Error Estimation via a Refined Shapley-Folkman Lemma

Work in Summer 2023, at Cornell

Haoyu Wu[1], A. Kevin Tang[2]

[1]Department of Mathematics
Hong Kong University of Science and Technology
[2]School of Eletrical and Computer Engineering
Cornell University

SCIE 3900 Prensentation

Figure: Lloyd Shapley (2012 Nobel Memorial Prize in Economic Sciences)

Figure: Ross Starr

Table of Contents

Measurement of Sets in Euclidean Space

Definition

To state the conclusion, we define the measurement of the size of set $S \subset \mathbb{R}^m$, i.e. (outer) radius $\mathrm{rad}(S)$ and diameter $D(S)$,

$$\mathrm{rad}(S) = \inf_{y \in \mathbb{R}^m} \sup_{x \in S} |x - y|, \; D(S) = \sup_{x,y \in S} |x - y|$$

And the Hausdorff distance of sets,

$$d_H(X, Y) = \max\{\sup_{x \in X} \inf_{y \in Y} |x - y|, \sup_{y \in Y} \inf_{x \in X} |x - y|\}$$

Remark

When $X \subset Y$, $d_H(X, Y) = \sup_{y \in Y} \inf_{xinX} |x - y|$.

Introduction

Shapley-Folkman Lemma

Let S_i, $1 \leq i \leq n$ be subsets of $\mathbb{R}^m$, $S = \sum_{i=1}^n S_i$ be the Minkowski sum of all S_i. For any $z \in \text{conv}S$, there exists $z_i \in \text{conv}S_i$ such that $z = \sum_{i=1}^n z_i$ and $z_i \in S_i$ except for at most $\min\{m, n\}$ values of i.

Shapley-Folkman-Starr Theorem

With the same assumption above and let d_H denote the Hausdorff distance.

$$d_H^2(S, \text{conv}S) \leq \min\{m, n\} R^2 := \min\{m, n\} \max_{1 \leq i \leq n} R_i^2$$

Remarks

With additional compactness assumption, for any $z \in \text{conv}S$, there is $x \in S$ such that

$$|x - z|^2 \leq \min\{m, n\} R^2$$

Applications of Shapley-Folkman Lemma

Shapley-Folkman Lemma

Let S_i, $1 \leq i \leq n$ be subsets of $\mathbb{R}^m$, $S = \sum_{i=1}^n S_i$ be the Minkowski sum of all S_i. For any $z \in \text{conv}S$, there exists $z_i \in \text{conv}S_i$ such that $z = \sum_{i=1}^n z_i$ and $z_i \in S_i$ except for at most $\min\{m, n\}$ values of i.

Remark

In given dimension $\mathbb{R}^m$, as n increases, we consider the average,

$$d_H^2(\frac{S}{n}, \text{conv}\frac{S}{n}) \leq \frac{m}{n^2}R^2 \longrightarrow 0, \text{as } n \to \infty$$

- Nonconvex consumer preference by Ross Starr.
- Lyapunov convexity theorem in measure theory [Tar90].
- Law of large numbers [AV75].
- Estimation of the duality gap of separable optimization problems [AE76].

Definition

Denote the kth-convex hull of a set S, as $\text{conv}_k S$, the set of convex combination of at most k elements,

$$\text{conv}_k S = \left\{ \sum_{i=1}^{k} a_i v_i : v_i \in S, 0 \leq a_i \leq 1, 1 \leq i \leq k, \sum_{i=1}^{k} a_i = 1 \right\}$$

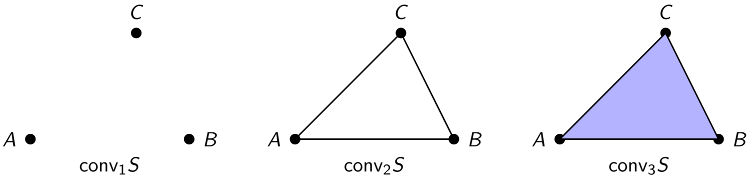

Figure: the kth-convex hull of set $S = \{A, B, C\}$

Strong Version of Shapley-Folkman Lemma

Strong Version SF Lemma [Zho93]

