Convex Analysis Notes

Walid Krichene Maximilian Balandat Roy Dong

1 Affine sets

- Affine subspace, dimension of an affine subspace
- Orthogonal complement of a subspace L. $L + L^\perp = \mathbb{R}^n$
- Hyperplanes, $\mathcal{H} = \{x : \langle x, b\rangle = \beta\}$. A hyperplane has "two sides".
- An affine set can be represented as $\{x : Bx = b\}$, $B \in \mathbb{R}^{m\times n}$. So it is the intersection of hyperplanes.
- Affine hull of S: smallest affine space containing S. $(b_0, \dots, b_m)$ is affinely independent if $\text{aff}(b_0, \dots, b_m) = b_0 + \text{aff}(0, b_1 - b_0, \dots, b_m - b_0)$ is m-dimensional, i.e. $b_1 - b_0, \dots, b_m - b_0$ are linearly independent.
- $x \in M = \text{aff}(b_0, \dots, b_m)$ can be expressed as $x = \lambda_0 b_0 + \dots + \lambda_m b_m$, with $\sum \lambda_i = 1$. The representation is unique if (b_i) are affinely independent. Then (λ_i) is a barycentric coordinate system.
- A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is affine if

  $$T((1-\lambda)x + \lambda y) = (1-\lambda)T(x) + \lambda T(y)$$

  then T is necessarily of the form $T(x) = Ax + a$.
- If M is affine then $T(M)$ is affine. So $\text{aff}(T(S)) = T(\text{aff}(S))$.

  Theorem (1.6)**.** *Let $\{b_0, \dots, b_m\}$, $\{b_0', \dots, b_m'\}$ be affinely independent in $\mathbb{R}^n$. Then $\exists$ affine T such that $T(b_i) = b_i'$ for all i. If $m = n$, T is unique.*

- Graphs
- The graph of a linear $A : \mathbb{R}^n \to \mathbb{R}^m$ is $\{(x, y) \in \mathbb{R}^{m+n} : Ax = y\}$. Its orthogonal is

  $$L^\perp = \{(x^*, y^*) \in \mathbb{R}^{n+m} : x^* = -A^* y^*\}$$

  proof: $(x^*, y^*) \in L^\perp$ iff it is orthogonal to all $(x, y) \in L$, i.e. $\forall x \in \mathbb{R}^n$,

  $$0 = \langle x, x^*\rangle + \langle Ax, y^*\rangle = \langle x, x^*\rangle + \langle x, A^* y^*\rangle = \langle x, x^* + A^* y^*\rangle$$

  i.e. $x^* + A^* y^* = 0$.
- Tucker representation: An affine space M in $\mathbb{R}^N$, of dimension n can be represented as the graph of an affine transformation T.

  $$M = \{x : Bx = \beta\}$$

  Then $\dim N(B) = n$ and $\text{rank}\, B = N - n = m$. Thus there is a permutation of indices such that

  $$\xi_{n+i} = \alpha_{i1}\xi_1 + \dots + \alpha_{in}\xi_n + \alpha_i, \ i \in \{1, \dots, n\}$$

2 Convex Sets

- Arbitrary intersection of convex sets is convex.
- Polyhedral convex set: intersection of finitely many half spaces.
- $\lambda_1 x_1 + \cdots + \lambda_m x_m$ is a convex combination if $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$
- C is convex iff it is closed under convex combinations.
- Convex hull of S: intersection of all convex sets containing S. Also the set of all convex combinations of S (Theorem 2.1).
- By Caratheodory's theorem, $\operatorname{conv} S$ is the set of convex combinations involving $n+1$ points.
- Convex hull of finitely many points: polytope.
- If the vertices of the polytope are affinely independent, it is called a simplex.
- Dimension of a convex set C: dimension of its affine hull. It is also the largest dimension of a simplex contained in C.

Convex cones

- K is a cone if $\lambda x \in K$ for all $x \in K$, $\lambda > 0$. It is the union of half lines.
- A convex cone is not necessarily pointed. E.g. subspaces, half-spaces.
- Arbitrary intersection of convex cones is a convex cone.

  Theorem (2.6). *C is a convex cone iff it is closed under addition and multiplication by a positive scalar, iff it is closed under positive linear combination.*

  $$\lambda K \subset K \qquad\qquad K + K \subset K$$

- If C is convex, $K = \{\lambda x,\ \lambda > 0, x \in C\}$ is the smallest convex cone which includes C.
- If S is an arbitrary set, the smallest convex cone which contains S is the set of positive linear combinations of S.
- A convex set C in $\mathbb{R}^n$ can be viewed as the intersection of a cone in $\mathbb{R}^{n+1}$ with a hyperplane. The cone can be defined as $K = \{(\lambda x, \lambda) : x \in C, \lambda > 0\}$.

  Theorem (2.7). *If K is a convex cone that contains 0, then $\operatorname{aff} K = K - K$, and the largest affine space contained in K is $K \cap -K$.*

Normal cone

- x^* is normal to C at a if $\langle x^*, x - a\rangle \leq 0$ for all $x \in C$.
- The set of all such vectors is the normal cone.

3 The Algebra of Convex Sets

- C_1, C_2 convex $\Rightarrow C_1 + C_2$ convex. Helps to think of it as the union $x + C_2$, for $x \in C_1$.

  Theorem (3.2)**.** *C is convex iff for all $\lambda_1, \lambda_2 \geq 0$, $(\lambda_1 + \lambda_2)C = \lambda_1 C + \lambda_2 C$.*

  follows from the fact that $\lambda_1 x + \lambda_2 y = (\lambda_1 + \lambda_2)\left(\frac{\lambda_1}{\lambda_1+\lambda_2}x + \frac{\lambda_2}{\lambda_1+\lambda_2}y\right)$ and convexity.

- The collection of convex subsets of $\mathbb{R}^n$ is a complete lattice under inclusion. The smallest convex set contained in $\{C_i\}_{i\in I}$ is their intersection. Also

  Theorem (3.3)**.**
  $$\text{conv}\{C_i\}_{i\in I} = \cup_{\textit{finite convex combinations}} \sum_{i\in I} \lambda_i C_i$$

- Given a linear transformation $A : \mathbb{R}^n \to \mathbb{R}^m$, define
  $$AC = \{Ax, x \in C\} \qquad\qquad A^{-1}D = \{x : Ax \in D\}$$

  Theorem (3.4)**.** *C is convex $\Rightarrow AC$ is convex. D is convex $\Rightarrow A^{-1}D$ is convex.*

  Example: $\{x : Ax \geq a\} = A^{-1}(a + \mathbb{R}^n_+)$ is convex.

  Theorem (3.5)**.** *If C and D are convex, then the direct sum*
  $$C \oplus D = \{(y, z),\ y \in C,\ z \in D\}$$
  is convex.

  Theorem (3.6)**.** *Let C_1, C_2 be convex subsets of $\mathbb{R}^{m+p}$, and let*
  $$C = \{(y, z_1 + z_2) \in \mathbb{R}^{m+p} : (y, z_1) \in C_1,\ (y, z_2) \in C_2\}$$
  Then C is convex.

  The operation is called partial addition. $m = 0$ corresponds to ordinary addition, and $p = 0$ corresponds to intersection.

- There is a bijection between convex sets of $\mathbb{R}^n$ and convex cones in $\mathbb{R}^{n+1}$ such that $K \cap \{(1, x)\lambda \leq 0\} = \{0\}$.
  $$\begin{aligned}\phi : \mathcal{K} &\to \mathcal{C}\\ K &\mapsto \{x : (x, 1) \in K\}\end{aligned}$$
  $$\begin{aligned}\psi : \mathcal{C} &\to \mathcal{K}\\ C &\mapsto \{\lambda(x, 1), x \in C\} = \text{convex cone generated by } \{(1, x), x \in C\}\end{aligned}$$
  An operation that preserves $\mathcal{K}$ can be translated into an operation that preserves $\mathcal{C}$.

