When Almost All Generalized Sumsets Are Difference-Dominated

Ginny Hogan (ginny6@stanford.edu) and
Steven J. Miller (sjm1@williams.edu)

http://www.williams.edu/Mathematics/sjmiller/public_html

CANT 2013: May 21, 2013

Introduction

Statement

A finite set of integers, $|A|$ its size. Form

- Sumset: $A + A = \{a_i + a_j : a_j, a_j \in A\}$.
- Difference set: $A - A = \{a_i - a_j : a_j, a_j \in A\}$.

Statement

A finite set of integers, $|A|$ its size. Form

- Sumset: $A + A = \{a_i + a_j : a_j, a_j \in A\}$.
- Difference set: $A - A = \{a_i - a_j : a_j, a_j \in A\}$.

Definition

We say A is difference dominated if $|A - A| > |A + A|$, balanced if $|A - A| = |A + A|$ and sum dominated (or an MSTD set) if $|A + A| > |A - A|$.

Questions

We expect a generic set to be difference dominated:

- Addition is commutative, subtraction isn't.
- Generic pair (a_i, a_j) gives 1 sum, 2 differences.

Questions

We expect a generic set to be difference dominated:

- Addition is commutative, subtraction isn't.
- Generic pair (a_i, a_j) gives 1 sum, 2 differences.

Questions

- What happens when we increase the number of summands?
- What happens if we let the probability of choosing elements decay with N?

Past Results

- Martin and O'Bryant, 2006: Positive percentage of sets are MSTD when chosen with uniform probability.

Past Results

- Martin and O'Bryant, 2006: Positive percentage of sets are MSTD when chosen with uniform probability.
- Iyer, Lazarev, Miller, Zhang, 2011: Generalized results above to an arbitrary number of summands.

Past Results

- Hegarty and Miller, 2008: When elements chosen with probability $p(N) \to 0$ as $N \to \infty$, then $|A - A| > |A + A|$ almost surely.

Past Results

- Hegarty and Miller, 2008: When elements chosen with probability $p(N) \to 0$ as $N \to \infty$, then $|A - A| > |A + A|$ almost surely.
- Found critical value of $\delta = \frac{1}{2}$ for probability $p(N) = cN^{-\delta}$, $\delta \in (0, 1)$.
- Critical value is a phase transition because at critical value, the number of repeated elements is on the same order as the number of distinct elements.

Generalized Sumsets

Definition

For $s > d$, consider the Generalized Sumset $A_{s,d} = A + \cdots + A - A - \cdots - A$ where we have s plus signs and d minus signs. Let $h = s + d$.

Goal: Study $A_{s,d}$ when $p(N) = cN^{-\delta}$.

Our Results

Let h be a positive integer, $c > 0$, and choose pairs of integers (s_i, d_i) with $s_i \geq d_i$ and $s_i + d_i = h$. Each element of I_N is independently chosen to be in A with probability $p(N) = cN^{-\delta}$.

- For $\delta > \frac{h-1}{h}$, the set A_{s_i,d_i} with the larger d_i is larger almost surely, and with probability one $|A_{s_1,d_1}|/|A_{s_2,d_2}| = (s_2!d_2!)/(s_1!d_1!) + o(1)$ as $N \to \infty$.
- Define $g(x; s, d) := \sum_{k=1}^{m} (-1)^{k-1} \frac{b_{h,k}}{(s!d!)^k} x^{(s+d)k}$. If $\delta = \frac{h-1}{h}$ then almost surely $|A_{s_i,d_i}| \sim Ng(c; s_i, d_i)$, and with probability one as $N \to \infty$ $|A_{s_1,d_1}|/|A_{s_2,d_2}| = g(c; s_1, d_1)/g(c; s_2, d_2) + o(1)$.

Cases for δ

- Fast Decay: $\delta > \frac{h-1}{h}$.
- Critical Decay: $\delta = \frac{h-1}{h}$.
- Slow Decay: $\delta < \frac{h-1}{h}$.

Fast Decay: $\delta > \frac{h-1}{h}$

- Set with more differences is larger 100% of the time.
- Ratio of the sizes of $A_{s,d}$ is a function of $\binom{h}{d}$.
- Proofs use scarcity of elements in A.

Proof of the ratio of sizes of $A_{s,d}$

- Compute the number of distinct h-tuples.

Proof of the ratio of sizes of $A_{s,d}$

- Compute the number of distinct h-tuples.
- Show that sum of repeated elements is close to its expectation.

