Convex Analysis

Functions Taking Infinite Values

We consider functions f mapping $\mathbf{E}$ to the extended-real-line $\overline{\mathbf{R}} = \mathbf{R} \cup \{\pm\infty\}$.

Care must be taken when working with $\pm\infty$. In particular, we set $0 \cdot \pm\infty = 0$ and will be careful to avoid the expressions $(+\infty) + (-\infty)$ throughout.

Since the primary focus of our discussion is convex functions, there is a bias between $+\infty$ and $-\infty$.

Given $f : \mathbf{E} \to \overline{R}$, the *effective domain* and *epigraph* of f are

$$\begin{aligned} \operatorname{dom} f &:= \{x \in \mathbf{E} : f(x) < +\infty\}, \\ \operatorname{epi} f &:= \{(x, r) \in \mathbf{E} \times \mathbf{R} : f(x) \leq r\}, \end{aligned}$$

respectively.

A function $f : \mathbf{E} \to \overline{\mathbf{R}}$ is called *proper* if it never takes the value $-\infty$ and $\operatorname{dom} f \neq \emptyset$.

Epigraphs

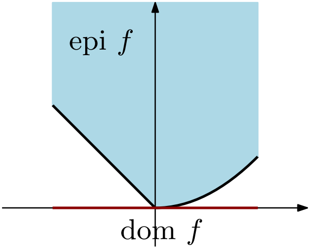


Figure: Epigraph and effective domain of the function whose value is $\max\{-x, \frac{1}{2}x^2\}$ for $x \in [-1, 1]$ and $+\infty$ elsewhere.

Epigraphs

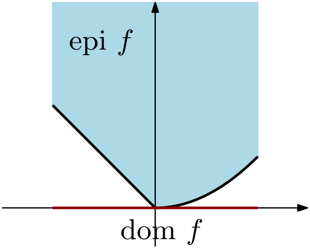


Figure: Epigraph and effective domain of the function whose value is $\max\{-x, \frac{1}{2}x^2\}$ for $x \in [-1, 1]$ and $+\infty$ elsewhere.

Lemma: A function $f : \mathbf{E} \to \overline{\mathbf{R}}$ is closed (lsc) if and only if epi f is a closed set.

Convex Functions

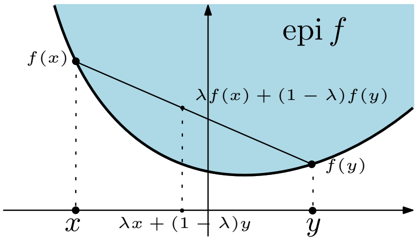
We say that the function $f : \mathbf{E} \to \overline{\mathbf{R}}$ is convex if epi f is a convex set.

Convex Functions

We say that the function $f : \mathbf{E} \to \overline{\mathbf{R}}$ is convex if epi f is a convex set.

Lemma: $f : \mathbf{E} \to \overline{\mathbf{R}}$ is **convex** if and only if

$$f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y) \qquad \forall\, x, y \in \mathbf{E} \text{ and } \lambda \in (0,1).$$

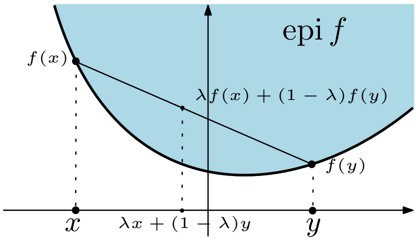


Convex Functions

We say that the function $f : \mathbf{E} \to \overline{\mathbf{R}}$ is convex if epi f is a convex set.

Lemma: $f : \mathbf{E} \to \overline{\mathbf{R}}$ is **convex** if and only if

$$f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y) \qquad \forall\, x, y \in \mathbf{E} \text{ and } \lambda \in (0,1).$$



Lemma: If $f : \mathbf{E} \to \overline{\mathbf{R}}$ is convex, then, for all $r \in \mathbf{R}$ the set $\{x \,|\, f(x) \leq r\}$ is convex.

3 Special Functions for $Q \subset \mathbf{E}$

The *indicator function* for Q:

$$\delta_Q(x) := \begin{cases} 0 & , \; x \in Q, \\ +\infty & , \; x \notin Q. \end{cases}$$

The *support function* for Q:

$$\delta_Q^*(x) := \sup_{v \in Q} \langle v, x \rangle \; .$$

The *gauge function* for Q:

$$\gamma_Q(x) := \inf \{ \lambda \in \mathbf{R}_+ \,|\, x \in \lambda Q \} \; .$$

3 Special Functions for $Q \subset \mathbf{E}$

The *indicator function* for Q:

$$\delta_Q(x) := \begin{cases} 0 & , \; x \in Q, \\ +\infty & , \; x \notin Q. \end{cases}$$

The *support function* for Q:

$$\delta_Q^*(x) := \sup_{v \in Q} \langle v, \, x \rangle \; .$$

The *gauge function* for Q:

$$\gamma_Q(x) := \inf \{ \lambda \in \mathbf{R}_+ \, | \, x \in \lambda Q \} \; .$$

(1) If $\mathbb{B} \subset \mathbf{E}$ is the closed unit ball for the norm $\|\cdot\|$, then

$$\|\cdot\| = \delta_{\mathbb{B}^\circ}^* = \gamma_{\mathbb{B}}.$$

3 Special Functions for $Q \subset \mathbf{E}$

The *indicator function* for Q:

$$\delta_Q(x) := \begin{cases} 0 & , \; x \in Q, \\ +\infty & , \; x \notin Q. \end{cases}$$

The *support function* for Q:

$$\delta^*_Q(x) := \sup_{v \in Q} \langle v, x \rangle \; .$$

The *gauge function* for Q:

$$\gamma_Q(x) := \inf \{ \lambda \in \mathbf{R}_+ \,|\, x \in \lambda Q \} \; .$$

(1) If $\mathbb{B} \subset \mathbf{E}$ is the closed unit ball for the norm $\|\cdot\|$, then

$$\|\cdot\| = \delta^*_{\mathbb{B}^\circ} = \gamma_{\mathbb{B}}.$$

(2) If $K \subset \mathbf{E}$ is a closed convex cone, then

$$\delta^*_{K^\circ} = \delta_K = \gamma_K.$$

Epigraphical Perspective

In our study of functions $f : \mathbf{E} \to \overline{\mathbf{R}}$ we take an epigraphical perspective, that is, we study properties of a function by studying properties of its epigraph.

Epigraphical Perspective

In our study of functions $f : \mathbf{E} \to \overline{\mathbf{R}}$ we take an epigraphical perspective, that is, we study properties of a function by studying properties of its epigraph.

For example, a function is closed (lsc) if its epigraph is a closed set. Similarly, a function is convex if its epigraph is a convex set.

The primary advantages of this perspective is that it allows us to discover properties of functions through properties of sets.

Epigraphical Perspective

In our study of functions $f : \mathbf{E} \to \overline{\mathbf{R}}$ we take an epigraphical perspective, that is, we study properties of a function by studying properties of its epigraph.

For example, a function is closed (lsc) if its epigraph is a closed set. Similarly, a function is convex if its epigraph is a convex set.

The primary advantages of this perspective is that it allows us to discover properties of functions through properties of sets.

A key observation in this regard is the fact that for every $x \in \operatorname{dom} f$,

$$f(x) = \inf_{(x,\mu)\in \operatorname{epi} f} \mu \ .$$

Epigraphs that are Cones

What are the functions whose epigraphs are cones?

Epigraphs that are Cones

What are the functions whose epigraphs are cones?

For $\lambda > 0$, $\lambda \text{epi}\, f = \text{epi}\, f$, i.e., if $(x, \mu) \in \text{epi}\, f$ so is $(\lambda x, \lambda\mu)$ for all $\lambda \geq 0$. Hence, we can relate the values of $f(\lambda x)$ to those of $f(x)$ as follows: for $\lambda > 0$,

$$\begin{aligned} f(\lambda x) &= \inf_{(\lambda x, \lambda\mu) \in \text{epi}\, f} \lambda\mu \\ &= \lambda \inf_{(x,\mu) \in \lambda^{-1}\text{epi}\, f} \mu \\ &= \lambda \inf_{(x,\mu) \in \text{epi}\, f} \mu \\ &= \lambda f(x)\ . \end{aligned}$$

From this, it is easy to show that epi f is a cone if and only if $f(\lambda x) = \lambda f(x)$ for all $x \in \text{dom}\, f$ and $\lambda \geq 0$.

Such functions are called *positively homogeneous*.

Epigraphs that are Convex Cones

If epi f is a convex cone, what can be said about f?

We have already shown that f must be positively homogeneous. But convexity tells us that epi f = epi f + epi f, i.e., for every pair $(x, \mu), (y, \tau) \in \text{epi}\, f$ we have

$$(x, \mu) + (y, \tau) = (x + y, \mu + \tau) \in \text{epi}\, f.$$

Consequently,

$$\{\mu + \tau \,|\, (x, \mu), (y, \tau) \in \text{epi}\, f\} \subset \{\omega \,|\, (x + y, \omega) \in \text{epi}\, f\},$$

and so, for all $x, y \in \text{dom}\, f$,

$$\begin{aligned} f(x+y) &= \inf_{(x+y,\omega)\in \text{epi}\, f} \omega \;\leq \inf_{(x,\mu),(y,\tau)\in \text{epi}\, f} \mu + \tau \\ &= \left(\inf_{(x,\mu)\in \text{epi}\, f} \mu\right) + \left(\inf_{(y,\tau)\in \text{epi}\, f} \tau\right) = f(x) + f(y). \end{aligned}$$

Since this inequality trivially holds if either x or y is not in dom f,

$$f(x + y) \leq f(x) + f(y) \qquad \forall\, x, y \in \mathbf{E}.$$

Such functions are called *subadditive*. Hence functions whose epigraphs are convex cones are both positively homogeneous and subadditive. Such functions are called *sublinear*.

Exercise

1) Show that a the epigraph of a positively homogeneous function is a cone.

2) Show that the epigraph of a sublinear function is a convex cone.