- Four "natural" partial additions on $\mathbb{R}^{n+1}$: addition, intersection, partial addition of x denoted $+^1$, and partial addition of λ denoted $+^2$.
  - $K_1 + K_2 = \{\lambda_1(x_1, 1) + \lambda_2(x_2, 1),\ x_1 \in C_1, x_2 \in C_2\}$. Then $\phi(K_1 + K_2)$ is the union of convex combinations of C_1, C_2, i.e. $\phi(K_1 + K_2) = \text{conv}(C_1 \cup C_2)$.
  - $K_1 +^1 K_2 = \{(\lambda(x_1 + x_2), \lambda),\ x_1 \in C_1,\ x_2 \in C_2,\ \lambda \geq 0\}$. Then $\psi(K_1 +^1 K_2) = C_1 + C_2$.
  - $\phi(K_1 \cap K_2) = C_1 \cap C_2$

– $K_1 +^2 K_2 = \{(x, \lambda_1 + \lambda_2) : \frac{x}{\lambda_1} \in C_1,\ \frac{x}{\lambda_2} \in C_2\}$. Then $\psi(K_1 +^2 K_2)$ is the union of all $\lambda_1 C_1 \cap \lambda_2 C_2$, $\lambda_1 + \lambda_2 = 1$. This operation is called inverse sum, and denoted $C_1 \# C_2$.

  Theorem (3.7)**.** *$C_1 \# C_2$ is convex.*

Umbra and penumbra

- $\{(1-\lambda)x + \lambda y,\ \lambda \geq 1\}$ can be thought of as the shadow of y cast by a light source at x.
- Umbra of C w.r.t. S

$$\cap_{x \in S} \cup_{\lambda \geq 1} (1-\lambda)x + \lambda C$$

  it is convex if C is convex

- Penumbra of C w.r.t. S

$$\cup_{x \in S} \cup_{\lambda \geq 1} (1-\lambda)x + \lambda C$$

  it is convex if both C and S are convex.

4 Convex functions

$f : C \to \mathbb{R} \cup \{\pm\infty\}$.

- f is a convex function iff $\operatorname{epi} f$ is a convex subset. $\operatorname{epi} f = \{(x, \mu), x \in S,\ \mu \geq f(x)\}$.
- f is convex iff its restrictions to line segments are convex. The restriction is defined for $x, y \in C$

$$\lambda \in [0,1] \mapsto f(x + \lambda(y - x))$$

Theorem (4.1)**.** *f is convex iff $f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y)$ for all $\lambda \in (0,1)$.*

Theorem (4.3, Jensen's inequality)**.** *f is convex iff $f(\sum_i \lambda_i x_i) \leq \sum_i \lambda_i f(x_i)$ for all $\lambda_i \geq 0$ with* $\sum_i \lambda_i = 1$

A cute consequence: by convexity of $-\log$, $-\log(\lambda_1 x_1 + \cdots + \lambda_m x_m) \leq -\lambda_1 \log x_1 - \cdots - \lambda_m \log x_m$, so for $\lambda_i = 1/m$, $(x_1 + \cdots + x_m)/m \geq (x_1 \ldots x_m)^{1/m}$.

- $\operatorname{dom} f$, the effective domain of f is $\{x : f(x) < \infty\}$. It is the projection of $\operatorname{epi} f$ on $\mathbb{R}^n$.
- f is proper if $\operatorname{epi} f \neq \emptyset$ and contains no vertical lines, i.e. $f(x) > -\infty$ everywhere.

Theorem (4.4)**.** *Let $f : I \to \mathbb{R}$ be C^2. Then f is convex iff $f'' \geq 0$ on I.*

- the Hessian of f at x is denoted Q_x.

Theorem (4.5)**.** *Let $f : C \to \mathbb{R}$ be C^2, C a convex subset of $\mathbb{R}^n$. Then f is convex iff $Q_x \succeq 0$ for all $x \in C$.*

Correspondences between convex functions and sets:

- Indicator functions

$$\delta(x|C) = \begin{cases} 0 & \text{fi } x \in C \\ +\infty & \text{if } x \notin C \end{cases}$$

- Support function

$$\delta^*(x|C) = \sup_{y \in C} \langle x, y \rangle$$

- Gauge, $C \neq \emptyset$

$$\gamma(x|C) = \inf\{\lambda \geq 0 : x \in \lambda C\}$$

- Euclidean distance to C

$$d(x, C) = \inf_{y \in C} |x - y|$$

- Level set of f

$$\{x : f(x) \leq \alpha\}$$

Theorem (4.6)**.** *The level sets of a convex f are convex.*

These are horizontal cross sections of $\operatorname{epi} f$ (projection of $\operatorname{epi} f \cap \{(x, \alpha), x \in \mathbb{R}^n\}$).

- Convexity can be obtained by a change of variable: Let $g(x) = \beta \xi_1^{\alpha_1} \ldots \xi_n^{\alpha_n}$. With the change of variable $\zeta_i = \log \xi_i$, g becomes

$$h(z) = \beta e^{\langle \alpha, z \rangle}$$

which is convex.

- f is positively homogeneous of degree 1 if for all x,

$$f(\lambda x) = \lambda f(x),\ \lambda > 0$$

equivalent to epi f is a cone. Example, $|x|$.

Theorem (4.7)**.** *A positively homogeneous f is convex iff $f(x+y) \leq f(x) + f(y)$ for all x, y.*

Corollary (4.7.1)**.** *If f is prop. conv. pos. homog., then $f(\sum_i \lambda_i x_i) \leq \sum_i \lambda_i f(x_i)$ for all $\lambda_i \geq 0$*

Corollary (4.7.2)**.** *If f is prop. conv. pos. homog., then $f(-x) \geq -f(x)$.*

Theorem (4.7)**.** *A prop. conv. pos. homog. f is linear on a subspace L iff $f(-x) = -f(x)$ for all $x \in L$, iff $f(-b_i) = -f(b_i)$ for a basis of L.*

5 Functional operations

- Composition

  Theorem (5.1). *$f : \mathbb{R}^n \to (-\infty, +\infty]$ convex, $\phi : \mathbb{R} \to (-\infty, +\infty]$ convex, increasing, then $\phi \circ f$ is convex.*

  Example, f convex positive, $p \geq 1$, then f^p is convex: take $\phi(x) = x^p 1_{x\geq 0}$.

- Sum

  Theorem (5.2). *If f_1, f_2 are proper convex then so is $f_1 + f_2$.*

  Example, $f + \delta(\cdot|C)$. This is equivalent to restricting the effective domain of f.

- Lower boundary (inverse operation of epi)

  Theorem (5.3). *If $F \subset \mathbb{R}^{n+1}$ is convex, then $f(x) = \inf\{\mu : (x, \mu) \in F\}$ is convex.*

- Addition of epigraphs, a.k.a. infimal convolution

  Theorem (5.4). *Let f_i be proper convex and let*

  $$f(x) = (\Box_i f_i)(x) = \inf\{\sum_i f_i(x_i) : \sum_i x_i = x\} = \inf\{\mu : (x, \mu) \in \sum_i \operatorname{epi} f_i\}$$

  Then f is convex. It is not necessarily proper (inf may be $-\infty$).

  It is called intimal convolution because

  $$f \Box g = \inf_y f(x - y) + g(y)$$

  It is commutative, associative on the set of convex functions. Its identity is $\delta(\cdot|0)$.
  Examples

  - $f \Box \delta(\cdot|a)(x) = \inf_y f(x - y)\delta(y|a) = f(x - a)$.
  - $|\cdot|\Box\delta(\cdot|C) = \inf_{y\in C} |x - y| = d(\cdot|C)$.

- Right scalar multiplication (multiplication of epigraphs), for $\lambda > 0$, define

  $$(f\lambda)(x) = \lambda f(x/\lambda)$$

  (and define $f0 = \delta(\cdot|0)$).
  This corresponds to Thm 5.3 with $F = \lambda = \operatorname{epi} f$.
  A function is pos. hom. if $f\lambda = f$ for all $\lambda > 0$.