Proof of the ratio of sizes of $A_{s,d}$

- Compute the number of distinct h-tuples.
- Show that sum of repeated elements is close to its expectation.
- Conclude that almost all h-tuples generate a distinct number as $N \to \infty$.

Proof of the ratio of sizes of $A_{s,d}$

- Compute the number of distinct h-tuples.
- Show that sum of repeated elements is close to its expectation.
- Conclude that almost all h-tuples generate a distinct number as $N \to \infty$.
- Using combinatorics, conclude that ratio is:

$$\frac{|A_{s_1,d_1}|}{|A_{s_2,d_2}|} = \frac{\binom{h}{d_1}}{\binom{h}{d_2}} = \frac{s_2!d_2!}{s_1!d_1!}.$$

Critical Decay: $\delta = \frac{h-1}{h}$

- Set with more minus signs is larger 100% of time.

Critical Decay: $\delta = \frac{h-1}{h}$

- Set with more minus signs is larger 100% of time.
- Result depends on counting number of repeated elements by counting number of ways to add to each integer in the interval.

Critical Decay: $\delta = \frac{h-1}{h}$

- Set with more minus signs is larger 100% of time.
- Result depends on counting number of repeated elements by counting number of ways to add to each integer in the interval.
- Proof uses strong concentration results.

Defining a Function to Count Ways to Generate n

First step: determine a tractable formula for $R(n, s, d)$, the number of $h_{(s,d)}$-tuples of integers drawn from $\{0, \ldots, N\}$ that generate n.

Defining a Function to Count Ways to Generate n

First step: determine a tractable formula for $R(n, s, d)$, the number of $h_{(s,d)}$-tuples of integers drawn from $\{0, \ldots, N\}$ that generate n.

Let $n' := n + dN$. We have:

$$R(n, s, d) = \sum_{i=0}^{\lfloor \frac{n'}{N} \rfloor - 1} (-1)^i \binom{h}{i} \binom{n' - i(N+1) + h - 1}{h - 1}.$$

Defining a Function to Count Ways to Generate n

First step: determine a tractable formula for $R(n, s, d)$, the number of $h_{(s,d)}$-tuples of integers drawn from $\{0, \ldots, N\}$ that generate n.

Let $n' := n + dN$. We have:

$$R(n, s, d) = \sum_{i=0}^{\lfloor \frac{n'}{N} \rfloor - 1} (-1)^i \binom{h}{i} \binom{n' - i(N+1) + h - 1}{h - 1}.$$

- Cookie Problem (or Stars and Bars Problem)
- Inclusion-Exclusion
- Considering Differences

Slow Decay: $\delta < \frac{h-1}{h}$

- Harder for a number of reasons.

Slow Decay: $\delta < \frac{h-1}{h}$

- Harder for a number of reasons.
- Focuses on number of elements missing from $A_{s,d}$.

Slow Decay: $\delta < \frac{h-1}{h}$

- Harder for a number of reasons.
- Focuses on number of elements missing from $A_{s,d}$.
- Analysis focuses on the fringe of $[-dN, sN]$, the integers very close to $-dN$ or sN.

Will be investigated by SMALL '13.

Conclusion

Key Techniques:

- Order of the size of the set of repeated elements.
- Counting the ways to generate an integer n.
- Using existing inequalities to bound true value near expectation.

Open question: Generalizing the case of slow decay.
Hopefully will be done for CANT 2014!

Acknowledgements

We would like to thank the National Science Foundation for supporting our research through NSF Grant DMS0850577 and NSF Grant DMS0970067, as well as Williams College.

Bibliography

Bibliography

- G. Iyer, O. Lazarev, S.J. Miller, L. Zhang. *Generalized More Sums Than Differences Sets*. Journal of Number Theory **132** (2012), no 5, 1054–1073. http://arxiv.org/pdf/1108.4500.pdf

- O. Lazarev, S.J. Miller, K. O'Bryant. *Distribution of Missing Sums in Sumsets*, Experimental Mathematics **22** (2013), no. 2, 132–156. http://arxiv.org/abs/1109.4700

- P. V. Hegarty and S. J. Miller, *When almost all sets are difference dominated*, Random Structures and Algorithms **35** (2009), no. 1, 118–136. http://arxiv.org/abs/0707.3417

- G. Martin, K. O'Bryant, *Many Sets Have More Sums Than Differences*, Additive Combinatorics, **43** (2007), no. 4, 287–305. http://arxiv.org/abs/math/0608131