Exploring Tensor Rank

Benjamin Weitz

October 27, 2011

Abstract

We consider the problem of tensor rank. We define tensor rank, discuss the motivations behind exploring the topic, and give some examples of the difficulties we face when trying to compute tensor rank. Some simpler lower and upper bounds for tensor rank are proven, and two techniques for giving lower bounds are explored. Finally we give one explicit example of a construction of an $n^k \times n^k \times n$ tensors of rank $2n^k - O(n^{k-1})$. As a corollary we obtain an $[n]^r$ shaped tensor with rank $2n^{\lfloor r/2 \rfloor} - O(n^{\lfloor r/2 \rfloor - 1})$ when r is odd, an improvement from the previously best-known construction of $n^{\lfloor r/2 \rfloor}$.

1 Introduction

In this talk we will discuss tensor rank, an important property extending the concept of matrix rank. In particular, we will discuss the difficulty of computing, and even bounding the ranks of arbitrary tensors, in stark contrast to matrix rank. We will discuss some motivation behind finding tensors of large rank, including the complexity of multilinear forms and the recent result of Ran Raz connecting high-rank tensors to functions with super-polynomial circuit lower bounds. Finally, we will discuss some techniques that have been used to prove some of the better known bounds, including partitioning.

2 Definitions

There are many ways to define tensors, but we need only the simplest. An $n_1 \times \cdots \times n_d$ tensor is simply an $n_1 \times \cdots \times n_d$-dimensional array of elements of a field F. We call d the *order* of the tensor, and $(n_1, \ldots, n_d)$ the *dimensions* of the tensor. The space of $n_1 \times \cdots \times n_d$ tensors is denoted $F^{n_1} \otimes \cdots \otimes F^{n_d}$, and is equipped with a natural addition where the sum of two tensors is the tensor of the element-wise sums, and a natural scalar multiplication where cA is the tensor whose elements have all been multiplied by c. To define a concept of rank, we first need the concept of a tensor product, a product that takes vectors and creates tensors:

Definition 1. *Let $x_i \in F^{n_i}$ be an n_i-length vector for each $1 \leq i \leq d$. Then the* tensor product *of $\{x_i\}_{1 \leq i \leq d}$ is the tensor $A \in F^{n_1} \otimes \cdots \otimes F^{n_d}$ whose $i_1 \ldots i_d$th entry is $x_1(i_1) \ldots x_d(i_d)$. We denote this by $A = x_1 \otimes \cdots \otimes x_d$. If there exist vectors such that $A = x_1 \otimes \cdots \otimes x_d$, we call A* simple.

Fact 1. *Not every tensor A is the tensor product of a set of vectors, but the simple tensors span $F^{n_1} \otimes \cdots \otimes F^{n_d}$.*

This allows us to define tensor rank in the following way:

Definition 2. *If $A = 0$, then rank$(A) = 0$. If A is simple, then rank$(A) = 1$. Otherwise, rank$(A) = r$ is the smallest integer such that there exist r simple tensors $B_1, \ldots, B_r$ such that $A = \sum_{i=1}^{r} B_i$.*

This coincides with the definition of rank for matrices, i.e. order two tensors. If a collection of simple tensors $B_1, \ldots, B_k$ satisfy $\sum_{i=1}^{k} B_i = A$, then we call $\{B_1, \ldots, B_k\}$ a *rank-k decomposition* for A. Tensor rank is invariant to any permutation of the indices of the tensor, since any rank-k decomposition of one extends to a rank-k decomposition of the other by just permuting the indices of the simple tensors. We can also characterize rank in another way, using slices of a tensor.

Definition 3. *Let A be a $n_1 \times \cdots \times n_d$ tensor. The $n_1 \times \cdots \times n_{d-1}$ tensor B_k that satisfies*

$$B_k(i_1, \ldots, i_{d-1}) = A(i_1, \ldots i_{d-1}, k)$$

is called the kth slice of A in the dth direction. A tensor A is called nondegenerate *if the slices of A in any direction are linearly independent. The tensor of indeterminates $A(s_1, \ldots, s_{n_d}) = \sum_{i=1}^{n_d} B_i s_i$ is called the* defining tensor. *Since* *we often work with order 3 tensors, this is often a matrix, and easier to work with.*

A tensor A has rank r iff there is a collection of r linearly independent slices $\{D_1, \dots, D_r\}$ that spans the slices of A (in any direction), and any such collection has at least r slices.

3 Motivations

The multiplicative complexity of multilinear forms is exactly the rank of the associated tensor. For example, in the bilinear case, we want to simultaneously evaluate the m bilinear forms

$$B_i = \sum_{j=1}^{p} \sum_{k=1}^{q} x_j \gamma_{ijk} y_k$$

where $T_{ijk} = \gamma_{ijk}$ is the associated tensor. Then the smallest number of multiplications required to compute all B_i is $\operatorname{rank}(T)$. Thus bounding the rank of various tensors lets us bound the complexity of computing multilinear forms. For example, Strassen's matrix multiplication algorithm, which seems very arbitrary at first glance, was discovered by finding a rank-7 decomposition of the tensor of 2×2 matrix multiplication, even though the only obvious decompositions had rank 8.