- Pos. hom. convex function generated by f: obtained by applying Thm 5.3 with $F = \operatorname{cone} \operatorname{epi} f$.

  $$f(x) = \inf_{\lambda\geq 0} (h\lambda)(x)$$

  Can also define the pos. hom. conv. func. generated by $h(\lambda, x) = \begin{cases} f(x) & \lambda = 1 \\ +\infty & \lambda \neq 1 \end{cases}$, it is given by $g(\lambda, x) = \begin{cases} (f\lambda)(x) & \lambda \geq 0 \\ +\infty & \lambda < 0 \end{cases}$.
  The gauge of C is the pos. hom. conv. func. generated by $\delta(x|C) + 1$.

  $$\gamma(x|C) = \inf\{\lambda \geq 0 : x \in \lambda C\}$$

- Supremum (intersection of epigraphs)

  Theorem (5.5)**.** *The pointwise supremum of an arbitrary collection convex functions is convex.*

  Examples

  - support function $\delta^*(x|C) = \sup_{y\in C} \langle x, y\rangle$.
  - $x \mapsto \max_i \xi_i$ (support function of the simplex)
  - Tchebycheff norm $k(x) = \max_i |\xi_i|$ (support function of the unit 1-ball, $B = \{x : \|x\|_1 \leq 1\}$, also the gauge of the cube $\{x : |x_i| \leq 1\}$)
  - Note: any non-negative support function is the gauge of a closed convex set.

- Convex hull of (non convex) g: greatest convex function majorized by g. Equivalent to $\text{epi}^{-1}(\text{conv epi } f)$ by Thm 5.3

$$f(x) = (\text{conv } g)(x) = \inf_{\sum_i \lambda_i x_i = x} \sum_i \lambda_i g(x_i)$$

  Convex hull of the pointwise infimum (greatest convex f majorized by all f_i)

$$\begin{aligned}
\text{conv}\{f_i\}_{i\in I} &= \text{conv}(\inf_{i\in I} f_i) \\
&= \text{epi}^{-1}(\text{conv} \cup_i \text{epi } f_i) \\
&= \inf_{\sum_i \lambda_i x_i = x} \sum_i \lambda_i f_i(x_i) \qquad\qquad \text{Theorem 5.6} \\
&= \inf_{\sum \lambda_i = 1, \lambda \geq 0} (\Box_i f_i \lambda_i)(x)
\end{aligned}$$

  Example: $f_i(x) = \delta(x|a_i) + \alpha_i$. Then the greatest convex function with $f(a_i) \leq \alpha_i$ is $f(x) = \inf_{\sum \lambda_i a_i = x} \sum \lambda_i \alpha_i$.
  In the ordered set of convex functions, with $f \leq g$ pointwise: for a family of convex functions

  - Least upper bound $\sup_{i\in I} f_i$
  - Greatest lower bound $\text{conv}\{f_i\}_{i\in I}$

- Linear transformation.

  Theorem. *Let g, h be convex and $A : \mathbb{R}^n \to \mathbb{R}^m$ linear. Then*

$$\begin{aligned}
(gA)(x) &= g(Ax) \\
(Ah)(y) &= \inf_{Ax=y} h(x)
\end{aligned}$$

  are convex.

- Partial addition of epigraphs: I am skipping this part.

Function	Epigraph
$f_1 \Box f_2$	$F_1 + F_2$
$f\lambda$	λF
$\inf_{\lambda\geq 0} f\lambda(x)$	$\text{cone } F$
$\sup_{i\in I} f_i$	$\cap_{i\in I} F_i$
$\text{conv } f$	$\text{conv } F$
$\text{conv}\{f_i\}_{i\in I}$	$\text{conv} \cup_{i\in I} F_i$

Table 1: The correspondence is $f(x) = \text{epi}^{-1}(F) = \inf_{(x,\mu)\in F} \mu$, and $F = \text{epi } f = \{(x, \mu) : f(x) \leq \mu\}$.

6 Relative Interiors of Convex Sets

- Euclidean distance metric $d(x, y) = |x - y| = \langle x - y, x - y\rangle^{1/2}$ is convex (composition of $|\cdot|$ with linear transformation). Topology induced by Euclidean metric $\equiv$ Topology induced by other norms (equiv. of norms in $\mathbb{R}^n$).
- Euclidean unit ball: $B = \{x \mid x \leq 1\}$
- For any set $C \in \mathbb{R}^n$,

$$\operatorname{cl} C = \bigcap \{C + \epsilon B \mid \epsilon > 0\} \tag{1}$$

$$\operatorname{int} C = \{x \mid \exists \epsilon > 0 : x + \epsilon B \subset C\} \tag{2}$$

- *Relative Interior:* Interior that results when regarding C as a subset of its affine hull aff C:

$$\operatorname{ri} C = \{x \in \operatorname{aff} C \mid \exists \epsilon > 0 : (x + \epsilon B) \cap (\operatorname{aff} C) \subset C\} \tag{3}$$

- $\operatorname{ri} C \subset C \subset \operatorname{cl} C$. *Relative boundary* $(\operatorname{cl} C) \setminus (\operatorname{ri} C)$. Say C is *relatively open* if $\operatorname{ri} C = C$.
- Any affine set is relatively open and closed at the same time. $\operatorname{cl} C \subset \operatorname{cl}(\operatorname{aff} C) = \operatorname{aff} C$.
- *Reminder:* Easy to simplify proofs by using that relative interiors are preserved under one-to-one affine transformations of $\mathbb{R}^n$ onto itself.

Theorem (6.1). *$C \subset \mathbb{R}^n$ convex, $x \in \operatorname{ri} C$ and $y \in \operatorname{cl} C$. Then $(1 - \lambda)x + \lambda y \in \operatorname{ri} C$ for $0 \leq \lambda < 1$.*

Theorem (6.2). *$C \subset \mathbb{R}^n$ convex. Then $\operatorname{cl} C$ and $\operatorname{ri} C$ are convex sets in $\mathbb{R}^n$ with the same affine hull (hence the same dimension) as C (in particular, $\operatorname{ri} C \neq \emptyset$ if $C \neq \emptyset$).*

- For any $C \subset \mathbb{R}^n$, not necessarily convex, $\operatorname{cl}(\operatorname{cl} C) = \operatorname{cl} C$ and $\operatorname{ri}(\operatorname{ri} C) = \operatorname{ri} C$.

Theorem (6.3). *$C \subset \mathbb{R}^n$ convex. Then $\operatorname{cl}(\operatorname{ri} C) = \operatorname{cl} C$ and $\operatorname{ri}(\operatorname{cl} C) = \operatorname{ri} C$.*

- Some Corollaries:
  - $C_1, C_2 \subset \mathbb{R}^n$ convex. Then $\operatorname{cl} C_1 = \operatorname{cl} C_2 \Leftrightarrow \operatorname{ri} C_1 = \operatorname{ri} C_2$. Equivalent to $\operatorname{ri} C_1 \subset C_2 \subset \operatorname{cl} C_1$.
  - $C \subset \mathbb{R}^n$ convex, then every open set which meets $\operatorname{cl} C$ also meets $\operatorname{ri} C$.
  - If $C_1 \subset (\operatorname{cl} C_2) \setminus (\operatorname{ri} C_2)$ for some $C_2 \subset \mathbb{R}^n$ non-empty and convex, then $\dim C_1 < \dim C_2$.

Theorem (6.4). *$C \subset \mathbb{R}^n$ non-empty, convex. Then $z \in \operatorname{ri} C$ if and only if, for every $x \in C$, $\exists \mu > 1$ such that $(1 - \mu)x + \mu z \in C$. That is, every line segment in C having z as one endpoint can be prolonged beyond z without leaving C.*

- If $C \subset \mathbb{R}^n$ convex, then $z \in \operatorname{int} C \Leftrightarrow \forall y \in \mathbb{R}^n, \exists \epsilon > 0 : z + \epsilon y \in C$.

