$$\min_{x \in D} f(x)$$

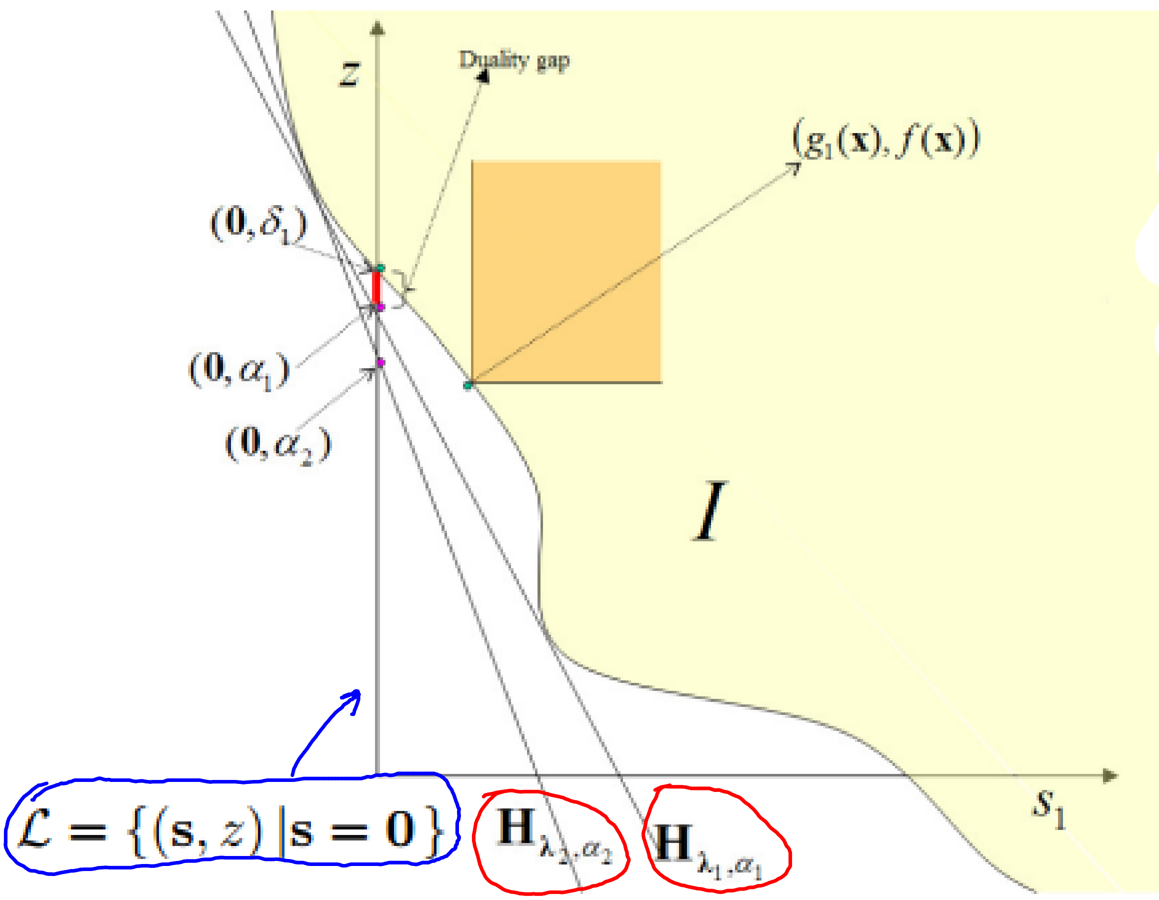
$$\text{s.t. } g_i(x) \le 0 \quad i=1\ldots m$$

(The general dual problem & its geometric interpretation)

pg 292, sec 4.4.3 of http://www.cse.iitb.ac.in/~cs709/notes/BasicsOfConvexOptimization.pdf

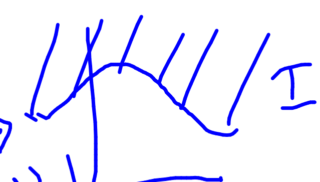
Consider the set:

$$\mathcal{I} = \{(\mathbf{s}, z) \mid \mathbf{s} \in \Re^m,\ z \in \Re,\ \exists \mathbf{x} \in \mathcal{D} \ with\ g_i(\mathbf{x}) \le s_i\ \forall 1 \le i \le m,\ f(\mathbf{x}) \le z\}$$



smallest z value in I for $s_1 \le 0$ will be for $s_1 = 0$ since $(s_1, z_1) \in I \Rightarrow (s_2, z_2) \in I$ $\forall s_2 \ge s_1$ & $z_2 \ge z_1$

$$\mathcal{H}_{\lambda,\alpha} = \{(\mathbf{s}, z) \mid \lambda^T.\mathbf{s} + z = \alpha\}$$



Thus ⇒ is not possible!

$$\begin{array}{ll} \max & \alpha \\ \text{subject to} & \mathcal{H}^{+}_{\lambda,\alpha} \supseteq \mathcal{I} \end{array}$$

$$\begin{array}{ll} \max & \alpha \\ \text{subject to} & \lambda^T.\mathbf{s} + z \geq \alpha \ \forall(\mathbf{s}, z) \in \mathcal{I} \end{array}$$

a

$$\begin{array}{ll} \max & \alpha \\ A_\alpha \left\{ \begin{array}{l} \text{subject to} \\ \end{array} \right. & \begin{array}{l} \lambda^T.\mathbf{s} + z \geq \alpha \ \forall(\mathbf{s}, z) \in \mathcal{I} \\ \lambda \geq \mathbf{0} \end{array} \end{array}$$

If $\exists$ α_b soln to b,
s.t $\lambda^T g(x_b) + f(x_b) = \alpha_b$
then $s = g(x_b)$ $z = f(x_b)$
is soln to a
complete proof of equivalence

$$\begin{array}{ll} \max & \alpha \\ B_\alpha \left\{ \begin{array}{l} \text{subject to} \\ \end{array} \right. & \begin{array}{l} \lambda^T.\mathbf{g}(\mathbf{x}) + f(\mathbf{x}) \geq \alpha \ \forall \mathbf{x} \in \mathcal{D} \\ \lambda \geq \mathbf{0} \end{array} \end{array}$$

b

Proof. $A_\alpha \subseteq B_\alpha$ ie $\{\alpha \mid \lambda^T s + z \geq \alpha \ \forall (s,z) \in \mathcal{I}\} \subseteq \{\alpha \mid \lambda^T g(x) + f(x) \geq \alpha \ \forall x \in D\}$

for if $\lambda^T g(x) + f(x) < \alpha$ for some α & x then for $(s,z) = (g(x), f(x)) \in \mathcal{I}$, $\lambda^T s + z < \alpha \Rightarrow \alpha \notin A_\alpha$

$B_\alpha \subseteq A_\alpha$ ie $\{\alpha \mid \lambda^T g(x) + f(x) \geq \alpha \ \forall x \in D\} \subseteq \{\alpha \mid \lambda^T s + z \geq \alpha \ \forall (s,z) \in \mathcal{I}\}$

for if $\lambda^T g(x) + f(x) \geq \alpha$ $\forall x \in D$ then since for each $(s,z) \in \mathcal{I}$, $\exists$ $(g(x), f(x)) \leq (s,z)$,
$\lambda^T s + z \geq \alpha$ ie $\alpha \in A_\alpha$

