On Lower Bounds on the Size of Sums-of-Squares Formulas

Daniel M. Kane
Massachusetts Institute of Technology
dankane@mit.edu

December 31, 2005

Abstract

For sums of squares formulas of the form

$$(x_1^2 + \cdots + x_r^2)(y_1^2 + \cdots + y_s^2) = z_1^2 + \cdots + z_t^2$$

where the z_i are bilinear functions of the x_i and y_i, we have two well-known lower bounds on the size of t given r and s. One was obtained independently by Hopf and Stiefel, and another by Atyiah. These bounds are given by requiring certain binomial coefficients be divisible by certain powers of 2. Although the behavior of the Hopf-Stiefel bound is fairly well understood, the Atyiah bound is not. In this paper we provide an efficient algorithm for computing the Atyiah bound and some results on which of the lower bounds is largest.

1 Introduction

In this paper we analyze the relative sizes of some lower bounds on the sizes of sums of squares formulae. For given r and s we wish to find lower bounds on the smallest t such that there exists a formula of the form

$$(x_1^2 + \cdots + x_r^2)(y_1^2 + \cdots + y_s^2) = z_1^2 + \cdots + z_t^2$$

where the z_i are bilinear functions of the $x's$ and $y's$. Such sums of squares formulas frequently have relations to important algebraic structures. For example, the multiplicativity of the norm of the complex numbers is equivalent to the sums of squares formula

$$(x_1^2 + x_2^2)(y_1^2 + y_2^2) = (x_1y_1 - x_2y_2)^2 + (x_1y_2 + x_2y_1)^2.$$

Also note that the multiplicativity of the quaternion norm is equivalent to a sums of squares formula with $r = s = t = 4$. Note that if we have any sums of squares formula we can increase the value of t by letting $z_{t+1} = 0$. Hence the

question of interest is to determine the smallest possible value of t for a given r and s.

As far as we know, there are two known lower bounds for the size of t as a function of r and s. In 1940, Hopf and Stiefel proved independently that for there to be a sum of squares formula, we must have that $\binom{t}{i}$ is a multiple of 2 for all $t-r < i < s$ (see [4] and [6]). Furthermore, Atiyah proved that if $c = \lfloor \frac{s-1}{2} \rfloor + 1$, $\binom{t}{i}$ must be divisible by 2^{c-i} for all $t - r < i < c$. These both provide natural lower bounds of the t required for a sum of squares formula with given r and s. We denote these bounds by $HS(r,s)$ and $A(r,s)$, respectively. In recent papers by Dugger and Isaksen ([1] and [2]), these conditions were generalized to sums of squares formulas over any field of characteristic not equal to 2.

The bound that Hopf and Stiefel's condition produces can be readily computed using a simple recursion. On the other hand, Atiyah's bound is more complicated to compute exactly. In this paper, we discuss ways of computing and approximating the bound on t implied by Atiyah's condition as well as ways to compute it exactly. Moreover, we discuss which of the given lower bounds on t is largest for any given r and s.

In Section 2, we introduce our notation. In Section 3, we will develop an algorithm compute $A(r,s)$ in $O(\log(rs))$ bit operations. This is significantly faster than the naive algorithm, and optimal in terms of bit operations. In Section 4, we prove some results about the relative sizes of $HS(r,s)$, $A(r,s)$ and $A(s,r)$. We determine the density of r and s so that $HS(r,s) \geq A(r,s)$ as well as provide conditions on r and s that usually determine when this is the case. We also show that if $r \geq s$ then $A(r,s) \geq A(s,r)$ unless r and s are very close to each other, and provide some conditions on when this does and does not hold. In Section 5, we discuss further directions of inquiry.

2 Notation

Definition. *Let $e(n)$ be the largest number m so that $2^m | n$. In other words $e(n)$ is the number of times that 2 divides n.*

Definition. *$HS(r,s)$ is the minimum t so that $e\left(\binom{t}{i}\right) \geq 1$ for all $t-r < i < s$.*

$HS(r,s)$ is the lower bound on t proven by Hopf and Stiefel.

It can be shown that we can compute HS using the following recursion:

$$HS(r,s) = HS(s,r).$$

$$HS(r,s) = \begin{cases} 2^n & \text{if } 2^{n-1} < r, s \leq 2^n \\ 2^n + HS(r, s-2^n) & \text{if } r \leq 2^n < s. \end{cases}$$

Notice that this gives us the following way to compute $HS(r,s)$. Consider the binary representations of r and s ending in repeating 1's after the decimal point. Find the most significant bit where both are 1. Replace all less significant bits in both numbers by 0, and add the resulting numbers to get t. It is easy to verify that the resulting function satisfies the same recursion as HS.