Theorem (6.5). *$C_i \subset \mathbb{R}^n$ convex for $i \in I$. Suppose the sets $\operatorname{ri} C_i$ have at least one point in common. Then $\operatorname{cl}(\cap_{i \in I} C_i) = \cap_{i \in I}(\operatorname{cl} C_i)$. If I is finite, then also $\operatorname{ri}(\cap_{i \in I} C_i) = \cap_{i \in I}(\operatorname{ri} C_i)$*

- If $C \subset \mathbb{R}^n$ convex, $M \subset \mathbb{R}^n$ affine such that $M \cap \operatorname{ri} C \neq \emptyset$, then $\operatorname{ri}(M \cap C) = M \cap \operatorname{ri} C$ and $\operatorname{cl}(M \cap C) = M \cap \operatorname{cl} C$
- If $C_1, C_2 \subset \mathbb{R}^n$ convex, $C_2 \subset \operatorname{cl} C_1$ but $C_2 \not\subset (\operatorname{cl} C_1) \setminus (\operatorname{ri} C_1)$, then $\operatorname{ri} C_2 \subset \operatorname{ri} C_1$.

Theorem (6.6). *$C \subset \mathbb{R}^n$ convex, $A : \mathbb{R}^n \to \mathbb{R}^m$ linear. Then $\operatorname{ri}(AC) = A(\operatorname{ri} C)$ and $\operatorname{cl}(AC) \supset A(\operatorname{cl} C)$.*

- Note that, more generally, $\operatorname{cl}(TC) \supset T(\operatorname{cl} C)$ for any set $C \subset \mathbb{R}^n$ for any continuous transformation T.

- If $C \subset \mathbb{R}^n$ convex, $\lambda \in \mathbb{R}$, then $\operatorname{ri}(\lambda C) = \lambda(\operatorname{ri} C)$.
- If $C_1, C_2 \subset \mathbb{R}^n$ convex, then $\operatorname{ri}(C_1 + C_2) = \operatorname{ri} C_1 + \operatorname{ri} C_2$ and $\operatorname{cl}(C_1 + C_2) \supset \operatorname{cl} C_1 + \operatorname{cl} C_2$.

Theorem (6.7)**.** *$A : \mathbb{R}^n \to \mathbb{R}^m$ linear, $C \subset \mathbb{R}^m$ convex s.t. $A^{-1}(\operatorname{ri} C) \neq \emptyset$. Then $\operatorname{ri}(A^{-1}C) = A^{-1}(\operatorname{ri} C)$ and $\operatorname{cl}(A^{-1}C) = A^{-1}(\operatorname{cl} C)$.*

Theorem (6.8)**.** *$C \subset \mathbb{R}^{m+p}$ convex. Let $C_y := \{z \in \mathbb{R}^p \,|\, (y,z) \in C\}$ and $D := \{y \,|\, C_y \neq \emptyset\}$. Then $(y,z) \in \operatorname{ri} C \;\Leftrightarrow\; y \in \operatorname{ri} D, z \in \operatorname{ri} C_y$.*

- If $C \subset \mathbb{R}^n$ non-empty convex, K the convex cone in $\mathbb{R}^{n+1}$ generated by $\{(1,x) \,|\, x \in C\}$, then $\operatorname{ri} K = \{(\lambda, x) \mid \lambda > 0, x \in \lambda \operatorname{ri} C\}$. More generally, the relative interior of the convex cone in $\mathbb{R}^n$ generated by a non-empty convex set C consists of the vectors of the form λx with $\lambda > 0$ and $x \in \operatorname{ri} C$.

Theorem (6.9)**.** *$C_1 \ldots, C_m \subset \mathbb{R}^n$ non-empty convex. Let $C_0 := \operatorname{conv}(C_1 \cup \ldots \cup C_m)$. Then $\operatorname{ri} C_0 = \bigcup \big\{\lambda_1 \operatorname{ri} C_1 \ldots + \lambda_m \operatorname{ri} C_m \;\mid\; \lambda_i > 0, \sum_i \lambda_i = 1\big\}$.*

7 Closures of Convex Functions

Definition (Lower Semicontinuity). *$f : S \to \mathbb{R}^*$ with $S \subset \mathbb{R}^n$ is* lower semi-continuous *(l.s.c.) at $x \in S$ if $f(x) \leq \lim_i f(x_i)$ for every sequence (x_i) s.t. $x_i \to x$ and the limit exists in $[-\infty, \infty]$. This can be expressed as $f(x) = \liminf_{y \to x} f(y) = \lim_{\epsilon \searrow 0} \inf\{f(y) \mid |x - y| \leq \epsilon\}$. f is said to be* upper semi-continuous *(u.s.c.) if $f(x) = \limsup_{y \to x} f(y) = \lim_{\epsilon \searrow 0} \sup\{f(y) \mid |x - y| \leq \epsilon\}$.*

- l.s.c. + u.s.c. $\Leftrightarrow$ continuity

Theorem (7.1). *For $f : \mathbb{R}^n \to \mathbb{R}^*$ the following are equivalent:*

1. *f is l.s.c. on $\mathbb{R}^n$*
2. *$\{x \mid f(x) \leq \alpha\}$ is closed for every $\alpha \in \mathbb{R}$*
3. *$\operatorname{epi} f$ is a closed set in $\mathbb{R}^{n+1}$*

- *lower semi-continuous hull* of f: function whose epigraph is the closure of $\operatorname{epi} f$ in $\mathbb{R}^{n+1}$

Definition (Closure of a convex function). *The closure $\operatorname{cl} f$ of a convex function $f : \mathbb{R}^n \to \mathbb{R}^*$ is the lower semicontinuous hull of f provided $f(x) \neq -\infty$ for all x, and $-\infty$ otherwise.*

- A convex function f is closed if $\operatorname{cl} f = f$.
- For a proper convex function, closedness is the same as lower semi-continuity.
- The only closed improper convex functions are $+\infty$ and $-\infty$.
- $\operatorname{epi}(\operatorname{cl} f) = \operatorname{cl}(\operatorname{epi} f)$ for a proper convex function. Thus $(\operatorname{cl} f)(x) = \liminf_{y \to x} f(y)$
- $\operatorname{cl} f \leq f$ and $f_1 \leq f_2 \implies \operatorname{cl} f_1 \leq \operatorname{cl} f_2$

Theorem (7.2). *f improper, convex $\Rightarrow f(x) = -\infty$ for every $x \in \operatorname{ri}(\operatorname{dom} f)$.*

- An improper convex function is necesarily infinite except perhaps at relative boundary points of its effective domain.
- A lower semi-continuous improper convex function can have no finite values.
- If f is improper convex then $\operatorname{cl} f$ is closed improper convex and agrees with f on $\operatorname{ri}(\operatorname{dom} f)$.
- If f is convex with $\operatorname{dom} f$ relatively open, then either $f(x) > -\infty$ for all x or $f(x)$ is infinite for all x

Lemma (7.3). *For any convex f, $\operatorname{ri}(\operatorname{epi} f) = \{(x, \mu) \ : \ x \in \operatorname{ri}(\operatorname{dom} f), f(x) < \mu < \infty\}$*

- If f convex and $f(x) < \alpha$ for some x, then $f(x) < \alpha$ for some $x \in \operatorname{ri}(\operatorname{dom} f)$.
- If f convex, C convex such that $\operatorname{ri} C \subset \operatorname{dom} f$ and $f(x) < \alpha$ for some $x \in \operatorname{cl} C$, then $f(x) < \alpha$ for some $x \in \operatorname{ri} C$.
- If f and g convex, $\operatorname{ri}(\operatorname{dom} f) = \operatorname{ri}(\operatorname{dom} g) =: C$, $f \equiv g$ on C, then $\operatorname{cl} f = \operatorname{cl} g$.

Theorem (7.4). *If $f : \mathbb{R}^n \to \mathbb{R}^*$ is proper convex, then $\operatorname{cl} f$ is proper convex. Moreover, $\operatorname{cl} f$ agrees with f except perhaps at relative boundary points of $\operatorname{dom} f$.*

- If f proper, convex, with $\operatorname{dom} f$ an affine set, then f is closed.

