The Landsberg-Ottaviani equations on foot (and in coordinates)

Markus Bläser, Saarland University

Draft

Chapter 1

Strassen's equation

1.1 Flattenings

Let A, B, and C be vector spaces over some field $\mathbb{F}$ of dimensions $\bar{a}$, $\bar{b}$, and $\bar{c}$. Let $T \in A \otimes B \otimes C$. Recall that $R(T) \leq r$ if there are $a_i \in A$, $b_i \in B$, and $c_i \in C$, $1 \leq i \leq r$ such that

$$T = \sum_{i=1}^{r} a_i \otimes b_i \otimes c_i. \tag{1.1.1}$$

Since $\mathsf{Hom}(B^*, A \otimes C) \cong (B^*)^* \otimes A \otimes C \cong A \otimes B \otimes C$, we can view T as a linear map $B^* \to A \otimes B$. The "usual" matrix rank $\mathrm{rk}(T)$ (with respect to this splitting of the spaces) is the minimum number s such that there are $b_i \in B$ and $x_i \in A \otimes C$ such that

$$T = \sum_{i=1}^{s} a_i \otimes x_i \tag{1.1.2}$$

By writing $a_i \otimes c_i$ in (1.1.1) as x_i, we see that

$$R(T) \geq \mathrm{rk}(T)$$

and, since the matrix rank is lower semi-continuous,

$$\underline{R}(T) \geq \mathrm{rk}(T).$$

The process is called *flattening*.

1.2 Strassen's equation

We now augment $T : B^* \to A \otimes C$ to a mapping $T_A = \mathrm{Id}_A \otimes T$ which is a mapping $A \otimes B^* \to A \otimes A \otimes C$. Recall that $A = \Lambda^2 A \oplus S^2 A$. The mapping $T_A^\wedge : A \otimes B^* \to \Lambda^2 A \otimes C$ is obtained by concatenating T_A with the natural projection of $A \otimes A \to \Lambda^2 A$.

Now assume that $\dim A = \bar{a} = 3$ and choose a basis a_1, a_2, a_3. Write

$$T = a_1 \otimes X_1 + a_2 \otimes X_2 + a_3 \otimes X_3$$

with $X_j : B^* \to C$ being a $\bar{c} \times \bar{b}$-matrix. T_A maps $a_j \otimes \beta$ to $a_j \otimes (a_1 \otimes X_1\beta + a_2 \otimes X_2\beta + a_3 \otimes X\beta)$. T_A looks like

$$
T_A = \begin{pmatrix}
X_1 & & \\
X_2 & & \\
X_3 & & \\
& X_1 & \\
& X_2 & \\
& X_3 & \\
& & X_1 \\
& & X_2 \\
& & X_3
\end{pmatrix}
$$

The matrix has a block structure, each block has size $\bar{c} \times \bar{b}$. The three block columns correspond to bases of the form $a_1 \otimes c_j$, $a_2 \otimes c_j$, and $a_3 \otimes c_j$, respectively. The nine block rows correspond to bases $a_1 \otimes a_1 \otimes \beta_j$, $a_1 \otimes a_2 \otimes \beta_j$, $a_1 \otimes a_3 \otimes \beta_j$, $a_2 \otimes a_1 \otimes \beta_j$, ..., $a_3 \otimes a_3 \otimes \beta_j$, respectively.

The projection $A \otimes A \to \Lambda^2 A$ maps $a_i \otimes a_j \to a_i \wedge a_j = \frac{1}{2}(a_i \otimes a_j - a_j \otimes a_i)$, that is, we identify $a_i \otimes a_j$ with $-a_j \otimes a_i$. Therefore, $T_A^\wedge$ looks like

$$
T_A^\wedge = \begin{pmatrix}
X_2 & -X_1 & \\
X_3 & & -X_1 \\
& X_3 & -X_2
\end{pmatrix}.
$$

Here, the three block rows correspond to the bases $a_1 \wedge a_2 \otimes \beta_j$, $a_1 \wedge a_3 \otimes \beta_j$, and $a_2 \wedge a_3 \otimes \beta_j$, respectively.

1.2.1 Exercise. *Prove that*

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

if W is invertible. (X, Y, Z, W are matrices of matching sizes.)

