac{8}{3}n^2-7n, \tag{0.3}$$

and when $p = 3$ to

$$\operatorname{rk}(M_{n,n,n}) \geqslant \frac{11}{4}n^2-17n. \tag{0.4}$$

The bound $\frac{8}{3}n^2-7n$ improves all previous bounds for all $n \geqslant 24$ (for $n \leqslant 84$ the best bound is in [1] and for $n > 84$ it is in [2]), while the bound $\frac{11}{4}n^2-17n$ improves $\frac{8}{3}n^2-7n$ for every $n \geqslant 120$.

The mistake in [3] was the calculation of the matrix associated to T_A^p. We find more convenient to consider the operator $(T_A^p)^*$, the transpose of T_A^p. The matrix associated to $(T_A^p)^*$ with respect to the basis $a_0 \wedge \cdots \wedge a_{p-1}, \dots, a_{p+1} \wedge \cdots \wedge a_{2p}$ of $\bigwedge^p A$, and $a_0 \wedge \cdots \wedge a_p, \dots, a_p \wedge \cdots \wedge a_{2p}$ of $\bigwedge^{p+1} A$ is of the form

$$Mat\big((T_A^p)^*\big) = \begin{pmatrix} Q & 0 \\ R & \overline{Q} \end{pmatrix} \tag{0.5}$$

where the matrix is blocked $\big(\binom{2p}{p+1}b, \binom{2p}{p}b\big) \times \big(\binom{2p}{p+1}b, \binom{2p}{p}b\big)$, the lower left block is given by

$$R = \begin{pmatrix} X_0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & X_0 \end{pmatrix}$$

and Q is a matrix having blocks $X_1, \dots, X_{2p}$ and zero.

The matrix is related to Q in the following way. Write $Q = (Q_{i,j})$, where the $Q_{i,j}$ are the $n \times n$ blocks of Q and let $Q_{(k)} = (Q_{k,1}, \dots, Q_{k,\binom{2p}{p}})$ be the k-th block-row of Q. Then $\overline{Q}$ is the matrix whose l-th block-column is $Q^{(l)} = (Q_{\binom{2p}{p},\binom{2p}{p+1}-l+1}, \dots, Q_{1,\binom{2p}{p+1}-l+1})$, with the convention that if $Q_{i,j} = X_h$, h odd, then the block is multiplied by -1. In general the matrix Q is as follows. Let us consider the entry (i, j) of Q corresponding to the basis vectors $a_{i_1} \wedge \cdots \wedge a_{i_{p+1}}$ of $\bigwedge^{p+1} A$ and $a_{j_1} \wedge \cdots \wedge a_{j_p}$ of $\bigwedge^{p} A$, and let $I = \{i_1, \ldots, i_{p+1}\}$, $J = \{j_1, \ldots, j_p\}$. Then

$$Q_{i,j} = \begin{cases} (-1)^{i+j} X_k & \text{if } I, J \text{ differ by just one element } k, \\ 0 & \text{otherwise.} \end{cases} \tag{0.6}$$

Remark 0.2. It follows from (0.6) that $Q\overline{Q}$ has either commutators or zeroes as entries and a lower left block $\mathcal{X}_{1,2} = \operatorname{diag}([X_1, X_2], \ldots, [X_1, X_2])$ with $\binom{2p-2}{p-1}$ blocks on the diagonal. Furthermore on the diagonal of $Q\overline{Q}$ if there is an entry $[X_i, X_j]$ then such entry appears at least twice. Finally on the diagonal all indices except $i = 1, 2p$ appear if $p \geqslant 3$ and in the case $p = 2$ all indices appear.

The proof of Theorem 0.1 will follow from the following lemma which is the correct version of [3, Lemma 3.2].

Lemma 0.3. *(See [4, Lemma 3.2].) Let $A = N^* \otimes L$, where $l = n$. Given any basis of A, there exists a subset of at least $h = n^2 - \left(n\left(2\binom{2p}{p+1} - \binom{2p-2}{p-1} + 2\right)\right)$ basis vectors, and elements $\alpha^0, \alpha^1, \ldots, \alpha^{2p}$ of A^*, such that*

- *α^0 is of maximal rank, and thus may be used to identify $L \simeq N$ and A as a space of endomorphisms. (I.e. in bases α^0 is the identity matrix.)*
- *choosing a basis of L, so the α^j become $n \times n$ matrices, the block matrix of (0.6) whose blocks are the α^i is such that $Q\overline{Q}$ has non-zero determinant, and*
- *the subset of at least h basis vectors annihilate $\alpha^0, \alpha^1, \ldots, \alpha^{2p}$.*

Proof of Theorem 0.1. The proof of (0.1) and (0.2) are the same as in [3], provided we use Lemma 0.3 instead of [3, Lemma 3.2]. The bound (0.3) can be proved using [3, Lemma 3.2], which works only for $p = 2$. Finally the bound (0.4) is obtained by the explicit calculation of the matrix $Q\overline{Q}$ for $p = 3$. □

References

[1] M. Bläser, A $\frac{5}{2}n^2$ lower bound for the rank of $n \times n$-matrix multiplication over arbitrary fields, in: 440th Annual Symposium on Foundations of Computer Science, New York, 1999, IEEE Computer Soc., Los Alamitos, CA, 1999, pp. 45–50. MR MR1916183.

[2] J.M. Landsberg, New lower bounds for the rank of matrix multiplication, SIAM J. Comput. (2013), in press, arXiv:1206.1530v2.

[3] A. Massarenti, E. Raviolo, The rank of $n \times n$ matrix multiplication is at least $3n^2 - 2\sqrt{2}n^{\frac{3}{2}} - 3n$, Linear Algebra Appl. 438 (11) (2013) 4500–4509.

[4] A. Massarenti, E. Raviolo, On the rank of $n \times n$ matrix multiplication, arXiv:1211.6320v2.