Explicit tensors

Markus Bläser

Abstract. This is an expository article the aim of which is to introduce interested students and researchers to the topic of tensor rank, in particular to the construction of explicit tensors of high rank. We try to keep the mathematical concepts and language used as simple as possible to address a broad audience. This article is thought to be an appetizer and does not provide by any means a complete coverage of this topic.

Mathematics Subject Classification (2010). 68Q17, 15A69.

Keywords. algebraic complexity theory, tensor rank, lower bounds.

1. Tensors and rank

Let U and V be vector spaces over some field k. It is a well known fact that every linear map $\alpha : U \to V$ is represented by a matrix $A = (a_{i,j}) \in k^{\ell \times m}$, where $\ell = \dim U$ and $m = \dim V$. The rank of the matrix A is the maximum number of rows that are linearly independent. There are a lot of equivalent definitions of the rank of a matrix, for instance,

- the maximum number of columns that are linearly independent,
- $\ell - \dim \ker \alpha$,
- $\dim \operatorname{im} \alpha$,
- the minimum number of matrices of the form $x^T \cdot y$ with $x \in k^\ell$ and $y \in k^m$, so called rank-one-matrices, such that A can be written as the sum of these matrices,

just to mention a few. The rank of linear maps and matrices is well understood, there are efficient algorithms to compute the rank, the most famous method is Gaussian elimination.

Let W be another vector space over k, $\dim W = n$. Let $\beta : U \times V \to W$ be a bilinear map. By choosing bases $u_1, \dots, u_\ell$ of U, $v_1, \dots, v_m$ of V and

This work was supported through funds provided by the Deutsche Forschungsgemeinschaft (BL 511/10-1) and by the Indo-German Max-Planck Center for Computer Science (IMPECS).

$w_1, \dots, w_n$ of W, we can associate structural constants $b_{h,i,j}$ with β:

$$\beta(u_h, v_i) = \sum_{j=1}^{n} b_{h,i,j} w_j, \qquad 1 \le h \le \ell,\ 1 \le i \le m. \tag{1}$$

We can view $B = (b_{h,i,j}) \in k^{\ell \times m \times n}$ as a three-dimensional matrix, a so-called *tensor*. As we have seen, there are many ways to define the rank of a matrix, which is nothing but a tensor in $k^{\ell \times m \times 1}$. From the many equivalent definitions of rank of a matrix given above, it turns out that the appropriate one for tensors is the last one. We call a tensor $S = (s_{h,i,j})$ a rank-one tensor or triad, if there are vectors $a = (a_1, \dots, a_\ell)^T \in k^\ell$, $b = (b_1, \dots, b_m)^T \in k^m$, and $c = (c_1, \dots, c_n)^T \in k^n$ such that

$$s_{h,i,j} = a_h b_i c_j, \qquad 1 \le h \le \ell,\ 1 \le i \le m,\ 1 \le j \le n.$$

We will write $S = a \otimes b \otimes c$. Then the *rank of a tensor* T is the minimum number r such that there are rank-one tensors $S_1, \dots, S_r$ with

$$T = S_1 + \cdots + S_r.$$

We denote the rank of T by $R(T)$. The question of the rank of tensors of order three or equivalently, of the rank of the corresponding bilinear mapping is a central question in algebraic complexity theory. The flagship problem is of course matrix multiplication, which is a bilinear mapping $k^{n \times n} \times k^{n \times n} \to k^{n \times n}$. The current best upper bounds are $O(n^{2.38})$, see [8, 19, 22], whereas the best lower bound is the recent $3n^2 - o(n^2)$ by Landsberg [11].

1.1. What is so special about matrices?

The rank of a matrix is a well-understood quantity. The maximum rank of a matrix in $k^{\ell \times m}$ is $\min\{\ell, m\}$ and it is easy to come up with a matrix that achieves this maximum. For instance, the identity matrix padded with rows or columns of zeroes does the job. Or Vandermonde matrices. The fact that we have a lot of equivalent characterizations of the rank of the matrix seems to be crucial for the fact that it is so easy to come up with explicit matrices of high rank. We can compute the rank of a matrix in polynomial time.

For tensors the situation is more complicated. Per se, it is even not clear what the maximum rank of tensors in $k^{\ell \times m \times n}$ is.[1] We will discuss this briefly in the next section. How to construct explicit tensors of high rank is a widely open problem. And finally, computing the rank of a tensor is an NP-hard problem.

Theorem 1 (Håstad [10]**).** *Let k be a field that can be represented over $\{0,1\}^*$. Let* `Tensor-Rank` *(over k) be the following problem: Given a tensor $T \in k^{\ell \times m \times n}$ and a bound b, decide whether $R(T) \le b$.*

1. *Over finite fields,* `Tensor-Rank` *is* NP*-complete.*
2. *Over $\mathbb{Q}$,* `Tensor-Rank` *is* NP*-hard.*

[1] Of course, ℓmn is an upper bound. However it is not clear—and not true—that this is necessary.

