DSA3102: Convex Optimization (Lecture 1)

Motivation for convex optimization.

consist of nodes/edges

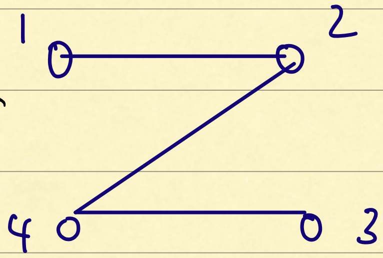
MAXCUT Problem in $\boxed{\text{graph}}$ theory

$G = (V, E)$ $\qquad V = \{1, \dots, n\}$, $E \subset \binom{V}{2}$

vertices nodes $\qquad$ edges.

unordered

undirected



(vertices) $V = \{1, 2, 3, 4\}$

set

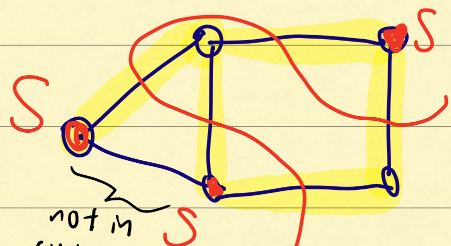
$E = \{\{1,2\}, \{2,4\}, \{3,4\}\}$.

(edge set)

Cut: Given a subset $S \subset V$, $\text{cut}(S) = \{e \in E : |e \cap S| = 1\}$.

↳ set of all edges in E, such that e intersect S = 1

red vertices form subset S



$|\text{cut}(S)| = 5$.

not in cut as touches 2 nodes belonging to S

Given a graph G, find $S \subset V$ s.t. $|\text{cut}(S)|$ is maximized.

eg. $2^{|E|}$

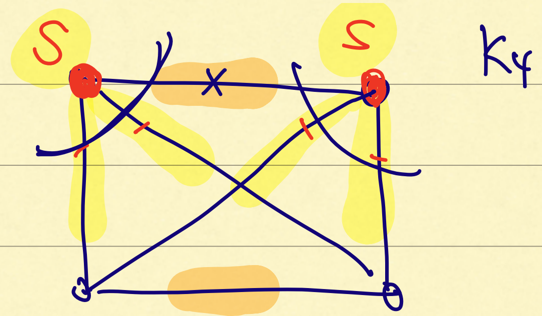
Ex: Complete graph $G = \boxed{K_n}$ (n even)

complete graph with n nodes

(edges that have intersection 1 with S)
exactly 1

$|S|=1$

$|\text{cut}(S)|=n-1$
$=1(n-1)$



$|S|=2$, $|\text{cut}(S)|=(n-2)+(n-2)=\underline{2(n-2)}$

In general, if $\underline{|S|=k}$, $\underline{|\text{cut}(S)|=k(n-k)}$

To maximize $k(n-k)$ over all $k\in\{0,1,\dots,n\}$,
by symmetry the optimal $\underline{k=n/2}$, e.g. $S=\{1,\dots,n/2\}$

$$\text{MAXCUT}(K_n)=\frac{n}{2}\left(n-\frac{n}{2}\right)=\left(\frac{n}{2}\right)^2.$$
n even

only activated when on diff sides of the cut

optimization prob: Write the MAXCUT problem $\forall\, G=(V,E)$ as.
(IQP Integer Quadratic Program)

$$\text{OPT}(G)=\max_{x_1,\dots,x_n}\ \underbrace{\sum_{\{i,j\}\in E}\frac{1-x_ix_j}{2}}_{\#\text{ edges on diff sides of the cut}},\quad x_i\in\{-1,+1\}$$
binary

$x_i=+1 \Leftrightarrow i\in S$, $x_i=-1 \Leftrightarrow i\in S^c=V\setminus S$.

$$\frac{1-x_ix_j}{2}=\begin{cases}0 & \text{if } i\ \&\ j \text{ are on the same sides of the cut}\\ 1 & \text{if } i\ \&\ j \text{ are on diff sides of the cut}\end{cases}$$

restricted to boolean values

<u>NP-hard</u>
(very difficult)

→ binary ⇒ continuous ?

Reformulate this to approximated opt problem.

inner product

L_2 norm

optimising over n vectors in $\mathbb{R}^n$

$$\max_{v_1,\dots,v_n} \sum_{\{i,j\}\in E} \frac{1-\langle v_i, v_j\rangle}{2} \qquad \|v_i\|=1,\ v_i\in\mathbb{R}^n,\ i=1,\dots,n.$$

If $n=1$, this problem reduces to the above IQP

This problem can be written as a convex opt problem.

semi definite programme **SDP**

$$\max_{\mathbb{R}^{n\times n}\ni X\succeq 0} \frac{1}{2}\sum_{\{i,j\}\in E} \frac{1-X_{ij}}{2}, \qquad X_{ii}=1 \quad i=1,\dots,n$$

over a matrix that is (positive semi-definite)

$$X_{ij}=\langle v_i, v_j\rangle, \qquad V=\begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}$$

matrix of inner products

$$\underbrace{V^TV}_{\succeq 0} = \begin{bmatrix} — v_1^T — \\ \vdots \\ — v_n^T — \end{bmatrix}\begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} = \begin{bmatrix} \langle v_1,v_1\rangle & \cdots & \langle v_1,v_n\rangle \\ \vdots & \ddots & \vdots \\ \langle v_n,v_1\rangle & \cdots & \langle v_n,v_n\rangle \end{bmatrix}$$

eigenvalues are all non-negative

diagonal elements all have to be norm 1

Suppose we can solve the SDP for $v_1, v_2, \dots, v_n\in\mathbb{R}^n$.