Definition. *We let $A(r,s)$ denote the smallest number t so that with $c = \lfloor \frac{s-1}{2} \rfloor + 1$, we have $e(\binom{t}{i}) \geq c - i$ for all $i > t - r$.*

$A(r,s)$ corresponds to the lower bound on t derived from the work of Atiyah.

This paper makes extensive use of asymptotic notation. Recall that $O(f)$ refers to some quantity whose absolute value is bounded above by some constant times f for all values of the input parameters. $\Omega(f)$ refers to some quantity whose absolute value is bounded below by some constant times f for all values of the input parameters. $\Theta(f)$ refers to some quantity that is both $O(f)$ and $\Omega(f)$.

3 An Algorithm for Computing $A(r,s)$

In this section we develop an algorithm for computing $A(r,s)$ in $O(\log(rs))$ bit operations.

Notice that $t = r + c \geq A(r,s)$ since then there is no i satisfying $t - r < i < c$.

Lemma 1. *$e(n!)$ equals n minus the number of 1's in the binary representation of n. Also $e(n!) = n - O(\log n)$.*

Proof. The first statement follows from induction on n. It clearly holds for $n = 0$. The number of 1's in the binary representation of n minus the number in $n+1$ is the number of trailing 0's in $n+1$ minus 1, or $e(n+1) - 1$. Hence, if $e(n!)$ is n minus the number of 1's in the binary representation of n, $e((n+1)!) = e(n+1) + e(n!)$ which equals n minus the number of 1's in the binary representation of n plus 1, plus the number of 1's in the binary representation of $n+1$ minus the number in this binary representation of n. This equals $n+1$ minus the number of 1's in the binary representation of $n+1$. This completes the proof of the first statement, from which the second statement easily follows. □

We will show that $A(r,s) = r + \frac{s}{2} - O(\log(rs))$. Letting $t = A(r,s)$ and $i = t - r + 1$, we have that $e(\binom{t}{i}) = O(\log t) = c + r - t$. Hence we have that $A(r,s) = r + \frac{s}{2} + O(\log(rs))$. This gives us a method for computing A.

Letting $c = \lfloor \frac{s-1}{2} \rfloor + 1$, we will show that $A(r,s)$ is the smallest value of t so that $e\left(\binom{t+i-1}{t-r+i}\right) \geq r + c - t - i$ for all $i \geq 1$. Let Condition 1 on c, r and t be that $e(\binom{t}{i}) \geq c - i$ for all $i \geq t - r$. Let Condition 2 on c, r and t be that $e(\binom{t+i-1}{t+i-r}) \geq c + r - t - i$ for all $i \geq 1$. We will now show that the Conditions 1 and 2 are equivalent. By the recursive formula for binomial coefficients, Condition 1 on c, r and t implies Condition 1 on c, r and $t+1$. Applying the inequalities implied by Condition 1 on c, r and $t+n$ for the correct values of n to the binomial coefficients we get by substituting $c, r, t+i-1$ and $t-r+i$ for c, r, t and i we get Condition 2. By the recursive formula for binomial coefficients Condition 2 for c, r and t implies Condition 2 for $c, r-1$ and t. Applying Condition 2 substituting $c, r-i, t$ and 1 for c, r, t and i we get Condition 1.

Hence to compute $A(r,s)$ all we need to do is compute the values of $e(\binom{r+c-1-n}{c-n})$, for $1 \le n \le O(\log(rs))$ and find the first one that is less than n. $A(r,s)$ is then $r+c-n$ for that value of n.

Since each binomial coefficient in this sequence is a rational number times the preceding one we can easily compute the next e value from the previous one. By above we only need to check $O(\log(rs))$ terms. To compute the next term from the preceding term we need to decrement t and $t-r+1$ and compute e of these numbers. This takes $O(1)$ bit operations per trailing zero. Since the total number of trailing zeroes in a sequence of m numbers is $O(m)$ plus the greatest number of trailing zeroes in any given term, the total number of trailing zeros is $O(\log(rs))$. Hence our total runtime is $O(\log(rs))$ bit operations to compute $A(r,s)$.

4 Comparing $HS(r,s), A(r,s)$ and $A(s,r)$

In this section, we look at the problem of computing the relative sizes of $HS(r,s), A(r,s)$ and $A(s,r)$.