Support Functions are Sublinear

Let $S \subset \mathbf{E}$ be nonempty and consider the support function $\delta_S^*(x) = \sup_{v \in S} \langle v, x \rangle$.

positive homogeneity: $\lambda \geq 0$,

$$\begin{aligned} \delta_S^*(\lambda x) &= \sup \{ \langle \lambda x, v \rangle \,|\, v \in S \} = \lambda \sup \{ \langle x, v \rangle \,|\, v \in S \} \\ &= \lambda \delta^\star(x \,|\, S) \quad \forall \, \lambda \geq 0. \end{aligned}$$

subadditivity: $x^1, x^2 \in \mathbf{E}$,

$$\begin{aligned} \delta_S^*(x^1 + x^2) &= \sup \left\{ \left\langle x^1 + x^2, v \right\rangle \,\middle|\, v \in S \right\} \\ &= \sup \left\{ \left\langle x^1, v^1 \right\rangle + \left\langle x^2, v^2 \right\rangle \,\middle|\, v^1 = v^2 \in S \right\} \\ &\leq \sup \left\{ \left\langle x^1, v^1 \right\rangle + \left\langle x^2, v^2 \right\rangle \,\middle|\, v^1, v^2 \in S \right\} \\ &\leq \sup \left\{ \left\langle x^1, v^1 \right\rangle \,\middle|\, v^1 \in S \right\} + \sup \left\{ \left\langle x^2, v^2 \right\rangle \,\middle|\, v^2 \in S \right\} \\ &= \delta^\star \left(x^1 \,\middle|\, S \right) + \delta^\star \left(x^2 \,\middle|\, S \right). \end{aligned}$$

Support Functions are Sublinear

Let $S \subset \mathbf{E}$ be nonempty and consider the support function $\delta^*_S(x) = \sup_{v \in S} \langle v, x \rangle$.

positive homogeneity: $\lambda \geq 0$,

$$\begin{aligned} \delta^*_S(\lambda x) &= \sup \{ \langle \lambda x, v \rangle \,|\, v \in S \} = \lambda \sup \{ \langle x, v \rangle \,|\, v \in S \} \\ &= \lambda \delta^\star(x \,|\, S) \quad \forall \lambda \geq 0. \end{aligned}$$

subadditivity: $x^1, x^2 \in \mathbf{E}$,

$$\begin{aligned} \delta^*_S(x^1 + x^2) &= \sup \left\{ \left\langle x^1 + x^2, v \right\rangle \,\middle|\, v \in S \right\} \\ &= \sup \left\{ \left\langle x^1, v^1 \right\rangle + \left\langle x^2, v^2 \right\rangle \,\middle|\, v^1 = v^2 \in S \right\} \\ &\leq \sup \left\{ \left\langle x^1, v^1 \right\rangle + \left\langle x^2, v^2 \right\rangle \,\middle|\, v^1, v^2 \in S \right\} \\ &\leq \sup \left\{ \left\langle x^1, v^1 \right\rangle \,\middle|\, v^1 \in S \right\} + \sup \left\{ \left\langle x^2, v^2 \right\rangle \,\middle|\, v^2 \in S \right\} \\ &= \delta^\star\left(x^1 \,\middle|\, S\right) + \delta^\star\left(x^2 \,\middle|\, S\right). \end{aligned}$$

Are sublinear functions support functions?

Convexity and Optimization

Strict Convexity: A convex function $f : \mathbf{E} \to \overline{\mathbf{R}}$ is said to be *strictly convex* if

$$f((1-\lambda)x + \lambda y) < (1-\lambda)f(x) + \lambda f(y) \quad \forall\, x, y \in \operatorname{dom} f,\ \lambda \in (0,1) \text{ with } x \neq y.$$

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be convex. If $\bar{x} \in \operatorname{dom} f$ is a local solution to the problem $\min f(x)$, then $\bar{x}$ is a global optimal solution. Moreover, if f is strictly convex, then the global optimal solution is unique.

Proof: If $f(\bar{x}) = -\infty$ we are done, so assume that $-\infty < f(\bar{x})$. Suppose there is a $\widehat{x} \in \mathbf{R}^n$ with $f(\widehat{x}) < f(\overline{x})$. Let $\epsilon > 0$ be such that $f(\overline{x}) \le f(x)$ whenever $\|x - \overline{x}\| \le \epsilon$.
Set $\lambda := \epsilon(2\|\overline{x} - \widehat{x}\|)^{-1}$ and $x_\lambda := \overline{x} + \lambda(\widehat{x} - \overline{x})$. Then $\|x_\lambda - \overline{x}\| \le \epsilon/2$ and

$$f(x_\lambda) \le (1 - \lambda)f(\overline{x}) + \lambda f(\widehat{x}) < f(\overline{x}).$$

This contradiction implies no such $\widehat{x}$ exists.

Proof: If $f(\bar{x}) = -\infty$ we are done, so assume that $-\infty < f(\bar{x})$. Suppose there is a $\widehat{x} \in \mathbf{R}^n$ with $f(\widehat{x}) < f(\overline{x})$. Let $\epsilon > 0$ be such that $f(\overline{x}) \leq f(x)$ whenever $\|x - \overline{x}\| \leq \epsilon$.
Set $\lambda := \epsilon(2\|\overline{x} - \widehat{x}\|)^{-1}$ and $x_\lambda := \overline{x} + \lambda(\widehat{x} - \overline{x})$. Then $\|x_\lambda - \overline{x}\| \leq \epsilon/2$ and

$$f(x_\lambda) \leq (1-\lambda)f(\overline{x}) + \lambda f(\widehat{x}) < f(\overline{x}).$$

This contradiction implies no such $\widehat{x}$ exists.

To see the second statement in the theorem, let x^1 and x^2 be distinct global minimizers of f. Then, for $\lambda \in (0, 1)$,

$$f((1-\lambda)x^1 + \lambda x^2) < (1-\lambda)f(x^1) + \lambda f(x^2) = f(x^1)\ ,$$

which contradicts the assumption that x^1 is a global minimizer.

The Directional Derivative

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be convex and let $x \in \operatorname{dom} f$.

(1) Given $d \in \mathbf{E}$ the difference quotient

$$\frac{f(x+td)-f(x)}{t}$$

is a non-decreasing function of t on $(0, +\infty)$.

The Directional Derivative

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be convex and let $x \in \operatorname{dom} f$.

(1) Given $d \in \mathbf{E}$ the difference quotient

$$\frac{f(x+td)-f(x)}{t}$$

is a non-decreasing function of t on $(0, +\infty)$.

(2) For all $d \in \mathbf{E}$, $f'(x; d)$ exists with

$$f'(x; d) := \inf_{t>0} \frac{f(x+td)-f(x)}{t}.$$

The Directional Derivative

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be convex and let $x \in \operatorname{dom} f$.

(1) Given $d \in \mathbf{E}$ the difference quotient

$$\frac{f(x+td)-f(x)}{t}$$

is a non-decreasing function of t on $(0, +\infty)$.

(2) For all $d \in \mathbf{E}$, $f'(x; d)$ exists with

$$f'(x; d) := \inf_{t>0} \frac{f(x+td)-f(x)}{t}.$$

(3) The "*subdifferential inequality*" holds for all $x \in \operatorname{dom} f$:

$$f(x) + f'(x : y - x) \leq f(y) \quad \forall\, y \in \mathbf{E}.$$

The Directional Derivative

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be convex and let $x \in \operatorname{dom} f$.

(1) Given $d \in \mathbf{E}$ the difference quotient

$$\frac{f(x+td)-f(x)}{t}$$

is a non-decreasing function of t on $(0, +\infty)$.

(2) For all $d \in \mathbf{E}$, $f'(x;d)$ exists with

$$f'(x;d) := \inf_{t>0} \frac{f(x+td)-f(x)}{t}.$$

(3) The "*subdifferential inequality*" holds for all $x \in \operatorname{dom} f$:

$$f(x) + f'(x : y - x) \leq f(y) \quad \forall\, y \in \mathbf{E}.$$

(4) The function $f'(x;\cdot)$ is *sublinear*. In particular, $f'(x;\cdot)$ is a convex function for all $x \in \operatorname{dom} f$.

$t \mapsto (f(x+td) - f(x))/t$ nondecreasing for $t > 0$

Let $x \in \operatorname{dom} f$ and $d \in \mathbf{E}$. If $x + td \notin \operatorname{dom} f$ for all $t > 0$, the result follows. So assume that

$$0 < \bar{t} = \sup\{t \,:\, x + td \in \operatorname{dom} f\}.$$

Let $0 < t_1 < t_2 < \bar{t}$. Then

$$\begin{aligned} f(x + t_1 d) &= f\left(x + \left(\frac{t_1}{t_2}\right) t_2 d\right) \\ &= f\left[\left(1 - \left(\frac{t_1}{t_2}\right)\right) x + \left(\frac{t_1}{t_2}\right)(x + t_2 d)\right] \\ &\leq \left(1 - \frac{t_1}{t_2}\right) f(x) + \left(\frac{t_1}{t_2}\right) f(x + t_2 d) \\ &= f(x) + t_1 \frac{f(x+t_2 d) - f(x)}{t_2}. \end{aligned}$$

Hence

$$\frac{f(x + t_1 d) - f(x)}{t_1} \leq \frac{f(x + t_2 d) - f(x)}{t_2}.$$

$f'(x;d) = \inf_{t>0}(f(x+td) - f(x))/t$

(2) If $x + td \notin \operatorname{dom} f$ for all $t > 0$, then the result is obviously true.
So assume there is a $\bar{t} > 0$ such that $x + td \in \operatorname{dom} f$ for all $t \in (0, \bar{t}]$. Since

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

and the difference quotient in the limit is non-decreasing in t on $(0, +\infty)$, the limit is necessarily given by the infimum of the difference quotient. This infimum always exists and so $f'(x;d)$ always exists and is given by the infimum.