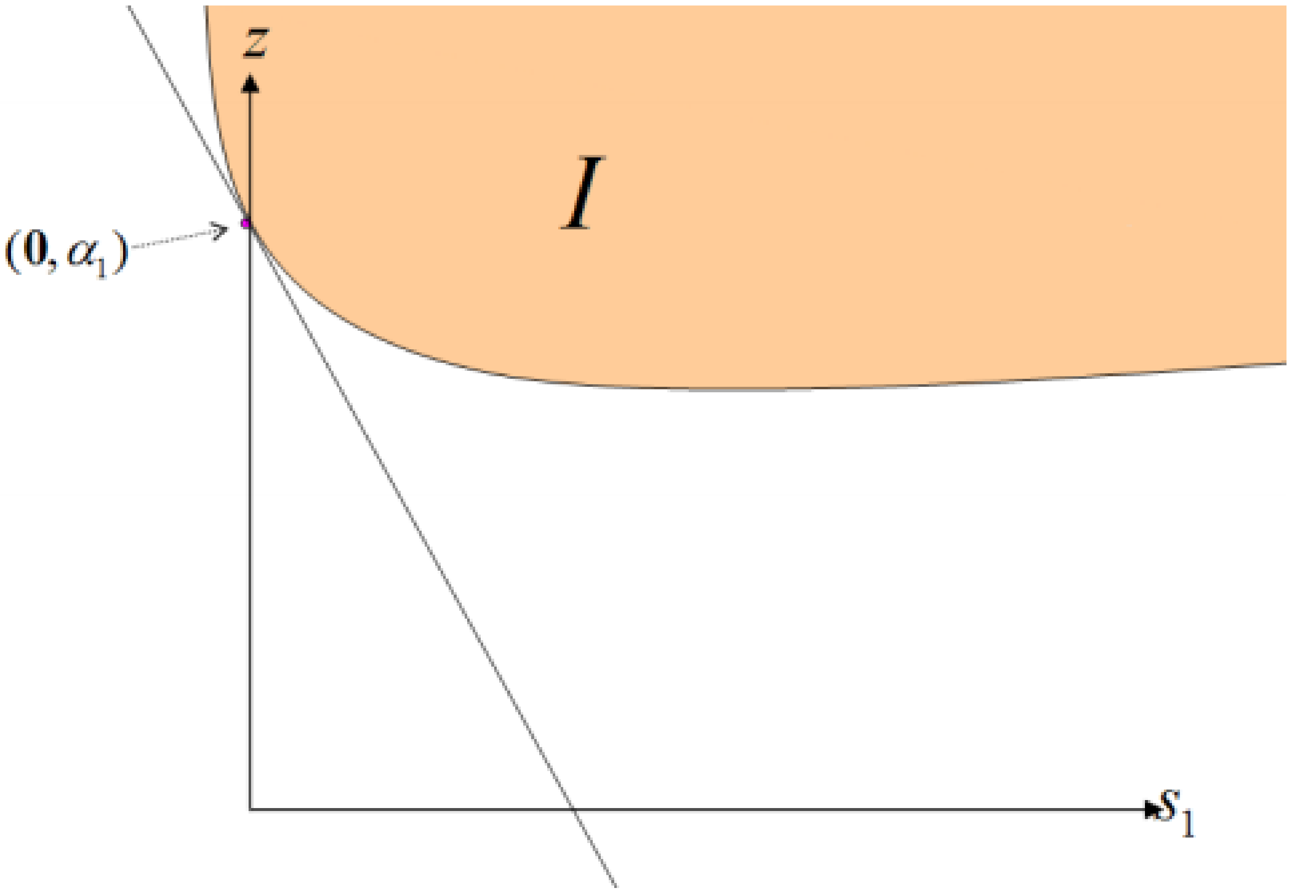
$$
\begin{array}{ll}
\max & \alpha \\
\text{subject to} & L(\mathbf{x}, \lambda) \geq \alpha \ \forall \mathbf{x} \in \mathcal{D} \\
& \lambda \geq \mathbf{0}
\end{array}
$$

Since, $L^*(\lambda) = \min\limits_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \lambda)$, we can deal with the equivalent

$$
\begin{array}{ll}
\max & \alpha \\
\text{subject to} & L^*(\lambda) \geq \alpha \\
& \lambda \geq \mathbf{0}
\end{array}
$$

This problem can be restated as

$$
\begin{array}{ll}
\max & L^*(\lambda) \\
\text{subject to} & \lambda \geq \mathbf{0}
\end{array}
$$



Q: What is desirable of the set $\mathcal{I}$ for zero duality gap?

Ans: $\exists\ (0,\alpha) \in \mathcal{I}$ and λ s.t. $\lambda^T s + z \geq \alpha\ \forall\ (s,z) \in \mathcal{I}$

& Intersection of $\mathcal{I}$ with z axis is closed below

$\Leftrightarrow$ $\exists$ a supporting hyperplane to $\mathcal{I}$ at $(0,\alpha)$ & Intersection of $\mathcal{I}$ with z axis is closed below (with $(0,\alpha)$ being boundary pt)

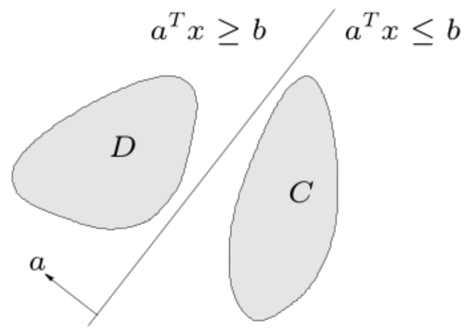
$\Leftarrow$ $\mathcal{I}$ is closed & $\exists$ a supporting hyperplane to $\mathcal{I}$ at every boundary point

kept aside for time being

Separating hyperplane theorem

if C and D are disjoint convex sets, then there exists $a \neq 0$, b such that

$$a^T x \leq b \text{ for } x \in C, \qquad a^T x \geq b \text{ for } x \in D$$



the hyperplane $\{x \mid a^T x = b\}$ separates C and D

strict separation requires additional assumptions (*e.g.*, C is closed, D is a singleton)

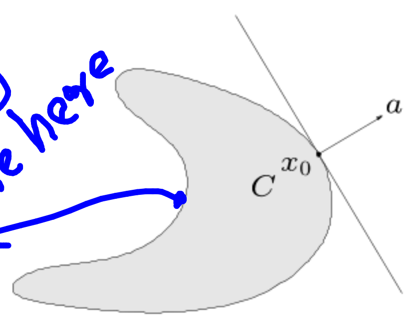
Consequence

Supporting hyperplane theorem

supporting hyperplane to set C at boundary point x_0:

$$\{x \mid a^T x = a^T x_0\}$$

where $a \neq 0$ and $a^T x \leq a^T x_0$ for all $x \in C$



no supporting hyperplane here

supporting hyperplane theorem: if C is convex, then there exists a supporting hyperplane at every boundary point of C

Proof (from separating hyperplane theorem):

(a) interior $(C) \neq \phi$ (so that $\text{int}(C) \cap \{x_0\} = \phi$)

Apply separating hyperplane theorem

to sets $C' = \{x_0\}$ and $D' = \text{interior}(C)$

Strict separation not applicable since $\text{int}(C)$ is open

$\exists\, a, b$ s.t. $a^T x \geq b \;\; \forall x \in D'$

$a^T x_0 \leq b$

This inequality extends to all boundary (limit pts) leading to $a^T x_0 \geq b$

combined: $a^T x_0 = b$

(b) $\text{int}(C) = \phi$

$\Rightarrow$ Any hyperplane containing that affine set contains C & x_0 (H/w: Prove)

$\Rightarrow$ This hyperplane is a trivial supporting hyperplane

H/W: Let $C \subseteq \mathbb{R}^n$ be a convex set. There exists a hyperplane H that contains X iff $\text{int}(C) = \phi$

if part

Proof: $\text{int}(C) = \phi \Rightarrow \exists$ hyperplane H s.t. $C \subseteq H$

We will prove by contradiction. Suppose there is no hyperplane H s.t. $C \subseteq H$. Let $x_0, x_1, x_2 \ldots x_n \in C$

Consider $y_i = x_i - x_0 \quad i = 1 \ldots n$. If y_i's were linearly dependent, then $\exists\ \alpha_1 \ldots \alpha_n \in \mathbb{R}$ not all 0 s.t.

