In Progress: Summary of Notation and Basic Results
Convex Analysis C&O 663, Fall 2009

Intructor: Henry Wolkowicz

November 26, 2009

University of Waterloo
Department of Combinatorics & Optimization

Abstract

This contains a list of definitions/hyperlinks and basic results in Convex Analysis. Please notify the instructor about any errors and/or missing content.

Contents

1 Euclidean Spaces, Linear Manifolds, Hyperplanes

Definition 1.1 *A* Euclidean space $\mathbb{E}$ *is a finite dimensional vector space over the reals,* $\mathbb{R}$*, equipped with an inner product,* $\langle \cdot, \cdot \rangle$*.*

> Definitions and basic results on the following well know results are available at e.g. wikipedia. Hyperlinks are provided for several of them here:
> linear manifold, polyhedral set/polyhedron, Minkowski-Weyl Theorem, hyperplanes and halfspaces, affine hull, span, linear transformation, adjoint, relative interior, closure, boundary, Bolzano-Weierstrass Theorem.

1.1 Basics for Background

- **Unit ball in** $\mathbb{E}$**.** $B = \{x \in \mathbb{E} : \|x\| \leq 1\}$
- **Open set** $S \subseteq \mathbb{E}$. $\forall x \in S, \exists \delta > 0, \{x\} + \delta B \subseteq S$
- **Interior of** $S \subseteq \mathbb{E}$. $\mathrm{int}(S) = \{x \in \mathbb{E} : \{x\} + \delta B \subseteq S$ for some $\delta > 0\}$ = union of all open sets contained in S
- **Closed set** $S \subseteq \mathbb{E}$. $\forall x \notin S, \exists \delta > 0, (\{x\} + \delta B) \cap S = \emptyset$
- **Closure of** $S \subseteq \mathbb{E}$. $\mathrm{cl}(S) = \{x \in \mathbb{E} : \forall \delta > 0, (\{x\} + \delta B) \cap S \neq \emptyset\}$ = intersection of all closed sets containing S
- **Linear subspace** $S \subseteq \mathbb{E}$. $\forall x, y \in S, \forall \lambda, \mu \in \mathbb{R}, \lambda x + \mu y \in S$
- **Linear function** $f : \mathbb{E} \to (-\infty, +\infty]$. $\forall x, y \in \mathrm{dom}(f), \forall \lambda, \mu \in \mathbb{R}, f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$
- **Lower semicontinuous function** $f : \mathbb{E} \to (-\infty, +\infty]$ **at** x. $\liminf_{y \to x} f(y) \geq f(x)$
- **Linear map** $L : \mathbb{E} \to \mathbb{Y}$. $\forall x, y \in \mathbb{E}, \forall \lambda, \mu \in \mathbb{R}, L(\lambda x + \mu y) = \lambda L(x) + \mu L(y)$
- **Adjoint of linear map** $A : \mathbb{E} \to \mathbb{Y}$. Linear map $A^{adj} : \mathbb{Y} \to \mathbb{E}$ satisfying $\forall x \in \mathbb{E}, \forall y \in \mathbb{Y}, \langle A^{adj} y, x \rangle_{\mathbb{E}} = \langle y, Ax \rangle_{\mathbb{Y}}$
- **Affine subspace** $S \subseteq \mathbb{E}$. (1) $\forall x, y \in S, \forall \lambda \in \mathbb{R}, \lambda x + (1 - \lambda) y \in S$
  (2) $S = V + \{x\}$ for some linear subspace V and vector x

- **Affine function** $a : \mathbb{E} \to (-\infty, +\infty]$. (1) $\forall x, y \in \mathrm{dom}(a), \forall \lambda \in \mathbb{R}, a(\lambda x + (1-\lambda)y) = \lambda a(x) + (1-\lambda)a(y)$
  (2) $a : x \mapsto f(x) + r$ for some linear function f and real number r
- **Affine map** $A : \mathbb{E} \to \mathbb{Y}$. (1) $\forall x, y \in \mathbb{E}, \forall \lambda \in \mathbb{R}, A(\lambda x + (1-\lambda)y) = \lambda A(x) + (1-\lambda)A(y)$
  (2) $A : x \mapsto L(x) + b$ for some linear map L and vector b
- **Affine hull of** $S \subseteq \mathbb{E}$. $\mathrm{Aff}(S) = \{\lambda x + (1-\lambda)y : x, y \in S, \lambda \in \mathbb{R}\}$ = intersection of all affine subspaces containing S
- **Cone** $K \subseteq \mathbb{E}$. $\forall x \in K, \forall \lambda > 0, \lambda x \in K$
- **Positively-homogeneous function** $f : \mathbb{E} \to [-\infty, +\infty]$. (1) $\forall x \in \mathbb{E}, \forall \lambda > 0, f(\lambda x) = \lambda f(x)$
  (2) $\mathrm{epi}(f)$ is a cone
- **Relatively open set** $S \subseteq \mathbb{E}$. $\forall x \in S, \exists \delta > 0, (\{x\} + \delta B) \cap \mathrm{Aff}(S) \subseteq S$
- **Relative interior of** S. $\mathrm{ri}(S) = \{x \in \mathrm{Aff}(S) : (\{x\} + \delta B) \cap \mathrm{Aff}(S) \subseteq S \text{ for some } \delta > 0\}$
- **Domain of** $f : \mathbb{E} \to [-\infty, +\infty]$. $\mathrm{dom}(f) = \{x \in \mathbb{E} : f(x) < +\infty\}$
- **Proper function** $f : \mathbb{E} \to [-\infty, +\infty]$. $\mathrm{dom}(f) \neq \emptyset$ and $f(x) > -\infty$ for all $x \in \mathbb{E}$
- **Epigraph of** $f : \mathbb{E} \to [-\infty, +\infty]$. $\mathrm{epi}(f) = \{(x, r) \in \mathbb{E} \oplus \mathbb{R} : f(x) \le r\}$
- **Sub-level set of** $f : \mathbb{E} \to [-\infty, +\infty]$ **at level** $r \in \mathbb{R}$. $S_r(f) = \{x \in \mathbb{E} : f(x) \le r\}$
- **Closure of** $f : \mathbb{E} \to [-\infty, +\infty]$. $\mathrm{cl}(f) : x \in \mathbb{E} \mapsto \liminf_{y \to x} f(y)$
- **Infimum convolution of** $f, g : \mathbb{E} \to (-\infty, +\infty]$. $f \odot g : x \in \mathbb{E} \mapsto \inf\{f(y) + g(x-y)\}$
- **Indicator function of** $S \subseteq \mathbb{E}$. $\delta_S : x \in \mathbb{E} \mapsto 0$ if $x \in S$, $+\infty$ otherwise