(3) The subdifferential inequality follows from (2) by taking $d := y - x$ and $t = 1$ in the infimum:

$$f'(x; y - x) \leq f(y) - f(x).$$

$f'(x;\cdot)$ is sublinear

Positive homogeneity:

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

Subadditivity:

$$\begin{aligned}
f'(x;u+v) &= \lim_{t\downarrow 0} \frac{f(x+t(u+v))-f(x)}{t} \\
&= \lim_{t\downarrow 0} \frac{f(x+\frac{t}{2}(u+v))-f(x)}{t/2} \\
&= \lim_{t\downarrow 0} 2\frac{f(\frac{1}{2}(x+tu)+\frac{1}{2}(x+tv))-f(x)}{t} \\
&\leq \lim_{t\downarrow 0} 2\frac{\frac{1}{2}f(x+tu)+\frac{1}{2}f(x+tv)-f(x)}{t} \\
&= \lim_{t\downarrow 0} \frac{f(x+tu)-f(x)}{t} + \frac{f(x+tv)-f(x)}{t} \\
&= f'(x;u)+f(x;v)\ .
\end{aligned}$$

Convexity and Optimality

Theorem: Let $f : \mathbf{E} \to \mathbf{R} \cup \{+\infty\}$ be convex, $\Omega \subset \mathbf{E}$ convex, $\bar{x} \in \operatorname{dom} f \cap \Omega$. Then $\bar{x}$ solves $\min_{x\in\Omega} f(x)$ if and only if $f'(\bar{x}; y - \bar{x}) \geq 0$ for all $y \in \Omega$.

Proof: $(\Rightarrow)$ Let $y \in \Omega$ so that $\bar{x} + t(y - \bar{x}) \in \Omega$ for all $t \in [0, 1]$. Then $f(\bar{x}) \leq f(\bar{x} + t(y - \bar{x}))$ for all $t \in [0, 1]$. Therefore, $f'(\bar{x}; y - \bar{x}) = \lim_{t\downarrow 0} t^{-1}(f(\bar{x} + t(y - \bar{x})) - f(\bar{x})) \geq 0$.

$(\Leftarrow)$ For $y \in \Omega$,

$$0 \leq f'(\bar{x}; y - \bar{x}) = \inf_{t>0} \frac{f(x+t(y-\bar{x})-f(x)}{t} \overset{(t=1)}{\leq} f(y) - f(\bar{x}).$$

Convexity and Optimality

Theorem: Let $f : \mathbf{E} \to \mathbf{R} \cup \{+\infty\}$ be convex, $\Omega \subset \mathbf{E}$ convex, $\bar{x} \in \operatorname{dom} f \cap \Omega$. Then $\bar{x}$ solves $\min_{x \in \Omega} f(x)$ if and only if $f'(\bar{x}; y - \bar{x}) \geq 0$ for all $y \in \Omega$.

Proof: $(\Rightarrow)$ Let $y \in \Omega$ so that $\bar{x} + t(y - \bar{x}) \in \Omega$ for all $t \in [0, 1]$. Then $f(\bar{x}) \leq f(\bar{x} + t(y - \bar{x}))$ for all $t \in [0, 1]$. Therefore, $f'(\bar{x}; y - \bar{x}) = \lim_{t \downarrow 0} t^{-1}(f(\bar{x} + t(y - \bar{x})) - f(\bar{x})) \geq 0$.

$(\Leftarrow)$ For $y \in \Omega$,

$$0 \leq f'(\bar{x}; y - \bar{x}) = \inf_{t > 0} \frac{f(x+t(y-\bar{x})-f(x)}{t} \overset{(t=1)}{\leq} f(y) - f(\bar{x}).$$

Corollary: If f is differentiable at $\bar{x}$, $\bar{x}$ solves $\min_{x \in \Omega} f(x)$ if and only if $-\nabla f(\bar{x}) \in N_\Omega(\bar{x})$.

Proof: $0 \leq f'(\bar{x}; y - \bar{x}) = \langle \nabla f(\bar{x}), \, y - \bar{x} \rangle$ for all $y \in \Omega$ iff $-\nabla f(\bar{x}) \in N_\Omega(\bar{x})$.

Differential Tests for Convexity

The following are equivalent for a C^1-smooth function $f\colon U \to \mathbf{R}$ defined on a convex open set $U \subset \mathbf{E}$.

(a) **(convexity)** f is convex.

(b) **(gradient inequality)** $f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle$ for all $x, y \in U$.

(c) **(monotonicity)** $\langle \nabla f(y) - \nabla f(x), y - x\rangle \geq 0$ for all $x, y \in U$.

If f is C^2-smooth, then the following property can be added to the list:

(d) The relation $\nabla^2 f(x) \succeq 0$ holds for all $x \in U$.

Examples of Convex Functions

(1) Given a self-adjoint linear operator $\mathcal{A}\colon \mathbf{E} \to \mathbf{E}$, a point $c \in \mathbf{E}$, and $b \in \mathbf{R}$ the quadratic function $f(x) = \frac{1}{2}\langle \mathcal{A}x, x\rangle + \langle c, x\rangle + b$ is convex if and only if $\mathcal{A}$ is positive semidefinite.

(2) (Boltzmann-Shannon entropy)

$$f(x) = \begin{cases} x \log x & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ +\infty & \text{if } x < 0 \end{cases}$$

(3) (Fermi-Dirac entropy)

$$f(x) = \begin{cases} x \log(x) + (1-x)\log(1-x) & \text{if } x \in (0,1) \\ 0 & \text{if } x \in \{-1, 1\} \\ +\infty & \text{otherwise} \end{cases}$$

Examples of Convex Functions

(4) (Hellinger)

$$f(x) = \begin{cases} -\sqrt{1-x^2} & \text{if } x \in [-1,1] \\ +\infty & \text{otherwise} \end{cases}$$

(5) (Exponential) $f(x) = e^x$

(6) (Log-exp) $f(x) = \log(1 + e^x)$

Bounds for β-Smooth Convex Functions

Let $f : \mathbf{E} \to \overline{\mathbf{R}}$. TFAE (the following are equivalent)

(1) f is β-smooth.

(2) $0 \leq f(y) - f(x) - \langle \nabla f(x), y - x \rangle \leq \frac{\beta}{2}\|x - y\|^2$

(3) $f(x) + \langle \nabla f(x), y - x \rangle + \frac{1}{2\beta}\|\nabla f(x) - \nabla f(y)\|^2 \leq f(y)$

(4) $\frac{1}{\beta}\|\nabla f(x) - \nabla f(y)\|^2 \leq \langle \nabla f(x) - \nabla f(y), x - y \rangle$

(5) $0 \leq \langle \nabla f(x) - \nabla f(y), x - y \rangle \leq \beta\|x - y\|^2$

Epigraphical Operations

Recall that for a convex function f and $x \in \operatorname{dom} f$,

$$f(x) = \inf_{(x,\mu)\in \operatorname{epi} f} \mu \ .$$

This construction fact can be extended to by defining the lower envelope for *any* subset Q of $\mathbf{E} \times \mathbf{R}$:

$$E_Q(x) := \inf_{(x,r)\in Q} r \ .$$

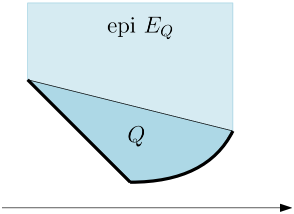


Figure: Lower envelope of Q.

Hence $\operatorname{epi} E_Q = Q + (\{0\} \times \mathbf{R}_+)$ when the infimum is attained when finite.

Example: $\lambda \operatorname{epi} f,\ \lambda > 0$

Epi-multiplication

$$\inf_{(x,r)\in\lambda \operatorname{epi} f} r = \inf\left\{ r \,\middle|\, (\lambda^{-1}x, \lambda^{-1}r) \in \operatorname{epi} f \right\}$$

$$= \lambda \inf\left\{ \lambda^{-1}r \,\middle|\, (\lambda^{-1}x, \lambda^{-1}r) \in \operatorname{epi} f \right\}$$

$$= \lambda \inf\left\{ \tau \,\middle|\, (\lambda^{-1}x, \tau) \in \operatorname{epi} f \right\}$$

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

Example: epi f_1 + epi f_2

Epi-addition or *infimal convolution*

$$
\begin{aligned}
\inf_{(x,r)\in \operatorname{epi} f_1 + \operatorname{epi} f_2} r &= \inf \{ r \mid (x,r) = (x_1,r_1) + (x_2,r_2),\ (x_i,r_i) \in \operatorname{epi} f_i \} \\
&= \inf \{ r_1 + r_2 \mid (y,r_1) \in \operatorname{epi} f_1,\ (x-y,r_2) \in \operatorname{epi} f_2 \} \\
&= \inf_y \inf_{r_1,r_2} \{ r_1 + r_2 \mid (y,r_1) \in \operatorname{epi} f_1,\ (x-y,r_2) \in \operatorname{epi} f_2 \} \\
&= \inf_y f_1(y) + f_2(x-y) \\
&=: (f_1 \Box f_2)(x)\ .
\end{aligned}
$$

Inverse Linear Image

Let $A \in \mathbf{L}[\mathbf{Y}, \mathbf{E}]$. Recall

$$E_Q(x) := \inf_{(x,r)\in Q} r \ .$$

What is E_Q when $Q = [A \times I]\text{epi}\, f$?

$$E_Q(x) = \inf\{r \,|\, x = Ay,\ (y,r) \in \text{epi}\, f\}$$

$$= \inf_{x=Ay} \inf_{(y,r)\in\text{epi}\, f} r$$

$$= \inf_{x=Ay} f(y) \ .$$

Infimal Projection

Let $g : \mathbf{E} \times \mathbf{Y} \to \overline{\mathbf{R}}$ and consider the projection $P \in \mathbf{L}[\mathbf{E} \times \mathbf{Y} \times \mathbf{R}]$ given by $P(x, y) = x$.

What is $E_{[P \times I]\mathrm{epi}\, g}$?

$$E_{[P \times I]\mathrm{epi}\, g}(x) = \inf \{\mu \,|\, x = P(z, y),\ (z, y, \mu) \in \mathrm{epi}\, g\}$$

$$= \inf_{x = P(z,y)} g(z, y)$$

$$= \inf_{y} g(x, y)\ .$$

The Perspective mapping

Let $Q := \mathbf{R}_+(\{1\} \times \operatorname{epi} f)$. What is $E_Q(\lambda, x)$ for $\lambda \geq 0$?