$$\sum_{i=1}^{n} \alpha_i y_i = 0 \quad \text{ie} \quad \text{for every } x_i,\ a^T x_i = b \text{ for some } a \,\&\, b$$

why?

ie $x_0 \ldots x_n$ are affinely dependent

$\Rightarrow x_1 \ldots x_n$ & their convex combinations all lie on the hyperplane $a^T x = b$

$\Rightarrow$ Contradicts assumption that $\exists$ no hyperplane $H \supseteq C$

$\therefore y_1 \ldots y_n$ are linearly independent & span $\mathbb{R}^n$

Now for $\alpha_0 \ldots \alpha_n \in [0,1]$ & $\sum_{i=0}^{n} \alpha_i = 1$, $z = \sum_{i=0}^{n} \alpha_i x_i \in X$

$z = x_0 + \sum_{i=1}^{n} \alpha_i y_i$ which is a representation for any $z \in C$ with $\sum_{i=1}^{n} \alpha_i \leq 1$ & $\alpha_i \in [0,1]$

Fix a point $\hat{z} = x_0 + \sum_{i=1}^{n} \hat{\alpha}_i y_i$

For any z near $\hat{z}$, it has representation

No constraints on β_i ← $z = x_0 + \sum_{i=1}^{n} \beta_i y_i$ (since $z - x_0 \in \mathbb{R}^n$ & $y_1 \dots y_n$ is a basis for $\mathbb{R}^n$)

with β_i's close to α_i's $\forall\, i$

As $z \to \hat{z}$, $\beta_i \to \alpha_i$ for each i & $\sum \beta_i < 1$ } $\because y_1 \dots y_n$ is basis for $\mathbb{R}^n$

$\Rightarrow$ In other words, C contains a ball with center at $\hat{z}$ and sufficiently small radius.

$\Rightarrow$ $z \in \text{int}(C)$ contradicting that $\text{int}(C) = \phi$

$\text{Int}(C) \neq \phi \Rightarrow$ No hyperplane H contains C → only if part

Let $x \in \text{Int}(C)$. Then $\exists\ \delta > 0$ s.t

$$B_{2\delta}(x) = \{ y \mid \|y - x\|_2 \leq 2\delta \} \subseteq C$$

$\Rightarrow \exists\ n$ points $x + \delta e_i$ $e_i = (0, \dots 1, 0 \dots 0)$ s.t $x + \delta e_i \in C$

$\Rightarrow$ If Hyperperplane $H_{\lambda, \alpha}$ passes through x (ie $\lambda^T x = \alpha$) then if it contains $x + \delta e_i$

$$\lambda^T (x + \delta e_i) = \lambda^T x + \delta(\lambda^T e_i) = \alpha + \delta \lambda_i = \alpha$$

which requires $\lambda_i = 0\ \forall\, i$. $\therefore$ That is $\exists$ no hyperplane in $\mathbb{R}^n$ that contains $B_{2\delta}(x)$ or even C

Reason why $\sum \alpha_i y_i = 0 \Rightarrow x_1 \dots x_n$ lie on a hyperplane (not all α_i's are zero) & $y_i = x_i - x_0$

$$\underbrace{\sum_{i=1}^{n} \alpha_i y_i = 0 \text{ (not all } \alpha_i = 0)}_{\text{Linear dependence of } y_i\text{'s}} \Leftrightarrow \underbrace{\underbrace{\sum_{i=1}^{n} \alpha_i x_i}_{\text{Not all } \alpha_i = 0} = x_0 \underbrace{\sum_{i=1}^{n} \alpha_i}_{\neq 0}}_{\text{Affine dependence of } x_0, x_1 \dots x_n} \Leftrightarrow (x_1 \dots x_n) \text{ lie on a hyperplane}$$

(ie an affine subspace of dim $n-1$ in $\mathbb{R}^n$)

since $\sum_{i=0}^{n} \beta_i x_i = 0$

where $\sum_{i=0}^{n} \beta_i = 0$

& $\beta_i = \alpha_i$ for $i = 1..n$

& $\beta_0 = -\sum \alpha_i$

$V_0 .. V_n$ are affinely dependent if one of them can be expressed as an affine combination of the others

$$\Leftrightarrow V_0 = \sum_{i=1}^{n} \lambda_i V_i \quad \left(\sum_{i=1}^{n} \lambda_i = 1\right) \Leftrightarrow \sum_{i=0}^{n} \beta_i V_i = 0 \quad \left(\beta_i = \lambda_i \ i=1..n \ \& \ \beta_0 = -1 = \sum_{i=1}^{n} \lambda_i\right)$$

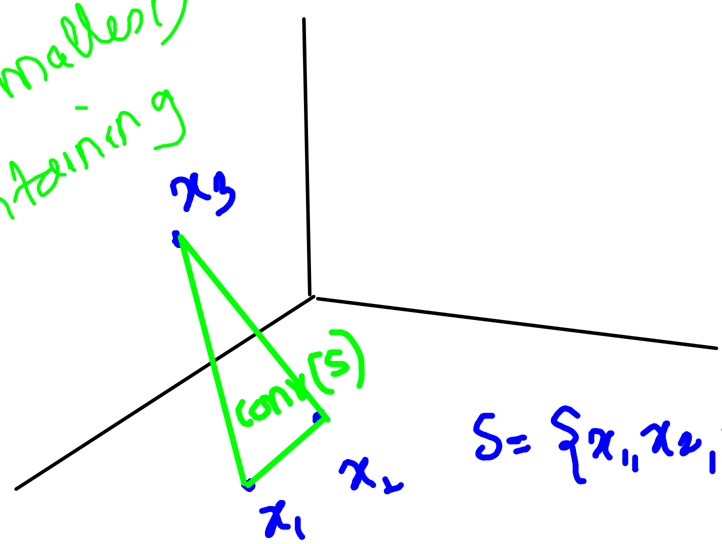
http://www.math.bgu.ac.il/~shakhar/teaching/combinatorial_geometry_files/lec-notes.pdf

Aside: We just saw connection between "linear dependence" & "affine dependence"

Is there a "convex dependence?"

Caratheodory theorem: Let $S \subset \mathbb{R}^n$ & let $\dim(\mathrm{conv}(S)) = m$. Then, every point $x \in \mathrm{conv}(S)$ is a convex combination of at most $m+1$ points from S (Proof in section B.2.1 of Nemirovski)

dim of tightest (smallest?) shifted vector space containing C



$S = \{x_1, x_2, x_3\}$

$$\dim(C) = \dim(\mathrm{aff}(C)) = \dim(\underbrace{V})$$

$C \subseteq \mathbb{R}^n$

$\mathrm{aff}(C) = V + a$

$$\min_{x \in \mathcal{D}} f(x)$$

(Going back to $\mathcal{I}$)