2 Convex Sets and Functions

2.1 Convex Sets

Definition 2.1 *The set* $S \subset \mathbb{E}$ *is a* convex set *if*

$$\lambda x + (1-\lambda)y \in S, \forall \lambda \in (0,1), \forall x, y \in S.$$

Proposition 2.2 *For a nonempty convex set* C*:*

1. *We have* $\mathrm{relint}\, C \neq \emptyset$ *and the affine hulls* $\mathrm{aff}\, C = \mathrm{aff}\, \mathrm{relint}\,(C)$. *Moreover, for any* $x \in \mathrm{relint}\, C$ *and* $y \in \mathrm{cl}\, C$, *the line segment* $[x, y) \subset \mathrm{relint}\, C$ *and thus* $\mathrm{relint}\, C$ *is convex. Furthermore,*

$$\mathrm{cl}\, C = \mathrm{cl}\, \mathrm{relint}\, C, \qquad \mathrm{relint}\, C = \mathrm{relint}\, \mathrm{cl}\, C.$$

2. $\mathrm{relint}\, C \subset C \subset \mathrm{cl}\, C$.

Include definitions and basic results on: Basic (strong, strict) separation theorems; convex hull; convex combination, recession cones, Caratheodory Theorem.

2.2 Convex Functions

Definition 2.3 *The* epigraph *of a function* $f : \mathbb{R}^n \to (-\infty, +\infty]$ *is defined as*

$$\text{epi}\,(f) = \{(x, r) : f(x) \leq r\}.$$

Definition 2.4 *The function* $f : \mathbb{R}^n \to (-\infty, +\infty]$ *is* a convex function *if*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y), \forall x, y \in \mathbb{R}^n, \forall \lambda \in [0, 1].$$

Definition 2.5 *The* convex hull *or convex envelope of a function* $f : \mathbb{R}^n \to \mathbb{R}$ *is defined as*

$$\text{conv}\,(f)(x) = \inf\{t : (x, t) \in \text{conv}\,\text{epi}\,f\}.$$

Proposition 2.6 *A convex function* f *is locally Lipschitz on the interior of its domain.*

Include definitions and basic results on: composing convex functions, convex growth conditions, locally Lipschitz

2.3 Basics for Convex Functions and Convex Sets

- **Convex set**. $\forall x, y \in S, \forall \lambda \in (0, 1), \lambda x + (1-\lambda)y \in S$
- **Convex function** $f : \mathbb{E} \to [-\infty, +\infty]$. $\mathrm{epi}(f)$ is convex; note that f is a *proper convex function* if its domain is nonempty and it does not take on the value $-\infty$.
- **Sublinear function** $f : \mathbb{E} \to [-\infty, +\infty]$. f is positively-homogeneous and convex
- **Subadditive function** $f : \mathbb{E} \to [-\infty, +\infty]$. $\forall x, y \in \mathrm{dom}(f), f(x+y) \leq f(x) + f(y)$
- **Convex hull of** $S \subseteq \mathbb{E}$. $\mathrm{conv}(S) = \{\lambda x + (1-\lambda)y : x, y \in S, \lambda \in (0, 1)\}$
- **Convex hull of** $f : \mathbb{E} \to [-\infty, +\infty]$. $\mathrm{conv}(f) : x \mapsto \inf\{r : (x, r) \in \mathrm{conv}(\mathrm{epi}(f))\}$
- **Locally Lipschitz** f **at** $x \in \mathrm{dom}(f)$. $\exists K > 0, \exists \delta > 0, \forall y, z \in \{x\} + \delta B, |f(y) - f(z)| \leq K\|y - z\|$

3 Duality of Functions and Sets

3.1 Conjugate, Positively Homogeneous, Sublinear Functions

Definition 3.1 *The Fenchel conjugate of* $h : \mathbb{E} \to [-\infty, +\infty]$ *is*

$$h^*(\phi) := \sup_{x \in \mathbb{E}} \{\langle \phi, x \rangle - h(x)\}.$$

Proposition 3.2 *1.* $f \geq g \Rightarrow f^* \leq g^*$

3.2 Indicator Functions, Support Functions and Sets, Closures

Definition 3.3 *The* indicator function of a set $S \subset \mathbb{E}$ *is*

$$\delta_S(x) := \begin{cases} 0 & \textit{if } x \in S \\ \infty & \textit{otherwise} \end{cases}$$