Now assume that $B = C$ and that X_1 is invertible. By performing a base change, we can assume that $X_1 = \mathrm{Id}$, this does not change $\mathrm{rk}(T_A^\wedge)$. After some row operations, we get that

$$
T_A^\wedge = \begin{pmatrix}
& X_3 & -X_2 \\
X_2 & \mathrm{Id} & \\
X_3 & & \mathrm{Id}
\end{pmatrix}.
$$

$T_A^\wedge$ has full rank iff $\det(T_A^\wedge) \neq 0$. By Exercise 1.2.1,

$$
\det(T_A^\wedge) = \det \begin{pmatrix} X_3 & -X_2 \end{pmatrix} \begin{pmatrix} X_2 \\ X_3 \end{pmatrix} = \det(X_3 X_2 - X_2 X_3).
$$

1.2.2 Proposition (Strassen [Str83])**.** *Let $T \in \mathbb{F}^3 \otimes \mathbb{F}^{\bar{b}} \otimes \mathbb{F}^{\bar{b}}$ be given by slices (Id, X, Y). If $\mathrm{rk}(XY - YX) = \bar{b}$, then $\underline{R}(T) \geq \frac{3}{2}\bar{b}$.*

1.2.3 Exercise. *Generalize this to: $\underline{R}(T) \geq \bar{b} + \frac{1}{2}\mathrm{rk}(XY - YX)$ (when $\mathrm{rk}(XY - YX) < \bar{b}$).*

Proof. By flattening, we have

$$
\mathrm{rk}(T_A^\wedge) \leq \underline{R}(T_A^\wedge).
$$

We have $\mathrm{rk}(T_A^\wedge) = 3\bar{b}$, since

$$
\det(T_A^\wedge) = \det(XY - YX) \neq 0.
$$

Let $T_\epsilon = \sum_{i=1}^{r} a_{\epsilon,i} \otimes b_{\epsilon,i} \otimes c_{\epsilon,i}$ such that $T_\epsilon \to T$. Let r be minimal such that such a sequence T_ϵ exists, that is, $r = \underline{R}(T)$. We have

$$(T_\epsilon)_A^{\wedge} = \sum_{i=1}^{r} (a_{\epsilon,i} \otimes b_{\epsilon,i} \otimes c_{\epsilon,i})_A^{\wedge},$$

since the transformation is linear. By Lemma 1.2.4 below (recall that $\bar{a} = 3$) and the subadditivity of rank, we have

$$\mathrm{rk}((T_\epsilon)_A^{\wedge}) \leq 2r.$$

Since rk is lower semi-continuous,

$$\mathrm{rk}((T_\epsilon)_A^{\wedge}) \geq \mathrm{rk}(T_A^{\wedge}) = 3\bar{b}.$$

Hence $3\bar{b} \leq 2r = 2\underline{R}(T)$ and the claim follows. $\qquad\square$

1.2.4 Lemma. $\mathrm{rk}((a \otimes b \otimes c)_A^{\wedge}) \leq \bar{a} - 1$.

Proof. Let $a = a_1, a_2, \ldots, a_{\bar{a}}$ be a basis of A.

$$a \otimes b \otimes c : B^* \to A \otimes C$$

maps

$$\beta \mapsto \beta(b) \cdot a \otimes c.$$

Furthermore,

$$(a \otimes b \otimes c)_A : A \otimes B^* \to A \otimes A \otimes C$$

maps

$$a_i \otimes \beta \mapsto \beta(b) \cdot a_i \otimes a \otimes c,$$

and finally,

$$(a \otimes b \otimes c)_A^{\wedge} : A \otimes B^* \to \Lambda^2 A \otimes C$$

maps

$$a_i \otimes \beta \mapsto \beta(b) \cdot (a_i \wedge a) \otimes c.$$

Since $a_1 = a$, we have $a_1 \wedge a = 0$ and the image of $(a \otimes b \otimes c)_A^{\wedge}$ is at most $(\bar{a} - 1)$-dimensional (and in fact, the dimension equals $\bar{a} - 1$). $\qquad\square$

1.3   The slices of the matrix multiplication tensor

Let $E_{i,j}$ be the matrix that has a 1 in position (i, j) and 0's elsewhere. The matrix multiplication map maps

$$(E_{i,j}, E_{\bar{j},k}) \mapsto E_{i,j} E_{\bar{j},k} = \begin{cases} E_{i,k} & \text{if } j = \bar{j}, \\ 0 & \text{otherwise.} \end{cases}$$

If we order the first basis as $E_{1,1}, \ldots, E_{1,n}, E_{2,1}, \ldots$ and the second as $E_{1,1}, \ldots, E_{n,1}, E_{2,1}, \ldots$, then the 3-slices of the matrix multiplication tensor corresponding to $E_{i,k}$ is of the form $E_{i,k} \otimes \mathrm{Id}_n$, since exactly n pairs of matrices contribute to this slice, namely the one above for $j = \bar{j}$.

1.4 Lower bound for the border rank of matrix multiplication

Let $t \in A \otimes B \otimes C$ be a tensor and let $E : A \to A'$, $F : B \to B'$, and $G : C \to C'$ be endomorphisms. Recall that $(A \otimes B \otimes C)t$ is defined on decomposable elements by $(E \otimes F \otimes G)(a \otimes b \otimes c) = E(a) \otimes F(b) \otimes G(c)$ and then extended to arbitrary elements by linearity. Furthermore, we have

$$R((E \otimes F \otimes G)t) \leq R(t)$$

and the same is true for border rank.