Another motivation for bounding the rank of tensors is the following theorem proven by Ran Raz:

Theorem 2. *Let $A : [n]^r$ be a tensor such that $r \leq O(\log n / \log \log n)$. If the tensor rank of A is $\geq n^{r \cdot (1 - o(1))}$ then there is no polynomial size formula for the polynomial*

$$f_A(x_{1,1}, \dots, x_{r,n}) = \sum_{i_1, \dots i_r \in [n]} A(i_1, \dots, i_r) \cdot \prod_{j=1}^{r} x_{j, i_j}.$$

Proving super-polynomial circuit lower bounds has been an extremely intractable field for complexity theorists, and has large implications for **P** vs. **NP**. In particular, if an **NP** function can be shown to have super-polynomial circuit lower bounds, then $\mathbf{P} \neq \mathbf{NP}$. Thus if we can find enough high-rank tensors, hopefully one of them will give some insight into this area.

4 Difficulties

Unfortunately, the general problem of computing tensor rank has been shown to be very difficult. Hastad proved that, given an order-3 tensor A and a positive integer r, asking whether $\operatorname{rank}(A) \leq r$ is **NP**-complete. This means that unless $\mathbf{P} = \mathbf{NP}$, something thought to be very unlikely, the problem of bounding tensor rank does not have polynomial-time algorithms. This approximately means that brute-forcing every possible decomposition is the only way we know how to compute tensor rank, and this is an exponential-time algorithm. Hastad's proof is a little too complicated to completely reproduce, but he gives a reduction from the satisfiability problem, i.e. from a logical 3-CNF ϕ, he constructs a tensor $A(\phi)$ that has rank at least r iff ϕ is satisfiable. Thus if we could compute tensor rank in polynomial-time, we could determine if ϕ was satisfiable in polynomial-time, which would imply $\mathbf{P} = \mathbf{NP}$.

5 Some Known Bounds

A simple counting argument gives an implicit argument for the existence of tensors of high rank.

Theorem 3. *Consider tensors of shape $n_1 \otimes \cdots \otimes n_d$, and let $S = \sum_{i=1}^{d} n_i$ and $P = \prod_{i=1}^{d} n_i$. Then there exist tensors of rank at least $r \geq P/S$.*

This theorem is proven by simply considering that there are only at most q^{rS} tensors of rank less than or equal to r, and there are q^P tensors total. However, this theorem is unhelpful for applications to Raz's theorem, because a similar counting argument holds for circuit lower bounds, but we want explicit functions that require super-polynomial circuits, so we can try to argue that those functions are in **NP**. Similarly, we require explicit tensors so we can argue that their associated functions are in **NP**. A similar upper bound can be shown as well:

Theorem 4. *Let A be a tensor of shape $n_1 \otimes \cdots \otimes n_d$, and let $P_i = \prod_{j \neq i} n_j$. Then rank$(A) \leq \min_i P_i$.*

Furthermore, Joseph JaJa' proved that for $n \times n \times 2$ tensors, the maximum achievable rank is $3n/2$. This is one of the only non-trivial shapes that we have a complete result for.

6 Some Techniques

There are two main techniques we have for lower bounding tensor rank. The first is based on mapping the tensor into a tensor space with smaller order, but larger field size.

Theorem 5. *Let F be a field and U, V, and W be F-vector spaces. Let $Y = U \otimes V$. There is an isomorphism ϕ between $U \otimes V \otimes W$ and $Y \otimes W$ such that $u \otimes v \otimes w \mapsto (u \otimes v) \otimes w$. If $A \in U \otimes V \otimes W$, and $R(A)$ denotes the rank of A, then $R(\phi(A)) \leq R(A)$.*

In particular, we can use this theorem to construct $\underbrace{n \times \cdots \times n}_{r \text{ times}}$ tensors of rank at least $n^{\lfloor r/2 \rfloor}$, and if r is odd, tensors of rank $3n^{\lfloor r/2 \rfloor}/2$. The second technique is based on partitioning the tensor along a direction into two different tensors, and if the two parts don't "overlap very much" then the original tensor has rank at least the rank of one plus the number of slices of the other.

Theorem 6. *If $A(s_1, \ldots, s_{n_d})$ is a nondegenerate defining tensor, and $A(s_1, \ldots, s_{n_d}) = B(s_1, \ldots, s_k) + C(s_{k+1}, \ldots, s_{n_d})$, then*

$$\text{rank}(A) \geq \min_T \text{rank}(B(s_1, \ldots, s_k) + C(T(s_1, \ldots, s_k))) + (n_d - k)$$

where T is a $(n_d - k) \times k$ matrix.

7 One Tensor of Medium Rank

While we have not yet been able to give any explicit tensors of high rank, using partitioning and the isomorphism theorem, we can give a tensor of high-ish rank, better than the previous best-known by a constant factor.

Theorem 7. *There is an* $\underbrace{n \times \cdots \times n}_{r\ times}$ *tensor* A *that is constructible in polynomial time such that rank*$(A) \geq 2n^{\lfloor r/2 \rfloor} - O(n^{\lfloor r/2 \rfloor - 1})$ *when* r *is odd.*

This tensor is essentially recursively constructed by taking the tensor A_{i+1} with slices $\begin{bmatrix} A_i & 0 \\ 0 & A_i \end{bmatrix}$ and $\begin{bmatrix} 0 & A_i \\ 0 & 0 \end{bmatrix}$, and taking the image of this tensor from $n^{(r-1)/2} \times n^{(r-1)/2} \times n$ into $n \times \cdots \times n$.