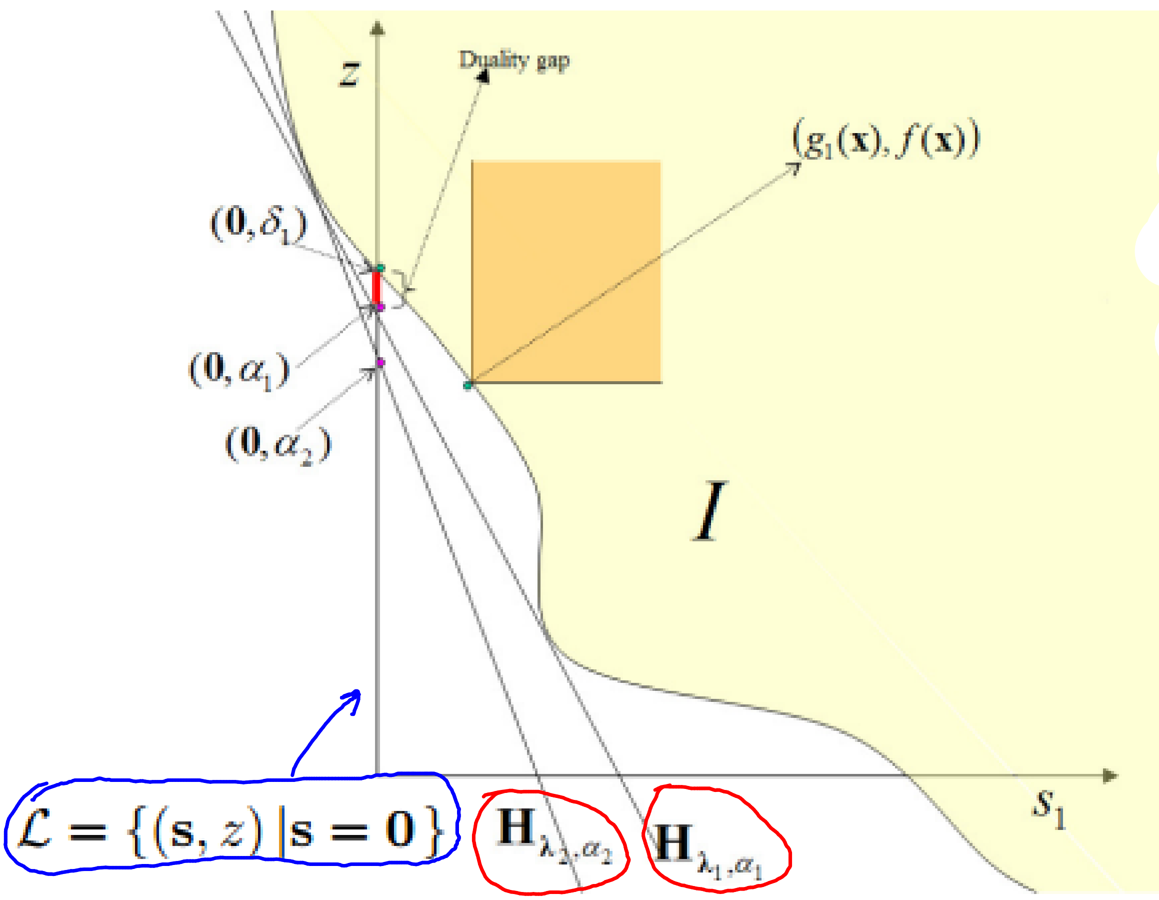
s.t. $g_i(x) \le 0 \quad i = 1 \dots m$

(The general dual problem & its geometric interpretation)

pg 292, sec 4.4.3 of http://www.cse.iitb.ac.in/~cs709/notes/BasicsOfConvexOptimization.pdf

Consider the set:

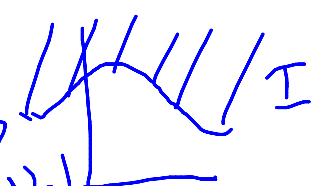
$$\mathcal{I} = \{(\mathbf{s}, z) \mid \mathbf{s} \in \Re^m,\ z \in \Re,\ \exists \mathbf{x} \in \mathcal{D} \text{ with } g_i(\mathbf{x}) \le s_i\ \forall 1 \le i \le m,\ f(\mathbf{x}) \le z\}$$



Smallest z value in $\mathcal{I}$ for $s_1 \le 0$ will be for $s_1 = 0$

Since $(s_1, z_1) \in \mathcal{I}$
$\Rightarrow (s_2, z_2) \in \mathcal{I}$
$\forall s_2 \ge s_1 \ \& \ z_2 \ge z_1$

$$\mathcal{H}_{\lambda,\alpha} = \{(\mathbf{s}, z) \mid \lambda^T.\mathbf{s} + z = \alpha\}$$



Thus $\Rightarrow$ is not possible!

Q: When is I convex?

Ans: $I = \left\{ (s,z) \mid s \in \mathbb{R}^m, z \in \mathbb{R}, \exists x \in D \text{ s.t. } g_i(x) \le s_i \; i=1..m, \; f(x) \le z \right\}$

I is projection/restriction of the epigraph of the vector valued fn $\bar{f}(x) = \begin{bmatrix} f(x) \\ g_1(x) \\ \vdots \\ g_m(x) \end{bmatrix}$

$$\text{epi}(\bar{f}) = \left\{ (x,s,z) \mid x \in D, f(x) \le z, g_i(x) \le s_i \; i=1..m \right\}$$

(In general, w.r.t. a generalized inequality $\preceq_K$,

$$\text{epi}_K(\bar{f}) = \left\{ (x,t) \mid x \in D, \bar{f}(x) \preceq_K t \right\}$$
)

$\boxed{\begin{array}{l}\text{Based on midsem Q2(a),} \\ \text{OR closure under affine transform} \\ \text{.... More properties on following slides}\end{array}}$

$\text{epi}(\bar{f})$ is convex

$\Downarrow$ (Take $A = \begin{bmatrix} I & 0 \end{bmatrix}$ $b=0$

$(s,z) \times (s,z)$ $\quad$ $(s,z) \times \{x\}$

I is convex

Converse does not hold. Inverse image of $A\,\text{epi}(\bar{f})$ is NOT $\text{epi}(\bar{f})$

Q: When is $\text{epi}(\bar{f})$ convex?

BEGIN: SUPPLEMENTARY NOTES FOR CONVEX SETS

•

Operations that preserve convexity

practical methods for establishing convexity of a set C

1. apply definition

$$x_1, x_2 \in C, \quad 0 \le \theta \le 1 \quad \Longrightarrow \quad \theta x_1 + (1-\theta)x_2 \in C$$

2. show that C is obtained from simple convex sets (hyperplanes, halfspaces, norm balls, . . .) by operations that preserve convexity
   - intersection
   - affine functions
   - perspective function
   - linear-fractional functions

Intersection

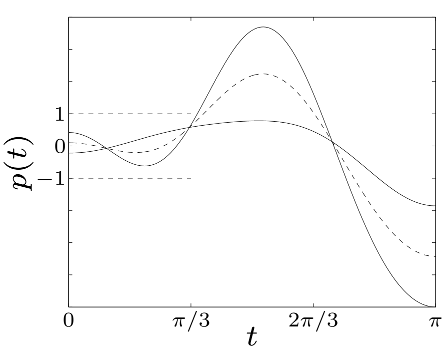
the intersection of (any number of) convex sets is convex

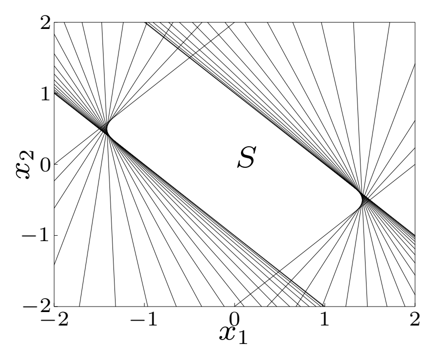
example:

$$S = \{x \in \mathbf{R}^m \mid |p(t)| \le 1 \text{ for } |t| \le \pi/3\}$$

where $p(t) = x_1 \cos t + x_2 \cos 2t + \cdots + x_m \cos mt$

for $m = 2$:





Operations that preserve convexity

practical methods for establishing convexity of a set C

1. apply definition

show that

$$\forall x_1, x_2 \in C, \forall 0 \leq \theta \leq 1 \quad \Longrightarrow \quad \theta x_1 + (1-\theta)x_2 \in C$$

2. show that C is obtained from simple convex sets (hyperplanes, halfspaces, norm balls, . . .) by operations that preserve convexity
   - intersection
   - affine functions
   - perspective function
   - linear-fractional functions

3. Empirical / Experimental [Homework]

Look for "smart" ideas

You may want to sample x_1 & x_2 along boundary instead of randomly

Intersection

the intersection of (any number of) convex sets is convex

↳ show that S^n_+ is convex using this property

example:

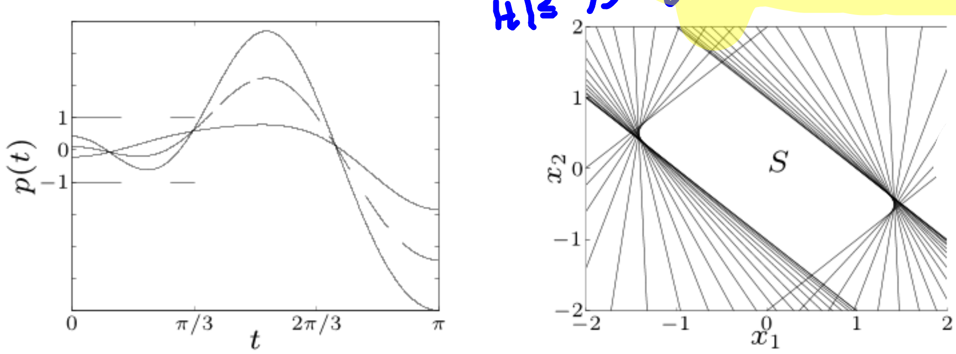
$$S = \{x \in \mathbf{R}^m \mid |p(t)| \leq 1 \text{ for } |t| \leq \pi/3\}$$

where $p(t) = x_1 \cos t + x_2 \cos 2t + \cdots + x_m \cos mt$

$$S = \bigcap_{|t| \leq \pi/3} \{x \in R^m \mid |\sum_{i=1}^{m} x_i \cos(it)| \leq 1\}$$

this is convex for fixed t

for $m = 2$:

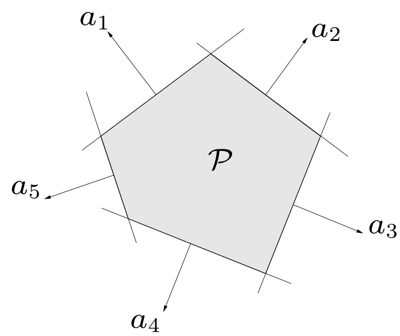


Polyhedra

solution set of finitely many linear inequalities and equalities

$$Ax \preceq b, \qquad Cx = d$$

($A \in \mathbf{R}^{m \times n}$, $C \in \mathbf{R}^{p \times n}$, $\preceq$ is componentwise inequality)



polyhedron is intersection of finite number of halfspaces and hyperplanes

Positive semidefinite cone

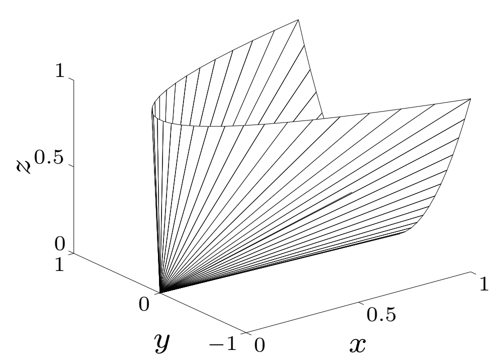
notation:

- $\mathbf{S}^n$ is set of symmetric $n \times n$ matrices
- $\mathbf{S}^n_+ = \{X \in \mathbf{S}^n \mid X \succeq 0\}$: positive semidefinite $n \times n$ matrices

$$X \in \mathbf{S}^n_+ \iff z^T X z \geq 0 \text{ for all } z$$

  $\mathbf{S}^n_+$ is a convex cone
- $\mathbf{S}^n_{++} = \{X \in \mathbf{S}^n \mid X \succ 0\}$: positive definite $n \times n$ matrices

example: $\begin{bmatrix} x & y \\ y & z \end{bmatrix} \in \mathbf{S}^2_+$

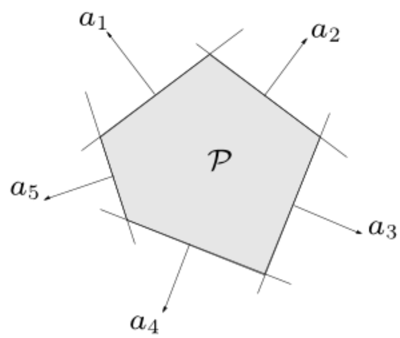


Polyhedra

solution set of finitely many linear inequalities and equalities

$$Ax \preceq b, \qquad Cx = d$$

($A \in \mathbf{R}^{m \times n}$, $C \in \mathbf{R}^{p \times n}$, $\preceq$ is componentwise inequality)



polyhedron is intersection of finite number of halfspaces and hyperplanes

The Hahn Banach Thm: Any closed convex set in $\mathbb{R}^n$ is equivalent to intersection of all halfspaces that contain it

Positive semidefinite cone

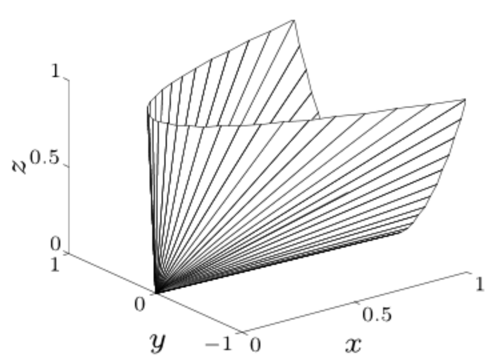
notation:

- $\mathbf{S}^n$ is set of symmetric $n \times n$ matrices
- $\mathbf{S}^n_+ = \{X \in \mathbf{S}^n \mid X \succeq 0\}$: positive semidefinite $n \times n$ matrices

$$X \in \mathbf{S}^n_+ \quad \Longleftrightarrow \quad z^T X z \geq 0 \text{ for all } z$$

  $\mathbf{S}^n_+$ is a convex cone
- $\mathbf{S}^n_{++} = \{X \in \mathbf{S}^n \mid X \succ 0\}$: positive definite $n \times n$ matrices

example: $\begin{bmatrix} x & y \\ y & z \end{bmatrix} \in \mathbf{S}^2_+$



Primal OR Edge representation

ⓐ $\underbrace{\text{Convex hull}(S)}_{\text{denoted } conv(S)}$ = set of all convex combinations of pts in S

ⓑ $\underbrace{\text{Convex hull}(S)}_{\text{denoted } conv(S)}$ = Smallest convex set that contains S [Prove as h/w]

Also: The idea of a convex combination can be generalised to include infinite sums, integrals, and, in the most general form, probability distributions

Primal OR Edge representation

ⓐ $\underbrace{\text{Conic/Affine hull}(S)}_{conic(S) \text{ or } aff(S)}$ = set of all conic/affine combination of pts in S

ⓑ $\underbrace{\text{Conic/Affine hull}}_{conic(S) \text{ or } aff(S)}$ (S) = Smallest conic/affine set that contains S