Strassen's lower bound is for tensors with three slices. To apply it to the matrix multiplication tensor $\langle n, n, n \rangle$, we will define a projection $\pi : \mathbb{F}^{n \times n} \to S$, where S is subspace of dimension three, and then apply the lower bound to $(\mathrm{Id} \otimes \mathrm{Id} \otimes \pi) \langle n, n, n \rangle$.

Let

$$P = \begin{pmatrix} 0 & 1 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & 0 & \ldots & 0 \end{pmatrix},$$

be the permutation matrix that corresponds to the cyclic shift and let

$$L \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix},$$

be a diagonal matrix with pairwise distinct entries on the main diagonal.

It is easy to see that

$$LP - PL = \begin{pmatrix} 0 & \lambda_1 - \lambda_2 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \lambda_{n-1} - \lambda_n \\ \lambda_n - \lambda_1 & 0 & \ldots & 0 \end{pmatrix},$$

therefore $LP - PL$ has full rank, since the λ_i are pairwise distinct.

As the three slices of the matrix multiplication tensor, we choose $\mathrm{Id}_n \otimes \mathrm{Id}_n$, $L \otimes \mathrm{Id}_n$, and $P \otimes \mathrm{Id}_n$. Note that

$$(L \otimes \mathrm{Id}_n)(P \otimes \mathrm{Id}_n) - (P \otimes \mathrm{Id}_n)(L \otimes \mathrm{Id}_n) = (LP - PL) \otimes \mathrm{Id}_n$$

has full rank. Let π be any projection from $\mathbb{F}^{n^2 \times n^2}$ onto $\langle \mathrm{Id}_n \otimes \mathrm{Id}_n, L \otimes \mathrm{Id}_n, P \otimes \mathrm{Id}_n \rangle$. Then we have

$$\begin{aligned} \underline{R}(\langle n, n, n \rangle) &\geq \underline{R}((\mathrm{Id} \otimes \mathrm{Id} \otimes \pi) \langle n, n, n \rangle) \\ &= \underline{R}(\mathrm{Id}_n \otimes \mathrm{Id}_n, L \otimes \mathrm{Id}_n, P \otimes \mathrm{Id}_n) \\ &\geq \frac{3}{2} n^2 \end{aligned}$$

where the last inequality follows from Proposition 1.2.2.

Chapter 2

Lower bounds for rank

Strassen's lower bound is of the following form: We apply some projection π to the tensor in such a way, that a certain polynomial does not vanish on the slices of the resulting tensor. Note that for π, we only defined the image. There is a second degree of freedom, namely, choosing its kernel. By choosing the kernel properly, we can get improved bounds for the rank.

2.1 A technical lemma

2.1.1 Lemma. *Let $p \in \mathbb{F}[X_{1,1}, X_{1,2}, \ldots, X_{n,n}]$ be a nonzero polynomial of degree d over an infinite field $\mathbb{F}$. Let $A_1, \ldots, A_N$ be a basis of $\mathbb{F}^{n \times n}$, $N = n^2$. Then there are indices $i_1, \ldots, i_d$ such that $\langle A_{i_1}, \ldots, A_{i_d} \rangle$ contains a matrix B such that $p(B) \neq 0$.*

Proof. Consider the polynomial $q(\alpha_1, \ldots, \alpha_N) := p(\alpha_1 A_1 + \cdots + \alpha_N A_N)$. q is nonzero, since p is nonzero: Let $C \in \mathbb{F}^{n \times n}$ such that $p(C) \neq 0$ and let $C = \gamma_1 A_1 + \cdots + \gamma_N A_N$. Then

$$q(\gamma_1, \ldots, \gamma_N) = p(C) \neq 0.$$

Furthermore, q is of degree d in $\alpha_1, \ldots, \alpha_N$. Choose $i_1, \ldots, i_{d'}$, $d' \leq d$ such that $\alpha_{i_1} \cdots \alpha_{i_{d'}}$ appears as a monomial in q. (The indices need not be pairwise distinct. Set all other α_j to zero. Let $\hat{q}(\alpha_{i_1}, \ldots, \alpha_{i_{d'}})$ be the resulting polynomial. $\hat{q}$ is nonzero, since it contains the monomial $\alpha_{i_1} \cdots \alpha_{i_{d'}}$ (with some nonzero coefficient). Therefore, we can find an assignment to $\beta_{i_1}, \ldots, \beta_{i_{d'}}$ such that $\hat{q}(\beta_{i_1}, \ldots, \beta_{i_{d'}}) \neq 0$. $B := \beta_{i_1} A_{i_1} + \cdots + \beta_{i_{d'}} A_{i_{d'}}$ is the matrix we are looking for. $\quad\square$

2.2 Substitution method

2.2.1 Lemma (Substitution method)**.** *Let $t \in A \otimes B \otimes C$ be a tensor and let*

$$t = \sum_{i=1}^{r} a_i \otimes b_i \otimes c_i$$

be an optimal decomposition into rank-one tensors. ("Optimal" here means $r = R(t)$.) Let $\pi : C \to C$ be a projection. Then

$$R(t) \geq R((\mathrm{Id} \otimes \mathrm{Id} \otimes \pi)t) + \#\{i \mid c_i \in \ker \pi\}.$$

