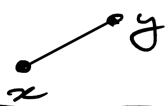
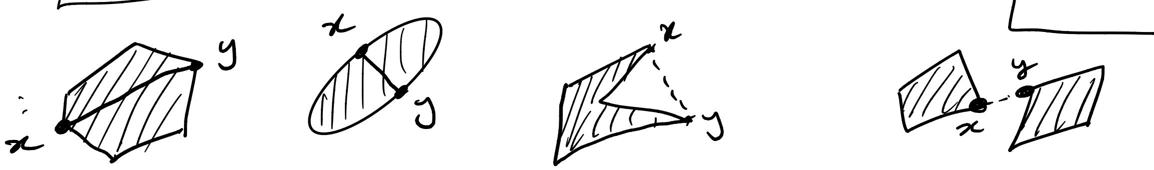
LECTURE 2

- One more example of how probability can help in high dim's: computational geometry

Def A set $T \subset \mathbb{R}^n$ is <u>convex</u> if $\forall\, x, y \in T$, $[x, y] \in T$
↑ interval



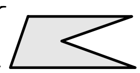
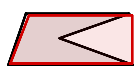
- **Ex**:

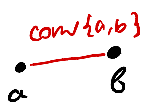


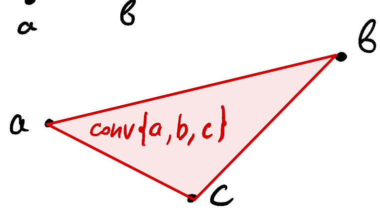
convex   convex   non-convex   non-convex

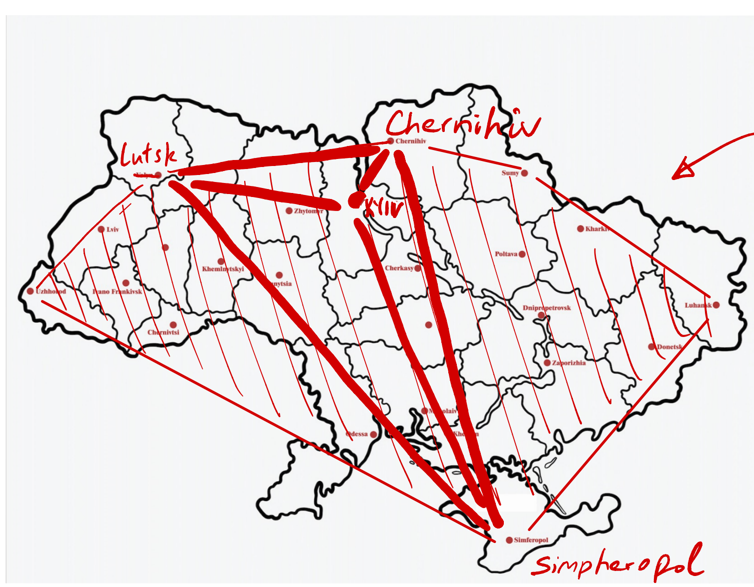
- How can one transform nonconvex set → convex?

Def The <u>convex hull</u> of a set $T \subset \mathbb{R}^n$, denoted $\boxed{\text{conv}(T)}$, is the smallest convex set that contains T.

Ex: (a) conv() = 

(b) 

(c) 

(d) 

T = cities in Ukraine

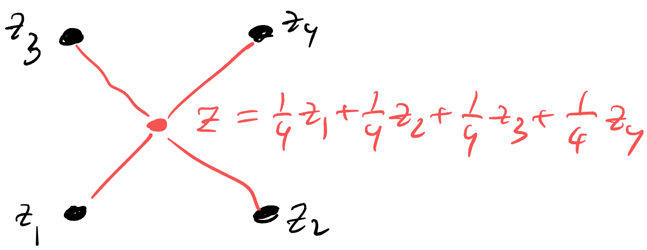
conv(T)

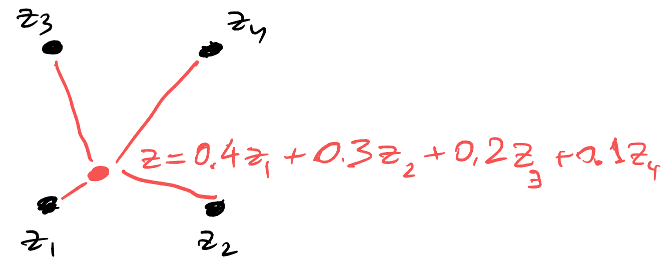
FACT (CONVEX COMBINATION)