What does it mean that “a field can be represented over $\{0, 1\}^*$”? Traditional complexity classes like P or NP are defined over some fixed alphabet, so we need to be able to encode the field elements by $\{0, 1\}$-strings. For instance, elements from finite fields can be represented in binary and rational numbers by tuples of integers represented in binary. The actual encoding does not matter as long as it is “reasonably nice”, that is, all operations like addition, multiplication, etc. can be performed in polynomial time.

The hardness proof is the same over finite fields and over $\mathbb{Q}$. Over finite fields, the problems is also NP-easy; we just have to guess b rank-one tensors and check whether their sum is T. Over $\mathbb{Q}$, it is not clear whether this is possible, since we do not know an upper bound on the number of bits of the representation of the entries of the rank-one tensors. It could be the case that the rank of T is b, but all sums of b rank-one tensors involve rational numbers with a huge number of bits. (“Huge” means superpolynomial in the size of the representation of the input tensor.) To the best of my knowledge, it is even not known whether `Tensor-Rank` over $\mathbb{Q}$ is decidable.

Over $\mathbb{R}$, the situation is somewhat better: $\mathbb{R}$ itself is not representable over $\{0, 1\}$, but since $\mathbb{Q} \subseteq \mathbb{R}$, we can look at tensors T over $\mathbb{Q}$ and ask what is the minimum number of rank-one tensor with entries from $\mathbb{R}$ such that T is the sum of these rank-one tensors. This problem is decidable and even in PSPACE, since it can be reduced to the existential theory over the reals [7, 18].

Open Problem 1. *What can you say about the approximability of* `Tensor-Rank`*? Is there a constant factor approximation algorithm? A PTAS? As far as I know, nothing in this direction is known.*

Open Problem 2. *What is the complexity of* `Tensor-Rank` *over* $\mathbb{Q}$*?*

Another important property of matrix rank is that it is semicontinuous. If $(M_i) \in k^{n\times n}$ is a sequence of matrices that converges to a matrix M (componentwisely), then

$$R(M_i) \leq r \text{ for all } i \qquad \Rightarrow \qquad R(M) \leq r.$$

Why does this hold? The fact that the rank of $R(M_i) \leq r$ is equivalent to the fact that all $(r + 1) \times (r + 1)$ minors vanish. These minors are polynomials and hence continuous functions. Therefore all $(r + 1) \times (r + 1)$ minors of M vanish, too.

For tensors, this is not the necessarily true. Consider the following tensor t given by the following two slices:

$$(t_{1,i,j}) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad (t_{2,i,j}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{2}$$

Define

$$t_\epsilon = \frac{1}{\epsilon} \cdot (\epsilon, 1) \otimes (1, \epsilon) \otimes (1, \epsilon) - \frac{1}{\epsilon} \cdot (0, 1) \otimes (1, 0) \otimes (1, 0)$$

The two slices of t_ϵ are

$$\begin{pmatrix} 1 & \epsilon \\ \epsilon & \epsilon^2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & \epsilon \end{pmatrix}.$$

So $t_\epsilon \to t$ when $\epsilon \to 0$. And $R(t_\epsilon) \leq 2$ for all $\epsilon > 0$. On the other hand, $R(t) = 3$. We can prove this using the so-called *substitution method*. This method was first introduced by Pan [16] to prove the optimality of the Horner scheme. See [3] for more applications of this method and more references.

Let

$$t = \sum_{i=1}^{r} u_i \otimes v_i \otimes w_i \tag{3}$$

with

$$u_i = (u_{i,1}, u_{i,2}), \quad v_i = (v_{i,1}, v_{i,2}), \quad w_i = (w_{i,1}, w_{i,2}), \quad 1 \leq i \leq r,$$

be an optimal decomposition of t into rank-one tensors. Since $t_{1,1,1} = 1$, there is an i_0 such that $u_{i_0,1} \neq 0$, w.l.o.g. $i_0 = r$. Think of t as consisting of two slices as in (2). From the decomposition (3), we will construct a new tensor in $k^{1\times 2\times 2}$, which is a linear combination of the two slices, in such a way that the rank drops by one. Specifically,

$$\sum_{i=1}^{r} (\alpha u_{i,1} + u_{i,2}) \otimes v_i \otimes w_i = \begin{pmatrix} \alpha & 1 \\ 1 & 0 \end{pmatrix} := t'.$$

If we set $\alpha = -u_{r,2}/u_{r,1}$, this kills the rth rank-one tensor. Therefore,

$$R(t) \geq R(t') + 1.$$

But t' is a matrix whose rank is obviously two. Therefore, $R(t) \geq 3$. Since t has only three entries that are nonzero, there is a trivial decomposition of t of length 3.