Goeman-Williamson alg.

Let $r\in\mathbb{R}^n$, $\|r\|=1$ norm 1 be a random vector drawn from $\{y\in\mathbb{R}^n : \|y\|=1\}$.

random rounding

generate cut

non-negative inner product with random vector

$$S=\{i\in\{1,\dots,n\} : \langle v_i, r\rangle \geq 0\}.$$

↳ acute angle ⇒ include in S

Thm: $\underbrace{OPT^*(G)}_{\text{efficient}} \geq \underbrace{0.87856}_{\alpha}\ \underbrace{OPT(G)}_{\text{cannot attain}}$

efficient solved using SDP. (with convex programme)

one is inside the cut,
one is outside the cut

Def: $i, j \in V = \{1, \dots, n\}$ are separated if $\langle v_i, r\rangle\langle v_j, r\rangle \le 0$.

$|\{i,j\} \cap S| = 1$

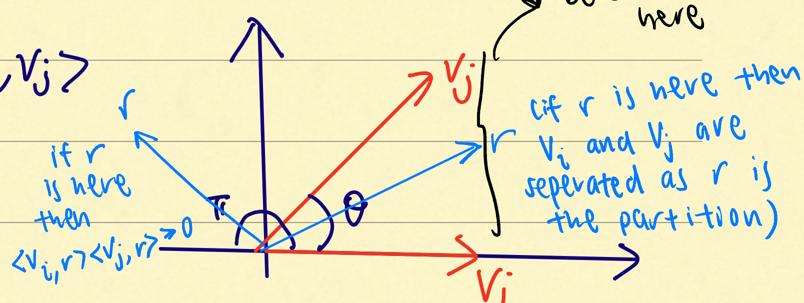
probability

Lem: $\Pr(i, j \text{ are separated}) \ge \alpha \left(\frac{1 - \langle v_i, v_j\rangle}{2}\right)$

Pf: Suppose θ is the angle between v_i & v_j, i.e.,

$$\cos\theta = \langle v_i, v_j\rangle$$

$$\Pr(i, j \text{ separated}) = \frac{\theta}{\pi}.$$



$$\frac{\Pr(i,j \text{ sep})}{(1 - \langle v_i, v_j\rangle/2)} = \frac{\theta/\pi}{(1 - \cos\theta)/2} = \left(\frac{2}{\pi}\right)\left(\frac{\theta}{1 - \cos\theta}\right) \ge \frac{2}{\pi} \min_{\theta \in [0, \pi]} \left(\frac{\theta}{1 - \cos\theta}\right)$$

plot this?

$$= \alpha \cong 0.87856.$$

$A(G)$: our cut obtained by SDP and random rounding.

$$\underbrace{\mathbb{E}[A(G)]}_{\text{our cut}} = \sum_{\{i,j\} \in E} \mathbb{E}\left[1\{ |\{i,j\} \cap S| = 1 \} \right]$$

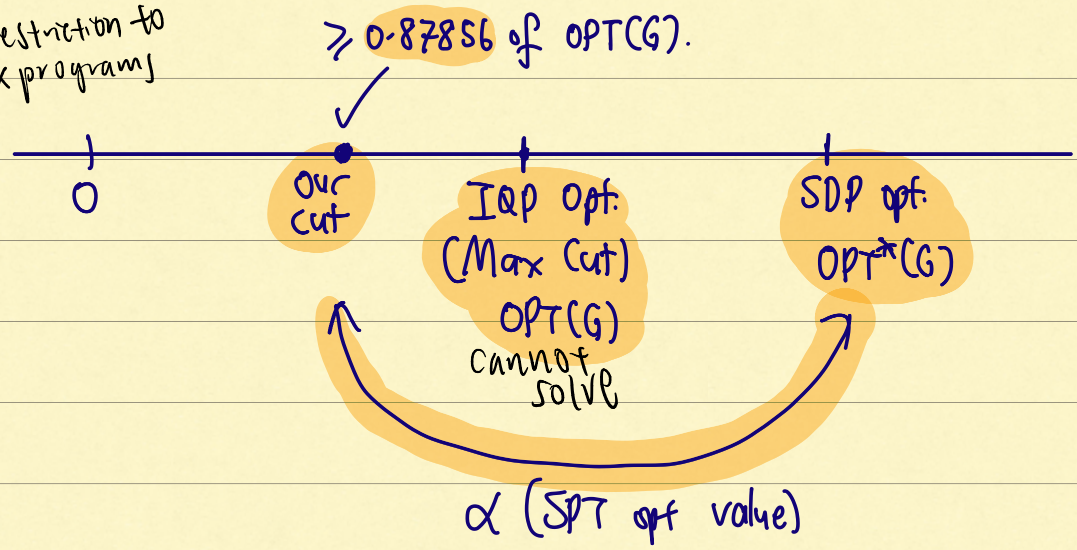
$\hookrightarrow$ count the number of edges where one is inside S and 1 is outside S