Definition 3.4 *The* support function of a set $S \subset \mathbb{E}$ *is*

$$\sigma_S(\phi) := \sup_{x \in S}\{\langle \phi, x \rangle\}.$$

Definition 3.5 *A function* f *is* positively homogeneous *if*

$$f(\lambda x) = \lambda f(x), \forall \lambda > 0, \forall x \in \mathbb{E}.$$

Remark 3.6 *Equivalently, the function* f *is positively homogeneous if*

$$f(\lambda x) \leq \lambda f(x), \forall \lambda > 0, \forall x \in \mathbb{E}.$$

And, a support function is positively homogeneous.

Definition 3.7 *A function is* sublinear *if it is subadditive and positively homogeneous, equivalently, if*

$$f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y), \forall \alpha > 0, \beta > 0, \forall x, y \in \mathbb{E}.$$

Definition 3.8 *The set* $S_f := \{\phi : \langle \phi, x \rangle \leq f(x), \forall x\}$ *is the* set supported by f.

Proposition 3.9 *Suppose that the function* f *is positively homogeneous. Then the conjugate function*

$$f^* = \delta_{S_f}.$$

3.2.1 Closures of Sets and Functions

Proposition 3.10 $\delta_S^{**} = \delta_S$ *iff* S *is closed and convex.*

Proposition 3.11 *The second conjugate function* $f^{**} = f$ *iff* f *is a closed and convex function.*

Definition 3.12 *The* closure of a function f *is defined as*

$$\mathrm{cl}\,(f)(x) = \inf\left\{\lim_{k\to\infty} f(x^k) : x^k \to x \textit{ and } \lim_{k\to\infty} f(x^k) \textit{ exists}\right\}$$

Proposition 3.13 *The second conjugate functions:*

$$\begin{aligned} \delta_S^{**} &= \delta_{\mathrm{cl}\,(\mathrm{conv}\,(S))} \\ \sigma_S^{**} &= \sigma_{\mathrm{cl}\,(\mathrm{conv}\,(S))} \\ f^{**} &= \mathrm{cl}\,(\mathrm{conv}\,(f)) \end{aligned}$$

Proposition 3.14 *The second polar* $S^{\circ\circ} = \mathrm{cl}\,(\mathrm{conv}\,(S \cup \{0\}))$.

3.2.2 Convex Cones

Proposition 3.15 *If* K *is a nonempty cone, then* $\mathsf{K}^{--} = \mathrm{cl}\,(\mathrm{conv}\,(\mathsf{K}))$.

3.2.3 More on Support Functions

Theorem 3.16

1. *If* $\emptyset \neq \mathsf{S} \subset \mathbb{E}$ *is a closed, convex set, then the support function* σ_{S} *is a proper, closed, sublinear function.*
2. *Moreover, if* f *is a proper, closed and sublinear function, then*
$$\mathsf{f} = \sigma_{\mathsf{S}_{\mathsf{f}}},$$
i.e. it is the support function of the set supported by f.
3. *Thus* $\mathsf{S} \leftrightarrow \sigma_{\mathsf{S}}$ *is a bijection between* {*closed, convex sets*} *and* {*closed, sublinear functions*}.

3.3 Gauge Functions, Polar of a Function, Norms and Dual Norms

Definition 3.17 *The function defined by* $\gamma_{\mathsf{S}}(x) := \inf\{\lambda \geq 0 : x \in \lambda \mathsf{S}\}$ *is called the* gauge of S.

Definition 3.18 *The* polar of a function g *is*

$$g^{\circ}(\phi) := \inf\{\lambda > 0 : \langle \phi, x \rangle \leq \lambda g(x), \forall x\}$$

Proposition 3.19

1. *The support function of the polar set of S,* $\sigma_{S^{\circ}}$*, is majorized by the gauge function of* S, γ_{S}.
2. $\gamma_{\mathsf{S}} \geq 0$ *and* $\gamma(0) = 0$.
3. γ_{S} *is positively homogeneous.*
4. *If* S *is convex, then* γ_{S} *is sublinear.*
5. *If* S *is closed and convex, then* γ_{S} *is closed and sublinear.*
6. $$\gamma_{\mathsf{S}} = \gamma_{\mathsf{S}}^{**} = \delta_{\mathsf{S}^{\circ}}^{*} = \sigma_{\mathsf{S}^{\circ}}.$$
7. *A gauge function is a non-negative sublinear function which maps the origin to* 0.
8. *A norm is a gauge function. Conversely, the gauge function of a closed, convex set containing* 0 *is a norm.*

Proposition 3.20 *Given a norm* $\|\cdot\|$*, then the polar function* $\|\cdot\|^{\circ}$*, is also a norm, called the dual norm. Moreover,*

$$\mathsf{S}_{\|\cdot\|} = \{\phi : \|\phi\|^{\circ} \leq 1\}, \qquad \mathsf{S}_{\|\cdot\|^{\circ}} = \{x : \|x\| \leq 1\} = \mathsf{S}_{\|\cdot\|}^{\circ}.$$

3.4 Subdifferentials, Directional Derivatives, Set Constrained Optimization

3.4.1 Subdifferentials and Directional Derivatives

Theorem 3.21 *Let* f *be a differentiable function on an open convex subset* $S \subset \mathbb{E}$. *Each of the following conditions is necessary and sufficient for* f *to be convex on* S*:*

1. $f(x) - f(y) \geq \langle x - y, \nabla f(y) \rangle, \forall x, y \in S.$
2. $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0, \forall x, y \in S.$
3. $\nabla^2 f(x)$ *is positive semidefinite for all* $x \in S$ *whenever* f *is twice differentiable on* S.