Proof. By definition,

$$(\mathrm{Id} \otimes \mathrm{Id} \otimes \pi)t = \sum_{i=1}^{r} a_i \otimes b_i \otimes \pi(c_i).$$

(This is well-defined.) Permute the summands in such a way that $c_1, \ldots, c_{r'} \notin \ker \pi$ and $c_{r'+1}, \ldots, c_r \in \ker \pi$. Then

$$(\mathrm{Id} \otimes \mathrm{Id} \otimes \pi)t = \sum_{i=1}^{r'} a_i \otimes b_i \otimes \pi(c_i).$$

Therefore,

$$
\begin{aligned}
R((\mathrm{Id} \otimes \mathrm{Id} \otimes \pi)t) &\leq r' \\
&= r - \#\{i \mid c_i \in \ker \pi\} \\
&= R(t) - \{i \mid c_i \in \ker \pi\},
\end{aligned}
$$

and the assertion of the lemma follows. $\qquad\square$

2.3 Lower bounds for matrix multiplication

Let

$$\langle n, n, n \rangle = \sum_{i=1}^{r} a_i \otimes b_i \otimes c_i.$$

We claim that $c_1, \ldots, c_r$ are linearly independent. Otherwise, there would be a nonzero linear map $E : \mathbb{F}^{n \times n} \to \mathbb{F}^{n \times n}$, such that

$$E(c_i) = 0$$

for all $1 \leq i \leq r$. Therefore

$$\langle (\mathrm{Id} \otimes \mathrm{Id} \otimes E)\langle n, n, n \rangle = 0.$$

But this means that there is a linear dependence among the 3-slices of $\langle n, n, n \rangle$, a contradiction, since we know that the 3-slices are $E_{i,j} \otimes \mathrm{Id}$, $1 \leq i, j, \leq n$.[1] Therefore, we can assume w.l.o.g. that $c_1, \ldots, c_N$ are a basis of $\mathbb{F}^{n \times n}$, $N = n^2$. Let $c_1^*, \ldots, c_N^*$ denote a dual basis of $c_1, \ldots, c_N$. Then

$$(\mathrm{Id} \otimes \mathrm{Id} \otimes c_i^*)t, \quad 1 \leq i \leq N$$

span the space of 3-slices of t. In particular, every $(\mathrm{Id} \otimes \mathrm{Id} \otimes c_i^*)t$ is of the form $s_i \otimes \mathrm{Id}$ for some $s_i \in \mathbb{F}^{n \times n}$ and $s_1, \ldots, s_N$ form a basis of $\mathbb{F}^{n \times n}$.

We have

$$\det(X_1 s_1 \otimes \mathrm{Id} + \cdots + X_N s_N \otimes \mathrm{Id}) = (\det(X_1 s_1 + \cdots + X_N s_n))^n.$$

The determinant on the lefthand side is nonzero iff the determinant on the righthand side is. The determinant on the righthand side has degree n and since $s_1, \ldots, s_N$ form a basis, there are values $\xi_1, \ldots, \xi_N$ such that $\xi_1 s_1 + \cdots + \xi_N s_N = \mathrm{Id}$. By Lemma 2.1.1, there are indices $i_1, \ldots, i_n$ and scalars $\alpha_1, \ldots, \alpha_n$ such that $\det(\alpha_1 s_{i_1} + \cdots + \alpha_n s_{i_n}) \neq 0$. Let $f = \alpha_1 s_{i_1} + \cdots + \alpha_n s_{i_n}$ and $F = f \otimes \mathrm{Id}$. By replacing $t := \langle n, n, n \rangle$ by $(F^{-1} \otimes \mathrm{Id} \otimes \mathrm{Id})t$, we can assume that $f = \mathrm{Id}$. Above, F^{-1} becomes an endomorphism of $\mathbb{F}^{N \times N}$ by left multiplication with F^{-1}. Since F is invertible, this does not change the rank of t. Note that $F^{-1} = f^{-1} \otimes \mathrm{Id}$, therefore, after the transformation, the 3-slices still form a basis of the space $\mathbb{F}^{n \times n} \otimes \mathrm{Id}$. To simplify notations, we call the new tensor again t and the slices $s_i \otimes \mathrm{Id}$, $1 \leq i \leq N$. We now have that $\mathrm{Id} \in \langle s_{i_1}, \ldots, s_{i_n} \rangle$.

Next, consider

$$\det[X_1 s_1 \otimes \mathrm{Id} + \cdots + X_N s_N \otimes \mathrm{Id}, Y_1 s_1 \otimes \mathrm{Id} + \cdots + Y_N s_N \otimes \mathrm{Id}] = (\det[X_1 s_1 + \cdots + X_N s_N, Y_1 s_1 + \cdots + Y_n s_n])^n.$$

---

[1] Tensors with linearly independent 3-slices are also called *3-concise*.