$$= \sum_{\{i,j\} \in E} \Pr(i, j \text{ sep})$$

$$\overset{\text{Lem.}}{\ge} \alpha \sum_{\{i,j\} \in E} \left(\frac{1 - \langle v_i, v_j\rangle}{2}\right)$$

$$= \alpha(\text{SDP opt. value})$$

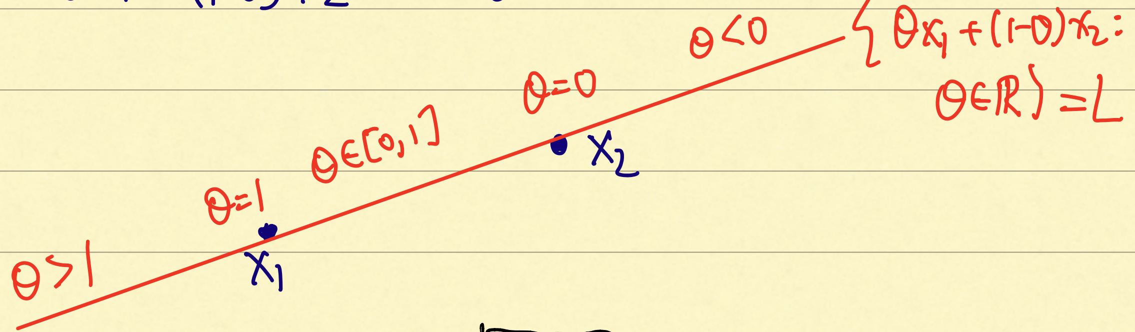
this number is the best that you can achieve given restriction to convex programs



Section 2.1 & 2.2 of BV.

$x_1, x_2 \in \mathbb{R}^n$ two distinct points in n-dim Euclidean space.

$y = \theta x_1 + (1-\theta) x_2 \qquad \theta \in \mathbb{R}$



Def: A set $C \subset \mathbb{R}^n$ is $\boxed{\text{affine}}$ if the line through any two distinct points of C belongs to C

$\forall x_1, x_2 \in C, \quad \theta x_1 + (1-\theta) x_2 \in C \quad \forall \theta \in \mathbb{R}.$

generalisation of the above

Ex: If C is an affine set, then $\forall x_1, \dots, x_k \in C$, then the point $\sum_{i=1}^{k} \theta_i x_i \in C$ if $\sum_{i=1}^{k} \theta_i = 1$

When $k=2$, it reduces to def. above.

Ex: If C is an $\boxed{\text{affine}}$ set and $x_0 \in C$, then the set

$$V = C - x_0 = \{x - x_0 : x \in C\}$$

is a subspace, i.e., closed under sum & scalar multiplication.

Pf: Take $v_1, v_2 \in V$ & $\alpha, \beta \in \mathbb{R}$.

$$\Downarrow$$

$$v_1 + x_0,\ v_2 + x_0 \in C$$

$$\alpha v_1 + \beta v_2 + x_0 = \underbrace{\alpha}_{}(v_1 + x_0) + \underbrace{\beta}_{}(v_2 + x_0) + \underbrace{(1-\alpha-\beta)}_{} x_0 \in C.$$

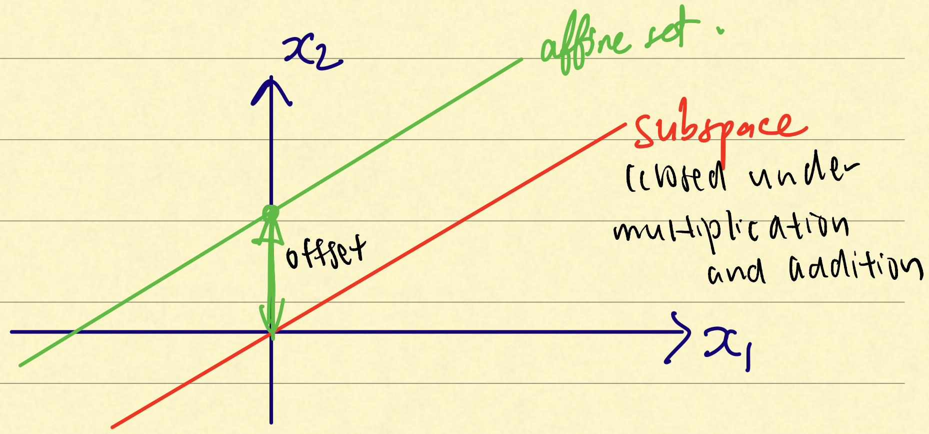
$\in C$

scalar
sum to one

$\because\ \alpha + \beta + (1-\alpha-\beta) = 1$

Thus $\alpha v_1 + \beta v_2 \in V$ since $\alpha v_1 + \beta v_2 + x_0 \in C$. ▨

Rmk: Any affine set C can be expressed as a subspace plus an offset



Ex: Solution of linear equation.

$C = \{x \in \mathbb{R}^n : Ax = b\}$ is affine set.

Take $x_1, x_2 \in C$. Then $Ax_1 = b$, $Ax_2 = b$. Take a.b. $\theta \in \mathbb{R}$.