It is straightforward to show that $E_Q(\lambda, x) = +\infty$ if $\lambda < 0$ and that $E_Q(0, x) = 0$. So we suppose $0 < \lambda$.

$$\begin{aligned}
E_Q(\lambda, x) &= \inf \{ r \,|\, (\lambda, x, r) \in \mathbf{R}_+(\{1\} \times \operatorname{epi} f) \} \\
&= \inf \{ r \,|\, \exists\, \tau \geq 0 \text{ s.t. } (\lambda, x, r) \in \tau(\{1\} \times \operatorname{epi} f) \} \\
&= \inf \{ r \,|\, (x, r) \in \lambda \operatorname{epi} f \} \\
&= \inf \left\{ r \,\middle|\, (\lambda^{-1} x, \lambda^{-1} r) \in \operatorname{epi} f \right\} \\
&= \lambda \inf \left\{ \lambda^{-1} r \,\middle|\, (\lambda^{-1} x, \lambda^{-1} r) \in \operatorname{epi} f \right\} \\
&= \lambda f(\lambda^{-1} x) \; .
\end{aligned}$$

Relative interiors of sets in a product space

pic

Theorem: Let $Q \subset \mathbf{X} \times \mathbf{Y}$. For each $x \in \mathbf{X}$ set

$$Q_x := \{y \in \mathbf{Y} \,|\, (x,y) \in Q\} \text{ and } D := \{x \in \mathbf{X} \,|\, Q_x \neq \emptyset\}.$$

Then

$$(x,y) \in \operatorname{ri} Q \iff x \in \operatorname{ri} D \text{ and } y \in \operatorname{ri} Q_x.$$

Proof: Let $\mathcal{P}(x,y) = x$ be the projection of $\mathbf{X} \times \mathbf{Y}$ onto $\mathbf{X}$, and set $\mathcal{A}_x := \{x\} \times \mathbf{Y}$. Then $\mathcal{P}Q = D$, so $\operatorname{ri} D = \operatorname{ri} \mathcal{P}Q = \mathcal{P} \operatorname{ri} Q$. Hence, $(x,y) \in \operatorname{ri} Q$ iff $x \in \operatorname{ri} D$ and

$$(x,y) \in \mathcal{A}_x \cap \operatorname{ri} Q = \operatorname{ri}(\mathcal{A}_x \cap Q) = \operatorname{ri}(\{x\} \times Q_x) = \{x\} \times \operatorname{ri} Q_x \;.$$

So, $(x,y) \in \operatorname{ri} Q$ if and only if $x \in \operatorname{ri} D$ and $y \in \operatorname{ri} Q_x$.

$\operatorname{ri}\operatorname{epi} f$

Lemma: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be convex. Then

$$\operatorname{ri}\operatorname{epi} f = \{(x, \mu) \,|\, x \in \operatorname{ri}\operatorname{dom} f \text{ and } f(x) < \mu\}.$$

Proof: Apply the previous result to $\operatorname{epi} f \subset \mathbf{E} \times \mathbf{R}$.

Then $D = \operatorname{dom} f$ and $(\operatorname{epi} f)_x = \{\mu \in \mathbf{R} \,|\, f(x) \leq \mu\}$.

Clearly, $\operatorname{ri}(\operatorname{epi} f)_x = \{\mu \in \mathbf{R} \,|\, f(x) < \mu\}$, which gives the result.

Local Boundedness of Cvx Func.s on $\operatorname{ri}\operatorname{dom}$

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be convex. Then, $\forall\, \bar{x} \in \operatorname{ri}\operatorname{dom} f$, there is a cvx nbhd U of $\bar{x}$ and an $M > 0$ s.t. $U \cap \operatorname{aff}\operatorname{dom} f \subset \operatorname{ri}\operatorname{dom} f$ and

$$f(x) \leq M \quad \forall\, x \in U \cap \operatorname{aff}\operatorname{dom} f.$$

Proof: Let $\bar{x} \in \operatorname{ri}\operatorname{dom} f$ and let $u_1, \ldots, u_n$ be an orthonormal basis for $\mathbf{E}$ with $u_1, \ldots, u_k$ an orthonormal basis for $\operatorname{par}\operatorname{dom} f$. Then $B_1 := \operatorname{intr}\operatorname{conv}\{\pm u_i \,|\, i = 1, \ldots, n\}$ is a sym. open nghd of the origin. Let $\epsilon > 0$ be s.t.

$$\bar{x} + \epsilon B_1 \cap \operatorname{par}\operatorname{dom} f = (\bar{x} + \epsilon B_1) \cap \operatorname{aff}\operatorname{dom} f \subset \operatorname{ri}\operatorname{dom} f.$$

Set $U := \bar{x} + \epsilon B_1$. Then, for every $x \in \bar{x} + \epsilon B_1 \cap \operatorname{par}\operatorname{dom} f$,

$$\exists\ \lambda_i, \mu_i \geq 0,\ i = 1, \ldots, n \text{ with } \textstyle\sum_{j=1}^{k}(\lambda_i + \mu_i) = 1$$

such that

$$x = \bar{x} + \epsilon[\textstyle\sum_{j=1}^{k} \lambda_i u_i + \mu_i(-u_i)] = \sum_{j=1}^{k} \lambda_i(\bar{x} + \epsilon u_i) + \mu_i(\bar{x} - \epsilon u_i).$$

Therefore,

$$\begin{aligned} f(x) &\leq \sum_{j=1}^{k} \lambda_i f(\bar{x} + \epsilon u_i) + \sum_{j=1}^{k} \mu_i f(\bar{x} - \epsilon u_i) \\ &\leq \max\{f(\bar{x} \pm \epsilon u_i) \,|\, i = 1, \ldots, k\} =: M. \end{aligned}$$

Local Lip. Cont. of Cvx Func.s on ri dom

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be convex. Then for every $\bar{x} \in \operatorname{ri}\operatorname{dom} f$ there is an $\epsilon > 0$ s.t. f is Lip. cont. on $B_\epsilon(\bar{x}) \cap \operatorname{aff}\operatorname{dom} f$.

Proof: Set $D := \operatorname{par}\operatorname{dom} f$. Let $\epsilon > 0$ and $M > 0$ be such that $B_{2\epsilon}(\bar{x}) \cap \operatorname{aff}\operatorname{dom} f \subset \operatorname{ri}\operatorname{dom} f$ with $f(x) \leq M \ \forall\, x \in B_{2\epsilon}(\bar{x}) \cap \operatorname{aff}\operatorname{dom} f$. Set $h(x) := (2M)^{-1}[f(x+\bar{x}) - f(\bar{x})]$. If h is Lip. cont. on D near 0, then f is Lip. cont. on $\operatorname{aff}\operatorname{dom} f$ near $\bar{x}$. Observe that $h(0) = 0$ and $h(x) \leq 1$ for all $x \in B_{2\epsilon}(0) \cap D$. Moreover, for every $x \in B_{2\epsilon}(0) \cap D$, $0 = h(0) = h(\frac{1}{2}x - \frac{1}{2}x) \leq \frac{1}{2}h(x) + \frac{1}{2}h(-x)$ so that $-1 \leq -h(x) \leq h(-x)$. That is, $-1 \leq h(x) \leq 1$ for all $x \in B_{2\epsilon} \cap D$. For $x, y \in B_\epsilon(0) \cap D$ with $x \neq y$ set $\alpha := \|x - y\|$ and $\beta := \epsilon/\alpha$. Define $w := y + \beta(y - x) \in B_{2\epsilon} \cap D$. Then

$$y = (1+\beta)^{-1}[w + \beta x] = \tfrac{1}{1+\beta}w + \tfrac{\beta}{1+\beta}x.$$

The convexity of h implies that

$$\begin{aligned} h(y) - h(x) &\leq \frac{1}{1+\beta}h(w) + \frac{\beta}{1+\beta}h(x) - h(x) = \frac{1}{1+\beta}[h(w) - h(x)] \\ &\leq \frac{2}{1+\beta} = \frac{2}{\alpha+\epsilon}\|x - y\| \leq 2\epsilon^{-1}\|x - y\|. \end{aligned}$$

Symmetric in x and y implies the local Lip. cont. of h.

Supporting hyperplanes to epigraphs

We apply the following separation theorem to $\operatorname{epi} f$.

Theorem: Let $Q \subset \mathbf{E}$ be convex with $\bar{x} \in \operatorname{rb} Q$. Then there exists $\overline{z} \in \mathbf{E}$ such that

$$\langle \overline{z}, x \rangle \leq \langle \overline{z}, \bar{x} \rangle \ \forall \, x \in \operatorname{cl} Q \text{ and } \langle \overline{z}, x \rangle < \langle \overline{z}, \bar{x} \rangle \ \forall \, x \in \operatorname{ri} Q \,.$$

Supporting hyperplanes to epigraphs

We apply the following separation theorem to $\operatorname{epi} f$.

Theorem: Let $Q \subset \mathbf{E}$ be convex with $\bar{x} \in \operatorname{rb} Q$. Then there exists $\overline{z} \in \mathbf{E}$ such that

$$\langle \overline{z},\, x\rangle \le \langle \overline{z},\, \bar{x}\rangle \;\forall\, x \in \operatorname{cl} Q \text{ and } \langle \overline{z},\, x\rangle < \langle \overline{z},\, \bar{x}\rangle \;\forall\, x \in \operatorname{ri} Q\,.$$

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be proper convex and let $\bar{x} \in \operatorname{ri} \operatorname{dom} f$. Then there is a $v \in \mathbf{E}$ such that

$$\sup_x [\langle v,\, x\rangle - f(x)] \le \langle v,\, \bar{x}\rangle - f(\bar{x}).$$

Supporting hyperplanes to epigraphs

Proof: Since $\bar{x} \in \mathrm{ri}\,\mathrm{dom}\, f$, f is cont. at $\bar{x}$ relative to $\mathrm{dom}\, f$ and so $\mathrm{cl}\, f(\bar{x}) = f(\bar{x})$. In particular, $(\bar{x}, f(\bar{x})) \in \mathrm{rb}\,\mathrm{epi}\, f$. Hence, there exists $(w, \tau) \in \mathbf{E} \times \mathbf{R}$ s.t.

$$\langle (w,\tau),\, (x,\mu)\rangle \leq \langle (w,\tau),\, (\bar{x}, f(\bar{x}))\rangle \;\forall\, (x,\mu) \in \mathrm{cl}\,\mathrm{epi}\, f \;\text{ and}$$
$$\langle (w,\tau),\, (x,\mu)\rangle < \langle (w,\tau),\, (\bar{x}, f(\bar{x}))\rangle \;\forall\, (x,\mu) \in \mathrm{ri}\,\mathrm{epi}\, f\,.$$

Hence,

$$\langle w,\, x - \bar{x}\rangle + \tau(\mu - f(\bar{x})) < 0 \quad \forall\, x \in \mathrm{ri}\,\mathrm{dom}\, f,\; \mu > f(x).$$

Taking $x = \bar{x}$, we see that $\tau < 0$. Dividing by $|\tau|$ and setting $v = w/|\tau|$ and $\mu = f(x)$, we obtain

$$\langle v,\, x\rangle - f(x) \leq \langle v,\, \bar{x}\rangle - f(\bar{x}) \quad \forall\, x \in \mathrm{dom}\, f.$$

The result follows since if $x \notin \mathrm{dom}\, f$ then the above inequality is trivially true.