Theorem (7.5). *Let $f : \mathbb{R}^n \to \mathbb{R}^*$ proper convex, $x \in \operatorname{ri}(\operatorname{dom} f)$. Then $(\operatorname{cl} f)(y) = \lim_{\lambda \nearrow 1} f((1 - \lambda)x + \lambda y)$ for every y.*

- If f proper, convex, closed, then $f(y) = \lim_{\lambda \nearrow 1} f((1-\lambda)x + \lambda y)$ for every $x \in \operatorname{dom} f$ and every y.

Theorem (7.6)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^*$ proper convex, $\alpha \in \mathbb{R}$, $\alpha > \inf f$. Then $\{x \mid f(x) \leq \alpha\}$ and $\{x \mid f(x) < \alpha\}$ have the same closure and the same relative interior, namely $\{x \mid (\operatorname{cl} f)(x) \leq \alpha\}$ and $\{x \in \operatorname{ri}(\operatorname{dom} f) \mid f(x) < \alpha\}$, respectively. Furthermore, they have the same dimensions as $\operatorname{dom} f$ (and f).*

- If f proper, convex, closed, with $\operatorname{dom} f$ relatively open, then for $\inf f < \alpha < +\infty$ it holds that $\operatorname{ri}\{x \mid f(x) \leq \alpha\} = \{x \mid f(x) < \alpha\}$ and $\operatorname{cl}\{x \mid f(x) < \alpha\} = \{x \mid f(x) \leq \alpha\}$.
- Convexity of f, rather than its level sets, is essential here (consider e.g. $f(x) = 1_{|x|>1}$).

8 Recession Cones and Unboundedness

9 Some Closedness Criteria

10 Continuity of Convex Functions

Definition: A function f is *continuous relative to* $S \subset \mathbb{R}^n$ if the restriction of f to S is a continuous function.

Theorem 10.1: A convex function f on $\mathbb{R}^n$ is continuous relative to any relatively open convex set C in its effective domain. (In particular, we can take $C = \operatorname{ri}(\operatorname{dom} f)$.)

Corollary 10.1.1: A convex function finite on all of $\mathbb{R}^n$ is necessarily continuous.

Skipped: This chapter goes on to give other conditions for a function to be upper semi-continuous. (If $\operatorname{cl} f$ is upper semi-continuous, then $\operatorname{cl} f$ is continuous.)

11 Separation Theorems

Definition: Let C_1, C_2 be non-empty sets in $\mathbb{R}^n$. A hyperplane H *separates* C_1 and C_2 if C_1 is contained in one of the closed half-spaces associated with H and C_2 lies in the opposite closed half-space.

Definition: H separates C_1 and C_2 *properly* if C_1 and C_2 are not both contained in H itself.

Definition: H separates C_1 and C_2 *strongly* if there exists $\epsilon > 0$ such that $C_1 + \epsilon B$ is contained in one of the open half-spaces associated with H and $C_2 + \epsilon B$ is contained in the opposite open half-space.

Definition: H separates C_1 and C_2 *strictly* if C_1 is contained in one of the open half-spaces associated with H and C_2 lies in the opposite open half-space.

Note: Proper separation and strong separation are the most useful because they correspond in a natural way to extrema of linear functions.

Theorem 11.1: Let C_1 and C_2 be non-empty sets in $\mathbb{R}^n$. There exists a hyperplane separating C_1 and C_2 properly if and only if there exists a vector b such that:

$$\inf\{\langle x, b\rangle : x \in C_1\} \geq \sup\{\langle x, b\rangle : x \in C_2\}$$

$$\sup\{\langle x, b\rangle : x \in C_1\} > \inf\{\langle x, b\rangle : x \in C_2\}$$

There exists a hyperplane separating C_1 and C_2 strongly if and only if there exists a vector b such that:

$$\inf\{\langle x, b\rangle : x \in C_1\} > \sup\{\langle x, b\rangle : x \in C_2\}$$

Note: One main application of separation theory is in the proofs of existence theorems. (Typically, the existence of vectors b with certain properties.)

Theorem 11.2: Let C be a non-empty relatively open convex set in $\mathbb{R}^n$, and let M be a non-empty affine set in $\mathbb{R}^n$ such that $M \cap C$ is empty. Then, there exists a hyperplane H such that $M \subset H$ and one of the open half-spaces associated with H contains C.

This is the main separation theorem: *Theorem 11.3:* Let C_1 and C_2 be non-empty convex sets in $\mathbb{R}^n$. There exists a hyperplane separating C_1 and C_2 properly if and only if $\operatorname{ri} C_1 \cap \operatorname{ri} C_2$ is empty.

Theorem 11.4: Let C_1 and C_2 be non-empty convex sets in $\mathbb{R}^n$. There exists a hyperplane separating C_1 and C_2 strongly if and only if:

$$\inf\{|x_1 - x_2| : x_1 \in C_1, x_2 \in C_2\} > 0$$

(In other words, $\operatorname{cl}(C_1 - C_2)$ does not contain 0.)

Definition: Let C be a non-empty convex set in $\mathbb{R}^n$. C *recedes in the direction* D if C includes all the half-lines in the direction D which start at points of C.

Corollary 11.4.1: Let C_1 and C_2 be non-empty disjoint closed convex sets in $\mathbb{R}^n$ having no common directions of recession. Then there exists a hyperplane separating C_1 and C_2 strongly.

Corollary 11.4.2: Let C_1 and C_2 be non-empty convex sets in $\mathbb{R}^n$ whose closures are disjoint. If either set is bounded, there exists a hyperplane separating C_1 and C_2 strongly.

Theorem 11.5: A closed convex set C is the intersection of the closed half-spaces which contain it.

Corollary 11.5.1: Let $S \subset \mathbb{R}^n$. Then $\operatorname{cl}(\operatorname{conv} S)$ is the intersection of all the closed half-spaces containing S.

Corollary 11.5.2: Let C be a convex subset of $\mathbb{R}^n$ with $C \neq \mathbb{R}^n$. Then there exists a closed half-space containing C. In other words, there exists some $b \in \mathbb{R}^n$ such that the linear function $\langle \cdot, b\rangle$ is bounded above on C.

Definitions: Let C be a convex set in $\mathbb{R}^n$. A *supporting half-space* to C is a closed half-space which contains C and has a point of C in its boundary. A *supporting hyperplane* to C is a hyperplane which is the boundary of a supporting half-space to C.

Comments: A supporting hyperplane to C can be associated with a linear function which achieves its maximum on C. The supporting hyperplanes passing through a given point $a \in C$ correspond to the vectors b *normal* to C at a.

We generally speak of *non-trivial* supporting hyperplanes, which do not contain C itself. (Consider when C is not n-dimensional.)

Theorem 11.6: Let C be a convex set, and let D be a non-empty convex subset of C. There exists a non-trivial supporting hyperplane to C containing D if and only if D is disjoint from $\operatorname{ri} C$.

Corollary 11.6.1: A convex set has a non-zero normal at each of its boundary points.

Corollary 11.6.2: Let C be a convex set. An $x \in C$ is a relative boundary point of C if and only if there exists a linear function h not constant on C such that h achieves its maximum over C at x.

Theorem 11.7: Let $C_1 \subset \mathbb{R}^n$ be a non-empty cone, and $C_2 \subset \mathbb{R}^n$ be non-empty. If there exists a hyperplane which separates C_1 and C_2 properly, then there exists a hyperplane which separates C_1 and C_2 properly and passes through the origin.

Corollary 11.7.1: A non-empty closed convex cone in $\mathbb{R}^n$ is the intersection of the homogenous closed half-spaces which contain it. (A homogenous half-space is a half-space with the origin on its boundary.)

Corollary 11.7.2: Let $S \subset \mathbb{R}^n$, and let $K = \operatorname{cl}(\operatorname{cone} S)$. Then K is the intersection of all the homogenous closed half-spaces containing S.

Corollary 11.7.3: Let K be a convex cone in $\mathbb{R}^n$ such that $K \neq \mathbb{R}^n$. Then K is contained in some homogenous closed half-space of $\mathbb{R}^n$. In other words, there exists a $b \neq 0$ such that $\langle x, b \rangle \leq 0$ for every $x \in K$.