Let S_i, $1 \leq i \leq n$, be subsets of $\mathbb{R}^m$, $S = \sum_{i=1}^n S_i$ be the Minkowski sum of all S_i. For any $z \in \text{conv}S$, there exists $k_i \in \mathbb{N}$ and $\sum_{i=1}^n k_i \leq n + m$ such that $z = \sum_{i=1}^n z_i$ and $z_i \in \text{conv}_{k_i} S_i$.

[BR20] Improved bound

With the same assumption above and additional assumption that $n > m$ and all S_i are compact sets, for any $z \in \text{conv}S$, there is $x \in S$ such that

$$|x - z|^2 \leq \frac{mD^2}{4}$$

Remark

For any set $S \subset \mathbb{R}^m$, $\sqrt{\frac{2(m+1)}{m}} R \leq D \leq \frac{R}{2}$, which implies [BR20] is a quantative improvement up to 30%.

Table of Contents

Concept for Convexity Categorization II

Definition

A point z in a convex set S is called a k-extreme point of S if there NOT exist $(k+1)$ linear independent vectors $d_1, d_2, \cdots, d_{k+1}$ such that $z \pm d_i \in S$ for $i = 1, 2, \cdots, k+1$.

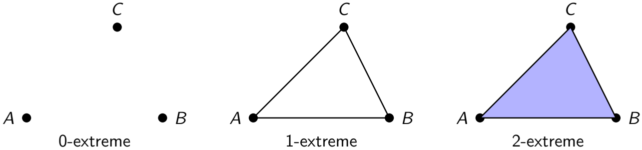

Figure: the kth-convex hull of set $S = \{A, B, C\}$

- A point is extreme point if only if it's 0-extreme.
- A point is on the boundary if only if it's $m-1$-extreme
- Any point in the convex hull in $\mathbb{R}^m$ is m extreme.

Refined Shapley-Folkman Lemma

Shapley-Folkman Lemma

Let S_i, $1 \le i \le n$ be subsets of $\mathbb{R}^m$, $S = \sum_{i=1}^n S_i$ be the Minkowski sum of all S_i. For any $z \in \text{conv}S$, there exists $z_i \in \text{conv}S_i$ such that $z = \sum_{i=1}^n z_i$ and $z_i \in S_i$ except for at most $\min\{m, n\}$ values of i.

Refined Shapley-Folkman Lemma[BT20]

Let S_i, $1 \le i \le n$ be subsets of $\mathbb{R}^m$, $S = \sum_{i=1}^n S_i$ be the Minkowski sum of all S_i. Assume z is a k-extreme point of $\text{conv}S$, then there exist integers $1 \le k_i \le k+1$ with $\sum_{i=1}^n k_i \le n + k$ and points $z_i \in \text{conv}_{k_i} S_i$ such that $z = \sum_{i=1}^n z_i$.

- k-extreme to categorize convexity in $\text{conv}S$,
- k-convex hull to distinguish $\text{conv}S_i$

Table of Contents

Main Theorem

Main Theorem

Let S_i, $1 \leq i \leq n$, be compact subsets of $\mathbb{R}^m$, $S = \sum_{i=1}^n S_i$ be the Minkowski sum of all S_i. Assume z is a k-extreme point of conv$S$, then there exist $x \in S$ such that,

$$|z - x|^2 \leq \frac{1}{2} \min \left\{ \frac{nkD^2}{n+k}, \sum_{i=1}^{n} D_i^2 - \frac{\left(\sum_{i=1}^n D_i\right)^2}{n+k} \right\}$$

Where $D_i = \sup_{x,y \in S_i} |x - y|$ is the diameter of S_i and $D = \max_{1 \leq i \leq n} D_i$.

- The first bound has same degree $O(\min\{n, m\} D^2)$ compared to the origin bound.
- The latter bound jointly consider all k_i, D_i instead directly taking $\max D$.