The determinant on the righthand side is a polynomial of degree $2n$, it is nonzero, since we can instantiate the variables such that we get the matrices L and P from the preceeding chapter. Note that every monomial of the determinant has degree n in the X-variables and degree n in the Y-variables as well. It is very easy to extend Lemma 2.1.1 to this situation and prove that there indices $j_1, \ldots, j_n$ and $k_1, \ldots, k_n$ such that there are matrices $\ell \in \langle s_{j_1}, \ldots, s_{j_n} \rangle$ and $p \in \langle s_{k_1}, \ldots, s_{k_n} \rangle$ such that $\det[\ell, p] \neq 0$. Now let $C : \mathbb{F}^{n \times n} \to \mathbb{F}^{n \times n}$ be the projection along $\langle s_i \mid i \notin \{i_1, \ldots, i_n, j_1, \ldots, j_n, k_1, \ldots, k_n\} \rangle$ onto $\langle s_{i_1}, \ldots, s_{i_n}, s_{j_1}, \ldots, s_{j_n}, s_{k_1}, \ldots, s_{k_n} \rangle$.

Now

$$(\mathrm{Id} \otimes \mathrm{Id} \otimes C)\langle n, n, n \rangle = \sum_{i=1}^{r} a_i \otimes b_i \otimes C(c_i).$$

There are $N - 3n$ of the c_i in the kernel of C, therefore,

$$R(\langle n, n, n \rangle) \geq R((\mathrm{Id} \otimes \mathrm{Id} \otimes C)t) + N - 3n. \tag{2.3.1}$$

Furthermore $\mathrm{Id} \otimes \mathrm{Id}, \ell \otimes \mathrm{Id}, p \otimes \mathrm{Id}$ are in span of the slices of $(\mathrm{Id} \otimes \mathrm{Id} \otimes C)t$, by the definition of C.

By Strassen's equation,

$$R((\mathrm{Id} \otimes \mathrm{Id} \otimes C)t) \geq \frac{3}{2}N.$$

(In fact, this lower bound even holds for the border rank, but we cannot make use out of this, since (2.3.1) only holds for the rank. To apply Strassen's equation, we formally have to project the third component onto our three slices $\mathrm{Id} \otimes \mathrm{Id}, \ell \otimes \mathrm{Id}, p \otimes \mathrm{Id}$.) Plugging the last equation into the first, we obtain:

2.3.2 Theorem ([Blä00])**.**

$$R(\langle n, n, n \rangle) \geq \frac{5}{2}n^2 - 3n$$

Chapter 3

Landsberg–Ottaviani equations

Now, we take our tensor $T : B^* \to A \otimes C$ and tensor it this time with $\mathrm{Id}_{\Lambda^p A}$ for some $p \leq n-1$ and then project $\Lambda^p A \otimes A \to \Lambda^{p+1} A$. We obtain a map

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

Like before, we want to bound $\underline{R}(T)$ in terms of $\mathrm{rk}\, T_A^{\wedge^p}$.

We start with a generalisation of Lemma 1.2.4:

3.0.1 Lemma. $\mathrm{rk}((a \otimes b \otimes c)_A^{\wedge^p}) \leq \binom{\bar{a}-1}{p}$.

Proof. Let $a = a_1, \ldots, a_{\bar{a}}$ be a basis of A. $(a \otimes b \otimes c)_A^{\wedge^p}$ sends

$$a_{i_1} \wedge \cdots \wedge a_{i_p} \otimes \beta \quad \mapsto \quad \beta(b) a_{i_1} \wedge \cdots \wedge a_{i_p} \wedge a \otimes c$$

$a_{i_1} \wedge \cdots \wedge a_{i_p} \wedge a$ is 0 iff there is a j such that $i_j = 1$. Thus the image of has dimension at most $\binom{\bar{a}-1}{p}$ (and is in fact equal to $\binom{\bar{a}-1}{p}$). $\qquad\square$

3.0.2 Lemma. $\underline{R}(T) \geq \frac{\mathrm{rk}(T_A^{\wedge^p})}{\binom{\bar{a}-1}{p}}$.

Proof. Let $r = \underline{R}(T)$. Let T_ϵ be a sequence of tensors such that $T_\epsilon \to T$ and $R(T_\epsilon) = r$. Write

$$T_\epsilon = \sum_{i=1}^r a_{i,\epsilon} \otimes b_{i,\epsilon} \otimes c_{i,\epsilon}.$$

Then

$$(T_\epsilon)_A^{\wedge^p} = \sum_{i=1}^r (a_{i,\epsilon} \otimes b_{i,\epsilon} \otimes c_{i,\epsilon})_A^{\wedge^p}.$$