Consider $A(\underbrace{\theta x_1 + (1-\theta)x_2}_{\in C}) = \theta A x_1 + (1-\theta) A x_2$

$= \theta b + (1-\theta) b = b.$

by linearity of matrix operations

$\Rightarrow$ Set C is affine.

Def: Set of all affine combinations of pts in a set $C \subset \mathbb{R}^n$ is called the affine hull of C

$$\text{aff}(C) = \left\{ \theta x_1 + \dots + \theta_k x_k : x_1, \dots, x_k \in C, \sum_{i=1}^{k} \theta_i = 1 \right\}$$

no need to be non-negative

Rmk: Affine hull of any set C is the smallest affine set that contains C.

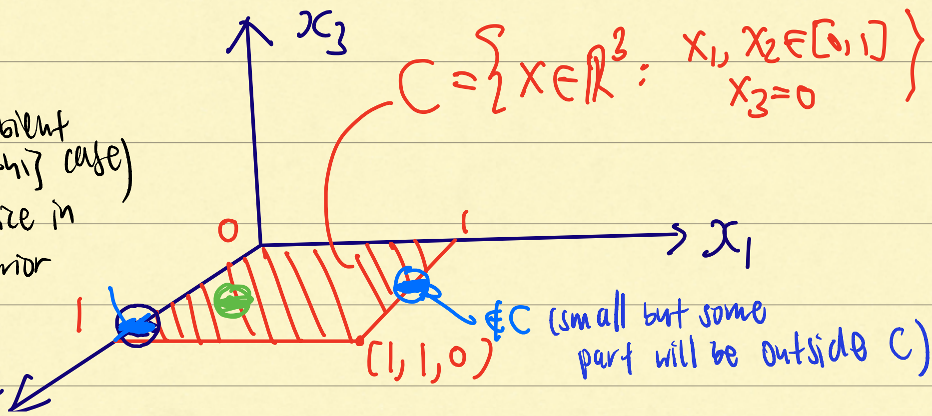
Relative interior of C.

$$\text{relint}(C) = \{ x \in C \mid \underbrace{B(x,r)}_{\text{ball centered at } x \text{ with radius } r} \cap \text{aff}(C) \subseteq C \text{ for some } r > 0 \}.$$

$\text{int}(C) = \emptyset$

interior $\Rightarrow$ relative to ambient space ($\mathbb{R}^3$ in this case)

$\hookrightarrow$ red portion only 1 slice in 3 dimension thus interior is empty



If you choose a small enough r, you will be strictly contained inside C, after done intersection with affine hull

Take any 2 points in red square and take all possible affine combinations → affine hull

$\mathrm{aff}(C) = x_1, x_2$ plane $= \{x \in \mathbb{R}^3 : x_3 = 0\}$.

$\mathrm{relint}(C) = \{ x \in \mathbb{R}^3 : x_1, x_2 \in (0,1),\ x_3 = 0 \}$

↳ open interval (do not contain sides)

Recall: Norm. Fix $x \in \mathbb{R}^n$.

$\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$

$\|x\|_1 = |x_1| + \cdots + |x_n|$

$\|x\|_\infty = \max_{1 \le i \le n} |x_i|$.

$\|x\|_p = \left(\sum |x_i|^p \right)^{1/p}$

$p \geq 1$

Ex: $0 \le p < 1$,

Is $\|x\|_p$ a norm?

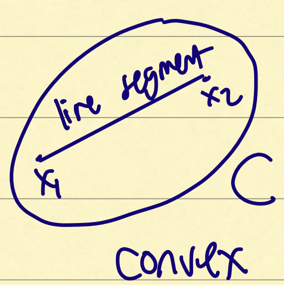
homogeneity, triangle inequality, $\|x\|_p = 0 \Rightarrow x = 0$

Convex Sets.

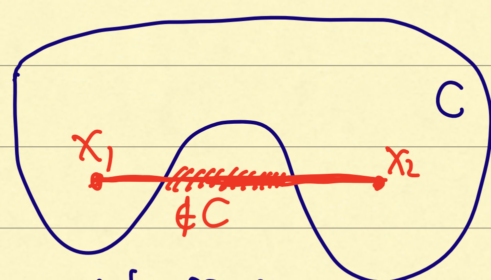
Def: A set $C \subseteq \mathbb{R}^n$ is convex if the line segment between any 2 points in C lies in C.

$\forall x_1, x_2 \in C$ and any $\theta \in [0,1]$, $\theta x_1 + (1-\theta) x_2 \in C$.

need to test $\boxed{\text{every}}$ 2 points



convex



not convex.

Def: <u>Convex combination</u> of points $x_1, \ldots, x_k \in C$ is a point of the form $\theta_1 x_1 + \cdots + \theta_k x_k$, $\theta_i \geq 0$, $\sum_{i=1}^{k} \theta_i = 1$

$i \in [k]$

coefficients (must all be non-negative and add up to 1)

(probability mass function)

Def: A set C is a <u>cone</u> if $\forall x \in C$, $\theta \geq 0$, $\theta x \in C$

C is a convex cone if it is convex & a cone.

$\Rightarrow \forall x_1, x_2 \in C$, $\theta_1, \theta_2 \geq 0$, $\theta_1 x_1 + \theta_2 x_2 \in C$.