To extend results as in Theorem 3.21 to the nondifferentiable case, we use the following.

Definition 3.22 *The vector* ϕ *is called a subgradient of* f *at* x *if*

$$f(y) - f(x) \geq \langle \phi, y - x \rangle, \forall y \in \mathcal{E}.$$

The subdifferential of f *at* x *is*

$$\partial f(x) = \{\phi : f(y) - f(x) \geq \langle \phi, y - x \rangle, \forall y \in \mathcal{E}.$$

$\partial f(x) = \emptyset$, *if* $x \notin \operatorname{dom}(f)$.

Proposition 3.23 *Suppose that* f *is convex. Then* $\partial f(x)$ *is a closed convex set. And,* $x \in \operatorname{argmin}_x f(x)$ *if and only if* $0 \in \partial f(x)$.

Proposition 3.24 *Suppose that* $f : \mathbb{E} \to (-\infty, +\infty]$ *is convex. Let*

$$g(t) := \frac{f(x + td) - f(x)}{t}.$$

Then for all $x, d \in \mathbb{E}$, $x \in \operatorname{dom}(f)$, *the function* g *is monotonically nondecreasing for* $t > 0$ *(and for* $t < 0$*).*

Definition 3.25 *The directional derivative of* f *at* x *(in* $\operatorname{dom}(f)$*) along* d *is*

$$f'(x; d) := \lim_{t \downarrow 0} \frac{1}{t} f(x + td) - f(x)$$

if it exists.

Theorem 3.26 *Suppose that* f *is convex. Then for all* $x, d \in \mathbb{E}$, $x \in \operatorname{dom}(f)$, *the directional derivative*

$$f'(x; d) = \lim_{t \downarrow 0} \frac{f(x + td) - f(x)}{t}$$

exists in $[-\infty, +\infty]$.

3.4.2 Properties of $f'(x;d), \partial f(x)$

Proposition 3.27 *Let* f *be convex and* $x \in \mathrm{dom}\,(f)$. *Then* ϕ *is a subgradient of* f *at* x *iff* $f'(x;d) \geq \langle \phi, d \rangle, \forall d \in \mathbb{E}$.

Proposition 3.28 *Let* f, g *be proper convex functions.*

1. $f'(x;\cdot)$ *is positively homogeneous.*
2. *If* f *is convex, then* $f'(x;\cdot)$ *is convex; hence it is sublinear.*
3. *If* f *is convex, then* $\forall x \in \mathrm{dom}\,(f)$ *we have*
$$\partial f(x) = S_{f'(x;\cdot)}.$$
4. $$\partial(f+g)(x) \supset \partial f(x) + \partial g(x).$$
5. *With* $f(x)$ *finite:*
   (a) $\partial f(x) \neq \emptyset \Rightarrow f(x) = f^{**}(x)$.
   (b) $f(x) = f^{**}(x) \Rightarrow \partial f(x) = \partial f^{**}(x)$.
   (c) $y \in \partial f(x) \Rightarrow x \in \partial f(y)$.

Example 3.29 *Let* $X \in \mathbb{S}^n$, $f(X) := \lambda_{\max}(X)$ *denote the largest eigenvalue of* X, *and let* V *be the corresponding eigenspace, i.e. the subspace of eigenvectors* $V = \{v : Xv = \lambda_{\max}(X)v\}$. *Then the directional derivative in the direction* $D \in \mathbb{S}^n$ *is*

$$f'(X;D) = \max_{\|v\|=1, v \in V} v^T D v = \sigma_{\partial f(X)}.$$

Therefore, f *is differentiable if* $\partial f(X)$ *is a singleton, i.e. if the eigenvalue* $\lambda_{\max}(X)$ *is a singleton so the dimension of the eigenspace* V *is* 1.

3.5 Basics for Duality of Functions and Sets

- **Polar set of** $S \subseteq \mathbb{E}$. $S^\circ = \bigcap_{x \in S} \{\phi \in \mathbb{E} : \langle \phi, x \rangle \leq 1\}$
- **Polar cone of** $K \subseteq \mathbb{E}$. $K^- = \bigcap_{x \in K} \{\phi \in \mathbb{E} : \langle \phi, x \rangle \leq 0\}$
- **Fenchel conjugate of** $f : \mathbb{E} \to [-\infty, +\infty]$. $f^* : \phi \in \mathbb{E} \mapsto \sup_{x \in \mathrm{dom}(f)} \{\langle \phi, x \rangle - f(x)\}$
- **Support function of** $S \subseteq \mathbb{E}$. $\sigma_S = \delta_S^* : \phi \mapsto \sup\{\langle \phi, x \rangle : x \in S\}$
- **Set supported by** $f : \mathbb{E} \to [-\infty, +\infty]$. $S_f = \{\phi \in \mathbb{E} : \forall x \in \mathbb{E}, \langle \phi, x \rangle \leq f(x)\}$

4 Optimization

4.1 Set Constrained Optimization and Normal Cones

Proposition 4.1 *Suppose that* f *is a differentiable convex function and* S *is an open convex set. Then* $\bar{x} \in \operatorname{argmin}_{x \in S} f(x)$ *iff* $\nabla f(\bar{x}) = 0$.