2. Explicit tensors of high rank imply circuit lower bounds

2.1. Higher order tensors

We can generalize the concept of tensors of order three to higher orders. Let $V_1, \dots, V_n$ be vector spaces, $\dim V_i = d_i$, $1 \leq i \leq n$. The tensor product $V_1 \otimes \cdots \otimes V_n$ can be built as follows: Choose bases $v_{i,j}$, $1 \leq j \leq d_i$, for each V_i. Then we formally built the elements $v_{1,j_1} \otimes \cdots \otimes v_{n,j_n}$, $1 \leq j_1 \leq d_1, \dots, 1 \leq j_n \leq d_n$. They form a basis of the vector space $V_1 \otimes \cdots \otimes V_n$. If we have arbitrary vectors $x_i = \alpha_{i,1} v_{i,1} + \cdots \alpha_{i,d_i} v_{i,d_i} \in V_i$, $1 \leq i \leq d_i$, then

$$x_1 \otimes \cdots \otimes x_n = \sum_{j_1=1}^{d_1} \cdots \sum_{j_n=1}^{d_n} \alpha_{1,j_1} \cdots \alpha_{n,j_n} v_{1,j_1} \otimes \cdots \otimes v_{n,j_n}.$$

An element $v_1 \otimes \cdots \otimes v_n$ with $v_i \in V_i$ is a rank-one tensor. As before, the rank of a tensor $t \in V_1 \otimes \cdots \otimes V_n$ is minimum number of rank-one tensors $s_1, \dots, s_r$ such that

$$t = s_1 + \cdots + s_r.$$

The definition above is a coordinate-free definition of tensor rank. You can also think in coordinates, if you prefer that: V_i is isomorphic to k^{d_i}, simply choose bases. These isomorphims naturally extend to an isomorphism between $V_1 \otimes \cdots \otimes V_n$ and $k^{d_1} \otimes \cdots \otimes k^{d_n}$. There is also a way of defining tensor products without choosing bases at all by the universal property of turning multilinear mappings into linear ones.

2.2. Basic properties

Let $\pi \in S_n$ be a permutation of $\{1, \dots, n\}$. If $v_i \in V_i$, then $v_{\pi(1)} \otimes \cdots \otimes v_{\pi(n)} \in V_{\pi(1)} \otimes \cdots \otimes V_{\pi(n)}$. So π identifies the rank-one tensors of $V_1 \otimes \cdots \otimes V_n$ with the rank-one tensors of $V_{\pi(1)} \otimes \cdots \otimes V_{\pi(n)}$. We can extend this mapping to a linear mapping $V_1 \otimes \cdots \otimes V_n \to V_{\pi(1)} \otimes \cdots \otimes V_{\pi(n)}$. This mapping clearly is surjective and by comparing dimensions, we see that is is in fact an isomorphism. The image of any tensor t under this mapping is denoted by t^π.

If you think in coordinates, then the entries $t'_{i_1,\dots,i_n}$ of t^π are defined by $t'_{i_1,\dots,i_n} = t_{i_{\pi^{-1}(1)},\dots,i_{\pi^{-1}(n)}}$, where $t = (t_{j_1,\dots,j_n})$.

Fact 1. *$R(t) = R(t^\pi)$.*

Let $U_1, \dots, U_n$ be vector spaces. Let $h_i : V_i \to U_i$ be homomorphism of vector spaces, $1 \le i \le n$. We get a mapping that maps the rank-one tensors of $V_1 \otimes \cdots \otimes V_n$ to the rank-one tensors of $U_1 \otimes \cdots \otimes U_n$ by

$$v_1 \otimes \cdots \otimes v_n \mapsto h_1(v_1) \otimes \cdots \otimes h_n(v_n).$$

Again, we can extend this to a linear mapping $V_1 \otimes \cdots \otimes V_n \to U_1 \otimes \cdots \otimes U_n$. We denote this mapping by $h_1 \otimes \cdots \otimes h_n$.

Fact 2. *$R(t) \ge R(h_1 \otimes \cdots \otimes h_n(t))$ for any tensor t. If all h_i are isomorphisms, then equality holds.*

Let $t \in V_1 \otimes \cdots \otimes V_n$ and $s \in U_1 \otimes \cdots \otimes U_n$. We can embed both tensors into the larger space $(V_1 \oplus U_1) \otimes \cdots \otimes (V_n \oplus U_n)$ as follows: Since each V_i is a subspace of $V_i \oplus U_i$, each rank-one tensor of $V_1 \otimes \cdots \otimes V_n$ is also a rank-one tensor of $(V_1 \oplus U_1) \otimes \cdots \otimes (V_n \oplus U_n)$. Every tensor t in $V_1 \otimes \cdots \otimes V_n$ is a sum of rank-one tensors, so t embeds into $(V_1 \oplus U_1) \otimes \cdots \otimes (V_n \oplus U_n)$ as well. The same works for s. $t \oplus s$ denotes the tensor that we get by viewing t and s as tensors in $(V_1 \oplus U_1) \otimes \cdots \otimes (V_n \oplus U_n)$ and forming their sum. The following fact follows immediately.