12 Conjugates of Convex Functions

Two ways to view a curve/surface (e.g. a conic): a locus of points or an envelope of tangents.

For convexity: a closed convex set is the intersection of the closed half-spaces which contain it.

Conjugate of a function grows out of the fact the epigraph of a closed proper convex function on $\mathbb{R}^n$ is the intersection of the closed half-spaces in $\mathbb{R}^{n+1}$ that contain it.

Hyperplanes in $\mathbb{R}^{n+1}$ can be represented by the linear functions $\mathbb{R}^{n+1}$, i.e.

$$(x, \mu) \mapsto \langle x, b \rangle + \mu\beta_0$$

Hyperplane is unaffected by non-zero scaling of these linear functions, so there are essentially 2 cases: $\beta_0 = 0$ and $\beta_0 = -1$.

Definition: A hyperplane is *vertical* if it is of the form $\{(x, \mu) : \langle x, b \rangle = \beta\}$ for some $b \neq 0$.

Note: Let $h(x) = \langle x, b \rangle - \beta$.

Definition: A closed half-space is *vertical* if it is of the form $\{(x, \mu) : \langle x, b \rangle \leq \beta\}$ for some $b \neq 0$. Note this is $\{(x, \mu) : h(x) \leq 0\}$.

Definition: A closed half-space is *upper* if it is of the form $\{(x, \mu) : \mu \geq \langle x, b \rangle - \beta\}$. This is $\operatorname{epi} h$.

Definition: A closed half-space is *lower* if it is of the form $\{(x, \mu) : \mu \leq \langle x, b \rangle - \beta\}$.

Theorem 12.1: A closed convex function f is the pointwise supremum of the collection of all affine functions h such that $h \leq f$.

Note: Theorem 11.5 is a special case of Theorem 12.1.

Corollary 12.1.1: Let $f : \mathbb{R}^n \to [-\infty, +\infty]$. Then $\operatorname{cl}(\operatorname{conv} f)$ is the pointwise supremum of the collection of all affine functions on $\mathbb{R}^n$ majorized by f.

Corollary 12.1.2: Given any proper convex function f on $\mathbb{R}^n$, there exists some $b \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$ such that $f(x) \geq \langle x, b \rangle - \beta$ for every x.

Formulation:

Let f be any closed convex function on $\mathbb{R}^n$. By Theorem 12.1, we can describe f as the set F^* of all pairs $(x^*, \mu^*) \in \mathbb{R}^{n+1}$ such that the affine function $h(x) = \langle x, x^* \rangle - \mu^*$ is majorized by f.

We have $h(x) \leq f(x)$ for every x if and only if $\mu^* \geq \sup\{\langle x, x^* \rangle - f(x) : x \in \mathbb{R}^n\}$. So, F^* is actually the epigraph of f^* on $\mathbb{R}^n$ defined below.

Definition: For an arbitrary function f, f^* is the *conjugate* of f, defined as:

$$f^*(x^*) = \sup_x \{\langle x, x^* \rangle - f(x)\} = -\inf_x \{f(x) - \langle x, x^* \rangle\}$$

Comments:

Note that $f^*(x^*) = \sup\{\langle x, x^* \rangle - \mu : (x, \mu) \in \operatorname{epi} f\}$.

Note that the constant functions $+\infty$ and $-\infty$ are conjugate to each other. These are the only improper closed convex functions.

For arbitrary functions, f^* is the conjugate of $\operatorname{cl}(\operatorname{conv} f)$.

Conjugates reverse functional inequalities: $f_1 \leq f_2$ implies $f_1^* \geq f_2^*$.

Theorem 12.2: Let f be a convex function. The conjugate function f^* is a closed convex function. f^* is proper if and only if f is proper. Also: $(\operatorname{cl} f)^* = f^*$ and $f^{**} = \operatorname{cl} f$.

Corollary 12.2.1: The conjugacy operation $f \mapsto f^*$ induces a symmetric one-to-one correspondence in the class of all closed proper convex functions on $\mathbb{R}^n$.

Corollary 12.2.2: For any convex function f on $\mathbb{R}^n$, we have:

$$f^*(x^*) = \sup\{\langle x, x^* \rangle - f(x) : x \in \operatorname{ri}(\operatorname{dom} f)\}$$

Intuition:

The theory of conjugacy can be seen as the 'best' inequalities of the type:

$$\langle x, y \rangle \leq f(x) + g(y) \text{ for all } x, y$$

(Here, f and g map from $\mathbb{R}^n$ to $(-\infty, +\infty]$.)

Let W denote the set of all function pairs (f, g) satisfying this inequality. The 'best' pairs in W are those for which the inequality cannot be tightened, i.e. if $(f', g') \in W$ such that $f' \le f$ and $g' \le g$, then $f' = f$ and $g' = g$. Note that the 'best' pairs then are those such that $g = f^*$ and $f = g^*$. Thus, the 'best' inequalities correspond to the pairs of mutually conjugate closed proper convex functions.

Fenchel's inequality: For any proper convex function f:

$$\langle x, x^* \rangle \le f(x) + f^*(x^*) \text{ for all } x, x^*$$

The pairs (x, x^*) for which Fenchel's inequality is satisfied with equality form the graph of a multivalued mapping ∂f known as the *subdifferential*. (See §23-25.) The relationship between $f \mapsto f^*$ and the classical Legendre transformation is in §26.

More comments:

The identity $f^* = f$ has a unique solution $w(x) = (1/2)\langle x, x \rangle$. (Uniqueness can be shown by Fenchel's.)

There are many convex functions which satisfy $f^*(x^*) = f(-x^*)$.

If f is the indicator of a subspace L, then f^* is the indicator of the orthogonal complement $L^\perp$. Thus, $f^{**} = f$ corresponds to $L^{\perp\perp} = L$. (So, this is a special case, too.)

Formulation:

Let's generalize this. Consider a non-empty affine set, on which an affine function is given. These can be identified with *partial affine functions*, i.e. proper convex functions f such that $\operatorname{dom} f$ is an affine set, and f is affine on $\operatorname{dom} f$. These are necessarily closed (Corollary 7.4.2), so it is the conjugate of its conjugate.

Thus: partial affine functions come in dual pairs.

Any partial affine function can be expressed (non-uniquely):

$$f(x) = \delta(x|L + a) + \langle x, a^* \rangle + \alpha$$

L is a subspace, a, a^* are vectors, $\alpha \in \mathbb{R}$.

The conjugate partial affine function is:

$$f^*(x^*) = \delta(x^*|L^\perp + a^*) + \langle x^*, a \rangle + \alpha^*$$

Here, $\alpha^* = -\alpha - \langle a, a^* \rangle$.

Theorem 12.3: Let h be a convex function on $\mathbb{R}^n$, and let:

$$f(x) = h(A(x - a)) + \langle x, a^* \rangle + \alpha$$

where A is a one-to-one linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$, $a, a^* \in \mathbb{R}^n$, and $\alpha \in \mathbb{R}$. Then:

$$f^*(x^*) = h^*(A^{*-1}(x^* - a^*)) + \langle x^*, a \rangle + \alpha^*$$

where A^* is the adjoint of A and $\alpha^* = -\alpha - \langle a, a^* \rangle$.

(This is a simple proof.)

Tucker representation interpretation.

Formulation:

A proper convex function f is a *partial quadratic convex function* if it can be written:

$$f(x) = q(x) + \delta(x|M)$$

Here, q is a finite quadratic convex function on $\mathbb{R}^n$ and M is an affine set in $\mathbb{R}^n$.

Elementary partial quadratic convex functions. (The conjugate is also elementary.)

More generally, the conjugate of a partial quadratic convex function is a partial quadratic convex function. (Apply Theorem 12.3.)

Formulation:

Let f be any closed proper convex function. (Note $f^{**} = f$.) Note that: $\inf_x f(x) = -f^*(0)$ and $\inf_{x^*} f^*(x^*) = -f(0)$.