A point $\underline{\sum_{i=1}^{k} \theta_i x_i}$ with $\underline{\theta_i \geq 0}$, $i \in \{1, \ldots, k\}$, is called a <u>conic</u> <u>combination</u> of $x_1, \ldots, x_k$

$\theta_i \geq 0$

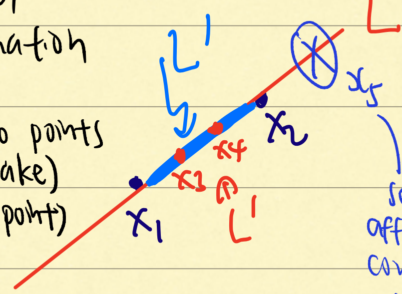
↳ adding up

$$\text{conichull}(C) = \Big\{ \underbrace{\theta_1 x_1 + \cdots + \theta_k x_k}_{\text{take any } k \text{ points in } C \text{ and look at the conic combination}} : \begin{matrix} x_i \in C \\ \theta_i \geq 0 \end{matrix} \ \forall i = 1, \ldots, k \Big\}$$

<u>Examples:</u>

1) ϕ: empty set is convex, affine (no two points to take)
2) $\{x_0\}$: singleton is convex, affine (same point)
3) $\mathbb{R}^n$: convex, affine (very big space)
4) Fix $x_1, x_2 \in \mathbb{R}^n$. $L = \{\theta x_1 + (1-\theta) x_2 : \theta \in \mathbb{R}\}$.

   convex, affine

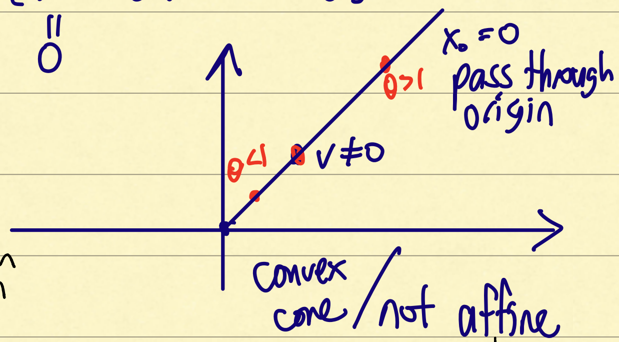


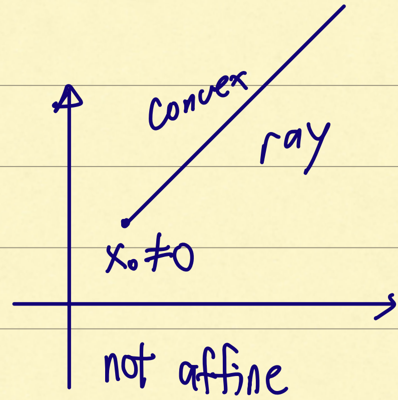
some affine combinations not in L' so L' is not affine

5) Line segment $L' = \{\theta x_1 + (1-\theta) x_2 : \theta \in [0, 1]\}$

   convex, not affine.

6) Ray $\{\underbrace{x_0}_{=0} + \theta v : \theta \geq 0\}$



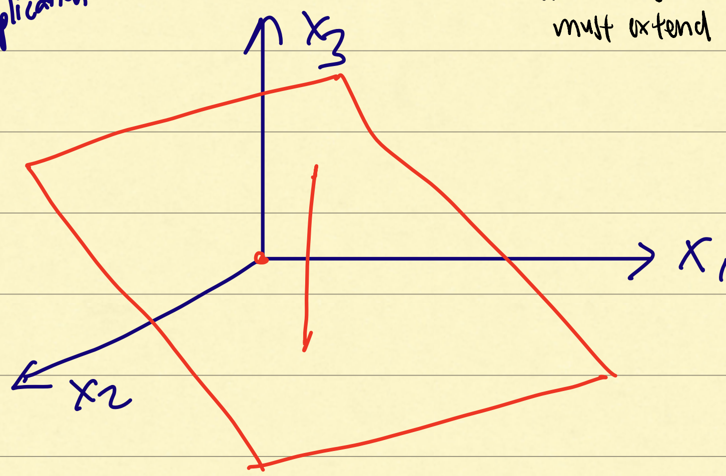


must contain the origin

7) A subspace is affine & a convex cone.

closed under addition & scalar multiplication

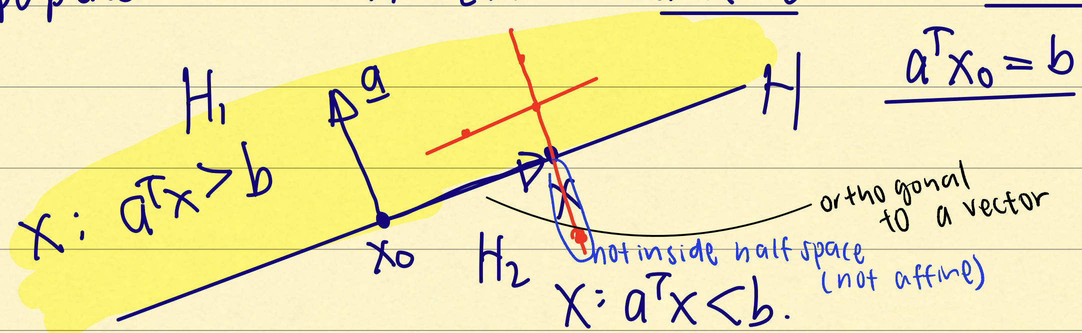
↳ roughly speaking, only extends in 1 direction (for something to be affine, must extend in all directions)



Hyperplanes & Halfspaces.