Fact 3. *$R(t \oplus s) \le R(t) + R(s)$*

Open Problem 3. *Does $R(t \oplus s) = R(t) + R(s)$ hold for all tensors t and s? This is known as* Strassen's additivity conjecture*.*

Finally, we define the tensor product of t and s, which is an element of $(V_1 \otimes U_1) \otimes \cdots \otimes (V_n \otimes U_n)$. For two rank-one tensors $x = v_1 \otimes \cdots \otimes v_n$ and $y = u_1 \otimes \cdots \otimes u_n$, their tensor product is defined as

$$x \otimes y = (v_1 \otimes u_1) \otimes \cdots \otimes (v_n \otimes u_n).$$

If we write $t = x_1 + \cdots + x_r$ as a sum of rank-one tensors and $s = y_1 + \cdots + y_p$, then we define their tensor product as

$$t \otimes s = \sum_{i=1}^{r} \sum_{j=1}^{p} x_i \otimes y_j. \tag{4}$$

It is easy to verify that this is well defined.

If you think in coordinates, then the tensor product of $t = (t_{i_1,\dots,i_n}) \in k^{d_1 \times \cdots \times d_n}$ and $s = (s_{j_1,\dots,j_n}) \in k^{e_1 \times \cdots \times e_n}$ is given by

$$t \otimes s = (t_{i_1,\dots,i_n} s_{j_1,\dots,j_n}) \in k^{d_1 e_1 \times \cdots \times d_n e_n}.$$

The pair (i_1, j_1) is interpreted as a number from $\{1, \dots, d_1 e_1\}$ and is used to index the first coordinite, (i_2, j_2) for the second, and so on.

From (4), the next fact follows easily.

Fact 4. *$R(t \otimes s) \leq R(t)R(s)$.*

Note that in this case, the inequality may be strict. For instance, the rank of 2×2-matrix multiplication is 7, however, the rank of $2^m \times 2^m$-matrix matrix multiplication is strictly less than 7^m for large enough m, since there are algorithms for matrix multiplication that are asymtotically faster than Strassen's algorithm.

2.3. From tensor rank bounds to formula size bounds

With a tensor $t = (t_{i_1,\dots,i_d}) \in k^n \otimes \cdots \otimes k^n$, we associate the following polynomial in the nd variables $X_{i,j}$, $1 \leq i \leq d$, $1 \leq j \leq n$:

$$f_t = \sum_{i_1=1}^{n} \cdots \sum_{i_d=1}^{n} t_{i_1,\dots,i_d} X_{1,i_1} \cdots X_{n,i_n}.$$

Raz [17] proved the following result:

Theorem 2. *For any family of tensors t_n of order $d(n)$ such that $\omega(1) \leq d(n) \leq o(\log n / \log \log n)$ and $R(t_n) \geq n^{(1-o(1))d(n)}$, the polynomials f_{t_n} have superpolynomial formula size.*

Note that $R(t_n) \leq n^{d(n)}$ by the trivial decomposition. Therefore, the family t_n has "almost" highest rank possible. It is a major open problem to find a family of polynomials with superpolynomial formula size. So finding high rank tensors might be a way of doing so. The best lower bounds we have are due to Kalorkoti [15] and are quadratic.

Several decades earlier, Strassen [20] proved the following result.

Theorem 3. *For any family of tensors t_n of order 3, the circuit complexity of the trilinear forms f_{t_n} are bounded by $\Omega(R(t_n))$.*

This means that a family of tensors of superlinear rank yields a family of polynomials with superlinear circuit complexity, something which we do not know for general circuit models.

But there is a catch, as we will see in the next section. Finding some family of tensors/polynomials with the desired properties is easy, a random choice does the job. So what we really want is an *explicit* tensor. We call a family of tensors $t_n = (t_{n;i_1,\dots,i_d})$ explicit if the mapping $(n; i_1, \dots, i_d) \mapsto t_{n;i_1,\dots,i_d}$ can be computed by an arithmetic circuit of size polynomial in d and $\log n$. One can think also of other notion of explicitness. For the purpose of this appetizer, any notion that prevents random tensors is fine. If the entries of the tensors are rational, we could also require that the mapping is computable in $\mathsf{P/poly}$. Then, by using Valiant's criterion, we can use high rank tensors to separate classes in Valiants model, in particular, we could show that the permanent does not have polynomial size formulas, see [4]

2.4. Random tensors

Let V be a vector space of dimension n. A generic rank-one tensor in $V^{\otimes d}$ is described by dn variables,

$$(x_{1,1}, \dots, x_{n,1}) \otimes \cdots \otimes (x_{1,d}, \dots, x_{n,d}).$$

The sum of r generic rank-one tensors is described by rdn varibles. Its entries are multilinear polynomials in these variables. A generic tensor in $V^{\otimes d}$ is described by n^d variables, all of them begin algebraically independent. Therefore, $rnd \geq n^d$ is required to write every tensor as a sum of r rank-one tensor, that is,

$$r \geq \frac{n^{d-1}}{d}.$$

This is a very simple argument, but sufficient for our needs and almost optimal. With more sophisticated ones, we can get tighter bounds, see the work of Lickteig and Strassen for three-dimensional tensors [14, 21], see Landsberg's book for the general case [12].