Thus $\mathrm{rk}((T_\epsilon)_A^{\wedge^p}) \leq r\binom{\bar{a}-1}{p}$. Since rk is lower semicontinuous, we get

$$\mathrm{rk}(T_A^{\wedge^p}) \leq r \binom{\bar{a}-1}{p}.$$

Now the claim follows, as $r = \underline{R}(T)$. $\qquad\square$

Now assume that $\bar{a} = 2p+1$. If this is not the case, then we can restrict to some subspace of A. Write

$$T = a_0 \otimes X_0 + \cdots + a_{2p} \otimes X_{2p}$$

with $X_i : B^* \to C$ and $a_0, \ldots, a_{2p}$ being a basis. $T_A^{\wedge^p} : \Lambda^p A \otimes B^* \to \Lambda^{p+1} A \otimes C$ maps
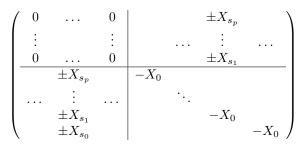
$$a_{i_1} \wedge \cdots \wedge a_{i_p} \otimes \beta \mapsto \sum_{j=0}^{2p} a_{i_1} \wedge \cdots \wedge a_{i_p} \wedge a_j \otimes X_j(\beta). \tag{3.0.3}$$

We now define two bases for $\Lambda^p A$ and $\Lambda^{p+1} A$. For $I \subseteq \{0, \ldots, 2p\}$, set $a_I = \bigwedge_{i \in I} a_i$.

$\Lambda^p A$: a_I with $|I| = p$,

         first all I with $0 \in I$, $\binom{2p}{p-1}$ many

         then all other, $\binom{2p}{p}$ many

$\Lambda^{p+1} A$: a_I with $|I| = p+1$,

         first all I with $0 \notin I$, $\binom{2p}{p+1} = \binom{2p}{p-1}$ many

         then all other, $\binom{2p}{p}$

Every basis consists of two blocks. The second block of the first basis consists of a_I with $|I| = p$ and $0 \notin I$. The second block of the second basis consists of A_J with $|J| = p+1$ and $0 \in J$. We can write $J = \{0\} \cup I$ with $0 \notin I$ and $|I| = p$. So both blocks are indexed by sets I and we want that both blocks have the same order with respect to I. Now we write the matrix of (3.0.3) with respect to the chosen bases:

$$\begin{pmatrix} 0 & \cdots & 0 & & \pm X_{s_p} & \\ \vdots & & \vdots & \cdots & \vdots & \cdots \\ 0 & \cdots & 0 & & \pm X_{s_1} & \\ \hline \pm X_{s_p} & & & -X_0 & & \\ \cdots & \vdots & \cdots & & \ddots & \\ & \pm X_{s_1} & & & -X_0 & \\ & \pm X_{s_0} & & & & -X_0 \end{pmatrix}$$

The first block column corresponds to elements $a_{i_1} \wedge \cdots \wedge a_{i_p} \otimes \beta$ with $i_1 = 0$, the second block columns to elements with $i_1 > 0$. In each column we record the coefficients of the image of such an element under the map (3.0.3). The first block row corresponds to $a_{i_1} \wedge \cdots \wedge a_{i_{p+1}} \otimes c$ with $i_1 > 0$ and the second block row corresponds to $i_1 = 0$. The two columns with the $\pm X_{s_j}$ entries exemplify the image of $a_{i_1} \wedge \cdots \wedge a_{i_p} \otimes \beta$ under (3.0.3).[1] Note that in the sum on the righthand side of (3.0.3), only $p+1$ summands are nonzero. The column on the lefthand side is the case $i_1 = 0$. Then all nonzero coefficients in the matrix are in the lower block row, since every summand on the righthand side of (3.0.3) contains the index 0. We have $s_j \notin \{i_1, i_2, \ldots, i_p\}$. The column on the righthand side is the case $i_1 > 0$. One summand that we get is $-a_0 \wedge a_{i_1} \wedge \cdots \wedge a_{i_p} \otimes X_0$ and this produces the $-X_0$ in the lower right block. Every other summand does not contain 0 and therefore, the coefficients appear in the upper block row. Write the matrix above as

$$\begin{pmatrix} 0 & Q \\ \tilde{Q} & R \end{pmatrix}.$$

By an appropriate change of bases, we can achieve that $X_0 = -\,\mathrm{Id}$ and henceforth, $R = \mathrm{Id}$. (The identity matrices have different sizes.) Therefore,

$$\det \begin{pmatrix} 0 & Q \\ \tilde{Q} & \mathrm{Id} \end{pmatrix} = \det(Q\tilde{Q})$$

---

[1] Strictly speaking, β runs over some basis, since each X_i is itself a matrix.