On the basis of computations performed by Armira Shkembi and Matthew Buckman, Isaksesn conjectured that $HS(r,s) \ge A(r,s)$ and $A(r,s) \ge HS(r,s)$ each occur infinitely often but that $HS(r,s) \ge A(r,s)$ occurs more frequently. We prove both of these conjectures with the following theorem:

Theorem 2. *If r and s are picked randomly, uniformly and independently from the set $\{1,2,\ldots,2^n\}$, then the probability that $HS(r,s) \ge A(r,s)$ is $\frac{5}{8} + O(n^2 2^{-n})$.*

Proof. $HS(r,s)$ is $(r-1)+(s-1)+O(1)$ minus the sums of the values of the bits that are less significant than the first bit where $r-1$ and $s-1$ are both 1. $A(r,s)$ is $(r-1)+(s-1)-\frac{s-1}{2}+O(n)$.

Let $s-1 = 2^m + f_s$ and $r-1 = k \cdot 2^m + f_r$ where k is an integer and $0 \le f_r, f_s < 2^m$ (2^m is the largest power of 2 less than s). If k is even, then $HS(r,s) \ge (r-1)+(s-1)-\frac{f_r}{2}-\frac{f_s}{2}$. Hence $HS(r,s)-A(r,s) \ge 2^{m-1}-\frac{f_r}{2}+O(n)$, which is positive with probability $1-O(n2^{-m})$ since f_r is equally likely any number in the range $0,\ldots,2^m-1$.

If k is odd, then $HS(r,s) = (r-1)+(s-1)-f_r-f_s$. So $HS(r,s) - A(r,s) = 2^{m-1} - f_r - \frac{f_s}{2} + O(n)$. For fixed f_s, this is positive with probability $\frac{1}{2} - \frac{f_s}{2^{m+1}} + O(n2^{-m})$. So averaging over f_s, $HS(r,s) - A(r,s)$ is positive with probability $\frac{1}{4} + O(n2^{-m})$.

Since for fixed m, k is even with probability $\frac{1}{2}$, we have for fixed m, that $HS(r,s) > A(r,s)$ with probability $\frac{1}{2}(1+\frac{1}{4}+O(n2^{-m})) = \frac{5}{8}+O(n2^{-m})$. Therefore, since $m = m_0$ with probability $O(2^{m_0-n})$, averaging over all possible values of m, the probability that $HS(r,s) > A(r,s)$ equals $\frac{5}{8}+\sum_{m=0}^{n-1} O(n2^{-m}2^{m-n}) = \frac{5}{8}+O(n^2 2^{-n}) \to \frac{5}{8}$ as $n \to \infty$. □

Lemma 3. *$e(\binom{n}{m})$ is the number of 1's in the binary representation of m plus the number of 1's in the binary representation of $(n-m)$ minus the number of 1's in the binary representation of n.*

Proof. This follows easily from the fact that $e(n!)$ is n minus the number of 1's in the binary representation of n. □

Lemma 4. $e((a+1)(a+2)\cdots(a+n)) = n + O(\log n) + \max_{1\le i\le n} e(a+i)$.

Proof. Let $k = \max_{1\le i\le n} e(a+i)$. Then the expression we are looking for is

$$\sum_{i=1}^{k}\left(\lfloor\frac{a+n}{2^i}\rfloor - \lfloor\frac{a}{2^i}\rfloor\right) = \sum_{i=1}^{\lfloor\log_2 n\rfloor}\left(\frac{n}{2^i}+O(1)\right) + \sum_{i=\lfloor\log_2 n\rfloor+1}^{k} 1 = n + O(\log n) + k.$$

□

Theorem 5. *Let $s = k\cdot 2^N + 1$ where $N = \Omega(\log\log k)$ and $r = k\cdot 2^N + N$. If k has some pair of repeated 1's in its binary representation, then $A(s,r) - A(r,s) \ge \frac{N}{2} + O(\log N + \log k)$.*

Proof. We will assume without loss of generality that N is even.

First we show that $A(s,r) \ge 3k2^{N-1} + \frac{N}{2} + O(\log N + \log k)$. By Lemma 3 it follows that $e(\binom{3k2^{N-1}+\frac{N}{2}-n+1}{k2^{N-1}+\frac{N}{2}-n+1}) \le O(\log N + \log k)$ for $n = O(N)$. Hence if $n = \Omega(\log N + \log k)$, then $A(s,r) \ge s + r/2 - n$.

Next we show that $A(r,s) \le 3k2^{N-1} + O(\log N + \log k)$. We will look at two cases.

First we note that since k has two consecutive 1's in its binary representation, we need to carry once when we add $k2^{N-1}$ and $k2^N + l$ for $l = O(N)$. Hence $e(\binom{3k2^{N-1}+N-l}{k2^{N-1}}) \ge 1$. This implies that for $t > 3k2^{N-1}$ we have $e(\binom{t}{c-1}) \ge 1$.

