Lecture 2

Approximate Caratheodory's Theorem

Convex sets and convex hulls

Definition 2.0.1: Convex sets

Let $\mathcal{S} \subseteq \mathbb{R}^d$. $\mathcal{S}$ is a **convex set** if

$$\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in \mathcal{S},$$

for any $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{S}$ and $\lambda \in [0,1]$.

We call $\boldsymbol{z}$ is a **convex combination** of $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\}$ if there exists $\boldsymbol{\lambda} = (\lambda_i) \in \mathbb{R}^n$ such that

$$\boldsymbol{z} = \sum_{i=1}^{n} \lambda_i \boldsymbol{x}_i$$

satisfying $\sum_{i=1}^{n} \lambda_i = 1$ and $\lambda_i \geq 0$ for all i.

Definition 2.0.2: Convex hulls

Let $\mathcal{S} \subseteq \mathbb{R}^d$. We call a set **convex hull** of $\mathcal{S}$, denoted by $\operatorname{conv}(\mathcal{S})$ if any element of this set, can be expressed as a convex combination of points from $\mathcal{S}$. Mathematically, for any $\boldsymbol{z} \in \operatorname{conv}(\mathcal{S})$, there exists a sequence $\{\boldsymbol{x}_i\}_{i=1}^n \subseteq \mathcal{S}$ for $n \in \mathbb{N}$ such that

$$\boldsymbol{z} = \sum_{i=1}^{n} \lambda_i \boldsymbol{x}_i$$

satisfying $\sum_{i=1}^{n} \lambda_i = 1$ and $\lambda_i \geq 0$ for all i.

Remark.

> The convex hull $\text{conv}(\mathcal{S})$ is the smallest convex set that contains $\mathcal{S}$, in the sense that $\text{conv}(\mathcal{S})$ is a subset of any convex set that contains $\mathcal{S}$.

Caratheodory's theorem and approximated Caratheodory's theorem

Theorem 2.0.3: Caratheodory's theorem

Every point in the convex hull of a set $\mathcal{S} \subseteq \mathbb{R}^d$ can be expressed as a convex combination of at most $d+1$ points from $\mathcal{S}$.

Caratheodory's theorem tells us the worst-case number of points needed to represent an element of a convex hull. Such worst-case number is dimensional-dependent and apparently cannot be improved.

What if we approximate $\boldsymbol{z} \in \text{conv}(\mathcal{S})$ rather than exactly represent it as a convex combination of points from $\mathcal{S}$. We show that such approximation lead to the number of points needed for representation does not depend not the dimension d.

Theorem 2.0.4: Approximate Caratheodory's theorem

Consider a bounded set $\mathcal{S} \subset \mathbb{R}^d$, i.e., there exists $r > 0$ for any $\boldsymbol{z} \in \mathcal{S}$, $\|\boldsymbol{z}\| \leq r$. For every point $\boldsymbol{x} \in \text{conv}(\mathcal{S})$ and every integer k, there exists a sequence of points $(\boldsymbol{x}_j)_{j=1}^k \subset \mathcal{S}$ such that

$$\left\| \boldsymbol{x} - \frac{1}{k}\sum_{j=1}^k \boldsymbol{x}_j \right\|_2 \leq \frac{r}{\sqrt{k}}.$$

Proof. Without loss of generality we assume that $\|\boldsymbol{z}\| \leq 1$ for any $\boldsymbol{z} \in \mathcal{S}$.

Let $\boldsymbol{x} \in \text{conv}(\mathcal{S})$, then there exists a sequence of points $(\boldsymbol{z}_i)_{i=1}^n$ for $n \in \mathbb{N}$ and $n \leq d+1$ such that

$$\boldsymbol{x} = \sum_{i=1}^n \lambda_i \boldsymbol{z}_i$$

satisfying $\sum_{i=1}^n \lambda_i = 1$ and $\lambda_i \geq 0$.

Since $\boldsymbol{\lambda} = (\lambda_i)$ belongs to the probability simplex, we define a discrete probability distribution of a random variable Z as follows

$$\mathbb{P}\{Z = \boldsymbol{z}_i\} = \lambda_i,$$

with expectation $\mathbb{E}Z = \boldsymbol{x}$.