Hyperplane $\quad H = \{x \in \mathbb{R}^n : \underline{a^T x = b}\} = \{x : \underline{a^T(x - x_0)} = 0\}$

$\underline{a^T x_0 = b}$



If H is a hyperplane, it divides $\mathbb{R}^n$ into two halfspaces
↳ convex

$H_1 = \{x : a^T x > b\}$ open halfspace.

$H_1 = \{x : a^T x \geq b\}$ closed halfspace

$H_2 = \{x : a^T x < b\}$ open

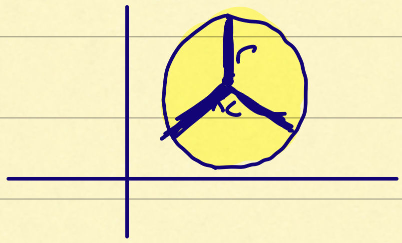
H_1 halfspace: convex, not affine.

Euclidean balls. $\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$, $\|x\|_2^2 = x^T x$

$B(x_c, r) = \{x : \|x - x_c\| \leq r\}$

↑ center ↑ radius

$= \{x : (x - x_c)^T (x - x_c) \leq r^2\}$



$B(x_c, r) = \{x_c + ru : \|u\| \leq 1\}$ all balls with radius r ⇒ translating it into location x_c and scaling it

Fact: $B(x_c, r)$ is $\boxed{\text{convex}}$

Pf: Take $x_1, x_2 \in B(x_c, r)$. $\|x_i - x_c\| \leq r \quad \forall i = 1, 2.$

Fix $\theta \in [0,1]$. Consider $\theta x_1 + (1-\theta) x_2$ convex combination

$$\|\theta x_1 + (1-\theta) x_2 - x_c\|$$

$\|a+b\| \leq \|a\| + \|b\|$

$$= \|\theta x_1 + (1-\theta) x_2 - \theta x_c - (1-\theta) x_c\|$$

we want to prove $\|\theta x_1 + (1-\theta) x_2 - x_c\| \leq r$

$$= \|\theta (x_1 - x_c) + (1-\theta)(x_2 - x_c)\|$$

$$\leq \|\theta (x_1 - x_c)\| + \|(1-\theta)(x_2 - x_c)\| \quad \text{(use } \triangle \text{ equality)}$$

$$= \theta \|x_1 - x_c\| + (1-\theta) \|x_2 - x_c\| \quad \text{(homogeneity property)}$$

$$\leq \theta r + (1-\theta) r = r$$

$\Rightarrow \theta x_1 + (1-\theta) x_2 \in B(x_c, r) \Rightarrow B(x_c, r)$ convex.

Ellipsoid: $E = \{x \in \mathbb{R}^n : (x - x_c)^T P^{-1} (x - x_c) \leq 1\}$.

$P = P^T > 0$: P is a symmetric positive definite matrix.
↳ invertible (eigenvalues all bigger than 0)

x_c: center

length of the semi-axes of E are given by $\sqrt{\lambda_i}$

λ_i: eigenvalues of P.

Euclidean ball = Ellipsoid $P = \underbrace{r^2 I}_{\text{proportional to identity matrix}}$ (I: identity)

$$(x - x_c)^T (r^2 I)^{-1} (x - x_c) \leq 1$$

$$\iff \|x - x_c\|^2 \leq r^2 \iff \|x - x_c\| \leq r.$$

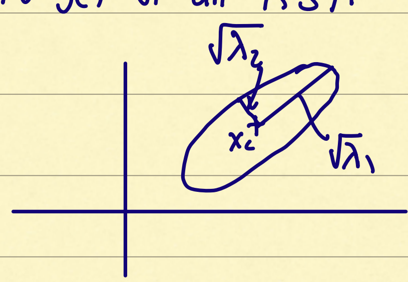
(equation of a sphere)

P : PD. $\exists A = P^{1/2}$. The ellipsoid is the set of all x s.t.

$$(x - x_c)^T (A^T A)^{-1} (x - x_c) \leq 1$$

$$\iff (x - x_c)^T A^{-1} (A^T)^{-1} (x - x_c) \leq 1$$

$$\|\underbrace{(A^{-1})^T (x - x_c)}_{u}\| \leq 1$$



Any point in the ellipsoid can be written as
$x = x_c + Au$ where $\|u\| \leq 1$

Norm: $\{x : \|x - x_c\| \leq r\}$ is convex.
(ball)

Norm cone: $\{(x, t) \in \mathbb{R}^{n+1} : \|x\| \leq t\}$ is a convex cone.
(x: n-dim; t is a scalar)