From the argument above, it follows that there is a tensor with rank at least n^{d-1}/d. But even random tensors have at least this rank with high probability. The entries of tensors that can be written as the sum of fewer rank-one tensors are algebraically dependent, since they can we written as polynomials in less than n^d variables. Therefore, these entries fulfill some polynomial relation. It is well known that random points do not fufill polynomial relations with high probability. In theoretical computer science, this fact is known as the Schwartz-Zippel lemma.[2]

Lemma 1 (Schwartz–Zippel). *Let F be a field. Let p be a nonzero polynomial in $F[X_1, \dots, X_n]$ of total degree d. Let $S \subseteq F$. Then*

$$\Pr_{r_1,\dots,r_n \in S}[p(r_1, \dots, r_n) = 0] \leq |S|/d.$$

Even if we do not know a bound on the polynomial describing the algebraic dependence, if the underlying field is large enough or even infinite, a random tensor will have rank $\geq n^{d-1}/d$ with high probability.

Although it sounds very simple, it is a major open problem to find a tensor of high rank that is explicit, i.e., whose entries can be constructed by a *deterministic* polynomial time algorithm.

---

[2] The name of the lemma is justified because Schwartz and independently Zippel were the *last* to prove this lemma.

3. Explicit tensors from bilinear mappings

This section shows the present (poor) knowledge of how to construct explicit tensors of high rank.

3.1. The rank of bilinear mappings and algebras

Let $\phi : U \times V \to W$ be a bilinear mapping. Every bilinear map corresponds in a unique way to a tensor t_ϕ in $U^* \otimes V^* \otimes W$, see (1). Since a vector space and its dual are isomorphic, we can also think if t_ϕ living in $U \otimes V \otimes W$. We define the rank of a bilinear map ϕ to be the rank of the corresponding tensor t_ϕ. If A is an finite dimensional associative algebra with unity, that is, A is a ring which is also a finite dimensional vector space over some field k, then the multiplication map in A is a bilinear mapping $A \times A \to A$. The rank $R(A)$ of A is the rank of its multiplication map.

If we think in coordinates, we get the tensor that corresponds to A as follows. Choose a basis $x_1, \dots, x_n$ of A. The product of any two elements of A is again an element of A and can be written as a linear combination of $x_1, \dots, x_n$. In particular

$$x_i \cdot x_j = \sum_{k=1}^{n} \alpha_{i,j,k} x_k.$$

The so-called *structural constants* $\alpha_{i,j,k}$ are the entries of the tensor (with respect to the chosen basis). Since a change of basis is an isomorphim of vector spaces, we get that the rank of this tensor is independent of the chosen basis.

The best lower bounds for the rank of an algebra and for any other tensor of order three are of the form $3 \dim A - o(\dim A)$. Very recently, Landsberg proved this for the algebra $k^{n \times n}$ of $n \times n$-matrices. An earlier example with an easier proof is the algebra $k[X_1, \dots, X_n]/I_d$ where I_d is the ideal generated by all monomials of degree d, see [2]. Because the families of algebras have a "regular" structure, it is clear that the corresponding tensors are explicit. We just have to compute the structural constants. For instance, in the case of the algebra $k^{n \times n}$ with the standard basis, we have

$$e_{i,i'} e_{j,j'} = \begin{cases} e_{i,j'} & \text{if } i' = j, \\ 0 & \text{otherwise.} \end{cases}$$

(Note that we use double indices since $\dim k^{n \times n} = n^2$.) In the second case, if we take all monomials of degree $< d$ as a basis, we get a similar expression.

It is an major open problem to find explicit tensors or explicit families of algebras with a larger rank.

Open Problem 4.
1. *Is there an explicit family of tensors* $t_n \in k^n \otimes k^n \otimes k^n$ *with* $R(t_n) \geq (3 + \epsilon)n$ *for some* $\epsilon > 0$.
2. *Can we even achieve this for tensors corresponding to the multiplication in an algebra, i.e., is there an explicit family of algebras* A_n *with* $R(A_n) \geq (3 + \epsilon) \dim A_n$ *for some* $\epsilon > 0$. *Of course,* $\dim A_n$ *should go to infinity.*

3.2. From tensors of order three to higher order tensors

We can use the lower bounds of the rank of tensors of order three to obtains bounds for the rank of higher order tensors. Up to lower order terms, they match the current best lower bounds (see the next section).