What are the block[2] entries of Q and $\tilde{Q}$. Columns of Q correspond to elements $a_{J'} = a_{j'_1} \wedge \cdots \wedge a_{j'_p}$ with $j'_1 > 0$. Rows correspond to elements $a_J = a_{j_1} \wedge \cdots \wedge a_{j_{p+1}}$ with $j_1 > 0$. The block entry $q_{J,J'}$ is nonzero, if J is by one element larger thant J', that is,

$$q_{J,J'} = \begin{cases} \pm X_k & \text{if } J' \cup \{k\} = J \\ 0 & \text{otherwise} \end{cases}$$

In the same way, columns of $\tilde{Q}$ correspond to elements $a_0 \wedge a_{i_2} \wedge \cdots \wedge a_{i_p}$ and rows to elements $a_0 \wedge a_{i'_2} \wedge \cdots \wedge a_{i'_{p+1}}$. Let $I = \{i_2, \ldots, i_p\}$ and $I' = \{i'_2, \ldots, i'_{p=1}\}$. We remove the 0 from the index sets to get a nicer condition of the $m_{J,I}$ below. We have

$$\tilde{q}_{I',I} = \begin{cases} \pm X_\ell & \text{if } I \cup \{\ell\} = I' \\ 0 & \text{otherwise} \end{cases}$$

For the entries of $Q\tilde{Q} = (m_{J,I})$ we have

$$m_{J,I} = \begin{cases} \pm(X_k X_\ell - X_\ell X_k) & \text{if } I \cup \{k, \ell\} = J \\ 0 & \text{otherwise} \end{cases}$$

See [MR12] for some explicit examples of the matrix $Q\tilde{Q}$.

3.1 Matrix multiplication

Now we consider the case of matrix multiplication $\langle n, n, n \rangle : \mathbb{F}^{n \times n} \times \mathbb{F}^{n \times n} \to \mathbb{F}^{n \times n}$. We choose $p = n - 1$ and will define a subspace $A \subseteq \mathbb{F}^{n \times n}$ of dimension $2p + 1 = 2n - 1$. $A = \langle a_0, \ldots, a_{2n-2} \rangle$ will be the space of *Hankel matrices*,

$$a_k = \sum_{i+j=k+2} e_{i,j}, \qquad 0 \le k \le 2n - 2$$

where $e_{i,j}$ is the matrix that has a 1 in position (i,j) and 0s elsewhere. Note that

$$a_k : e_j \mapsto e_{k+2-j}, \qquad 1 \le j \le k + 1$$

for all $0 \le k \le n - 1$ and

$$a_k : e_j \mapsto e_{k+2-j}, \qquad n + 2 - k \le j \le n$$

for all $n \le k \le 2n - 2$. Here e_i is the ith unit vector.

We consider the restriction of $\langle n, n, n \rangle$ to $A \otimes \mathbb{F}^{n \times n}$ and T denotes the flattening $(\mathbb{F}^{n \times n})^* \to A \otimes \mathbb{F}^{n \times n}$. The slices of $\langle n, n, n \rangle$ corresponding to a_i are $X_i = a_i \otimes \text{Id}_n$. (This was proven in GCT 1.) We can write $T = a_0 \otimes X_0 + \ldots a_{2n-2} \otimes X_{2n-2}$. We want to prove that $T_A^{\wedge^{n-1}}$ has full rank. It is enough to prove that $t_A^{\wedge^{n-1}}$ has full rank, where $t = a_0 \otimes x_0 + \cdots + a_{2n-2} \otimes x_{2n-2}$ and $x_i = a_i$. We prefer to write x_i instead of a_i to distinguish between the different spaces.

3.1.1 Lemma. *The map $t_A^{\wedge^{n-1}}$ is injective.*

Proof. We follow the proof given in [Lan17]. $t_A^{\wedge^{n-1}}$ maps

$$a_S \otimes e_j \mapsto \sum_{j \notin S} a_S \wedge a_j \otimes x_j(e_j),$$

---

[2]Block now refers to the division induced by the X_i, not the two blocks induced by the structure of the bases.

where $S \subseteq \{0, \dots, 2n-2\}$, $|S| = n-1$ and $a_S = \bigwedge_{i \in S} a_i$. We will prove by induction that every $a_P \otimes e_i$ is in the image of $T_A^{\wedge^{n-1}}$, where $P \subseteq \{0, \dots, 2n-2\}$, $|P| = n$, and $i \in \{1, \dots, n\}$. For the induction, we will define a partial ordering among the $a_P \otimes e_i$. We have

$$a_P \otimes e_i < a_Q \otimes e_j$$

if for $\ell = \min\{i, j\}$, the ℓ smallest elements of Q are smaller or equal to the ℓ smallest elements of P and one of them is strictly smaller, or these ℓ elements are the same and $i < j$.

The minimal element in this ordering is $a_{\{n-1,\dots,2n-2\}} \otimes e_1$. We have

$$t_A^{\wedge^{n-1}} \left(a_{\{n,\dots,2n-2\}} \otimes e_n \right) = a_{\{n-1,\dots,2n-2\}} \otimes e_1,$$

since $x_j(e_n) = 0$ for $j < n-1$. This is the induction base.