Generating a family of iid random variables from this distribution $(Z_1, Z_2, \ldots, Z_k)$ for

$k \in \mathbb{N}$, we obtain their sample average as

$$\frac{1}{k}\sum_{i=1}^{k} Z_i,$$

and

$$\mathbb{E}\left[\frac{1}{k}\sum_{i=1}^{k} Z_i\right] = \boldsymbol{x}.$$

We measure the derivation of the sample average from its expectation by

$$\mathbb{E}\left\|\frac{1}{k}\sum_{i=1}^{k} Z_i - \boldsymbol{x}\right\|_2^2$$

Since Z_i are iid and deriving from the inequality $\mathbb{E}\|Z - \mathbb{E}Z\|_2^2 \leq \mathbb{E}\|Z\|_2^2$ (Check by Yourself) for a d-dimensional random variable, we have

$$\begin{aligned}
\mathbb{E}\left\|\frac{1}{k}\sum_{i=1}^{k} Z_i - \boldsymbol{x}\right\|_2^2 &= \mathbb{E}\left\|\frac{1}{k}\sum_{i=1}^{k}(Z_i - \boldsymbol{x})\right\|_2^2 \\
&= \frac{1}{k^2}\mathbb{E}\left\|\sum_{i=1}^{k}(Z_i - \boldsymbol{x})\right\|_2^2 \\
&= \frac{1}{k^2}\sum_{i=1}^{k}\mathbb{E}\|Z_i - \boldsymbol{x}\|_2^2 \\
&\leq \frac{1}{k^2}\sum_{i=1}^{k}\mathbb{E}\|Z_i\|_2^2 \\
&\leq \frac{1}{k} \qquad (\|Z_i\|_2 \leq 1)
\end{aligned}$$

By concavity of $\sqrt{\cdot}$ and Jensen's inequality, we obtain

$$\mathbb{E}\left\|\frac{1}{k}\sum_{i=1}^{k} Z_i - \boldsymbol{x}\right\|_2 \leq \sqrt{\mathbb{E}\left\|\frac{1}{k}\sum_{i=1}^{k} Z_i - \boldsymbol{x}\right\|_2^2} \leq \frac{1}{\sqrt{k}}.$$

Here there exists a realization of $(Z_1, Z_2, \ldots, Z_k)$ denotes as $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k)$ such that

$$\left\|\frac{1}{k}\sum_{i=1}^{k} \boldsymbol{x}_i - \boldsymbol{x}\right\|_2 \leq \frac{1}{\sqrt{k}}.$$

□

Remark.

> Such realization is built from components of a convex combination of $\boldsymbol{x}$ and repetitions are allowed.

An application of approximated Caratheodory's theorem

Definition 2.0.5: Covering numbers

The covering number of a set $\mathcal{T} \subset \mathbb{R}^d$ is the smallest number of Euclidean balls of radius ϵ needed to cover $\mathcal{T}$, denoted by $N(\mathcal{T}, \epsilon)$.

Remark.

- Covering numbers measure the complexity of $\mathcal{T}$.
- They suffer from the dimension d.
- Let the centers of a set of Euclidean balls with radius ϵ be $\{\boldsymbol{c}_i\}$. Mathematically, we say these Euclidean balls cover $\mathcal{T}$, if for any $x \in \mathcal{T}$, there exists $\boldsymbol{c}_k \in \{\boldsymbol{c}_i\}$ such that
$$\|\boldsymbol{x} - \boldsymbol{c}_k\| \leq \epsilon.$$

Proposition 2.0.6

Let $\mathcal{B} := \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_2 \leq 1\}$ be a unit Euclidean ball, for any $0 < \epsilon < 1$ we have

$$N(\mathcal{B}, \epsilon) \geq \left(\frac{1}{\epsilon}\right)^d.$$

Proof. Let vol($\mathcal{B}$) denote the volume of $\mathcal{B}$ and vol($\epsilon\mathcal{B}$) denote the volume Euclidean ball of radius ϵ.

By the definition of covering numbers, we have

$$\text{vol}(\mathcal{B}) \leq N(\mathcal{B}, \epsilon) \cdot \text{vol}(\epsilon\mathcal{B}) = N(\mathcal{B}, \epsilon) \cdot \epsilon^d \cdot \text{vol}(\mathcal{B}).$$

Hence, $N(\mathcal{B}, \epsilon) \geq \left(\frac{1}{\epsilon}\right)^d$. □

The proposition shows that the covering numbers of the unit Euclidean ball grows exponentially in dimensionality. Is there any set whose covering number is dimension-free?