Pf convex: $(x_1, t_1), (x_2, t_2) \in C$

$\|x_1\| \le t_1, \quad \|x_2\| \le t_2$

column vector of length n+1

Consider $\theta(x_1, t_1) + (1-\theta)(x_2, t_2) \in C \qquad \theta \in [0,1]$

$\|\theta x_1 + (1-\theta) x_2\| \le \theta \|x_1\| + (1-\theta)\|x_2\|$

norm of the first n coordinates less than last coordinate

$\le \theta t_1 + (1-\theta) t_2.$

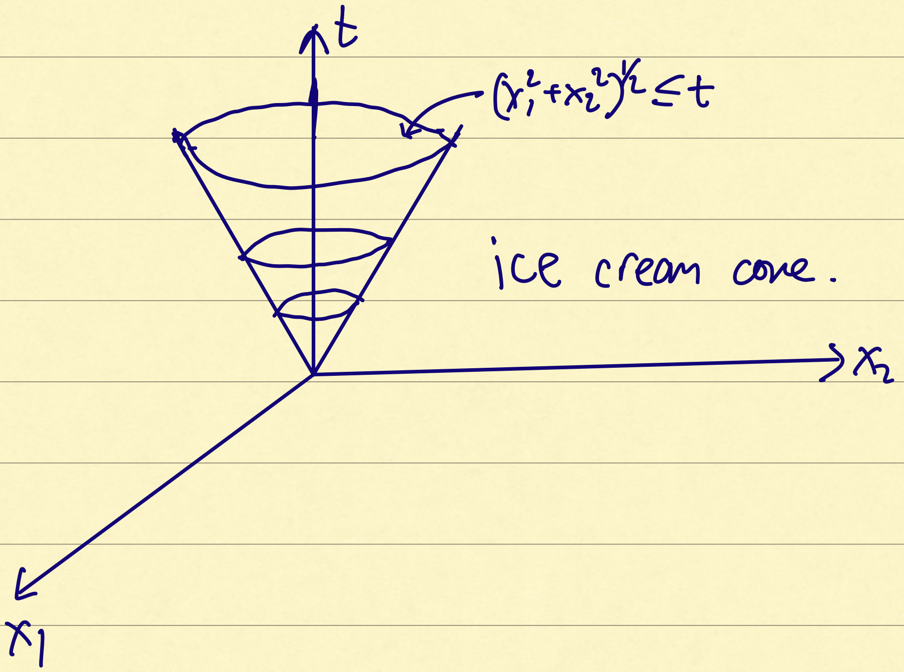
show that its convex

Pf cone: Take any point $(x,t) \in C$. Fix $\theta \ge 0$.

WTS: $\theta \underbrace{(x,t)}_{\in \mathbb{R}^{n+1}} \in C$. $\quad \theta \begin{bmatrix} x \\ t \end{bmatrix} \in C$

$\|\theta x\| \le \theta \|x\| \le \theta t$ homogeneity //

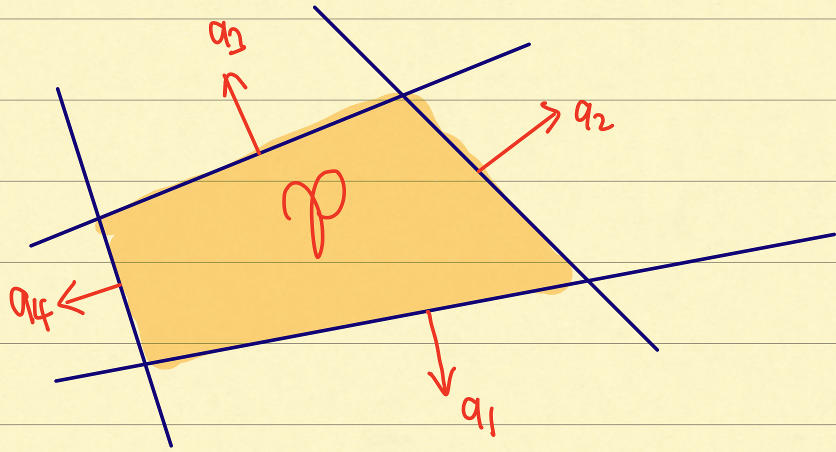
$\{(\underbrace{x_1, x_2}_{\underline{x}}, t) \in \mathbb{R}^3 : \underbrace{\sqrt{x_1^2 + x_2^2}}_{\text{norm}} \le t\}$



ice cream cone.

Polyhedra (Plural of polyhedron)

$$\mathcal{P} = \left\{ x \in \mathbb{R}^n : \begin{array}{l} a_j^T x \leq b_j, \ j = 1, \dots, m \\ c_j^T x = d_j, \ j = 1, \dots, p \end{array} \right\}$$



Can be written more concisely as.

$$\mathcal{P} = \{ x : \underbrace{Ax \leq b}, \ Cx = d \}.$$

vector inequality

vector componentwise.

$$A = \begin{bmatrix} — a_1^T — \\ \vdots \\ — a_m^T — \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \quad C = \begin{bmatrix} — c_1^T — \\ \vdots \\ — c_p^T — \end{bmatrix}$$

$$d = \begin{bmatrix} d_1 \\ \vdots \\ d_p \end{bmatrix}$$

Simplexes (plural of simplex)