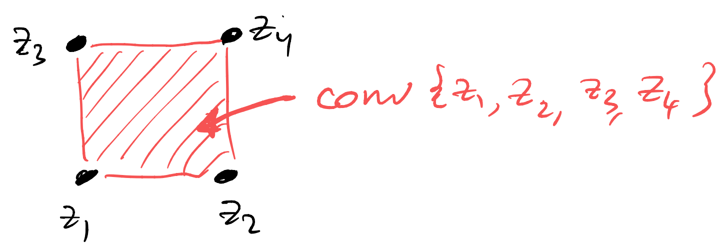
$$\boxed{\forall\ z \in \mathrm{conv}(T) \text{ can be expressed as } z = \sum_{i=1}^{m} \lambda_i z_i, \text{ where } \lambda_i \geq 0,\ \sum_{i=1}^{m} \lambda_i = 1,\ z_i \in T \qquad (*)}$$

- PROOF: DIY (do it yourself).
- (*) is called a "convex combination". It is a linear combination, similar to a basis expansion of z, but non-unique (if $m > n$)

Ex







Q: Computing convex hull: how large is m?

Caratheodory Thm $\forall$ point in $\mathrm{conv}(\overset{\mathbb{R}^n}{\overset{\cup}{T}})$ can be expressed as a convex combination of $\leq n+1$ points from T.

Ex Kyiv $\in \mathrm{conv}\{$Lutsk, Chernihiv, Simpheropol$\}$ $\qquad n = 2$.

Remark: $n+1$ is unimprovable (e.g. in the Kyiv example above)

But: if we allow to approximate x, then a dramatic improvement!

THM (Approx. Caratheodory) Let $T \subset \mathbb{R}^n$, $\mathrm{diam}(T) \le 1$. ← $\max\{\|x-y\|_2 : x, y \in T\}$

Then $\forall x \in \mathrm{conv}(T)$, $\forall k \in \mathbb{N}$ $\exists x_1, \dots, x_k \in T$:

$$\left\| x - \frac{1}{k} \sum_{i=1}^{k} x_i \right\|_2 \le \frac{1}{\sqrt{2k}}.$$

↑ Euclidean norm in $\mathbb{R}^n$.

Remark Dimension-free! Does NOT depend on n or geometry of T.

PROOF by a probabilistic method: the "empirical method of Maurey"

WILL USE STANDARD FACTS OF PROBABILITY:

① Def of expectation of a discrete r.v. X: if X takes values x_i with prob. p_i,

$$\mathbb{E}X \overset{\text{def}}{=\!=} \sum_i p_i x_i$$

(analogous to continuous case where $\mathbb{E}X = \int_{-\infty}^{\infty} x\, p(x)\, dx$)

② $\mathrm{Var}(X) = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}(X^2) - (\mathbb{E}X)^2 \overset{(*)}{=\!=} \frac{1}{2}\mathbb{E}(X - X')^2$

where X' is an independent copy of X (i.e. X, X' are independent and have the same distribution)

Proof of $(*)$: $\mathbb{E}(X-X')^2 = \mathbb{E}\left[X^2 - 2XX' + (X')^2\right]$

$$= \mathbb{E}[X^2] - 2\,\mathbb{E}(X)\underbrace{\mathbb{E}(X')}_{=\mathbb{E}X} + \underbrace{\mathbb{E}(X')^2}_{=\mathbb{E}X^2} \quad \text{(linearity \& independence) (same distribution)}$$

$$= 2\left(\mathbb{E}(X^2) - (\mathbb{E}X)^2\right) = 2\,\mathrm{Var}(X). \quad \text{QED}$$

③ Extend $(*)$ for random vectors X in $\mathbb{R}^n$:

$\mathbb{E}\|X - \mathbb{E}X\|_2^2 = \frac{1}{2}\mathbb{E}\|X - X'\|_2^2$ where X' is an indep. copy of X (HW2)

Proof of Approximate C.T.

- Fix $x \in \mathrm{conv}(T)$, express it as a convex combination

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

potentially large

- Interpret λ_i as probabilities, the sum as expectation

Let r.vector Z take value z_i with prob. λ_i

$$\Rightarrow \quad x = \mathbb{E} Z$$

- Consider indep. copies $Z_1, Z_2, \dots$ of Z.

