Monday, August 10 8.00 a.m.

<u>Definition:</u> A convex combination of points $z_1, z_2, \ldots, z_m \in \mathbb{R}^n$ is a linear combination w/ non-negative coefficients that sum to $\underline{\underline{1}}$,

i.e $\sum_{i=1}^{m} \lambda_i z_i$, $\lambda_i \geq 0$ $\sum_{i=1}^{m} \lambda_i = \underline{1}$.

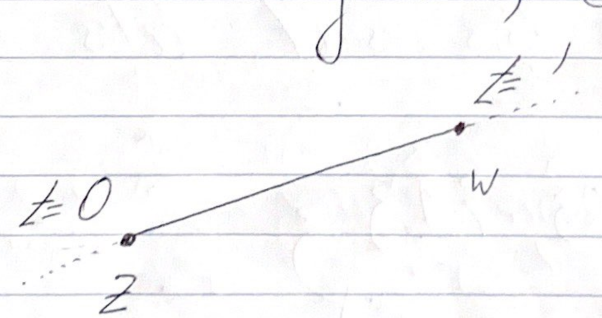
Example: $z_1 = z$, $z_2 = w$ $m = 2$

$\lambda_2 = t$, $\lambda_1 = 1-t$

Then $\sum_{i=1}^{m} \lambda_i z_i = (1-t)z + tw$

Observe that if $t=0$, $(1-t)z + tw = z$,

and if $t=1$, $(1-t)z + tw = w$



In fact, the set $S \subseteq \mathbb{R}^n$ is called <u>convex</u> if for any $z, w \in S$, $(1-t)z + tw \in S$,
$t \in [0,1]$

The convex hull of a set $T \subseteq \mathbb{R}^n$ is defined as

$$\operatorname{conv}(T) = \left\{ \begin{array}{c} \text{convex combinations of} \\ z_1, \dots, z_m \in T, \quad m \in \underbrace{\mathbb{N}}_{\text{natural numbers}} \end{array} \right\}$$

Theorem: (Caratheodory) Every point in the convex hull of $T \subseteq \mathbb{R}^n$ can be expressed as a convex combination of $\leq n+1$ points from T.

Before we prove this result, let's consider some examples:

$n = 1$, $T = \{\{0\}, \{1\}\}$

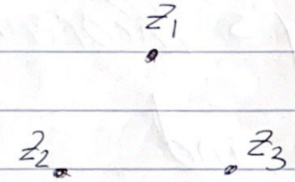
$\operatorname{conv}(T) =$ 

(3)

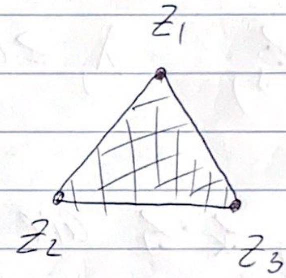
Every point in $[0,1]$ can be written in the form $x = (1-t)\cdot 0 + t \cdot 1$

choose $t = x$

Take $n=2$ $\quad T = \{z_1, z_2, z_3\}$



Then $\text{conv}(T) =$



Observe that there is no way to express every point in $\text{conv}(T)$ as a convex combination of only z_1, z_2, or z_1, z_3 or z_2, z_3. We need all three of those points.

If $n>2$, let $z_1, z_2, \dots, z_{n+1}$ be such that $|z_i - z_j| = 1$ if $i \neq j$ $\quad 1 \le i,j \le n+1$.

(4)

As an exercise, prove that such a set exists, and that its convex hull cannot be expressed using fewer than all $n+1$ z_j's.

Where are we going w/ this? Straight to the proof of Caratheodory's theorem.

Proof of Caratheodory's theorem:

Suppose that $K > d+1$ and

$$z = \sum_{j=1}^{K} \lambda_j z_j, \quad \lambda_j \geq 0, \quad \sum_{j=1}^{K} \lambda_j = 1$$

Then $z_2 - z_1, \ldots, z_K - z_1$ are linearly dependent, so

$\exists\ \mu_2, \mu_3, \ldots, \mu_K$, not all zero, such that

$$\sum_{j=2}^{K} \mu_j (z_j - z_1) = 0.$$

Define $\mu_1 = -\sum_{j=2}^{K} \mu_j$. Then

$$\sum_{j=1}^{K} \mu_j z_j = 0 \quad \text{since} \quad \sum_{j=1}^{K} \mu_j z_j = \sum_{j=2}^{K} \mu_j z_j + \mu_1 z_1$$

$$= z_1 \sum_{j=2}^{K} \mu_j + z_1 \mu_1 = 0 \text{ by definition.}$$

Also, $\sum_{j=1}^{K} \mu_j = 0$, once again, by definition.