$v_0, \dots, v_k \in \mathbb{R}^n$ affinely independent

$\iff$ $v_1 - v_0, v_2 - v_0, \dots, v_k - v_0$ are linearly independent

θ is a vector   component wise bigger than 0

$$C = \mathrm{conv}\{v_0, \dots, v_k\} = \{\theta_0 v_0 + \dots + \theta_k v_k \mid \underset{\in \mathbb{R}^{k+1}}{\theta} \succeq 0,\ \underbrace{\mathbf{1}^T\theta = 1}_{\sum_{i=0}^{k}\theta_i = 1}\}$$

adding up components

Ex: $v_0 = 0,\ v_1 = e_1,\ v_2 = e_2,\ \dots,\ v_k = e_k$

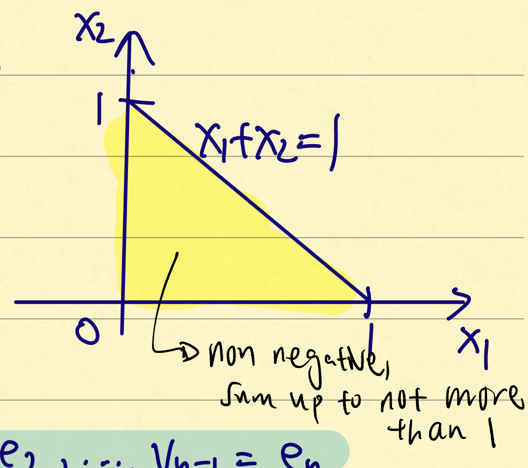
$e_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \leftarrow i^{th}$ location. (i^{th} unit vector)

sum of all comp. of x.

C = set of all x s.t. $x \succeq 0,\ \widetilde{\mathbf{1}^T x} \le 1$

Why? $x = \theta_0 \cdot \underbrace{0}_{v_0} + \theta_1 \underbrace{e_1}_{v_1} + \dots + \theta_k \underbrace{e_k}_{v_k} = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_k \end{bmatrix} \succeq 0$

$\sum_{i=0}^{k} \theta_i = 1 \Leftrightarrow \underbrace{\theta_0}_{\ge 0} + \underbrace{\theta_1 + \dots + \theta_k}_{\le 1} = 1$



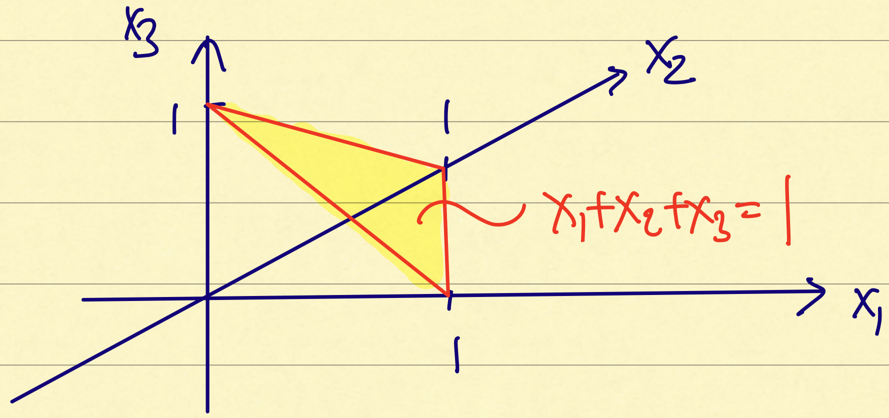
Ex: Probability simplex. $v_0 = e_1,\ v_1 = e_2,\ \dots,\ v_{n-1} = e_n$

C contains all x s.t.

$$x = \theta_0 e_1 + \theta_1 e_2 + \dots + \theta_{n-1} e_n = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_{n-1} \end{bmatrix} \succeq 0$$

$$\theta_i \ge 0,\quad \sum_{i=0}^{n-1} \theta_i = 1$$

x is s.t. $x \geq 0$, $\sum x_i = 1$



Q: Affine? No!   Q: Cone No!

Q: Convex? YES!

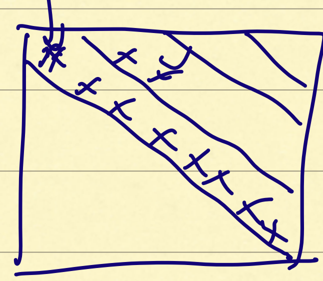
Positive Semidefinite Cone

S^n: set of $n \times n$ symmetric matrices

$X \in S^n$ iff $X = X^T$

vector space with dim $\frac{n(n+1)}{2}$,

$n + \binom{n}{2} = \frac{n(n+1)}{2}$



↳ need to describe diagonal entries and off diagonals

S_+^n: set of PSD symmetric matrices

$X \in S_+^n$ iff $\forall z \in \mathbb{R}^n \quad z^T X z \geq 0$

S_{++}^n: set of positive definite matrices

$X \in S_{++}^n$ iff $\forall z \in \mathbb{R}^n \setminus \{0\}, \quad z^T X z > 0$

S_+^n is a (convex cone)

Pf: $A, B \in S_+^n$. $\theta_1, \theta_2 \geq 0$. Fix $z \in \mathbb{R}^n$.

Consider $z^T(\theta_1 A + \theta_2 B) z$

$$= \theta_1 \underbrace{z^T A z}_{\geq 0} + \theta_2 \underbrace{z^T B z}_{\geq 0} \geq 0$$

////

Fix $\theta \in [0,1]$. Consider $\theta A + (1-\theta) B$