S is called basis of vector space V iff lin_span(S) = V

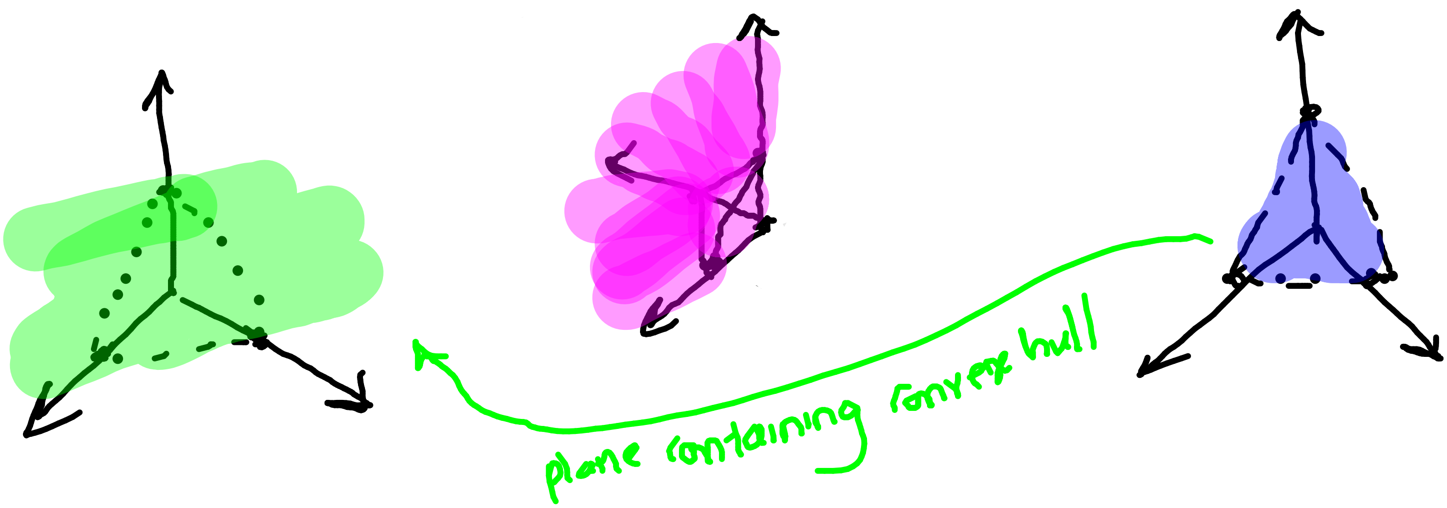
S is called affine basis of affine set A iff aff(S) = A

S is called conically spanning set of cone K iff conic(S) = K

S is called convexly spanning set of convex set C iff conv(S) = C

Let $S = \{(1,0,0), (0,1,0), (0,0,1)\}$

what is aff hull(S)? what is conichull(S)? What is convex hull(S)



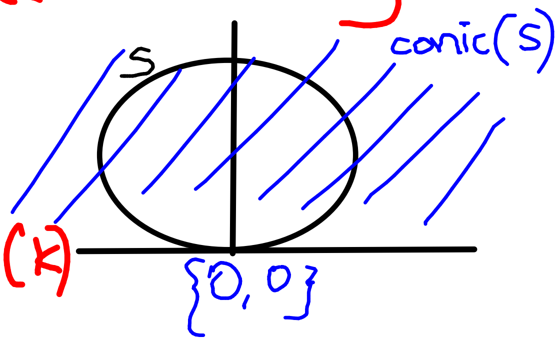
2–9

Further notes:

1. conic(S) is not always closed.

   Eg: S = circle passing through $(0,0)$
   $S \subseteq \mathbb{R}^2$

   conic(S) = open half space containing S
   $\cup \{0,0\}$



2. K is a convex cone iff $K = \text{conic}(K)$
3. Similarly conv(S) is not always closed

1) Recall that if S = linear vector space $\subseteq V$ & B is its basis, $S = \text{linear span}(B) = \{v \in V \mid \langle v, b\rangle_V = 0 \ \forall\, b \in B^\perp\}$

where $B^\perp$ is basis for $S^\perp$

Assuming V is an inner product space

Eg: if $V = \mathbb{R}^n$ & $\langle a,b\rangle = a^T b$

$\{v \in \mathbb{R}^n \mid \langle v, b\rangle_V = 0 \ \forall\, b \in B^\perp\}$

(can be written as) $\equiv \{v \in \mathbb{R}^n \mid Pv = 0\}$

s.t. $\text{rank}(P) + \dim(S) = n$

2) Recall that if A = affine set $\subseteq V$ & B is its basis, $A = \text{affinespan}(B) = \{v \in V \mid \langle v, b\rangle_V = c_b \ \forall\, b \in B^\perp\}$

where $B^\perp$ is basis for $S^\perp$ where

$A = a + S$

eg: If $V = \mathbb{R}^n$ & $\langle a,b\rangle = a^T b$

$\{v \in \mathbb{R}^n \mid \langle v, b\rangle_V = c \ \forall\, b \in B^\perp\}$

(can be written as) $\equiv \{v \in \mathbb{R}^n \mid Pv = c\}$

s.t. $\text{rank}(P) + \dim(A) = n$

Q: What about dual representations of conic sets?

Summary of dual descriptions

If V is a vector space with inner product $\langle\, , \rangle$ then

ⓐ $L \subseteq V$ is a (linear) subspace with finite dimension

Primal description: $L = \text{lin\_span}(\text{basis}(L))$

Dual description: ① Describe in terms of dual of its dual L^* to which it is isomorphic (see page 10 of 2015-5.pdf)

② Describe in terms of basis of its orthogonal complement $L^\perp$

$$L^\perp = \{ v \in V \mid \langle v, u \rangle = 0 \ \forall u \in L \}$$

$$L = \{ u \in V \mid \langle u, v \rangle = 0 \ \forall v \in \text{basis}(L^\perp) \}$$

$\boxed{\text{Claim: } L^\perp = \text{dual\_cone}(L)}$

Note: Linear subspace is a cone

(b) $A \subseteq V$ is affine set with finite basis (of the subspace shifted to give you A)

Let $A = L + b$ (shifted linear subspace, $b \in V$)

Primal description: $A = L + b = \mathrm{aff}(\text{affine\_basis}(A))$

Dual description: $A = \{ v \in V \mid \underbrace{\langle v, u \rangle = \alpha_u \ \forall u \in \text{basis}(L^\perp)}\}$

Recall $Ax = b$ description of affine sets ... $A = \begin{bmatrix} u_1 \in \text{basis}(L^\perp) \\ u_2 \in \text{basis}(L^\perp) \\ \vdots \\ u_k \in \text{basis}(L^\perp) \end{bmatrix}$

$b = \begin{bmatrix} \alpha_{u_1} \\ \alpha_{u_2} \\ \vdots \\ \alpha_{u_k} \end{bmatrix}$

(c) $C \subseteq V$ is a closed convex cone

Primal description: $C = \text{conic}(\text{conic\_span\_set}(C))$

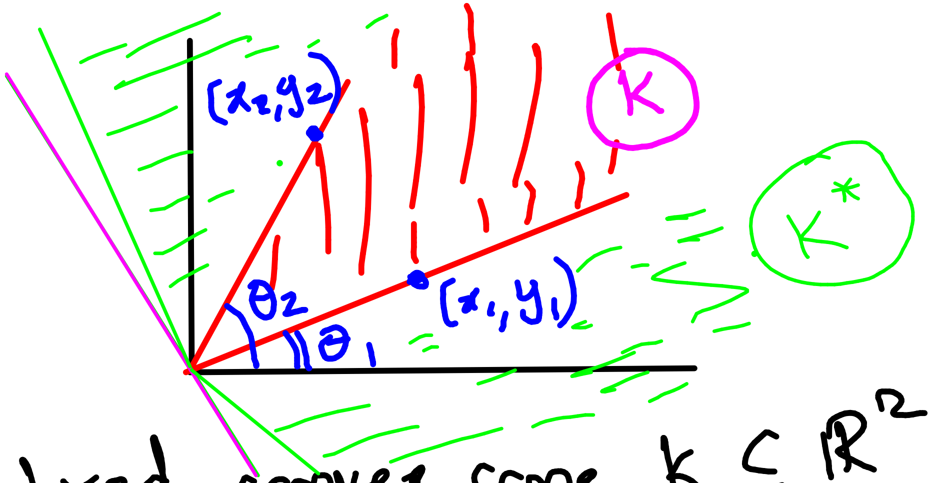
Dual description:- $C = \{v \in V \mid \langle v, u \rangle \geq 0 \ \forall\, u \in C^*\}$

OR

$C = \{v \in V \mid \langle v, u \rangle \geq 0 \ \forall\, u \in \text{conic\_span\_set}(C^*)\}$

Proof: By defn of C^*, $y \neq 0$ is the normal of a halfspace containing C iff $y \in C^*$

Eg:



consider closed convex cone $K \subseteq \mathbb{R}^2$

Primal descriptions:

① $K = \text{conic}\left((x_1,y_1),(x_2,y_2)\right)$ → cartesian coordinates

② $K = \left\{(r\cos\theta, r\sin\theta) \mid r \geq 0, \theta_1 \leq \theta \leq \theta_2\right\}$ → Polar coordinates

K is pointed as long as $\theta_2 - \theta_1 < 180^\circ$

Dual description:

$$K^* = \left\{y \in \mathbb{R}^2 \mid x^T y \geq 0 \ \forall\ x \in K\right\}$$

= { intersection of all half spaces passing through origin and containing K } = { intersection of extreme half spaces defined by extreme rays }

$$= \left\{y \mid y_1\cos\theta_1 + y_2\sin\theta_1 \geq 0 \text{ AND } y_1\cos\theta_2 + y_2\sin\theta_2 \geq 0\right\}$$

In fact, K is also "proper": What will $x \preceq_K y$ mean?