The Subgradient Inequality

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be proper convex and let $\bar{x} \in \operatorname{ri} \operatorname{dom} f$. Then there is a $v \in \mathbf{E}$ such that

$$f(\bar{x}) + \langle v, x - \bar{x} \rangle \leq f(x) \quad \forall\, x \in \mathbf{E}.$$

The Subgradient Inequality

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be proper convex and let $\bar{x} \in \operatorname{ri} \operatorname{dom} f$. Then there is a $v \in \mathbf{E}$ such that

$$f(\bar{x}) + \langle v, x - \bar{x} \rangle \leq f(x) \quad \forall\, x \in \mathbf{E}.$$

Proof: The Theorem tells us that there exist $v \in \mathbf{E}$ such that

$$\langle v, x \rangle - f(x) \leq \langle v, \bar{x} \rangle - f(\bar{x}) \quad \forall\, x \in \mathbf{E},$$

which gives the result.

The Subdifferential

Definition: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be convex and let $\bar{x} \in \operatorname{dom} f$. We say that f is subdifferentiable at $\bar{x}$ if there exists $v \in \mathbf{E}$ such that

$$f(\bar{x}) + \langle v, x - \bar{x} \rangle \le f(x) \quad \forall\, x \in \mathbf{E}.$$

We call v a *subgradient* for f at $\bar{x}$. The set of all subgradients at $\bar{x}$ is called the *subdifferential* of f at $\bar{x}$, denoted

$$\partial f(\bar{x}) := \{ v \,|\, f(\bar{x}) + \langle v, x - \bar{x} \rangle \le f(x) \quad \forall\, x \in \mathbf{E} \}.$$

For $x \notin \operatorname{dom} f$, we define $\partial f(x) = \emptyset$. The domain of ∂f is $\operatorname{dom} \partial f := \{ x \,|\, \partial f(x) \neq \emptyset \}$.

The Subdifferential

Definition: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be convex and let $\bar{x} \in \operatorname{dom} f$. We say that f is subdifferentiable at $\bar{x}$ if there exists $v \in \mathbf{E}$ such that

$$f(\bar{x}) + \langle v, x - \bar{x} \rangle \le f(x) \quad \forall\, x \in \mathbf{E}.$$

We call v a *subgradient* for f at $\bar{x}$. The set of all subgradients at $\bar{x}$ is called the *subdifferential* of f at $\bar{x}$, denoted

$$\partial f(\bar{x}) := \{ v \,|\, f(\bar{x}) + \langle v, x - \bar{x} \rangle \le f(x) \quad \forall\, x \in \mathbf{E} \}.$$

For $x \notin \operatorname{dom} f$, we define $\partial f(x) = \emptyset$. The domain of ∂f is $\operatorname{dom} \partial f := \{ x \,|\, \partial f(x) \ne \emptyset \}$.

Properties:
(1) $\operatorname{ri} \operatorname{dom} f \subset \operatorname{dom} \partial f \subset \operatorname{dom} f$
(2) $\partial f(x)$ is a nonempty closed convex set for all $x \in \operatorname{ri} \operatorname{dom} f$.
(3) If $x \in \operatorname{intr} \operatorname{dom} f$, then $\partial f(x)$ is compact.

Optimization and the Subdifferential

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be proper convex. Then $\bar{x} \in \mathbf{E}$ is a global solution to $\min f(x)$ if and only if $0 \in \partial f(\bar{x})$.

Proof: Apply the subgradient inequality:

$$f(\bar{x}) + \langle v, x - \bar{x} \rangle \leq f(x) \quad \forall\, x \in \mathbf{E}.$$

The Convex Conjugate

Recall that by applying the separation theorem to the epigraph of a proper convex function f, we found that for every $\bar{x} \in \operatorname{ri} \operatorname{dom} f$ there exists $v \in \mathbf{E}$ such that

$$\begin{aligned}
\delta^*_{\operatorname{epi} f}(v, -1) &= \sup_{x \in \operatorname{dom} f} [\langle v, x \rangle - f(x)] \\
&= \sup_x [\langle v, x \rangle - f(x)] \\
&\leq \langle v, \bar{x} \rangle - f(\bar{x}).
\end{aligned}$$

This relationship indicates that $f^* : \mathbf{E} \to \overline{\mathbf{R}}$ given by

$$f^*(v) := \sup_x [\langle v, x \rangle - f(x)]$$

plays a special in our study of convex functions.

We call f^* the *convex conjugate* of f.

The Convex Conjugate

Recall that by applying the separation theorem to the epigraph of a proper convex function f, we found that for every $\bar{x} \in \operatorname{ri} \operatorname{dom} f$ there exists $v \in \mathbf{E}$ such that

$$\begin{aligned} \delta^*_{\operatorname{epi} f}(v, -1) &= \sup_{x \in \operatorname{dom} f} [\langle v, x \rangle - f(x)] \\ &= \sup_x [\langle v, x \rangle - f(x)] \\ &\leq \langle v, \bar{x} \rangle - f(\bar{x}). \end{aligned}$$

This relationship indicates that $f^* : \mathbf{E} \to \overline{\mathbf{R}}$ given by

$$f^*(v) := \sup_x [\langle v, x \rangle - f(x)]$$

plays a special in our study of convex functions.

We call f^* the *convex conjugate* of f.

Note that $f^* = (\operatorname{cl} f)^*$ since $\delta^*_{\operatorname{epi} f} = \delta^*_{\operatorname{cl} \operatorname{epi} f}$.

The Bi-Conjugate and the Subdiffential

$$f^*(v) := \sup_x [\langle v,\, x\rangle - f(x)] = \delta^*_{\operatorname{epi} f}(v, -1) = \delta^*_{\operatorname{epi}\operatorname{cl} f}(v, -1)$$

By definition, f^* is a closed proper convex function whenever f is a proper convex function.

Theorem: [Fenchel-Young Inequality] Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be a proper convex function. Then

$$f^*(v) + f(x) \geq f^*(v) + \operatorname{cl} f(x) \geq \langle v,\, x\rangle \quad \forall\; x, v \in \mathbf{E}$$

with equality throughout if and only if $v \in \partial f(x)$.

The Bi-Conjugate and the Subdiffential

Consequently, for all $x \in \mathbf{E}$,

$$\begin{aligned} \operatorname{cl} f(x) &\geq \sup_{v \in \operatorname{dom} f^*} [\langle v, x \rangle - f^*(v)] \\ &= \sup_{v} [\langle v, x \rangle - f^*(v)] \\ &= (f^*)^*(x). \end{aligned}$$

The Bi-Conjugate and the Subdiffential

Consequently, for all $x \in \mathbf{E}$,

$$\begin{aligned} \operatorname{cl} f(x) &\geq \sup_{v \in \operatorname{dom} f^*} [\langle v, x \rangle - f^*(v)] \\ &= \sup_v [\langle v, x \rangle - f^*(v)] \\ &= (f^*)^*(x). \end{aligned}$$

Therefore,

$$\operatorname{cl} f(x) + f^*(v) \geq (f^*)^*(x) + f^*(v) \geq \langle v, x \rangle \quad \forall\, x, v \in \mathbf{E}$$

with equality throughout iff $x \in \partial f^*(v)$ iff $v \in \partial \operatorname{cl} f(x)$.

The Bi-Conjugate and the Subdiffential

Consequently, for all $x \in \mathbf{E}$,

$$\begin{aligned} \mathrm{cl}\, f(x) &\geq \sup_{v \in \mathrm{dom}\, f^*} [\langle v,\, x\rangle - f^*(v)] \\ &= \sup_v [\langle v,\, x\rangle - f^*(v)] \\ &= (f^*)^*(x). \end{aligned}$$

Therefore,

$$\mathrm{cl}\, f(x) + f^*(v) \geq (f^*)^*(x) + f^*(v) \geq \langle v,\, x\rangle \quad \forall\, x, v \in \mathbf{E}$$

with equality throughout iff $x \in \partial f^*(v)$ iff $v \in \partial \mathrm{cl}\, f(x)$.

Theorem: For every proper convex function $f : \mathbf{E} \to \overline{\mathbf{R}}$,

$$\mathrm{cl}\, f = (f^*)^* = f^{**}, \quad (\partial(\mathrm{cl}\, f))^{-1} = \partial f^*,$$

and

$$\partial(\mathrm{cl}\, f)(x) = \{v \,|\, \mathrm{cl}\, f(x) + f^*(v) \leq \langle v,\, x\rangle\}\ ,$$

with $\partial(\mathrm{cl}\, f)(x) = \partial f(x)$ whenever $x \in \mathrm{dom}\, \partial f$.

Proof: $\mathrm{cl}\, f$ coincides with f on $\mathrm{ri}\, \mathrm{dom}\, f = \mathrm{ri}\, \mathrm{dom}\, (\mathrm{cl}\, f)$ and $\mathrm{ri}\, \mathrm{dom}\, f \subset \mathrm{dom}\, \partial f$.

Support Functions Revisited

Let $Q \subset \mathbf{E}$ be nonempty closed and convex. Then

$$(\delta_Q(\cdot))^*(v) = \sup_x [\langle v, x \rangle - \delta_Q(x)] = \delta_Q^*(x).$$

Support Functions Revisited

Let $Q \subset \mathbf{E}$ be nonempty closed and convex. Then

$$(\delta_Q(\cdot))^*(v) = \sup_x [\langle v, x\rangle - \delta_Q(x)] = \delta_Q^*(x).$$

Recall that support functions are subadditive. We now address the question of whether a proper subadditive function can be written as a support function.