Now consider some $a_Q \otimes e_j$ and assume that all $a_P \otimes e_i$ with $a_P \otimes e_i < a_Q \otimes e_j$ are already in the image. Let $Q = \{q_1, \dots, q_n\}$ with $q_\nu < q_{\nu+1}$ for all ν. Let $Q' = Q \setminus \{q_j\}$ and consider the image of $a_{Q'} \otimes e_{2+q_j-j}$. (Note that $1 \leq 2 + q_j - j \leq n$.) The image is

$$\sum_{i \notin Q'} a_{Q'} \wedge a_i \otimes x_i(e_{1+q_j-j}).$$

If we consider $i = q_j$, then we see that $a_Q \otimes e_j$ is in the sum, since $x_{q_j}(e_{2+q_j-j}) = e_{q_j+2-2-q_j+j} = e_j$.

If $i < q_j$, then the summand is $a_{Q' \cup \{i\}} \otimes e_{i+j-q_j}$. Since $i + j - q_j < j$ and the first $j-1$ elements of Q' and Q are the same, we have that $a_{Q' \cup \{i\}} \otimes e_{i+j-q_j} < a_Q \otimes e_j$.

If $i > q_j$, then the summand is $a_{Q' \cup \{i\}} \otimes e_{i+j-q_j}$. Now $i + j - q_j > j$ and the first j elements of $Q' \cup \{i\}$ are larger than the first j elements of Q. Therefore, $a_{Q' \cup \{i\}} \otimes e_{i+j-q_j} < a_Q \otimes e_j$ in this case, too.

Since all other elements in the sum are smaller than $a_Q \otimes e_j$, they are in the image of $t_A^{\wedge^{n-1}}$ by the induction hypothesis. Therefore, $a_Q \otimes e_j$ is also in the image. □

3.1.2 Theorem (Landsberg & Ottaviani [LO11]). $\underline{R}(\langle n, n, n \rangle) \geq 2n^2 - n$.

Proof. By Lemma 3.1.1, $\mathrm{rk}(T_A^{\wedge^{n-1}}) \geq \binom{2n-1}{n-1}n^2$, where A is as chosen in the lemma. Therefore, by Lemma 3.0.2,

$$\underline{R}(\langle n, n, n \rangle) \geq \frac{\binom{2n-1}{n-1}}{\binom{2n-2}{n-1}} n^2 = \frac{2n-1}{n} n^2.$$

□

3.2 Extension to rank

In a similar way as above, we can prove that for $p \leq n-1$, $T_A^{\wedge^p}$ is injective for some appropriate space A of dimension $2p + 1$. The lower bound that you get in this case is

$$\underline{R}(\langle n, n, n \rangle) \geq \frac{\binom{2p+1}{p}}{\binom{2p}{p}} n^2 = (2 - \frac{1}{p+1})n^2.$$

When $p = \omega(1)$, then the righthand side is $2n^2 - o(n^2)$.

Now the idea is to choose p quite small, say $\log \log n$ and combine the bound for the border rank with the substitution method. $T_A^{\wedge^p}$ is injective if $\det(Q\tilde{Q})$ is nonzero. $\det(Q\tilde{Q})$ is a polynomial of degree $\binom{2p}{p-1}n^2$. However, Q can be written as $q \otimes \mathrm{Id}_n$ and $\tilde{Q}$ as $\tilde{q} \otimes \mathrm{Id}_n$. Therefore, $\det(Q\tilde{Q})$ is the nth power of a polynomial of degree $\binom{2p}{p-1}n$. We know that the polynomial is nonzero by the particular construction of A mentioned above.

If we now take any basis of $\mathbb{F}^{n \times n}$, we know that there are $\binom{2p}{p-1} n$ elements among them such that the corresponding polynomial that we get from $\det(Q\tilde{Q})$ by replacing the concrete choices by generic combinations of the chosen elements is nonzero. Since $\binom{2p}{p-1} = o(n^2)$, we get the followng theorem.

3.2.1 Theorem (Landsberg [Lan14]). $R(\langle n, n, n \rangle) \geq 3n^2 - o(n^2)$.

Bibliography

[Blä00] Markus Bläser. Lower bounds for the bilinear complexity of associative algebras. *Computational Complexity*, 9(2):73–112, 2000.

[Lan14] J. M. Landsberg. New lower bounds for the rank of matrix multiplication. *SIAM J. Comput.*, 43(1):144–149, 2014.

[Lan17] J. M. Landsberg. *Geometry and Complexity Theory*. Cambridge University Press, 2017.

[LO11] J. M. Landsberg and Giorgio Ottaviani. New lower bounds for the border rank of matrix multiplication. *CoRR*, abs/1112.6007, 2011.

[MR12] Alex Massarenti and Emanuele Raviolo. The rank of n×n matrix multiplication. *CoRR*, abs/1211.6320, 2012.

[Str83] Volker Strassen. Rank and optimal computation of generic tensors. *Linear Algebra Appl.*, 52-53:645–685, 1983.