Corollary

Let S_i, $1 \leq i \leq n$, be compact subsets of $\mathbb{R}^m$, $S = \sum_{i=1}^n S_i$ be the Minkowski sum of all S_i. Assume z is a k-extreme point of conv$S$, then there exist $x \in S$ such that,

$$|z - x|^2 \leq \frac{1}{2} \min \left\{ \frac{nkD^2}{n+k}, \sum_{i=1}^{n} D_i^2 - \frac{\left(\sum_{i=1}^{n} D_i\right)^2}{n+k} \right\}$$

Corollary 1

If $n \geq k$, the estimated bound $\frac{nkD^2}{2(n+k)}$ can be modified to $\frac{kD^2}{4}$.

Remark

In general, we are unable to determine whether the first term is greater than or less than the latter term. But the latter term is not larger than the first term in degree in the particular case $k \in O(n)$.

Comparing to Previous Results

Conditions \ Results	Origin Bound	[BR20]	Our Results
In general	$\min\{n, m\}R^2$	N.A.	$\frac{1}{2}\min\left\{\frac{nmD^2}{n+m}, \sum_{i=1}^n D_i^2 - \frac{\left(\sum_{i=1}^n D_i\right)^2}{n+k}\right\}$
$n \geq m$	mR^2	$mD^2/4$	$\min\left\{mD^2/4, \frac{1}{2}\left[\sum_{i=1}^n D_i^2 - \frac{\left(\sum_{i=1}^n D_i\right)^2}{n+k}\right]\right\}$
$n \geq m$ & k-extreme	mR^2	$mD^2/4$	$\min\left\{kD^2/4, \frac{1}{2}\left[\sum_{i=1}^n D_i^2 - \frac{\left(\sum_{i=1}^n D_i\right)^2}{n+k}\right]\right\}$
m is $O(n)$	$O(nD^2)$	N.A.	$\min\left\{O(kD^2), O\left(\sum_{i=1}^n D_i^2\right)\right\}$

$D_i's$ follow i.i.d.

$$|z-x|^2 \le \frac{1}{2}\min\left\{\frac{nkD^2}{n+k}, \sum_{i=1}^{n} D_i^2 - \frac{\left(\sum_{i=1}^{n} D_i\right)^2}{n+k}\right\}$$

Corollary 2

Consider the case that D_i's are i.i.d. and we estimate the bound of expectation. The first term is $\frac{nk}{2(n+k)}\mathbb{E}(D^2)$. The latter term is

$$\frac{nk}{n+k}\mathbb{E}(D_i^2) + \frac{n(n-1)\mathrm{Var}(D_i)}{n+k}$$

Remark

In case $k \in O(n)$ or $n\mathrm{Var}(D) \in O(k\mathbb{E}(D_i^2))$, the latter term has degree improvement $\frac{\mathbb{E}(D_i^2)}{\mathbb{E}(D^2)} = \frac{\mathbb{E}(D_i^2)}{\mathbb{E}(\max_{1\le i\le n} D_i^2)}$.

Qualitative Improved Cases

Corollary

Consider the case that D_i's are i.i.d. and we estimate the bound of expectation. The first term is $\frac{nk}{2(n+k)}\mathbb{E}(D^2)$. The latter term is

$$\frac{nk}{n+k}\mathbb{E}(D_i^2) + \frac{n(n-1)\mathrm{Var}(D_i)}{n+k}$$

In case $k \in O(n)$ or $n\mathrm{Var}(D) \in O(k\mathbb{E}(D_i^2))$, the latter has degree improvement $\frac{\mathbb{E}(D_i^2)}{\mathbb{E}(D^2)} = \frac{\mathbb{E}(D_i^2)}{\mathbb{E}(\max_{1\leq i\leq n} D_i^2)}$.