Support Functions Revisited

Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be proper subadditive. Then, for $\lambda > 0$,

$$\begin{aligned} f^*(v) &= \sup_{x \in \operatorname{dom} f} [\langle v, x\rangle - f(x)] \\ &= \sup_{x \in \operatorname{dom} f} [\langle v, \lambda x\rangle - f(\lambda x)] \\ &= \lambda \sup_{x \in \operatorname{dom} f} [\langle v, x\rangle - f(x)] \; = \; \lambda f^*(v). \end{aligned}$$

Therefore, $f^*(v) = 0$ for all $v \in \operatorname{dom} f^*$ and so $f^* = \delta_{\operatorname{dom} f^*}$.

Since f is proper convex, $\operatorname{cl} f = f^{**} = \delta^*_{\operatorname{dom} f^*}$.

Support Functions Revisited

Theorem: The class closed proper subadditive functions on $\mathbf{E}$ equals the class of support functions on $\mathbf{E}$. In particular, if $f : \mathbf{E} \to \overline{\mathbf{R}}$ is closed proper subadditive, then f is the support function of the set $\operatorname{dom} f^* = \{v \mid \langle v, x \rangle \leq f(x) \ \forall\, x \in \mathbf{E}\}$.

Support Functions Revisited

Theorem: The class closed proper subadditive functions on $\mathbf{E}$ equals the class of support functions on $\mathbf{E}$. In particular, if $f : \mathbf{E} \to \overline{\mathbf{R}}$ is closed proper subadditive, then f is the support function of the set $\text{dom}\, f^* = \{v \,|\, \langle v,\, x\rangle \leq f(x)\ \forall\, x \in \mathbf{E}\}$.

Proof: Since f is positively homogeneous,

$$\begin{aligned}
\text{dom}\, f^* &= \{v \,|\, \exists\, \mu > 0 \text{ s.t. } f^*(v) \leq \mu\} \\
&= \{v \,|\, \exists\, \mu > 0 \text{ s.t. } \langle v,\, x\rangle - f(x) \leq \mu\ \forall\, x \in \mathbf{E}\} \\
&= \{v \,|\, \exists\, \mu > 0 \text{ s.t. } \langle v,\, \lambda x\rangle - f(\lambda x) \leq \mu\ \forall\, x \in \mathbf{E},\ \lambda > 0\} \\
&= \left\{v \,\middle|\, \exists\, \mu > 0 \text{ s.t. } \langle v,\, x\rangle - f(x) \leq \frac{\mu}{\lambda}\ \forall\, x \in \mathbf{E},\ \lambda > 0\right\} \\
&= \{v \,|\, \langle v,\, x\rangle - f(x) \leq 0\ \forall\, x \in \mathbf{E}\}.
\end{aligned}$$

The result follows since we have shown that $f = \delta^*_{\text{dom}\, f^*}$.

$f'(x;\cdot)$ and ∂f

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be a proper convex function and let $\bar{x} \in \operatorname{dom} \partial f$. Then the closure of $f'(\bar{x};\cdot)$ is $\delta^*(\cdot \,|\, \partial f(\bar{x}))$. Moreover, if $\bar{x} \in \operatorname{ri} \operatorname{dom} f$, then $f'(\bar{x};\cdot)$ is closed and proper.

Proof: Let $v \in \partial f(\bar{x})$ and let φ be the closure of $f'(\bar{x};\cdot)$. Then, for $t > 0$ and $d \in \mathbf{E}$, $\langle v,\, d\rangle \le \frac{f(\bar{x}+td)-f(\bar{x})}{t}$ so $\langle v,\, d\rangle \le f'(\bar{x};d)$. Hence $f'(\bar{x};\cdot)$ is proper, and φ is closed proper and subadditive. Therefore, φ is the support function of the set

$$\begin{aligned}\{v \,|\, \langle v,\, d\rangle \le \varphi(d)\ \forall\, d \in \mathbf{E}\} &= \left\{ v \;\middle|\; \langle v,\, d\rangle \le \frac{f(\bar{x}+td)-f(\bar{x})}{t}\ \forall, d \in \mathbf{E},\, t > 0 \right\} \\ &= \{v \,|\, f(\bar{x}) + \langle v,\, d\rangle \le f(\bar{x}+d)\ \forall, d \in \mathbf{E}\} \\ &= \{v \,|\, f(\bar{x}) + \langle v,\, x-\bar{x}\rangle \le f(x)\ \forall, x \in \mathbf{E}\} \\ &= \partial f(\bar{x}).\end{aligned}$$

If $\bar{x} \in \operatorname{ri} \operatorname{dom} f$, then $\operatorname{dom} f'(\bar{x};\cdot) = \operatorname{par} \operatorname{dom} f = \operatorname{ri} \operatorname{dom} f'(\bar{x};\cdot)$ so that $f'(\bar{x};\cdot)$ is locally Lip. on its domain and so closed and proper.

$\partial f(x) = \{v\}$ implies differentiability

Corollary: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be a proper convex function. If $\bar{x} \in \operatorname{dom} \partial f$, then $(\operatorname{par} \operatorname{dom} f)^\perp \subset \partial f(\bar{x})$.

Proof: Let $v \in \partial f(\bar{x})$ and $w \in (\operatorname{par} \operatorname{dom} f)^\perp$. Then for every $y \in \operatorname{dom} f$,

$$f(\bar{x}) + \langle v + w,\, y - x \rangle = f(\bar{x}) + \langle v,\, y - x \rangle \leq f(y).$$

Corollary: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be a proper convex function. If $\bar{x} \in \operatorname{dom} \partial f$ is such that $\partial f(\bar{x}) = \{v\} + (\operatorname{par} \operatorname{dom} f)^\perp$, then f is differentiable relative to the affine manifold $S := \operatorname{aff} \operatorname{dom} f$ with gradient $\nabla_S f(\bar{x}) = v$. In particular, if $\bar{x} \in \operatorname{intr} \operatorname{dom} f$, then f is differentiable at $\bar{x}$ with $\nabla f(\bar{x}) = v$.

Proof: For $d \in \operatorname{par} \operatorname{dom} f$, $f'(\bar{x}; d) = \langle v,\, d \rangle$ is linear on the subspace $\operatorname{par} \operatorname{dom} f$. Hence, f is Gateaux differentiable relative to $\operatorname{aff} \operatorname{dom} f$ with Gateaux derivative v.

Computing the Subdifferential

Proposition: Let $Q \subset \mathbf{E}$ be a nonempty closed convex set. Then

$$\partial \delta_Q(x) = \begin{cases} \emptyset & , \ x \notin Q, \\ N_Q(x) & , \ x \in Q. \end{cases}$$

Computing the Subdifferential

Proposition: Let $Q \subset \mathbf{E}$ be a nonempty closed convex set. Then

$$\partial \delta_Q(x) = \begin{cases} \emptyset & , \ x \notin Q, \\ N_Q(x) & , \ x \in Q. \end{cases}$$

Note that this result implies that $N_Q(x) = [\operatorname{par} Q]^\perp$ when $x \in \operatorname{ri} Q$ since δ_Q is differentiable on $\operatorname{ri} Q$ relative to the affine manifold $\operatorname{aff} Q$ with derivative $\nabla_{\operatorname{aff} Q} \delta_Q(x) = 0$ for $x \in \operatorname{ri} Q$.

Computing the Subdifferential

Proposition: Let $Q \subset \mathbf{E}$ be a nonempty closed convex set. Then

$$\partial \delta_Q(x) = \begin{cases} \emptyset & , \ x \notin Q, \\ N_Q(x) & , \ x \in Q. \end{cases}$$

Note that this result implies that $N_Q(x) = [\operatorname{par} Q]^\perp$ when $x \in \operatorname{ri} Q$ since δ_Q is differentiable on $\operatorname{ri} Q$ relative to the affine manifold $\operatorname{aff} Q$ with derivative $\nabla_{\operatorname{aff} Q} \delta_Q(x) = 0$ for $x \in \operatorname{ri} Q$.

Proof: Given $\bar{x} \in Q$ and $v \in N_Q(\bar{x})$, we have

$$\langle v, x - \bar{x} \rangle \leq 0 \quad \forall\, x \in Q \ .$$

Computing the Subdifferential

Proposition: Let $Q \subset \mathbf{E}$ be a nonempty closed convex set. Then

$$\partial\delta_Q(x) = \begin{cases} \emptyset & , \ x \notin Q, \\ N_Q(x) & , \ x \in Q. \end{cases}$$

Note that this result implies that $N_Q(x) = [\operatorname{par} Q]^\perp$ when $x \in \operatorname{ri} Q$ since δ_Q is differentiable on $\operatorname{ri} Q$ relative to the affine manifold $\operatorname{aff} Q$ with derivative $\nabla_{\operatorname{aff} Q}\delta_Q(x) = 0$ for $x \in \operatorname{ri} Q$.

Proof: Given $\bar{x} \in Q$ and $v \in N_Q(\bar{x})$, we have

$$\delta_Q(\bar{x}) + \langle v, x - \bar{x} \rangle \leq \delta_Q(x) \ \forall \, x \in \mathbf{E} \ .$$

Computing the Subdifferential

Proposition: Let $Q \subset \mathbf{E}$ be a nonempty closed convex set. Then

$$\partial \delta_Q^*(x) = \operatorname*{argmax}_{v \in Q} \langle v, x \rangle \ .$$

Proof: For any closed proper convex function f, we have shown that

$$\partial f(x) = \{ v \,|\, f^*(v) + f(x) \leq \langle v, x \rangle \} \ .$$

Since both δ_Q and δ_Q^* are closed proper convex, we have

$$\partial \delta_Q^*(x) = \left\{ v \,\middle|\, \delta_Q(v) + \delta_Q^*(x) \leq \langle v, x \rangle \right\} = \operatorname*{argmax}_{v \in Q} \langle v, x \rangle \ .$$

The Subdifferential of a Norm

Corollary: Let $\|\cdot\|$ be any norm on $\mathbf{E}$ with closed unit ball $\mathbb{B}$. Then

$$\partial \|x\| = \begin{cases} \mathbb{B}^\circ & ,\ x = 0, \\ \{v \mid \|v\|_* = 1 \text{ and } \langle v,\, x\rangle = \|x\|\} & ,\ x \neq 0. \end{cases}$$

Proof: The result follows since $\|\cdot\| = \delta^*_{\mathbb{B}^\circ}(\cdot)$ where $\|\cdot\|_*$ is the dual norm for $\|\cdot\|$ whose closed unit ball is $\mathbb{B}$.

Computing Conjugates

Computing the conjugate f^* at v reduces to solving for x in the equation $v \in \partial f(x)$.