Definition 4.2 *The normal cone to the convex set* C *in* $\mathbb{E}$ *at* $\bar{x} \in C$ *is*

$$N_C(\bar{x}) := \{d \in \mathbb{E} : \langle d, x - \bar{x} \rangle \leq 0, \forall x \in C\}.$$

Definition 4.3 *The (convex) tangent cone to the convex set* C *in* $\mathbb{E}$ *at* $\bar{x} \in C$ *is*

$$T_C(\bar{x}) := \operatorname{cl} \operatorname{cone} (C - \bar{x}).$$

Definition 4.4 *The set of feasible directions to the convex set* C *in* $\mathbb{E}$ *at* $\bar{x} \in C$ *is*

$$\mathcal{D}_C(\bar{x}) := \operatorname{cone} (C - \bar{x}).$$

Proposition 4.5 *Suppose that* C *is a convex set and* $f : C \to \mathbb{R}$. *If* $\bar{x}$ *is a* <u>*local*</u> *minimum of* f *on* C*, then*

$$f'(\bar{x}; x - \bar{x}) \geq 0, \forall x \in C. \tag{4.1}$$

If f *is differentiable, this is equivalent to* $\nabla f(\bar{x}) \in -N_C(\bar{x})$.
If, in addition, f *is convex on* C*, then the condition (4.1) is sufficient for* $\bar{x}$ *to be a minimum of* f *on* C*, i.e. we get (if* f *is lsc on* S*) that*

$$\bar{x} \in \operatorname*{argmin}_{x \in C} f(x) \text{ iff } \exists \phi \in (-N_C(\bar{x})) \cap \partial f(\bar{x}).$$

4.2 Basics for Optimization

- **Subdifferential of** f **at** $x \in \operatorname{dom}(f)$. $\partial f(x) = \{\phi \in \mathbb{E} : \forall y \in \operatorname{dom}(f), \langle \phi, y - x \rangle \leq f(y) - f(x)\}$
- **Subgradient of** f **at** $x \in \operatorname{dom}(f)$. $\phi \in \partial f(x)$
- **Directional derivative of** f **at** $x \in \operatorname{dom}(f)$ **in direction** $d \in \mathbb{E}$. $f'(x; d) = \lim_{t \downarrow 0} \frac{1}{t}[f(x + td) - f(x)]$, if exists
- **Differentiability of** f **at** $x \in \operatorname{dom}(f)$. $\exists \nabla f(x) \in \mathbb{E}, \forall d \in \mathbb{E}, f'(x; d) = \langle \nabla f(x), d \rangle$; $\nabla f(x)$ is called the gradient
- **Normal cone to convex set** S **at** $x \in S$. $N_S(x) = \bigcap_{y \in S} \{\phi \in \mathbb{E} : \langle \phi, y - x \rangle \leq 0\} = \partial \delta_S(x)$

5 Theorems

5.1 Convexity

- **Relative interior**.
  S convex $\implies \emptyset \neq \operatorname{ri}(S) = \{x \in S : \forall y \in S, \exists \delta > 0, x + \delta(x - y) \in S\} = \{x \in S : \bigcup_{t \geq 0} t(S - \{x\})$ is a linear subspace$\}$

- **Convexity preserving operations**. Suppose that $\{S_t : t \in T\}$ is a collection of convex sets, $\{f_t : t \in T\}$ is a collection of convex functions, and A is an affine map. Then the following are convex:

  –: $\bigcap_{t\in T} S_t$ – $\bigoplus_{t\in T} S_t$ (T finite) –: $\sum_{t\in T} S_t$ (T finite)

  –: $A(S_t)$ $(t \in T)$ – $A^{-1}(S_t)$ $(t \in T)$ – $\mathrm{ri}(S_t)$ $(t \in T)$ – $\mathrm{cl}(S_t)$ $(t \in T)$

  –: $\sup_{t\in T} f_t$ – $\sum_{t\in T} f_t$ (T finite) –: $\bigodot_{t\in T} f_t$ (T finite) – $f_t \circ A$ $(t \in T)$

- **Monotonicity of gradient**. Suppose f continuous over $\mathrm{dom}(f)$ and differentiable over $\mathrm{int}(\mathrm{dom}(f))$, $\mathrm{dom}(f)$ convex, and $\mathrm{int}(\mathrm{dom}(f)) \neq \emptyset$.

  –: f convex $\iff \forall x, y \in \mathrm{int}(\mathrm{dom}(f)), \langle \nabla f(x) - \nabla f(y), x - y\rangle \geq 0$

  –: f strictly convex over $\mathrm{int}(\mathrm{dom}(f)) \iff \forall x, y \in \mathrm{int}(\mathrm{dom}(f)), \langle \nabla f(x) - \nabla f(y), x - y\rangle > 0$

- **Interior representation of convexity**. –: $S \subseteq \mathbb{E}$ is convex $\iff S = \mathrm{conv}(S)$
  –: $f : \mathbb{E} \to [-\infty, +\infty]$ is convex $\iff f = \mathrm{conv}(f)$

- **Basic separation**. S is a closed, convex set and $x \notin S \implies \exists a \in \mathbb{E}, \exists b \in \mathbb{R}, \forall y \in S, \langle a, x\rangle > b \geq \langle a, y\rangle$.
  If S is a cone, we may take $b = 0$.