- Consider further assumption $k \in O(n)$.
- If all D_i follow half-normal distribution, our result has improvement $O(1/\log n)$.
- If all D_i follow expotential distribution, our result has improvement $O(1/\log^2 n)$

Half-normal Distribution

Let X follow a normal distribution $N(0, \sigma^2)$ and $Y = |X|$ follows a half-normal distribution.

$$\mathbb{E}(Y) = \sigma\sqrt{\frac{2}{\pi}}, \mathbb{E}(Y^2) = \sigma^2, \ \mathbb{E}(\max_{i\in[n]} Y_i) = \sigma\sqrt{2\log n} + O(1)$$

Consider that all D_i's are i.i.d. of the random variable Y,

$$\mathbb{E}\left(\frac{nk}{n+k}D^2\right) \geq \frac{nk}{n+k}\mathbb{E}^2(D) \in O(\frac{nk\sigma^2}{n+k}\log n)$$

$$\mathbb{E}\left(\sum_{i=1}^{n} D_i^2 - \frac{\left(\sum_{i=1}^{n} D_i\right)^2}{n+k}\right) = \sigma^2 \frac{\left(1-\frac{2}{\pi}\right)(n^2-n)+nk}{n+k} \in O\left(n\sigma^2\right)$$

With more assumption that k is $O(n)$, the first term is least $O(n \log n\, \sigma^2)$ and the latter term has qualitative improvement $\frac{1}{\log n}$ compared to the first term.

Exponential Distribution

In case k is $O(n)$ and D_i's are i.i.d. following distribution $\exp(1)$, i.e. the probability density function $f(x) = e^{-x}$ for $x \geq 0$ and equal 0 else.

$$\mathbb{E}(D_i) = \mathrm{Var}(D_i) = 1, \mathbb{E}(D_i^2) = \mathbb{E}^2(D_i) + \mathrm{Var}(D_i) = 2$$

$$\mathbb{E}(D) = \sum_{i=1}^{n} \frac{1}{i}, \mathrm{Var}(D) = \sum_{i=1}^{n} \frac{1}{i^2}$$

$$\mathbb{E}(D^2) = \left(\sum_{i=1}^{n} \frac{1}{i} \right)^2 + \sum_{i=1}^{n} \frac{1}{i^2} \in O(\log^2 n)$$

.

- The latter bound has an qualitative improvement $\frac{1}{\log^2 n}$ compared to the first bound.

Conclusion

- Jointly consider all component employing constraint from the refined SF lemma.
- Concept for convexity categorization.
- Qualitative improvement in some cases.

Prospect

- $\mathbf{k} = (k_1, \cdots, k_n)^T \in (\mathbb{N} \cup \{0\})^n, ||\mathbf{k}||_1 \leq n + k$.
- $A(\mathbf{k}) = \sum_{i=1}^{n} \text{conv}_{k_i} S \subset \text{conv} S$.
- For any $z \in \text{conv}S$, $\mathbf{k}$ s.t. $z \in A(\mathbf{k})$
- $S \subset A(\mathbf{k}) \subset \text{conv}S$
- Relationship among all $A(\mathbf{k})$. (Hausdorff distance, overlap)

Table of Contents

References

J. P. Aubin and I. Ekeland, *Estimates of the duality gap in nonconvex optimization*, Mathematics of Operations Research **1** (1976), no. 3, 225–245.

Zvi Artstein and Richard A. Vitale, *A Strong Law of Large Numbers for Random Compact Sets*, The Annals of Probability **3** (1975), no. 5, 879 – 882.

Eric Budish and Philip J. Reny, *An improved bound for the shapley–folkman theorem*, Journal of Mathematical Economics **89** (2020), 48–52.

Yingjie Bi and Ao Tang, *Duality gap estimation via a refined shapley–folkman lemma*, SIAM Journal on Optimization **30** (2020), no. 2, 1094–1118.

Fabio Tardella, *A new proof of the lyapunov convexity theorem*, SIAM Journal on Control and Optimization **28** (1990), no. 2, 478–481.

Lin Zhou, *A Simple Proof of the Shapley-Folkman Theorem*.