To see this, observe that

$$f^*(v) = \sup_x[\langle v, x\rangle - f(x)] = -\inf_x[f(x) - \langle v, x\rangle].$$

Since $f(x) - \langle v, x\rangle$ is convex, we need only solve $0 \in \partial[f - \langle v, \cdot\rangle](x) = \partial f(x) - v$ for x, then plug this x back into $\langle v, x\rangle - f(x)$ to find $f^*(v)$. This is especially useful when f is differentiable on its domain.

Computing Conjugates

Computing the conjugate f^* at v reduces to solving for x in the equation $v \in \partial f(x)$.

To see this, observe that

$$f^*(v) = \sup_x[\langle v, x\rangle - f(x)] = -\inf_x[f(x) - \langle v, x\rangle].$$

Since $f(x) - \langle v, x\rangle$ is convex, we need only solve $0 \in \partial[f - \langle v, \cdot\rangle](x) = \partial f(x) - v$ for x, then plug this x back into $\langle v, x\rangle - f(x)$ to find $f^*(v)$. This is especially useful when f is differentiable on its domain.

Example: $f(x) = e^x$. Then $v = \nabla f(x) = e^x$ iff $x = \ln v$, in which case

$$f^*(v) = \langle v, \ln v\rangle - f(\ln v) = \begin{cases} v\ln v - v &, v > 0, \\ +\infty &, v \leq 0 . \end{cases}$$

Computing Conjugates

Computing the conjugate f^* at v reduces to solving for x in the equation $v \in \partial f(x)$.

To see this, observe that

$$f^*(v) = \sup_x[\langle v,\, x\rangle - f(x)] = -\inf_x[f(x) - \langle v,\, x\rangle].$$

Since $f(x) - \langle v,\, x\rangle$ is convex, we need only solve $0 \in \partial[f - \langle v,\, \cdot\rangle](x) = \partial f(x) - v$ for x, then plug this x back into $\langle v,\, x\rangle - f(x)$ to find $f^*(v)$. This is especially useful when f is differentiable on its domain.

Example: $f(x) = e^x$. Then $v = \nabla f(x) = e^x$ iff $x = \ln v$, in which case

$$f^*(v) = \langle v,\, \ln v\rangle - f(\ln v) = \begin{cases} v\ln v \;-\; v & ,\; v > 0, \\ +\infty & ,\; v \leq 0\;. \end{cases}$$

Check $f^{**}(x) = e^x$.

Computing Conjugates: Dual Operations

General formulas for conjugates of convex functions generated from other convex functions using convexity preserving operations are very powerful tools in applications.

Computing Conjugates: Dual Operations

General formulas for conjugates of convex functions generated from other convex functions using convexity preserving operations are very powerful tools in applications.

Example: What is $(\lambda f)^*$ when $\lambda > 0$ and f proper convex?

Computing Conjugates: Dual Operations

General formulas for conjugates of convex functions generated from other convex functions using convexity preserving operations are very powerful tools in applications.

Example: What is $(\lambda f)^*$ when $\lambda > 0$ and f proper convex?

$$\begin{aligned}(\lambda f)^*(v) &= \sup_x \langle v, x\rangle - \lambda f(x)\\ &= \lambda \sup_x \left\langle \frac{v}{\lambda}, x\right\rangle - f(x)\\ &= \lambda f^*(\frac{v}{\lambda})\end{aligned}$$

That is, the dual operation to multiplying a function by a positive scalar is epi-multiplication.

What is $(\lambda f(\cdot/\lambda))^*$ for $\lambda > 0$?

$$
\begin{aligned}
(\lambda f(\cdot/\lambda))^*(v) &= \sup_x [\langle v,\, x\rangle - \lambda f(x/\lambda)] \\
&= \lambda \sup_x [\langle v,\, x/\lambda\rangle - f(x/\lambda)] \\
&= \lambda \sup_z [\langle v,\, z\rangle - f(z)] \\
&= \lambda f^*(v)\ .
\end{aligned}
$$

What is $(f_1 \square f_2)^*$?

$$
\begin{aligned}
(f_1 \square f_2)^*(v) &= \sup_x [\langle v, x\rangle - \inf_{x=x_1+x_2} [f_1(x_1) + f_2(x_2)]] \\
&= \sup_x \sup_{x=x_1+x_2} [\langle v, x\rangle - (f_1(x_1) + f_2(x_2))] \\
&= \sup_{x_1, x_2} [\langle v, x_1 + x_2\rangle - f_1(x_1) - f_2(x_2)] \\
&= \sup_{x_1, x_2} [(\langle v, x_1\rangle - f_1(x_1)) + (\langle v, x_2\rangle - f_2(x_12))] \\
&= \sup_{x_1} [\langle v, x_1\rangle - f_1(x_1)] \; + \; \sup_{x_2} [\langle v, x_2\rangle - f_2(x_2)] \\
&= f_1^*(v) + f_2^*(v)
\end{aligned}
$$

What is $(f_1 + f_2)^*$?

The first point to consider By the bi-conjugacy theorm,

$$
\begin{aligned}
(\operatorname{cl} f_1 + \operatorname{cl} f_2)^* &= ((f_1^*)^* + (f_2^*)^*)^* \\
&= ((f_1^* \square f_2^*)^*)^* \\
&= \operatorname{cl}(f_1^* \square f_2^*)
\end{aligned}
$$

What is $(f_1 + f_2)^*$?

The first point to consider By the bi-conjugacy theorm,

$$(\mathrm{cl}\, f_1 + \mathrm{cl}\, f_2)^* = ((f_1^*)^* + (f_2^*)^*)^*$$

$$= ((f_1^* \Box f_2^*)^*)^*$$

$$= \mathrm{cl}\,(f_1^* \Box f_2^*)$$

It can be shown that if $(\mathrm{ri}\,\mathrm{dom}\, f_1) \cap (\mathrm{ri}\,\mathrm{dom}\, f_2) \neq \emptyset$, then the closure operation can be removed from the above equivalence, i.e.

$$(f_1 + f_2)^* = f_1^* \Box f_2^*.$$

Application: Distance to a Convex Cone

Let $K \subset \mathbf{E}$ be a closed convex cone and let $\|\cdot\|$ be *any* norm on $\mathbf{E}$ with closed unit ball $\mathbb{B}$. Then

$$\begin{aligned}
\operatorname{dist}(z \mid K) &= \inf_{y \in K} \|z - y\| \\
&= \inf_{y} \|z - y\| + \delta_K(y) \\
&= \inf_{y} \delta^*_{\mathbb{B}^\circ}(z - y) + \delta^*_{K^\circ}(y) \quad = (\delta^*_{\mathbb{B}^\circ} \Box \delta^*_{K^\circ})(z).
\end{aligned}$$

Application: Distance to a Convex Cone

Let $K \subset \mathbf{E}$ be a closed convex cone and let $\|\cdot\|$ be *any* norm on $\mathbf{E}$ with closed unit ball $\mathbb{B}$. Then

$$\begin{aligned}\operatorname{dist}(z \mid K) &= \inf_{y\in K} \|z-y\| \\ &= \inf_{y} \|z-y\| + \delta_K(y) \\ &= \inf_{y} \delta^*_{\mathbb{B}^\circ}(z-y) + \delta^*_{K^\circ}(y) \;= (\delta^*_{\mathbb{B}^\circ} \square \delta^*_{K^\circ})(z).\end{aligned}$$

Consequently,

$$\begin{aligned}\operatorname{dist}(\cdot \mid K)^* &= (\delta^*_{\mathbb{B}^\circ} \square \delta^*_{K^\circ})^* \\ &= \delta^{**}_{\mathbb{B}^\circ} + \delta^{**}_{K^\circ} \\ &= \delta_{\mathbb{B}^\circ} + \delta_{K^\circ} \;= \delta_{\mathbb{B}^\circ \cap K^\circ}.\end{aligned}$$

Application: Distance to a Convex Cone

Let $K \subset \mathbf{E}$ be a closed convex cone and let $\|\cdot\|$ be *any* norm on $\mathbf{E}$ with closed unit ball $\mathbb{B}$. Then

$$\begin{aligned} \operatorname{dist}(z \mid K) &= \inf_{y \in K} \|z - y\| \\ &= \inf_{y} \|z - y\| + \delta_K(y) \\ &= \inf_{y} \delta^*_{\mathbb{B}^\circ}(z - y) + \delta^*_{K^\circ}(y) \ = (\delta^*_{\mathbb{B}^\circ} \square \delta^*_{K^\circ})(z). \end{aligned}$$

Consequently,

$$\begin{aligned} \operatorname{dist}(\cdot \mid K)^* &= (\delta^*_{\mathbb{B}^\circ} \square \delta^*_{K^\circ})^* \\ &= \delta^{**}_{\mathbb{B}^\circ} + \delta^{**}_{K^\circ} \\ &= \delta_{\mathbb{B}^\circ} + \delta_{K^\circ} \ = \delta_{\mathbb{B}^\circ \cap K^\circ}. \end{aligned}$$

Therefore,

$$\operatorname{dist}(z \mid K) = \delta^*_{\mathbb{B}^\circ \cap K^\circ}(z) \ .$$

An Alternative Approach to the Subdifferential

Eventually, we would like to extend the notion of subdifferential beyond convex functions. One proposal is to define the (regular) subdifferential by the inequality

$$\hat{\partial} f(x) := \{ v \mid f(x) + \langle v, y - x \rangle \leq f(y) + o(\|y - x\|) \} .$$

Proposition: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be proper convex. Then, for all $x \in \operatorname{dom} \partial f(x)$, $\hat{\partial} f(x) = \partial f(x)$.

Proof: Clearly, $\partial f(x) \subset \hat{\partial} f(x)$, so let $v \in \hat{\partial} f(x)$. Then, for all $d \in \mathbf{E}$ and $t > 0$,

$$\langle v, d \rangle \leq \frac{f(x + td) - f(x)}{t} + \frac{o(t\,\|d\|)}{t},$$

and so $\langle v, d \rangle \leq f'(x; d) = \delta^*_{\partial f(x)}(d)$. Therefore, $v \partial f(x)$.