Theorem 2.0.7: Covering numbers of Polytopes

Let $\mathcal{P} \subset \mathbb{R}^d$ be a polytope with m vertices and diam($\mathcal{P}$) $\leq r$.[a] Then $\mathcal{P}$ can be covered by *at most* $m^{\lceil \frac{r^2}{\epsilon^2} \rceil}$ Euclidean balls of radius $\epsilon > 0$, i.e.,

$$N(\mathcal{P}, \epsilon) \leq m^{\lceil \frac{r^2}{\epsilon^2} \rceil}.$$

[a] $\text{diam}(\mathcal{P}) = \sup_{\boldsymbol{x},\boldsymbol{y} \in \mathcal{P}} \|\boldsymbol{x} - \boldsymbol{y}\|_2$

Proof. Without loss of generality we assume that diam($\mathcal{P}$) ≤ 1 and $\mathbf{0}$ belongs to the inside of $\mathcal{P}$.

Let $\mathcal{T} = (\boldsymbol{z}_i)_{i=1}^m$ be the vertices of $\mathcal{P}$. Note that $\mathcal{P}$ is a subset of $\operatorname{conv}(\mathcal{T})$ and $\mathcal{T}$ is bounded by 1.

By the approximate Caratheodory's theorem, for any $\boldsymbol{x} \in \operatorname{conv}(\mathcal{T})$, there exists a sequence of vertices $(\boldsymbol{z}_{i_j})_{j=1}^k$ and $\boldsymbol{z}_{i_j} \in \mathcal{T}$, such that

$$\|\boldsymbol{x} - \frac{1}{k}\sum_{j=1}^{k} \boldsymbol{z}_{i_j}\|_2 \le \frac{1}{\sqrt{k}}, \quad \text{for any } k \in \mathbb{N}. \tag{ACT}$$

Let $\frac{1}{\sqrt{k}} \le \epsilon$, $k \ge \frac{1}{\epsilon^2}$, and take $K = \lceil \frac{1}{\epsilon^2} \rceil$.

Denote $\mathcal{N} = \left\{ \frac{1}{K}\sum_{j=1}^{K} \boldsymbol{z}_{i_j} : \boldsymbol{z}_{i_j} \in \mathcal{T} \right\}$. The cardinality of $\mathcal{N}$ takes $|\mathcal{N}| = m^K$.

Let all elements of $\mathcal{N}$ be the centers of Euclidean balls with radius of ϵ. By (ACT), the union of these balls covers $\operatorname{conv}(\mathcal{T})$, and hence $\mathcal{P}$.

Therefore the smallest number of Euclidean balls with radius ϵ needed to cover $\mathcal{P}$ is less than m^K.

In summary $N(\mathcal{P}, \epsilon) \le m^K = m^{\lceil \frac{1}{\epsilon^2} \rceil}$. □

The following theorem shows an application of the above reslut.

Theorem 2.0.8: Volume of Polytopes

Let $\mathcal{B} \subset \mathbb{R}^d$ be the unit Euclidean ball , and $\mathcal{P} \subset \mathcal{B}$ be a polytope with m vertices. Then

$$\frac{\operatorname{vol}(\mathcal{P})}{\operatorname{vol}(\mathcal{B})} \le \left(4\sqrt{\frac{\log m}{d}}\right)^d.$$

Proof. Let $\epsilon\mathcal{B}$ denote the Euclidean ball with radius $\epsilon > 0$. By the definition of covering numbers, we have

$$\operatorname{vol}(\mathcal{P}) \le N(\mathcal{P}, \epsilon) \cdot \operatorname{vol}(\epsilon\mathcal{B}) \overset{\text{cov. \# of poly.}}{=} \le m^{\lceil \frac{1}{\epsilon^2} \rceil} \cdot \epsilon^d \cdot \operatorname{vol}(\mathcal{B}).$$

Hence

$$\frac{\operatorname{vol}(\mathcal{P})}{\operatorname{vol}(\mathcal{B})} \le m^{\lceil \frac{1}{\epsilon^2} \rceil} \cdot \epsilon^d \le m^{\frac{2}{\epsilon^2}} \cdot \epsilon^d$$

Minimizing the right hand side w.r.t. ϵ gives

$$\frac{\operatorname{vol}(\mathcal{P})}{\operatorname{vol}(\mathcal{B})} \le e^{\frac{d}{2}} \left(\sqrt{\frac{4\log m}{d}}\right)^d \le \left(4\sqrt{\frac{\log m}{d}}\right)^d.$$

□

Remark.

> Let $\delta = 4\sqrt{\frac{\log m}{d}}$, then $\operatorname{vol}(\mathcal{P}) \le \delta^d \cdot \operatorname{vol}(\mathcal{B}) = \operatorname{vol}(\delta\mathcal{B})$. If the dimensionality d is sufficiently large, the volume of a convex polytope with its vertices on the surface of a unit Euclidean is smaller that a tiny ball $\delta\mathcal{B}$.