Let $t \in V_1 \otimes \cdots \otimes V_n$. Let $I_1, \ldots, I_m$ be a partition of $\{1, \ldots, n\}$, that is, the I_j are pairwise disjoint and their union is $\{1, \ldots, n\}$. Let $U_j = \bigotimes_{i \in I_j} V_i$ for $1 \le j \le m$. We can view t as an element of $U_1 \otimes \cdots \otimes U_m$. Note that the rank of t as an element of $U_1 \otimes \cdots \otimes U_m$ is a lower bound for the rank of t as an element of $V_1 \otimes \cdots \otimes V_n$. Why? Any rank-one tensor $v_1 \otimes \cdots \otimes v_n \in V_1 \otimes \cdots \otimes V_n$ induces a rank-one tensor $u_1 \otimes \cdots \otimes u_m \in U_1 \otimes \cdots \otimes U_m$ by setting $u_j = \bigotimes_{k \in I_j} v_k$. When it is not clear from context, whether we think of t being a tensor in $U_1 \otimes \cdots \otimes U_m$ or $V_1 \otimes \cdots \otimes V_n$, we add it as a subscript.

Lemma 2. $R_{U_1 \otimes \cdots \otimes U_m}(t) \le R_{V_1 \otimes \cdots \otimes V_m}(t)$. □

The rank can indeed become smaller. Consider $\langle n, n, n \rangle \in k^{n \times n} \otimes k^{n \times n} \otimes k^{n \times n}$, the tensor of matrix multiplication. If we consider it as a tensor in $(k^{n \times n} \otimes k^{n \times n}) \otimes k^{n \times n}$, then it is a matrix of size $n^4 \times n^2$. Its rank is at most n^2. However, we know a lower bound of $3n^2 - o(n^2)$ for the rank of $\langle n, n, n \rangle$ as a tensor in $k^{n \times n} \otimes k^{n \times n} \otimes k^{n \times n}$ [11]. In fact, it is an old open problem, whether the so-called exponent of matrix multiplication is two.

3.3. Explicit tensors of higher order

Let d be even and let $N = n^{d/2}$ Take any full rank matrix $M \in k^{N \times N}$, for instance the identity matrix. It has rank $n^{d/2}$. By Lemma 2,

$$R_{\bigotimes_{i=1}^{d} k^n}(M) \ge n^{d/2}. \tag{5}$$

The tensor M is obviously explicit, an entry $m_{i_1, \ldots, i_d} = 1$ if $(i_1, \ldots, i_{d/2}) = (i_{d/2+1}, \ldots, i_d)$ and 0 otherwise. Note that if we could achieve $n^{(1-o(1))d}$, then this will lead to formula lower bounds.

It is a sad state of affairs that (5) is the asymptotically best lower bound for an explicit tensor that we currently know; further improvements just concern the constant factor.

Here is one such improvement which uses a lower bound by Hartmann [9]: Let k be a field and K be an extension field of dimension n. Consider the multiplication of the K-left module $K^{1 \times m}$ as a bilinear map over k. (We take $x \in K$ and $(y_1, \ldots, y_m) \in K^{1 \times m}$ and map them to $(xy_1, \ldots, xy_m) \in K^{1 \times m}$). However, we view this as a k-bilinear map and not as a K-bilinear map. Let $\hat{s}$ be the corresponding tensor. Hartmann showed that

$$R(K^{1 \times m}) \ge (2n - 1)m = 2nm - m. \tag{6}$$

If we now set $m = n^{e-1}$ and let $s \in K^{\otimes(2e+1)}$ be the tensor corresponding to $\hat{s}$, we get

$$R(s) \ge 2n^e - n^{e-1}$$

for a tensor of order $d = 2e + 1$. Note that if d is odd, then the approach above using an invertible matrix only gives the lower bound n^e.

If we take $K = k[X]/(X^n)$ instead of an extension field, then we can show the same bound as (6) and get an other example of an explicit tensor. As a basis of K, we choose the basis $1, X, \dots, X^{n-1}$. This induces a basis of $K^{1,m}$ in the natural way. How does the tensor of the multiplication of $K^{1\times m}$ look like? First consider the case $m = 1$. The tensor looks like

$$\begin{pmatrix} 1 & 2 & 3 & \dots & n \\ 2 & 3 & & \iddots & 0 \\ 3 & & \iddots & 0 & 0 \\ \vdots & \iddots & \iddots & \vdots & \vdots \\ n & 0 & \dots & 0 & 0 \end{pmatrix} \tag{7}$$