(a) $C \subseteq V$ is a closed convex set

Primal description:

$$C = \text{conv}(\text{convex\_spanning\_set}(C))$$

Dual description:

$$C = \{v \in V \mid \langle v, u\rangle \leq 1 \ \forall\, u \in \text{Polar}(C)\}$$

$$\text{Polar}(C) = \{u \mid \langle u, v\rangle \leq 1 \ \forall\, v \in C\}$$

Claim: C is closed & convex containing origin

$$\Updownarrow$$

$$C = \text{polar}(\text{polar}(C))$$

By defn: $y \in \text{Polar}(C), x \in C \Rightarrow \langle y, x\rangle \leq 1$

$\Rightarrow C \subseteq \text{Polar}(\text{Polar}(C))$

Now let $\exists\ \bar{x} \in \text{Polar}(\text{Polar}(C)) \setminus C$

we need the separation thm which states

Since C is non-empty, convex & closed & $\bar{x} \notin C$ $\bar{x}$ can be "strongly separated" from C. That is, $\exists$ b s.t

$$\langle b, \bar{x} \rangle \geq \sup_{x \in C} \langle b, x \rangle$$

Now $0 \in C \Rightarrow$ the LHS is positive

Using $a = \lambda b$ for some $\lambda > 0$, we can get $\langle a, \bar{x} \rangle > 1 \geq \sup_{x \in C} \langle a, x \rangle$

This is a contradiction since $1 \geq \sup_{x \in C} \langle a, x \rangle$ implies that $a \in \text{Polar}(C)$
But $\langle a, \bar{x} \rangle > 1$ contradicts that $\bar{x} \in \text{Polar}(\text{Polar}(C))$

Helly's Theorem

Let C be a finite family of convex sets in $\mathbf{R}^n$ such that, for $k \leq n+1$, any k (set) members of C have a nonempty intersection. Then the intersection of all (set) members of C is nonempty.

↳ Intersection of any # of sets upto the dimension (n) of the space is non-empty

⇒ Intersection of all sets is non-empty

Affine function

suppose $f : \mathbf{R}^n \to \mathbf{R}^m$ is affine ($f(x) = Ax + b$ with $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$)

- the image of a convex set under f is convex

$$S \subseteq \mathbf{R}^n \text{ convex} \quad \Longrightarrow \quad f(S) = \{f(x) \mid x \in S\} \text{ convex}$$

- the inverse image $f^{-1}(C)$ of a convex set under f is convex

$$C \subseteq \mathbf{R}^m \text{ convex} \quad \Longrightarrow \quad f^{-1}(C) = \{x \in \mathbf{R}^n \mid f(x) \in C\} \text{ convex}$$

examples

- scaling, translation, projection
- solution set of linear matrix inequality $\{x \mid x_1A_1 + \cdots + x_mA_m \preceq B\}$ (with $A_i, B \in \mathbf{S}^p$)
- hyperbolic cone $\{x \mid x^TPx \le (c^Tx)^2,\ c^Tx \ge 0\}$ (with $P \in \mathbf{S}^n_+$)

Perspective and linear-fractional function

perspective function $P : \mathbf{R}^{n+1} \to \mathbf{R}^n$:

$$P(x,t) = x/t, \qquad \mathbf{dom}\, P = \{(x,t) \mid t > 0\}$$

images and inverse images of convex sets under perspective are convex

linear-fractional function $f : \mathbf{R}^n \to \mathbf{R}^m$:

$$f(x) = \frac{Ax+b}{c^Tx+d}, \qquad \mathbf{dom}\, f = \{x \mid c^Tx + d > 0\}$$

images and inverse images of convex sets under linear-fractional functions are convex

$A = \begin{bmatrix} 1 & & & 0 \\ & 2 & & \\ & & 3 & \\ 0 & & & 4 \end{bmatrix}$ → Rescales

$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ → Permutes x_1 & x_2

Affine function

suppose $f : \mathbf{R}^n \to \mathbf{R}^m$ is affine ($f(x) = Ax + b$ with $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$) → translation

Permutes, rotates & rescales

Matrix of real scalars

- the image of a convex set under f is convex

If dom is convex range will be: $S \subseteq \mathbf{R}^n$ convex $\implies$ $f(S) = \{f(x) \mid x \in S\}$ convex

- the inverse image $f^{-1}(C)$ of a convex set under f is convex

If range is convex domain will be: $C \subseteq \mathbf{R}^m$ convex $\implies$ $f^{-1}(C) = \{x \in \mathbf{R}^n \mid f(x) \in C\}$ convex

examples

$x^T\tilde{A} \le \tilde{b}$: An affine set (every affine set is cvx)

- scaling, translation, projection
- solution set of linear matrix inequality $\{x \mid x_1A_1 + \cdots + x_mA_m \preceq B\}$ (with $A_i, B \in \mathbf{S}^p$) — symmetric pxp matrices
- hyperbolic cone $\{x \mid x^TPx \le (c^Tx)^2,\ c^Tx \ge 0\}$ (with $P \in \mathbf{S}^n_+$)

[H/W]

inverse image

symmetric psd matrices

$x \in \mathbb{R}^n$

$A = \begin{bmatrix} P^{1/2} \\ c^T \end{bmatrix}$ $b = 0$ → Affine transform

$\begin{bmatrix} y \\ t \end{bmatrix} = \begin{bmatrix} P^{1/2}x \\ c^Tx \end{bmatrix}$

$\{(y,t) \mid y^Ty \le t^2\}$: a second order convex cone in the image of the affine transform

Perspective and linear-fractional function

perspective function $P : \mathbf{R}^{n+1} \to \mathbf{R}^n$:

$$P(x, t) = x/t, \qquad \mathbf{dom}\, P = \{(x, t) \mid t > 0\}$$

images and inverse images of convex sets under perspective are convex

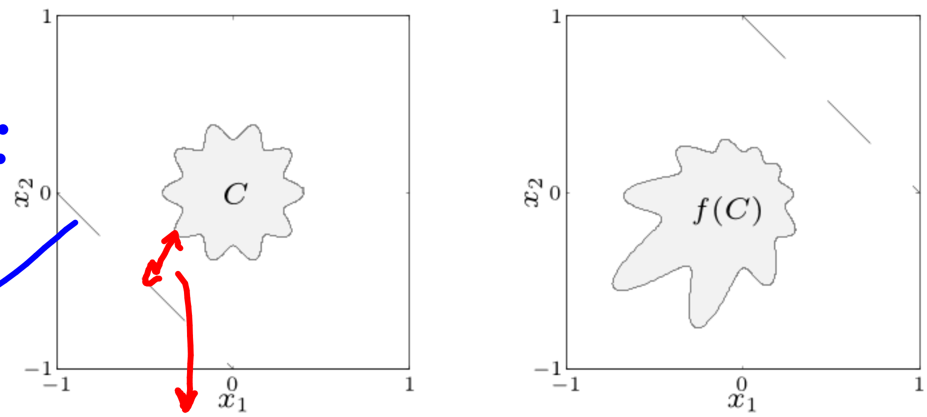
linear-fractional function $f : \mathbf{R}^n \to \mathbf{R}^m$:

$$f(x) = \frac{Ax + b}{c^Tx + d}, \qquad \mathbf{dom}\, f = \{x \mid c^Tx + d > 0\}$$

images and inverse images of convex sets under linear-fractional functions are convex

example of a linear-fractional function

$$f(x) = \frac{1}{x_1 + x_2 + 1} x$$



(In general, $c^T x + d = 0$:

$x_1 + x_2 + 1 = 0$

rescaling

based on inverse of distance

END: SUPPLEMENTARY NOTES FOR CONVEX SETS

Q: When is I convex?