- **Characterization of sublinearity**. f is sublinear $\iff$ f is positively-homogeneous and subadditive.
  f is proper, closed and sublinear $\iff S_f \neq \emptyset$ and $f = \sigma_{S_f}$

- **Continuity of convex functions**. f is proper and convex, and $x \in \mathrm{int}(\mathrm{dom}(f)) \implies f$ is locally Lipschitz at x

5.2 Duality

- **Exterior representation of convexity**.

  –: $S \subseteq \mathbb{E}$ is closed, convex and contains $0 \iff S = (S^\circ)^\circ$
  –: $f : \mathbb{E} \to [-\infty, +\infty]$ is closed and convex $\iff f = (f^*)^*$

- **Fenchel-Young inequality**. $\forall \phi, x \in \mathbb{E}, f(x) + f^*(\phi) \geq \langle \phi, x\rangle$, with equality iff $\phi \in \partial f(x)$

- **Polar Calculus**. Suppose S, T are nonempty sets, K a nonempty cone.

  –: $(S^\circ)^\circ = \mathrm{cl}(\mathrm{conv}(S \cup \{0\}))$

  –: $(K^-)^- = \mathrm{cl}(\mathrm{conv}(K))$

  –: $(S \cup T)^\circ = S^\circ \cap T^\circ$

  –: $(S \cap T)^\circ \supseteq \mathrm{cl}(\mathrm{conv}(S^\circ \cup T^\circ))$, with equality when S, T are closed, convex and contain 0

- **Conjugate calculus**. Suppose $f, g, f_1, \ldots, f_m$ are proper.

  –: $(f^*)^* = \mathrm{cl}(\mathrm{conv}(f))$

  –: f, g convex $\implies (f \odot g)^* = f^* + g^*$

  –: f, g convex $\implies (f + g)^* \leq f^* \odot g^*$. Equality holds when $\mathrm{int}(\mathrm{dom}(f)) \cap \mathrm{dom}(g) \neq \emptyset$

–: f convex, and A a linear map $\implies (f \circ A)^*(\phi) \le \inf\{f^*(\psi) : A^{adj}\psi = \phi\}$.
Equality holds when $\exists y, Ay \in \mathrm{int}(\mathrm{dom}(f))$, in which case infimum is attained when finite

–: $f_1, \dots, f_m$ convex with common domain $\implies (\max_i f_i)^*(\phi) \le \inf\{\sum_{i=1}^m \lambda_i f_i^*(\phi^i) : \sum_{i=1}^m \lambda_i(\phi^i, 1) = (\phi, 1), \lambda_i \ge 0\}$.
Equality holds when $\mathrm{int}(\mathrm{dom}(f_i)) \ne \emptyset$, in which case the infimum is attained when finite

- **Fenchel duality**. Suppose f, g are proper and convex

  –: $\inf\{f(x) + g(x)\} \ge \sup\{-f^*(-\phi) - g^*(\phi)\}$

  –: Equality holds when $\mathrm{int}(\mathrm{dom}(f)) \cap \mathrm{dom}(g) \ne \emptyset$, in which case the supremum is attained when finite

- **Convex conic duality**. Suppose K is a convex cone, and A, D are linear maps

  –: $\inf_x\{\langle c, x\rangle : b - Ax \in K, Dx = e\} \ge \sup_{\phi,\eta}\{\langle b, \phi\rangle + \langle e, \eta\rangle : A^{adj}\phi + D^{adj}\eta = c, \phi \in K^-\}$

  –: Equality holds when $\exists x, Dx = e, b - Ax \in \mathrm{int}(K)$, in which case the supremum is attained if finite

- **Lagrange duality**. Suppose $f, g_1, \dots, g_m$ are proper, $L : (x, \lambda) \in \mathbb{E} \oplus \mathbb{R}^m \mapsto f(x) + \sum_{i=1}^m \lambda_i g_i(x)$, and $D = \mathrm{dom}(f) \cap (\bigcap_{i=1}^m \mathrm{dom}(g_i))$.

  –: $\inf\{f(x) : g_i(x) \le 0, i = 1, \dots, m\} \ge \sup_{\lambda \ge 0} \inf_x L(x, \lambda)$

  –: $\exists x \in D, \lambda \ge 0, (g_i(x) \le 0, i = 1, \dots, m) \wedge$ (x minimizes $y \mapsto L(y, \lambda)$ over D) $\wedge (\lambda_i g_i(x) = 0, i = 1, \dots, m)$
  $\implies$ equality holds with x and λ attaining the infimum and supremum respectively

  –: ($f, g_1, \dots, g_m$ convex) $\wedge (\exists y \in \mathrm{dom}(f), \forall i \in \{1, \dots, m\} g_i(y) < 0) \wedge$ (x attains the infimum)
  $\implies$ equality holds and $\exists \lambda \ge 0$, (x minimizes $y \mapsto L(y, \lambda)$ over D) $\wedge (\lambda_i g_i(x) = 0, i = 1, \dots, m)$

  –: ($f, g_1, \dots, g_m$ closed and convex) $\wedge$ ($\exists \lambda \ge 0, x \mapsto L(x, \lambda)$ has bounded sub-level sets
  $\implies$ equality holds and the infimum is attained if finite