How to interpret this? It is a $\{0,1\}$-valued tensor of size $n \times n \times n$. An entry $k > 0$ in position (i, j) means that the entry in position (i, j, k) is 1. All other entries are 0 whether it is explicitly indicated or not. The tensor for arbitrary m looks like follows:

$$T = \begin{pmatrix} T_1 \\ T_2 \\ \vdots \\ T_m \end{pmatrix}.$$

Each T_j is a copy of the tensor in (7). However, these tensors T_j live in different slices. T is a tensor in $k^{n\times nm\times nm}$, each T_j lives in the slices $(j-1)n+1, \dots, jn$ in the third component. Now, as in the beginning, we want to apply substitution method. We will only work with the copy T_1, kill $2n-1$ products, and then we can simply apply induction. In an optimal decomposition of T into rank-one tensors

$$T = \sum_{i=1}^{r} u_i \otimes v_i \otimes w_i,$$

we can assume that $w_1, \dots, w_{n-1}$ restricted to the first $n - 1$ coordinates, are linearly independent. Let h be the projection along the linear span of $w_1, \dots, w_{n-1}$ onto $\langle e_n, e_{n+1}, \dots, e_{mn}\rangle$. Here $e_i \in k^{mn}$ is the ith unit vector and $\langle \dots \rangle$ the linear span. Applying h, we kill $n-1$ products in the decomposition of T. What happens to T under this homomorphism? Note that only the first n rows of T are affected, which just contain the copy T_1. In (7), h maps multiples of the slices $1, \dots, n - 1$ onto the nth one. The result is a lower triangular matrix with all 1s on the diagonal. Therefore, the matrix has full rank. Like before, we can now kill another n products and the tensor that is still computed is

$$T' = \begin{pmatrix} T_2 \\ T_3 \\ \vdots \\ T_m \end{pmatrix}.$$

Therefore, we can proceed by induction and get a lower bound of $(2n-1)m$.

4. Explicit tensors by combinatorial constructions

The currently best lower bound for an explicit tensor is due to [1]. It improves on the lower order term of the construction in the last section.

Let $\ell = \lfloor \log_2 n \rfloor$. For $i \in \{0, \dots, \ell\}$, we recursively define the following matrices:

1. $S_{1,0} = (1)$.
2. For even $n = 2m > 1$,

$$S_{n,i} = \begin{cases} \begin{pmatrix} 0 & 0 \\ S_{m,i} & 0 \end{pmatrix} & \text{if } i < \ell \\ \begin{pmatrix} I_m & 0 \\ 0 & I_m \end{pmatrix} & \text{otherwise} \end{cases}.$$

3. For odd $n = 2m + 1 > 1$,

$$S_{n,i} = \begin{cases} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ S_{m,i} & 0 & 0 \end{pmatrix} & \text{if } i < \ell \\ \begin{pmatrix} 0 & 0 & 0 \\ I_m & 0 & 0 \\ 0 & I_m & 0 \end{pmatrix} & \text{otherwise} \end{cases}.$$

Finally, let T_n be the tensor consisting of slices $S_{n,0}, \dots, S_{n,\ell}$. The format of T_n is $n \times n \times (\ell+1)$. T_n is certainly explicit, we can determine the entries by following the recursive structure.

Theorem 4. *$R(T_n) \geq 2n - 2h(n) + 1$ where $h(n)$ is the number of 1s in the binary expansion of n.*

Note that $h(n) \leq \log n$. We can use the substitution method to prove the theorem. Certainly $R(T_1) = 1$ holds. If $n = 2m$ is even, then we can "substitute away" the two identity matrices, killing n products. The remaining tensor is $T_{n/2}$. Therefore, we get the recurrence

$$R(T_{2m}) \geq R(T_m) + 2m.$$

In the same way, we get

$$R(T_{2m+1}) \geq R(T_m) + 2m.$$

It is easy to verify that the bound stated in the theorem is the solution of this recurrence. We can now extend this to a tensor of size $n \times (n+1) \times n$ by extending the slices above by one column and then add $n - (\ell + 1)$ linearly independent slices, which just have a one in the extra column. These slices can be substituted away and we get a tensor of size $n \times (n+1) \times n$ with rank bounded by $3n - \Theta(\log n)$. If we set $n = m^e$ and add just m extra slices, we get a tensor of size $m^e \times (m^e + 1) \times m^e$ the rank of which can be lower bounded by $2n + m - \Theta(\log n) = 2m^e + m - \Theta(d \log m)$. As before, we can interpret this tensor as a tensor of order $2e+1$. (The $+1$ in m^e+1 the second component disturbs this construction a little bit, to remedy this, we can start with a tensor of size $(n-1) \times (n-1) \times (n-1)$ and then extend this to a tensor of size $n \times n \times n$ by adding zeros.)