SLLN: $\frac{1}{k} \sum_{i=1}^{k} Z_i \to \mathbb{E} Z = x$ a.s.

$P(\mathbf{E})$

$\mathbf{E}$

- Error:

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

$x = (x_1, \dots, x_n) \in \mathbb{R}^n \qquad \|x\|_2^2 = x_1^2 + \dots + x_n^2$

$$= \frac{1}{k^2} \sum_{i=1}^{k} \underbrace{\mathbb{E} \| Z_i - \mathbb{E} Z_i \|_2^2}_{= \mathbb{E}\|Z - \mathbb{E}Z\|_2^2}$$

(variance of sum = sum of variances, for vectors) HW2

$$= \frac{1}{k} \mathbb{E} \| Z - \mathbb{E} Z \|_2^2 = \frac{1}{2k} \mathbb{E} \underbrace{\| Z - Z' \|_2^2}_{\le} \quad \text{(fact ③ previous page)}$$

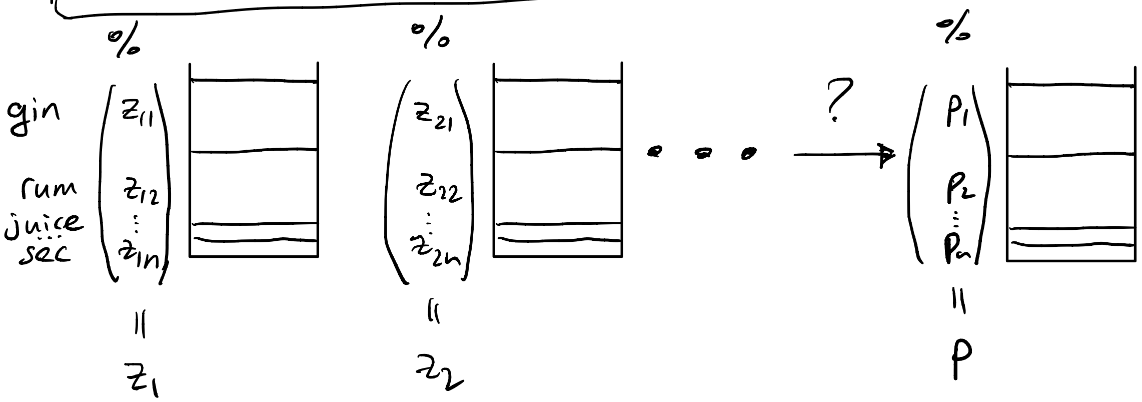
$$\le \frac{1}{2k}. \qquad \mathrm{diam}(T) \le 1 \text{ since } Z, Z' \in T$$

- $\Rightarrow \exists$ realization of the r.v's $Z_1, \dots, Z_k$ s.t.

$$\left\| x - \frac{1}{k} \sum_{i=1}^{k} Z_i \right\|_2^2 \le \frac{1}{2k}. \quad \text{Since } Z_i \in T, \quad \text{QED}$$

Application of Approx. Caratheodory Thm:

Cocktail Problem You are given N glasses with different cocktails, each made by mixing n ingredients in different proportions. Make a glass of cocktail with given proportions $p_1, \dots, p_n$.



Equivalently, we need to find λ_1 (portion of glass 1), λ_2 (port. of glass 2), ...

such that: $p = \sum_{i=1}^{N} \lambda_i z_i$, $\lambda_i \geq 0$, $\sum_{i=1}^{N} \lambda_i = 1$

proportions need to be nonnegative — and sum to 100%

- i.e. we need to express p as a convex combination of vectors $z_1, \dots, z_N$
- "Convex program" finds a solution in polynomial time.
- Approx. Caratheodory Thm transforms it into an approximate solution with few glasses, mixed in equal proportions; a fast randomized alg.

few: independent of n, N

Relevance of cocktail problem

(a) (Portfolio Building)

ingredients = stocks
glasses of cocktails = mutual funds
empty glass = portfolio

Problem: create a new mutual fund with a given combination of stocks by combining the mutual funds that are available on the market.

Solution: as above – a fast randomized algorithm builds a portfolio with few mutual funds.

(b) (Factor analysis)

$z_1, \dots, z_N$ = a dictionary of factors that need to explain z = behavior (consumer, animal, etc.)

Sol: behavior is explained by few factors.

$z = \sum \lambda_i z_i \Rightarrow$ factor 1 explains λ_i% of behavior, ...

$\Rightarrow$ A PARSIMONIOUS MODEL.