5.3 Optimization

- **Convex optimality conditions**. Suppose $f, g_1, \dots, g_m$ are proper and convex, and S is nonempty and convex

  –: x minimizes $f \iff 0 \in \partial f(x)$

  –: $0 \in \partial f(x) + N_S(x) \implies x$ minimizes f over S. The converse is true when $\mathrm{int}(\mathrm{dom}(f)) \cap S \ne \emptyset$

  –: (KKT condition) Suppose $S = \{y : g_i(y) \le 0, i = 1, \dots, m\}$, $x \in S$ and $f, g_1, \dots, g_m$ are differentiable at x.
  $\exists \lambda \ge 0, (\nabla f(x) + \sum_{i=1}^m \lambda_i \nabla g(x) = 0) \wedge (\lambda_i g_i(x) = 0, i = 1, \dots, m) \implies x$ minimizes f over S.
  The converse is true when $\exists y \in \mathrm{dom}(f), \forall i \in \{1, \dots, m\}, g_i(y) < 0$

- **Subdifferential calculus**. Suppose $f, g, f_1, \dots, f_m$ are proper and convex with $f_1, \dots, f_m$ sharing same domain, and A is a linear map.

  –: $\forall x \in \mathrm{dom}(f) \cap \mathrm{dom}(g), \partial(f + g)(x) \supseteq \partial f(x) + \partial g(x)$. Equality holds when $\mathrm{int}(\mathrm{dom}(f)) \cap \mathrm{dom}(g) \ne \emptyset$

–: $\forall x \in \mathrm{dom}(f \circ A)), \partial(f \circ A)(x) \supseteq A^{adj}\partial f(Ax)$. Equality holds when $\exists y, Ay \in \mathrm{int}(\mathrm{dom}(f))$

–: $\forall x \in \mathrm{dom}(f_i), \partial(\max_i f_i)(x) = \bigcup\{\partial(\sum_{i\in I}\lambda_i f_i)(x) : \sum_{i\in I}\lambda_i = 1, \lambda_i \geq 0\} \supseteq \mathrm{conv}(\bigcup_{i\in I}\partial f_i(x))$, where $I = \{i : f_i(x) = f(x)\}$. Equality holds when $\mathrm{int}(\mathrm{dom}(f_i)) \neq \emptyset$.

- **Sublinearity of directional derivatives**. Suppose f is proper and convex.

  –: $x \in \mathrm{dom}(f) \implies f'(x;\cdot) : d \in \mathbb{E} \mapsto f'(x;d)$ is sublinear and $\partial f(x) = S_{f'(x;\cdot)}$ is closed and convex

  –: $x \in \mathrm{int}(\mathrm{dom}(f)) \implies \partial f(x)$ is closed, convex and bounded, and $f'(x;\cdot) = \max_{\phi\in\partial f(x)}\langle\phi,\cdot\rangle$ is closed.

  –: $x \in \mathrm{dom}(f) \setminus \mathrm{int}(\mathrm{dom}(f)) \implies \partial f(x)$ is either empty or unbounded.

- **Directional derivatives of max-function**. $f_1, \dots, f_m$ are proper and convex, $x \in \bigcap_{i=1}^m \mathrm{int}(\mathrm{dom}(f_i))$ and $I = \{i : f_i(x) = f(x)\} \implies \forall d \in \mathbb{E}, f'(x;d) = \max_{i\in I}\{f_i'(x;d)\}$

- **Unique subgradient**. Suppose f is proper and convex and $x \in \mathrm{dom}(f)$

  –: f is differentiable at $x \iff \partial f(x)$ is a singleton

6 Nonsmooth (Nonconvex)

6.1 Definitions

- **Dini directional derivative of locally Lipschitz f at x in direction d.** $f^-(x;d) = \liminf_{t\downarrow 0} \frac{1}{t}[f(x+td) - f(x)]$
- **Michel-Penot directional derivative of locally Lipschitz f at x in direction d.** $f^\diamond(x;d) = \sup_{u\in\mathbb{E}} \limsup_{t\downarrow 0} \frac{1}{t}[f(x+tu+td) - f(x+tu)]$
- **Clarke directional derivative of locally Lipschitz f at x in direction d.** $f^\circ(x;d) = \limsup_{y\to x, t\downarrow 0} \frac{1}{t}[f(y+td) - f(y)]$
- **Dini subdifferential of locally Lipschitz f at x.** $\partial_- f(x) = \{\phi \in \mathbb{E} : \forall d \in \mathbb{E}, \langle \phi, d\rangle \le f^-(x;d)\}$
- **Michel-Penot subdifferential of locally Lipschitz f at x.** $\partial_\diamond f(x) = \{\phi \in \mathbb{E} : \forall d \in \mathbb{E}, \langle \phi, d\rangle \le f^\diamond(x;d)\}$
- **Clarke subdifferential of locally Lipschitz f at x.** $\partial_\circ f(x) = \{\phi \in \mathbb{E} : \forall d \in \mathbb{E}, \langle \phi, d\rangle \le f^\circ(x;d)\}$
- **Nonsmooth subgradients.** members of nonsmooth subdifferentials
- **Regularity of locally Lipschitz f at x.** $\forall d \in \mathbb{E}, f'(x;d) = f^\circ(x;d)$
- **Distance function of S.** $d_S : x \in \mathbb{E} \mapsto \inf\{\|x - y\| : y \in S\}$
- **Clarke normal cone of S at $x \in S$.** $N_S(x) = \mathtt{cl}(\bigcup_{t\ge 0} t\partial_\circ d_S(x))$
- **Clarke tangent cone of S at $x \in S$.** $T_S(x) = N_S(x)^- = \{d \in \mathbb{E} : d_S^\circ(x;d) = 0\}$