Thus: $\inf_x f(x) = 0 = f(0)$ if and only if $\inf_{x^*} f^*(x^*) = 0 = f^*(0)$.

Thus, the conjugacy correspondence preserves the class of non-negative closed convex functions which vanish at the origin.

Formulation:

Definition: A closed convex function f is *symmetric* if $f(-x) = f(x)$ for all x.

A closed convex function f is symmetric if and only if its conjugate is symmetric.

More generally: let G be any set of orthogonal linear transformations of $\mathbb{R}^n$ onto itself.

Definition: A closed convex function f is *symmetric with respect to* G if $f(Ax) = f(x)$ for all x and all $A \in G$.

Corollary 12.3.1: A closed convex function f is symmetric to a given set G of orthogonal linear transformations if and only if f^* is symmetric with respect to G.

The functions which are symmetric with respect to *all* orthogonal transformations of $\mathbb{R}^n$ are those of the form $f(x) = g(|x|)$, where $|\cdot|$ is the Euclidean norm and g is a function on $[0, +\infty)$.

Such an f is a closed proper convex function if and only if g is a non-decreasing lower semi-continuous convex function with $g(0)$ finite.

If so, the conjugate function is of the same type: $f^*(x^*) = g^+(|x^*|)$ where g^+ is a non-decreasing lower semi-continuous convex function on $[0, +\infty)$ with $g^+(0)$ finite.

Definition: The monotone conjugate of g (satisfying conditions above) is given by:

$$g^+(\xi^*) = \sup\{\xi\xi^* - g(\xi) : \xi \geq 0\}$$

Since $f^{**} = f$, we have $g^{++} = g$.

Thus: monotone conjugacy defines a symmetric one-to-one correspondence in the class of all non-decreasing lower semi-continuous convex functions on $[0, +\infty)$ which are finite at 0.

Note: The Euclidean norm can be replaced by any closed gauge function. (Formally seen in Theorem 15.3.)

Extended Formulation:

We can generalize monotone conjugacy to n dimensions.

Consider the class of functions f *symmetric in each coordinate*. f belongs to this class if and only if $f(x) = g(\text{abs } x)$, where g is a function on the non-negative orthant and abs is the coordinate-wise absolute value.

f is a closed proper convex function if and only if g is lower semi-continuous, convex, finite at the origin, and non-decreasing. (Here, non-decreasing is with respect to the partial ordering from element-wise inequalities.) Then, by Corollary 12.3.1:

$$f^*(x^*) = g^+(\text{abs } x^*)$$

Here, g^+ also satisfies all the conditions that g did. We can define the *monotone conjugate* of g:

$$g^+(z^*) = \sup\{\langle z, z^*\rangle - g(z) : z \geq 0\} \text{ for all } z^* \geq 0$$

Theorem 12.4: Let g be a non-decreasing lower semi-continuous convex function on the non-negative orthant of $\mathbb{R}^n$ such that $g(0)$ is finite. The monotone conjugate g^+ of g is another such function, and the monotone conjugate of g^+ is g.

There is a monotone conjugacy correspondence for concave functions as well.

13 Support Functions

Support function of a convex set C:

$$\delta^*(x^*|C) = \sup\{\langle x, x^* \rangle, x \in C\}$$

- Describes the half-spaces which contain C: $C \subset \{x : \langle x, x^* \rangle \leq \beta\} \Leftrightarrow \beta \geq \delta^*(x^*|C)$.
- $\operatorname{dom} \delta^*(\cdot|C)$ is the barrier cone of C.
- Can replace C by $\operatorname{ri} C$ or $\operatorname{cl} C$ in the definition (by continuity of linear operator).
- Consequences of separation theory

  Theorem (13.1)**.**

$$x \in \operatorname{cl} C \Leftrightarrow \langle x, x^* \rangle \leq \delta^*(x^*|C) \forall x^*$$
$$x \in \operatorname{ri} C \Leftrightarrow \langle x, x^* \rangle < \delta^*(x^*|C) \forall x^* \textit{ such that } \delta^*(x^*|C) \neq -\delta^*(-x^*|C)$$
$$x \in \operatorname{int} C \Leftrightarrow \langle x, x^* \rangle < \delta^*(x^*|C) \forall x^* \neq 0$$
$$\textit{If } C \neq \emptyset, x \in \operatorname{aff} C \Leftrightarrow \langle x, x^* \rangle = \delta^*(x^*|C) \forall x^* \textit{ such that } -\delta^*(x^*|C) = \delta^*(x^*|C)$$

- If C is closed and convex, then

$$C = \{x : \langle x, x^* \rangle \leq \delta^*(x^*|C)\ \forall x^*\}$$

- $\delta^*(\cdot|C_1 + C_2) = \delta^*(\cdot|C_1) + \delta^*(\cdot|C_2)$
- We have a correspondence between closed convex sets and functions $\mathbb{R}^n \to \mathbb{R}$. Which functions?

  Theorem (13.2)**.** *If C is closed convex, then $\delta^*(\cdot|C)$ and $\delta(\cdot|C)$ (the indicator of C) are convex conjugates of each other.*
  A function is the support function of some convex C iff it is a closed convex proper, positively homogeneous (CCPPH).

  proof: let f^* be a CCP and f its conjugate (also CCP). (note that we can restrict our attention to CCP functions because support functions of nonempty sets are CCP).
  f^* is the support function of some C iff f takes values in $\{0, \infty\}$, i.e. $f = \lambda f\ \forall \lambda > 0$.
  f^* is PH iff

$$f^*(x^*) = \lambda f^*(x^*/\lambda)$$

  but

$$\lambda f^*(x^*/\lambda) = \lambda \sup_x \langle x, x^* \lambda \rangle - f(x) = (\lambda f)^*$$

  so f^* is PH iff $f^* = (\lambda f)^*$ for all $\lambda > 0$.

- That set C is simply $C = \{x : \langle x, x^* \rangle \leq f^*(x^*)\}$. (Corollary 13.2.1)
- That set is bounded iff f^* is finite (Corollary 13.2.2)
- As a consequence, δ^* is lower semi continuous, and $\delta^*(x_1^* + x_2^*|C) \leq \delta_1(x_1^*|C) + \delta_2(x_2^*|C)$.
- Examples:

$$C = \{x \geq 0 : \sum x_i = 1\} \qquad \delta^*(x^*|C) = \max \xi_j^*$$
$$C = \{x : \|x\|_1 \leq 1\} \qquad \delta^*(x^*|C) = \max |\xi_j^*|$$
$$C = \{x : \xi_1 < 0,\ \xi_2 \leq \xi_1^{-1}\} \qquad \delta(x^*|C) = -2\sqrt{\xi_1^* \xi_2^*} + \delta^*(x^*|x^* \geq 0)$$

13.1 Recession cones and recession functions:

(A brief summary, since we skipped this chapter)

- C recedes in the direction y if $C + \lambda y \subset C$ for all $\lambda > 0$ (it is sufficient that this holds for $\lambda = 1$). These form a cone, called the recession cone, and denoted 0^+C.
- The notation comes from this fact: the recession cone of C can be obtained by forming the cone K (in $\mathbb{R}^{n+1}$) $K = \{\lambda(1, x), \lambda > 0, x \in C\}$, taking the closure, then intersecting with the hyperplane $\{(\lambda, x) : \lambda = 0\}$.
- The recession function of f, $f0^+$ is the convex function such that

$$\mathrm{epi}(f0^+) = 0^+(\mathrm{epi}\, f)$$

- It also satisfies

$$0^+f(y) = \sup\{f(x + y) - f(x), x \in \mathrm{dom}\, f\}$$

  to see this, we know that (y, v) is in $C = 0^+ \,\mathrm{dom}\, f$ iff $C + (y, v) \subset C$, which is equivalent to: $\forall x, f(x + y) \leq f(x) + v$.
- The lineality space of C is $(0^+C) \cap (-0^+C)$ (directions in which C is 'linear')
- Lineality space of f: $(0^+ \,\mathrm{epi}\, f) \cap (-0^+\pi f)$ (directions in which f is affine)

Theorem (13.3)**.** *Let f be CP. Then*

$$\delta^*(\cdot|\,\mathrm{dom}\, f) = f^*0^+$$

In particular, this means that $\mathrm{dom}\, f$ alone determines where f^* recedes.