Next, we consider $i < c-1$. Let $n = \Omega(\log N + \log k)$. Consider $e(\binom{3k2^{N-1}+n}{k2^{N-1}-l})$ where $l \le N - n$. Then by Lemma 3 this is at least $N + O(\log N + \log k)$ since $k2^{N-1} - l$ has at least $N + O(\log N)$ 1's in its binary representation, and $3k2^{N-1} + n$ has at most $O(\log N + \log k)$. Hence if n is a sufficiently large constant times $(\log N + \log k)$, then this is larger than l for all such l. This proves that $A(r,s) \le 3k2^{N-1} + O(\log N + \log k)$.

Hence $A(s,r) - A(r,s) \ge \frac{N}{2} + O(\log N + \log k)$. □

Theorem 5 proves that near most multiples of a power of 2 we find a counter-example to the conjecture (see [5]) that $r \ge s \Rightarrow A(r,s) \ge A(s,r)$ with both $r - s$ and $A(s,r) - A(r,s)$ asymptotically as large as possible since using the asymptotic form for $A(r,s)$, these are both clearly as bad as they can possibly get.

Theorem 5 implies that there exists one of these extreme counter-examples near most multiples of a power of 2. We next show that all extreme counter-examples are near large powers of 2.

Theorem 6. *If $r > s$ and $A(s,r) > A(r,s)$, then letting $N = \frac{r-s}{2} + A(s,r) - A(r,s)$ there is a multiple of $2^{(N-O(\log\log r))}$ within $O(\log r)$ of r.*

Proof. Let $t_1 = A(s,r)$, $t_2 = A(r,s)$, $c_1 = \lfloor \frac{r-1}{2} \rfloor + 1$ and $c_2 = \lfloor \frac{s-1}{2} \rfloor + 1$. Note that $t_1 - t_2$ and $r - s$ are both $O(\log r)$. For some $c_1 > i > t_1 - s - 1$, we have $e(\binom{t_1-1}{i}) < c_1 - t_1 + s$. Also notice that $e(\binom{t_2}{t_2-r+1}) \geq c_2 + r - t_2 - 1$. Hence we have that $e(\binom{t_2}{t_2-r+1}) - e(\binom{t_1-1}{i}) \geq \frac{r-s}{2} + (t_1 - t_2) + O(1)$. Therefore since $t_2 \leq t_1 - 1$, we have that $e(\frac{(i)!}{(t_2-r+1)!}) - e(\frac{(r-1)!}{(t_1-1-i)!}) - e(\frac{t_2!}{t_1!}) > N + O(1)$. Since $i < c_1$ and $r - 1 > s - 1 \geq t_1 - 1 - i$, we also have that $e(\frac{(c_1)!}{(t_2-r+1)!}) > N + O(1)$. By Lemma 4 this implies that 2^k is the greatest power of 2 dividing some number between c_1 and $t_2 - r$, then $k > N + O(\log((c_1) - (t_2 - r + 1))) = N + O(\log\log r)$. Since $r - s = O(\log r)$, both c_1 and $t_2 - r$ are within $O(\log r)$ of $r/2$, so there is a multiple of $2^{N+O(\log\log r)}$ within $O(\log r)$ of $r/2$. Doubling this we get a multiple of $2^{N+O(\log\log r)}$ within $O(\log r)$ of r. □

5 Conclusion

We have presented an efficient means of computing $A(r,s)$ and some asymptotic results on the relative sizes of $HS(r,s)$, $A(r,s)$ and $A(s,r)$. These results could be improved by coming up with a nice closed formula or recurrence relation for computing $A(r,s)$. It would also be nice to have a simple criterion to determine which of the three bounds is largest. Both of these goals seem to be quite difficult. Some slightly less ambitious goals would be to derive some of these results for special values of r and s such as $r = s$ or when r and s are both powers of 2.

Acknowledgements: I am thankful to Dan Isaksen for providing me with the topic for this paper and for providing reference [5]. This paper was written at the University of Minnesota, Duluth REU program under the supervision of Joseph Gallian, with support from the NSF (DMS 0447070) and NSA (H98230-04-1-0050).

References

[1] D. Dugger and D. C. Isaksen, The Hopf Condition for bilinear forms over arbitrary fields, preprint, 2003.

[2] D. Dugger and D. C. Isaksen, Algebriac K-theory and sums-of-squares formulas, *Doc. Math.*, to appear.

[3] M. F. Atiyah, Immersions and imbeddings of manifolds, *Topology* **1** (1962) 125-132.

[4] H. Hopf, Ein topologishcer Beitrag zur reelen Algebra, *Comment. Math. Helv.* **13** (1940/41), 219-239.

[5] Dan Isaksen, personal communication.

[6] E. Stiefel, Über Richtungsfelder in den projektiven Räumen und einen Satz aus der reelen Algebra, *Comment. Math. Helv.* **13** (1940/41), 201-218