6.2 Theorems

- **Bounded nonsmooth directional derivatives.** f locally Lipschitz at x with Lipschitz constant K
  $\implies \forall d \in \mathbb{E}, f^-(x;d) \le f^\diamond(x;d) \le f^\circ(x;d) \le K\|d\|$
- **Bounded nonsmooth subdifferentials.** f locally Lipschitz at x with Lipschitz constant K
  $\implies \partial_- f(x) \subseteq \partial_\diamond f(x) \subseteq \partial_\circ f(x) \subseteq KB$
- **Sublinarity of nonsmooth directional derivatives.** Suppose f locally Lipschitz at x
  - $f^\diamond(x;\cdot) : d \in \mathbb{E} \mapsto f^\diamond(x;d)$ and $f^\circ(x;\cdot) : d \in \mathbb{E} \mapsto f^\circ(x;d)$ are sublinear and finite everywhere
  - $f^\diamond(x;\cdot) = \sigma_{\partial_\diamond f(x)}$ and $f^\circ(x;\cdot) = \sigma_{\partial_\circ f(x)}$
- **Characterization of regularity.** Suppose f locally Lipschitz at x.
  f is regular at $x \iff \forall d \in \mathbb{E}, f^-(x;d) = f^\diamond(x;d) = f^\circ(x;d)$
- **Regularity of convex functions.** f is proper and convex, and $x \in \mathtt{int}(\mathtt{dom}(f)) \implies f$ is regular at x.

- **Nonsmooth subdifferential calculus**. Suppose $f_1, \ldots, f_m$ are locally Lipschitz at x. It holds $\partial_\diamond(\sum_{i=1}^m f_i)(x) \subseteq \sum_{i=1}^m \partial_\diamond f_i(x)$, $\partial_\circ(\sum_{i=1}^m f_i)(x) \subseteq \sum_{i=1}^m \partial_\circ f_i(x)$, $\partial_\diamond(\max_i f_i)(x) \subseteq \mathrm{conv}(\bigcup_{i\in I} \partial_\diamond f_i(x))$, and $\partial_\circ(\max_i f_i)(x) \subseteq \mathrm{conv}(\bigcup_{i\in I} \partial_\circ f_i(x))$, where $I = \{i : f_i(x) = f(x)\}$
Equalities hold throughout when $f_1, \ldots, f_m$ are regular at x.

- **Nonsmooth directional derivatives of max-function**. $f_1, \ldots, f_m$ locally Lipschitz at x and $I = \{i : f_i(x) = f(x)\} \implies \forall d \in \mathbb{E}, (f^\diamond(x; d) \leq \max_{i\in I}\{f_i^\diamond(x; d)\}) \wedge (f^\circ(x; d) \leq \max_{i\in I}\{f_i^\circ(x; d)\})$

- **Nonsmooth optimality conditions**. Suppose $f, g_1, \ldots, g_m$ locally Lipschitz at x, and S nonempty
  - x minimizes $f \implies 0 \in \partial_\diamond f(x) \subseteq \partial_\circ f(x)$
  - x minimizes f over $S \implies 0 \in \partial_\circ f(x) + N_S(x) \implies \forall d \in T_S(x), f^\circ(x; d) \geq 0$
  - x minimizes f over $\{y : g_i(y) \leq 0, i = 1, \ldots, m\}$
    $\implies \exists \lambda \geq 0, (\sum_{i=0}^m \lambda_i > 0) \wedge (\lambda_0 \partial_\circ f(x) + \sum_{i=1}^m \lambda_i \partial_\circ g(x) = 0) \wedge (\lambda_i g_i(x) = 0, i = 1, \ldots, m)$.
    We can take $\lambda_0 = 1$ when $\exists d \in \mathbb{E}, (g_i(x) < 0) \vee (g_i^\circ(x; d) < 0)$

Acknowledgement[1]

7 Bibliography

1. The main source of the content is from the course textbook on Convex Analysis [3] and the classic book by Rockafellar [6].
2. Another good source for information are the two books on Convex Analysis and Optimization: [2] and [5]; and the books on variational inequalities [1, 4].

References

[1] A. AUSLANDER and M. TEBOULLE. *Asymptotic cones and functions in optimization and variational inequalities.* Springer Monographs in Mathematics. Springer-Verlag, New York, 2003. 14

[2] D.P. BERTSEKAS. *Convex analysis and optimization.* Athena Scientific, Belmont, MA, 2003. With Angelia Nedić and Asuman E. Ozdaglar. 14

[3] J.M. BORWEIN and A.S. LEWIS. *Convex analysis and nonlinear optimization.* CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 3. Springer, New York, second edition, 2006. Theory and examples. 14

[4] J.M. BORWEIN and Q.J. ZHU. *Techniques of variational analysis.* CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 20. Springer-Verlag, New York, 2005. 14

[1] Thanks to Chek Beng Chua, National Technological University of Singapore, for parts of this document.

[5] S. BOYD and L. VANDENBERGHE. *Convex Optimization*. Cambridge University Press, Cambridge, 2004. 14

[6] R.T. ROCKAFELLAR. *Convex analysis*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks. 14

Index