Important: Not all the μ_j's are 0, so at least one $\mu_j > 0$.

We have $\alpha \equiv \min_{1 \le j \le K} \left\{ \frac{\lambda_j}{\mu_j} : \mu_j > 0 \right\} = \frac{\lambda_i}{\mu_i}$

definition

Note that $\alpha > 0$, and for every j between 1 & K,

$$\lambda_j - \alpha \mu_j \geq 0 \quad \text{since}$$

$$\lambda_j - \alpha \mu_j = \lambda_j - \frac{\lambda_i}{\mu_i} \mu_j \geq 0$$

since $\frac{\lambda_j}{\mu_j} \geq \frac{\lambda_i}{\mu_i}$ by construction.

Also, $\lambda_i - \alpha \mu_i = 0$. It follows that

$$x = \sum_{j=1}^{K} (\lambda_j - \alpha \mu_j) x_j \quad \text{w/ } \lambda_j - \alpha \mu_j \geq 0$$

$$\text{and } \lambda_i - \alpha \mu_i = 0$$

This implies that x is a linear combo (convex) of at most K-1 points & we're done!

⑥

We now prove an approximate form of Caratheodory:

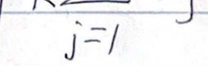
Theorem: Let $T \subseteq \mathbb{R}^n$ w/ diameter ≤ 1. Then for every point $x \in \operatorname{conv}(T)$ and every integer K, we can find points $x_1, x_2, \ldots, x_K \in T$ $\ni$

$$\left\| x - \frac{1}{K} \sum_{j=1}^{K} x_j \right\| \leq \frac{1}{\sqrt{K}}.$$

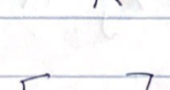
Note that $\underline{\underline{K}}$ <u>does</u> <u>not</u> depend on n.

Also, $\frac{1}{K}\sum_{j=1}^{K} x_j$ <u>is</u> <u>not</u> a general convex combination since all the coefficients are the same.

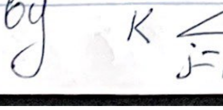
<u>Example:</u> $T \subseteq \mathbb{R}^1$, $T = \{0\} \cup \{1\}$



Then $\operatorname{conv}(T) = [0, 1]$



Then $x_j = 0$ or 1 & every $\frac{1}{K}\sum_{j=1}^{K} x_j$ is of the form $\frac{m}{K}$, $m = 0, 1, \ldots, K$



$K = 4$

Therefore, given K, every $x \in [0, 1]$ can be approximated by $\frac{1}{K}\sum_{j=1}^{K} x_j$ up to an error $\leq \frac{1}{2K}$.

The one-dimensional example raises the question of whether $\frac{1}{\sqrt{K}}$ error in the statement of the theorem is, in general, best possible.
Is it?

Proof of the approximate form of Caratheodory:
First, translate T, if necessary, so that

$$\|t\|_2 \leq 1 \quad \forall\, t \in T$$ (why is this always possible?)

$$\|t\|_2^2 = t_1^2 + \dots + t_n^2.$$

Fix $x \in \operatorname{conv}(T)$ & write it in the form

$$\sum_{i=1}^{m} \lambda_i z_i, \quad \lambda_i \geq 0 \quad z_i \in T$$

Define a random vector Z ∋

$$\mathbb{P}\{Z = z_i\} = \lambda_i, \quad i = 1, 2, \dots, m$$

recall that $0 \leq \lambda_i \leq 1$

(8)

Then $\underbrace{\mathbb{E}Z}_{\text{expected value}} = \sum_{i=1}^{m} \lambda_i z_i = \underline{\underline{x}}$

Consider independent copies $Z_1, Z_2, \dots,$ of Z.