If you look at the construction closely, we can view this again as a tensor related to an algebra, as pointed by Landsberg [13]. For simplicity, assume that n is a power of 2. Otherwise, the tensor will just have some additional zeros. It is quite easy to see, that up to permutations, the tensor T_n are just the slices $1, 2, \dots, 2^i, \dots, 2^\ell$ of the tensor of the algebra $k[X]/(X^n)$. For instance, T_8 has the form

$$T_8 = \begin{pmatrix} 1 & 2 & & 3 & & & & 4 \\ 2 & & 3 & & & & 4 & \\ & 3 & & & & 4 & & \\ 3 & & & & 4 & & & \\ & & & 4 & & & & \\ & & 4 & & & & & \\ & 4 & & & & & & \\ 4 & & & & & & & \end{pmatrix}$$

Looking at this, it is easy to see that we can project away the lower four rows and four columns to the righthand side. This reduces the rank by 8 (in general by n) and affects only the fourth slice. We can remove this slice and get T_4. Using induction, we get a lower bound of $1+2+4+8 = 15$ (in general $2n-1$, note that n is a power of 2).

5. Conclusions

In the examples we have seen, the bounds on the rank are proven via the substitution method. The bounds that are achievable with this method are usually limited by the sum of the dimension of the vector spaces. Up to lower order terms, our constructions reach this limit. To get better bounds, new lower bound techniques are needed. One promising approach for this is the geomtric complexity approach by Bürgisser and Ikenmeyer [5, 6].

References

[1] Boris Alexeev, Michael A. Forbes, Jacob Tsimerman: Tensor Rank: Some Lower and Upper Bounds. *IEEE Conference on Computational Complexity 2011*: 283-291.

[2] Markus Bläser: Improvements of the Alder-Strassen Bound: Algebras with Nonzero Radical. *ICALP 2001*: 79-91.

[3] Peter Bürgisser, Michael Clausen, M. Amin Shokrollahi: *Algebraic Complexity Theory*, Springer, 1997.

[4] Peter Bürgisser. *Completeness and Reduction in Algebraic Complexity Theory*, Springer, 2000.

[5] Peter Bürgisser, Christian Ikenmeyer: Geometric complexity theory and tensor rank. *STOC 2011*: 509-518.

[6] Peter Brgisser, Christian Ikenmeyer: Explicit Lower Bounds via Geometric Complexity Theory. *CoRR abs/1210.8368* (2012).

[7] John F. Canny: Some Algebraic and Geometric Computations in PSPACE. *STOC 1988*: 460-467.

[8] Don Coppersmith, Shmuel Winograd: Matrix Multiplication via Arithmetic Progressions. *J. Symb. Comput.* 9(3): 251-280 (1990).

[9] Werner Hartmann: On the Multiplicative Complexity of Modules Over Associative Algebras. *SIAM J. Comput.* 14(2): 383-395 (1985).

[10] Johan Håstad: Tensor Rank is NP-Complete. *J. Algorithms* 11(4): 644-654 (1990).

[11] J.M. Landsberg: New lower bounds for the rank of matrix multiplication. *arXiv:1206.1530v1 [cs.CC]*, 2012.

[12] J.M. Landsberg: *Tensors: Geometry and applications*, AMS, 2011.

[13] J.M: Landsberg: Nontriviality of equations and eplicit tensors in $\mathbb{C}^m \otimes \mathbb{C}^m \otimes \mathbb{C}^m$ of border rank at least $2m-1$. *arXiv:1209.1664v2*, 2013.

[14] Thomas Lickteig: Typical tensor rank. *Lin. Alg. Appl.* 69: 95–120 (1985).

[15] K. Kalorkoti: A Lower Bound for the Formula Size of Rational Functions. *SIAM J. Comput.* 14(3): 678-687 (1985).

[16] Victor Ya. Pan: Methods for computing values of polynomials. *Russ. Math. Surv.* 21: 105–136 (1966).

[17] Ran Raz: Tensor-rank and lower bounds for arithmetic formulas. *STOC 2010*: 659-666.

[18] James Renegar: On the Computational Complexity and Geometry of the First-Order Theory of the Reals, Part I: Introduction. Preliminaries. The Geometry of Semi-Algebraic Sets. The Decision Problem for the Existential Theory of the Reals. *J. Symb. Comput.* 13(3): 255-300 (1992).

[19] Andrew Stothers: *On the Complexity of Matrix Multiplication*, PhD thesis, University of Edinburgh, 2010.

[20] Volker Strassen: Vermeidung von Divisionen. *Crelle's J. Reine Angew. Math.* 264:184–202 (1973).

[21] Volker Strassen: Rank and optimal computation of generic tensors. *Lin. Alg. Appl.* 52: 645–685 (1983).

[22] Virginia Vassilevska Williams: Multiplying matrices faster than Coppersmith-Winograd. *STOC 2012*: 887-898.

Markus Bläser
Computer Science
Saarland University
Saarbrücken, Germany
e-mail: `mblaeser@cs.uni-saarland.de`