- A convex f is called co-finite if $\mathrm{epi}\, f$ only recedes vertically. True in particular if $\mathrm{dom}\, f$ is bounded.
- Corollary 13.3.1: $\mathrm{dom}\, f^* = \mathbb{R}^n \Leftrightarrow f$ is cofinite.
- Corollary 13.3.3: $\mathrm{dom}\, f^*$ is bounded $\Leftrightarrow f$ is finite everywhere and f is α Lipschitz for some $\alpha \geq 0$. Then

$$\alpha = \sup\{|x^*|, x^* \in \mathrm{dom}\, f^*\}$$

  proof: f is α-Lipschitz $\Leftrightarrow f(x+y) \leq f(x)+\alpha|y| \Leftrightarrow (f0^+)(y) \leq \alpha|y| \Leftrightarrow \mathrm{cl}\,\mathrm{dom}\, f^* \subset \alpha B$ where B is the unit Euclidean ball, because $\alpha|y|$ is the support function of αB.

Theorem (13.4)**.** *Let f be CP. The lineality space of f^* is $(lin\,\mathrm{aff}\,\mathrm{dom}\, f)^\perp$ and*

$$\mathit{lineality} f^* = n - \dim f$$

(recall the dimension of f is the affine dimension of its domain, so the second part is immediate).

L the lineality space of f^* is the space of vectors x^* such that $-(f^*0^+)(-x^*) = (f^*0^+)(x^*)$. These are the sup and inf of $\langle\cdot, x^*\rangle$ on $\mathrm{dom}\, f$. So $x^* \in L \Leftrightarrow \langle\cdot, x^*\rangle$ is constant on $\mathrm{dom}\, f$ (also on $\mathrm{aff}\,\mathrm{dom}\, f$: if it is constant on $\mathrm{dom}\, f$, it is also constant on hyperplanes containing $\mathrm{dom}\, f$). This is equivalent to $x^* \perp (lin\,\mathrm{aff}\,\mathrm{dom}\, f)$.

- Corollary 13.4.2: Let f be CP. Then $\mathrm{dom}\, f^*$ has nonempty interior if and only if $\mathit{lineality} f = 0$.

  Theorem (13.5)**.** *Let f be CCP. Then*

$$\delta^*(\cdot|\{x : f(x) \leq 0\}) = \mathrm{cl}\, g$$

  where g is the CPH function generated by f^*.

- Corollary 13.5.1: Let f be CCP. Then

$$k(\lambda, x) = \begin{cases} (f\lambda)(x) & \text{if } \lambda > 0, \\ (f0^+)(x) & \text{if } \lambda = 0, \\ +\infty & \text{if } \lambda < 0 \end{cases}$$

  is the support function of $C = (\lambda^*, x^*)|\lambda^* \leq -f^*(x^*)\}$.

14 Polars of Convex Sets

- Motivation of polars: We have shown that (Theorem 13.2) a function is the indicator of a convex set iff its conjugate is CCPPH. So the convex conjugate of a PH indicator is another PH indicator. But PH indicators are those of convex cones.

  So if K is a convex cone, the conjugate of $\delta(\cdot|K)$ is the indicator of some other convex cone, K°. It is given by

$$\begin{aligned} K^\circ &= \{x^*|\forall x,\ \langle x, x^*\rangle \le \delta(x|K)\} \\ &= \{x^*|\forall x \in K,\ \langle x, x^*\rangle \le 0\} \end{aligned}$$

- Examples:

  1. If K is a subspace, then $K^\circ = K^\perp$.
  2. If K is the closed cone generated by vectors $\{a_i\}_{i\in I}$ (i.e. K is the set of all non-negative linear combinations of the a_is), then

$$K^\circ = \{x^*|\forall x \in K,\ \langle x, x^*\rangle \le 0\} = \{x^*|\forall i,\ \langle a_i, x^*\rangle \le 0\}$$

Theorem (14.1)**.** *If K is a closed non-empty cone, then K° is a closed non-empty cone, and $K^{\circ\circ} = K$. The indicators of K and K° are conjugates of each other.*

Theorem (14.2)**.** *Let f be CP, and K the convex cone generated by* $\operatorname{dom} f$*. Then K° is the recession cone of f^*.*

proof: the recession cone of f^* is the recession cone of any sublevel set of f^*

$$C = \{x^*|f^*(x^*) \le \alpha\} = \{x^*|\ \langle x, x^*\rangle - f(x) \le \alpha\ \forall x \in \operatorname{dom} f\}$$

Then

$$\begin{aligned} y^* \in 0 + C &\Leftrightarrow \forall x^* \in C,\ x^* + \mathbb{R}_+ y^* \subset C \\ &\Leftrightarrow \langle x, x_0^* + \mathbb{R}_+ y^*\rangle - f(x) \le \alpha\ \forall x \in \operatorname{dom} f \\ &\Leftrightarrow \langle y^*, x\rangle \le 0\ \forall x \in \operatorname{dom} f \\ &\Leftrightarrow \langle y^*, x\rangle \le 0\ \forall x \in K,\ \text{the convex cone generated by } \operatorname{dom} f \end{aligned}$$

So

$$0^+C = K^\circ$$

- Corollary 14.2.1: The polar of the barrier cone of C (domain of $\delta(\cdot|C)$) is the recession cone of C
- Corollary 14.2.2: Let f be CCP. Then the sub level sets of f are bounded iff $0 \in \operatorname{int} \operatorname{dom} f^*$. (proof:

$$\begin{aligned} 0 \in \operatorname{int} \operatorname{dom} f^* &\Leftrightarrow K = \mathbb{R}^n (K \text{ is the convex cone generated by } \operatorname{dom} f^*) \\ &\Leftrightarrow K^\circ = \{0\} \\ &\Leftrightarrow \text{Recession cone of } f \text{ is } \{0\} \\ &\Leftrightarrow \text{bounded sub level sets.} \end{aligned}$$

Theorem (14.3)**.** *Let f be CCP with $f(0) > 0 > \inf f$. The closed convex cone generated by $\{x : f(x) \le 0\}$ and $\{x : f^*(x^*) \le 0\}$ are polar to each other.*

-

Theorem (14.4)**.** *Let f be CCP on $\mathbb{R}^n$, and let K be the convex cone generated by the vectors $(1, x, \mu) \in \mathbb{R}^{n+2}$ such that $\mu \geq f(x)$. Define K^* similarly for f^*. Then*

$$\operatorname{cl} K^* = \{(\lambda^*, x^*, \mu^*) | (-\mu^*, x^*, -\lambda^*) \in K^\circ\}$$

- So far we only defined polars of convex cones. This can be generalized to convex sets containing the origin. This is done by taking duals of the gauge of the set (instead of the indicator of the cone). Recall the definition of the gauge of a nonempty convex set:

$$\gamma(x|C) = \inf\{\lambda \geq 0 : x \in \lambda C\}$$

It is the CPH generated by $\delta(\cdot|C) + 1$. Its closure is the support function of $\{x^* | f^*(x^* \leq 0)$ (Theorem 13.5), and $f^* = \delta^*(\cdot|C) - 1$. So $\operatorname{cl} \gamma(\cdot|C) = \delta^*(\cdot|C^\circ)$ where

$$C^\circ = \{x^* | \delta^*(x^*|C) - 1 \leq 0\} = \{x^* | \forall x \in C, \langle x, x^* \rangle \leq 1\}$$

This is called the polar of C. It always contains the origin. If C is closed and contains the origin, then $C^{\circ\circ} = C$. We also have the following symmetric one-to-one correspondence:

Theorem (14.5)**.** *Let C be CC set with $0 \in C$. Then $\gamma(\cdot|C)$ is $\delta^*(\cdot|C^\circ)$ (and the dual is true).*