Then $\frac{1}{K}\sum_{j=1}^{K} Z_j \underbrace{\longrightarrow}_{\text{strong law of large numbers}} x$ almost surely as $K \longrightarrow \infty$

We need a quantitative version of this result

Recall that if X is a random variable,

$$\underbrace{Var(X)}_{\text{variance of } X} = \mathbb{E}(X - \mathbb{E}X)^2$$

We shall compute the variance of $\frac{1}{K}\sum_{j=1}^{K} Z_j$

$$\mathbb{E}\left\| x - \frac{1}{K}\sum_{j=1}^{K} Z_j \right\|_2^2 \underbrace{=}_{\text{why?}} \frac{1}{K^2}\mathbb{E}\left\| \sum_{j=1}^{K} (Z_j - x) \right\|_2^2$$

(9)

The previous identity is pretty straightforward since
$$x - \frac{1}{K}\sum_{j=1}^{K} z_j = \frac{1}{K}\sum_{j=1}^{K}(x - z_j)$$

We now observe that

$$\frac{1}{K^2}\mathbb{E}\left\|\sum_{j=1}^{K}(z_j - x)\right\|_2^2 = \frac{1}{K^2}\sum_{j=1}^{K}\mathbb{E}\|z_j - x\|_2^2$$

this is because the cross-terms cancel due to the fact that $z_1, z_2, \dots$ are independent copies and

$$\mathbb{E}(z_j - x) = 0 \text{ by construction.}$$

This is just a higher dimensional version of the fact that the variance of a sum of independent random variables equals the sum of the variances.

We have

$$\mathbb{E}\|z_j - x\|_2^2 = \underbrace{\mathbb{E}\|z - \mathbb{E}z\|_2^2}_{\text{why?}} = \mathbb{E}\|z\|_2^2 - \|\mathbb{E}z\|_2^2$$

basic identity (homework)

$$\le \mathbb{E}\|z\|_2^2 \le 1$$

Putting everything together, we have shown that

$$\mathbb{E}\left\|x - \frac{1}{K}\sum_{j=1}^{K} z_j\right\|_2^2 \leq \frac{1}{K} \quad \left(\text{why } \tfrac{1}{K} \text{ and } \underline{\text{not}}\ \tfrac{1}{K^2}\,?\right)$$

Therefore, there exists a realization of the random variables $z_1, z_2, \dots, z_K \ni$

$$\left\|x - \frac{1}{K}\sum_{j=1}^{K} z_j\right\|_2^2 \leq \frac{1}{K}$$

and we are done since each z_j takes values in T.

We shall now discuss an intriguing application of the result we just proved:

Given $P \subseteq \mathbb{R}^n$, how do we cover it by the smallest possible balls of radius $\underbrace{\epsilon > 0}_{\text{given}}$

How do we place these balls?

convex hull of finitely many points

Corollary: Let P be a polytope in $\mathbb{R}^n$ w/ N vertices, and diameter ≤ 1. Then P can be covered by

$$\leq N^{\lceil \frac{1}{\epsilon^2} \rceil}$$

Euclidean balls of radius $\epsilon > 0$.

Proof: Let $K = \lceil \frac{1}{\epsilon^2} \rceil$ and let

$$\mathcal{N} = \Big\{ \underbrace{\frac{1}{K}\sum_{j=1}^{K} x_j}_{} : x_j\text{'s are } \underbrace{\text{vertices of } P}_{T} \Big\}$$

we know from approximate Caratheodory that such expression can be used to approximate everything in P.

More precisely, given $x \in P$, we can find x_j's in T $\ni$

$$\Big\| x - \frac{1}{K}\sum_{j=1}^{K} x_j \Big\| \leq \frac{1}{\sqrt{K}} < \epsilon$$

by our choice of K above.

We now take the elements of $\mathcal{N}$ as centers of our balls. It remains to estimate the size of $\mathcal{N}$.

(12)

A direct (and very slightly wasteful) way to estimate the size of $\mathcal{N}$ is the following:

There are N^K ways to choose K vertices out of N vertices, so $|\mathcal{N}| \leq N^K = N^{\lceil \frac{1}{\varepsilon^2} \rceil}$ and we are done!

w/repetitions

This ends the Appetizer, but please don't be satisfied w/ just understanding the material above. Please make up new theorems, generate new ideas, write code to clarify your thoughts, and so on!