Ans: $I = \{(s,z) \mid s \in \mathbb{R}^m, z \in \mathbb{R}, \exists x \in D \text{ s.t } g_i(x) \le s_i \; i=1..m,$
$f(x) \le z\}$

I is projection/restriction of the epigraph of the vector valued fn $\bar{f}(x) = \begin{bmatrix} f(x) \\ g_1(x) \\ \vdots \\ g_m(x) \end{bmatrix}$

$$epi(\bar{f}) = \{(x,s,z) \mid x \in D, f(x) \le z, g_i(x) \le s_i \; i=1..m\}$$

{in general, w.r.t a generalized inequality $\preceq_K$,
$epi_K(\bar{f}) = \{(x,t) \mid x \in D, \bar{f}(x) \preceq_K t\}$ }

Based on midsem Q2(a)
OR closure under affine transform
.... More properties on following slides

$epi(\bar{f})$ is convex
$\Downarrow$
I is convex

Q: when is $epi(\bar{f})$ convex?

Ans: $epi(\bar{f})$ is convex iff $(f, g_1 \dots g_m) = \bar{f}$ is convex ie each of $f, g_1 \dots g_m$ are convex

3. Convex functions

- basic properties and examples
- operations that preserve convexity
- the conjugate function
- quasiconvex functions
- log-concave and log-convex functions
- convexity with respect to generalized inequalities

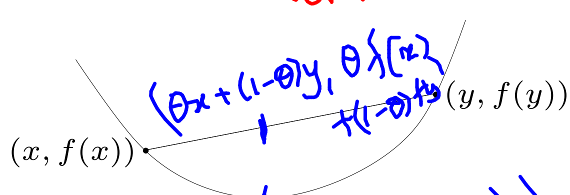
Definition

$f : \mathbf{R}^n \to \mathbf{R}$ is convex if $\mathbf{dom}\, f$ is a convex set and

→ so that $\theta x + (1-\theta) y \in \mathbf{dom}\, f$ $\forall\, x, y \in \mathbf{dom}\, f$

$$\underbrace{f(\theta x + (1-\theta)y)}_{\text{fn at convex combination}} \leq \underbrace{\theta f(x) + (1-\theta) f(y)}_{\text{convex combination of fns at } x \,\&\, y}$$

for all $x, y \in \mathbf{dom}\, f$, $0 \leq \theta \leq 1$



- f is concave if $-f$ is convex
- f is strictly convex if $\mathbf{dom}\, f$ is convex and

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta) f(y)$$

for $x, y \in \mathbf{dom}\, f$, $x \neq y$, $0 < \theta < 1$

What abt vector of functions ($x \in D$ where D is convex)
($f, g_1 \cdots g_m : D \to \mathbb{R}$)

$$\bar{f}(x) = \begin{bmatrix} f(x) \\ g_1(x) \\ \vdots \\ g_m(x) \end{bmatrix}$$

one notion is that each f, g_i are convex

ie $\bar{f}(\theta x_1 + (1-\theta)x_2) \leq \theta \bar{f}(x_1) + (1-\theta)\bar{f}(x_2)$

componentwise ineq.
ie generalised ineq with $K = \mathbb{R}^n_+$

(if $K = \mathbb{R}^n_{++}$ you get strict convexity)

Generalised notion of K-Convexity of $\bar{f}$

$$\bar{f}(\theta x_1 + (1-\theta)x_2) \preceq_K \theta \bar{f}(x_1) + (1-\theta)\bar{f}(x_2)$$

K: proper cone

Epigraph of K-convex fn $\bar{f} = \{(z, x) \mid \bar{f}(x) \preceq_K z\}$

Convexity with respect to generalized inequalities

$f : \mathbf{R}^n \to \mathbf{R}^m$ is K-convex if $\mathbf{dom}\, f$ is convex and

$$f(\theta x + (1-\theta)y) \preceq_K \theta f(x) + (1-\theta)f(y)$$

for x, $y \in \mathbf{dom}\, f$, $0 \le \theta \le 1$

example $f : \mathbf{S}^m \to \mathbf{S}^m$, $f(X) = X^2$ is $\mathbf{S}^m_+$-convex

proof: for fixed $z \in \mathbf{R}^m$, $z^T X^2 z = \|Xz\|_2^2$ is convex in X, *i.e.*,

$$z^T(\theta X + (1-\theta)Y)^2 z \le \theta z^T X^2 z + (1-\theta) z^T Y^2 z$$

for $X, Y \in \mathbf{S}^m$, $0 \le \theta \le 1$

therefore $(\theta X + (1-\theta)Y)^2 \preceq \theta X^2 + (1-\theta)Y^2$

Examples on R

convex:

- affine: $ax + b$ on **R**, for any $a, b \in$ **R**
- exponential: e^{ax}, for any $a \in$ **R**
- powers: x^α on $\mathbf{R}_{++}$, for $\alpha \geq 1$ or $\alpha \leq 0$
- powers of absolute value: $|x|^p$ on **R**, for $p \geq 1$
- negative entropy: $x \log x$ on $\mathbf{R}_{++}$

concave:

- affine: $ax + b$ on **R**, for any $a, b \in$ **R**
- powers: x^α on $\mathbf{R}_{++}$, for $0 \leq \alpha \leq 1$
- logarithm: $\log x$ on $\mathbf{R}_{++}$

Examples on $\mathbf{R}^n$ and $\mathbf{R}^{m \times n}$

affine functions are convex and concave; all norms are convex

examples on $\mathbf{R}^n$

- affine function $f(x) = a^T x + b$
- norms: $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ for $p \geq 1$; $\|x\|_\infty = \max_k |x_k|$

examples on $\mathbf{R}^{m \times n}$ ($m \times n$ matrices)

- affine function

$$f(X) = \mathbf{tr}(A^T X) + b = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} X_{ij} + b$$

- spectral (maximum singular value) norm

$$f(X) = \|X\|_2 = \sigma_{\max}(X) = (\lambda_{\max}(X^T X))^{1/2}$$

Epigraph and sublevel set

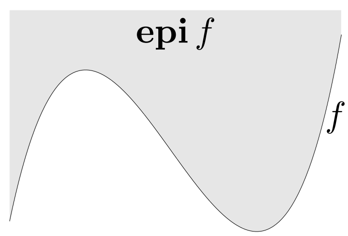
α-**sublevel set** of $f : \mathbf{R}^n \to \mathbf{R}$:

$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \leq \alpha\}$$

sublevel sets of convex functions are convex (converse is false)

epigraph of $f : \mathbf{R}^n \to \mathbf{R}$:

$$\mathbf{epi}\, f = \{(x, t) \in \mathbf{R}^{n+1} \mid x \in \mathbf{dom}\, f,\ f(x) \leq t\}$$



f is convex if and only if $\mathbf{epi}\, f$ is a convex set

More generally: $\bar{f}$ is K-convex iff epi $\bar{f}$ (wrt $\hat{K}$) is a convex set } Q1

Think: When is epi(f) closed?
When is epi$(\bar{f})$ closed? } Q2