For this reason, from now on we simply denote $\hat{\partial} f(x)$ by $\partial f(x)$ and call $\partial f(x)$ even when f is not necessarily convex. Again, $\operatorname{dom} \partial f := \{ x \mid \partial f(x) \neq \emptyset \}$

A simple subdifferential calculus rule

Proposition: Let $h : \mathbf{E} \to \overline{\mathbf{R}}$ be proper convex and $g : \mathbf{E} \to \overline{\mathbf{R}}$ be convex and differentiable on the open set U. Then, for all $x \in U \cap \operatorname{dom} \partial h$, $\partial(h+g)(x) = \partial h(x) + \nabla g(x)$.

Proof: We have already shown that $\partial g(x) = \{\nabla g(x)\}$ for all $x \in U$. Given $x \in U \cap \operatorname{dom} \partial h$ and $v \in \partial h(x)$, we have

$$\left.\begin{aligned} h(x) + \langle v, y - x\rangle &\le h(y) \\ g(x) + \langle \nabla g(x), y - x\rangle &\le g(y) \end{aligned}\right\} \ \forall\, y \in \mathbf{E} \ .$$

Adding these inequalities shows that $\partial h(x) + \nabla g(x) \subset \partial(h+g)(x)$.

Next let $w \in \partial(h+g)(x)$. Then

$$\begin{aligned} h(x) + g(x) + \langle w, y - x\rangle &\le h(y) + g(y) \\ &= h(y) + g(x) + \langle \nabla g(x), y - x\rangle + o(\|y - x\|). \end{aligned}$$

Hence,

$$h(x) + \langle w - \nabla g(x), y - x\rangle \le h(y) + o(\|y - x\|) \ \forall\, y \in \mathbf{E},$$

which implies that $w - \nabla g(x) \in \partial h(x)$.

Strong Convexity

Definition: A function $f\colon \mathbf{E} \to \overline{\mathbf{R}}$ is called *μ-strongly convex* (with $\mu \geq 0$) if the perturbed function $x \mapsto f(x) - \frac{\mu}{2}\|x\|^2$ is convex.

Theorem: Let $f\colon \mathbf{E} \to \overline{\mathbf{R}}$ be a μ-strongly convex function. Then for any $x \in \mathbf{E}$ and $v \in \partial f(x)$, the estimate holds:

$$f(y) \geq f(x) + \langle v, y - x\rangle + \frac{\mu}{2}\|y - x\|^2 \qquad \text{for all } y \in \mathbf{E}.$$

Proof: Apply the subdifferential inequality to the convex function $g := f - \frac{\mu}{2}\|\cdot\|^2$.

Strong Convexity

Definition: A function $f\colon \mathbf{E} \to \overline{\mathbf{R}}$ is called *μ-strongly convex* (with $\mu \geq 0$) if the perturbed function $x \mapsto f(x) - \frac{\mu}{2}\|x\|^2$ is convex.

Theorem: Let $f\colon \mathbf{E} \to \overline{\mathbf{R}}$ be a μ-strongly convex function. Then for any $x \in \mathbf{E}$ and $v \in \partial f(x)$, the estimate holds:

$$f(y) \geq f(x) + \langle v, y - x\rangle + \frac{\mu}{2}\|y - x\|^2 \qquad \text{for all } y \in \mathbf{E}.$$

Proof: Apply the subdifferential inequality to the convex function $g := f - \frac{\mu}{2}\|\cdot\|^2$.

Corollary: Any proper, closed, μ-strongly convex function $f\colon \mathbf{E} \to \overline{\mathbf{R}}$ is coercive and has a unique minimizer x satisfying

$$f(y) - f(x) \geq \frac{\mu}{2}\|y - x\|^2 \qquad \text{for all } y \in \mathbf{E}.$$

The Moreau Envelope

Definition: For any function $f\colon \mathbf{E} \to \overline{\mathbf{R}}$ and real $\alpha > 0$, define the *Moreau envelope* and the *proximal map*, respectively:

$$f_\alpha(x) := \left(f \square (\frac{1}{2\alpha}\|\cdot\|^2)\right)(x) = \min_y \; f(y) + \frac{1}{2\alpha}\|x-y\|^2$$

$$\mathrm{prox}_{\alpha f}(x) := \operatorname*{argmin}_y f(y) + \frac{1}{2\alpha}\|x-y\|^2.$$

Recall that $\mathrm{epi}\, f_\alpha = \mathrm{epi}\, f + \mathrm{epi}\, \left(\frac{1}{2\alpha}\|\cdot\|^2\right)$.

The Huber Function and Soft-Threshholding

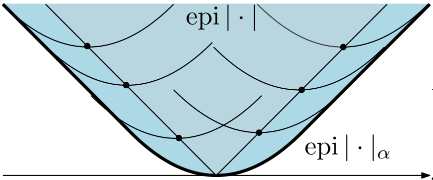
For $f(x) = |x|$,

$$f_\alpha(x) = \begin{Bmatrix} \frac{1}{2\alpha}|x|^2 & \text{if } |x| \leq \alpha \\ |x| - \frac{1}{2}\alpha & \text{otherwise} \end{Bmatrix}, \quad \text{prox}_{\alpha f}(x) = \begin{Bmatrix} x - \alpha & \text{if } x \geq \alpha \\ 0 & \text{if } |x| \leq \alpha \\ x + \alpha & \text{if } x \leq -\alpha \end{Bmatrix}.$$

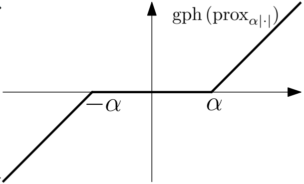
The Huber Function and Soft-Threshholding

For $f(x) = |x|$,

$$f_\alpha(x) = \begin{Bmatrix} \frac{1}{2\alpha}|x|^2 & \text{if } |x| \le \alpha \\ |x| - \frac{1}{2}\alpha & \text{otherwise} \end{Bmatrix}, \quad \text{prox}_{\alpha f}(x) = \begin{Bmatrix} x - \alpha & \text{if } x \ge \alpha \\ 0 & \text{if } |x| \le \alpha \\ x + \alpha & \text{if } x \le -\alpha \end{Bmatrix}.$$



(a) epi $|\cdot|_\alpha$ = epi $|\cdot|$ + epi $\frac{1}{2}|\cdot|^2$



(b) gph (prox$_{\alpha|\cdot|}$)

Figure: Moreau envelope and the proximal map of $|\cdot|$.

The Distance Function

Let $Q \subset \mathbf{E}$ be closed convex. Then

$$\begin{aligned}(\delta_Q)_\alpha(x) &= (\delta_Q \square \frac{1}{2\alpha}\|\cdot\|^2)(x) \\ &= \inf_{y \in Q} \frac{1}{2\alpha}\|x - y\|^2 \\ &= \frac{1}{2\alpha} d_Q^2(x)\end{aligned}$$

and

$$\mathrm{prox}_{\alpha\delta_Q}(x) = \mathrm{proj}_Q(x).$$

Prox is 1-Lipschitz

Theorem: Let $f\colon \mathbf{E} \to \overline{\mathbf{R}}$ be proper, closed, cvx. Then the set $\mathrm{prox}_f(x)$ is a singleton for every point $x \in \mathbf{E}$. Moreover,

$$\|\mathrm{prox}_f(x) - \mathrm{prox}_f(y)\|^2 \le \langle \mathrm{prox}_f(x) - \mathrm{prox}_f(y), x - y\rangle \quad \forall\, x, y \in \mathbf{E}.$$

Prox is 1-Lipschitz

Theorem: Let $f\colon \mathbf{E} \to \overline{\mathbf{R}}$ be proper, closed, cvx. Then the set $\mathrm{prox}_f(x)$ is a singleton for every point $x \in \mathbf{E}$. Moreover,

$$\|\mathrm{prox}_f(x) - \mathrm{prox}_f(y)\|^2 \le \langle \mathrm{prox}_f(x) - \mathrm{prox}_f(y), x - y\rangle \quad \forall\, x, y \in \mathbf{E}.$$

Proof: The map $z \mapsto f(z) + \frac{1}{2}\|z - x\|^2$ is proper, closed, and 1-strongly cvx, and hence $\mathrm{prox}_f(x)$ is the unique minimizer. Since $h(y) := f(y) + \frac{1}{2}\|y - x\|^2$ is 1-strongly cvx, for $x, y \in \mathbf{E}$,

$$\begin{aligned}
f(x^+) + \frac{1}{2}\|x^+ - x\|^2 &\le \left(f(y^+) + \frac{1}{2}\|y^+ - x\|^2\right) - \frac{1}{2}\|y^+ - x^+\|^2 \\
&= f(y^+) + \frac{1}{2}\|y^+ - y\|^2 - \frac{1}{2}\|y^+ - x^+\|^2 \\
&\quad + \frac{1}{2}\|y^+ - x\|^2 - \frac{1}{2}\|y^+ - y\|^2 \\
&\le \left(f(x^+) + \frac{1}{2}\|x^+ - y\|^2\right) - \|y^+ - x^+\|^2 \\
&\quad + \frac{1}{2}\|y^+ - x\|^2 - \frac{1}{2}\|y^+ - y\|^2,
\end{aligned}$$

so

$$\begin{aligned}
\|y^+ - x^+\|^2 &\le \frac{1}{2}\left(\|x^+ - y\|^2 - \|y^+ - y\|^2 + \|y^+ - x\|^2 - \|x^+ - x\|^2\right) \\
&= \langle x^+ - y^+, x - y\rangle \le \|y^+ - x^+\|\|x - y\|.
\end{aligned}$$

The Moreau Decomposition

Theorem: For any proper, closed, convex function $f\colon \mathbf{E} \to \overline{\mathbf{R}}$,

$$\mathrm{prox}_f(x) + \mathrm{prox}_{f^\star}(x) = x \qquad \forall x \in \mathbf{E}.$$

Proof Using the definition of the proximal map,

$$\begin{aligned}
z = \mathrm{prox}_f(x) &\iff 0 \in \partial\left(f + \frac{1}{2}\|\cdot - x\|^2\right)(z) \\
&\iff x - z \in \partial f(z) \\
&\iff z \in \partial f^\star(x - z) \\
&\iff 0 \in \partial f^\star(x - z) - z \\
&\iff 0 \in \partial\left(f^\star + \frac{1}{2}\|\cdot - x\|^2\right)(x - z) \\
&\iff x - z = \mathrm{prox}_{f^\star}(x).
\end{aligned}$$

∇f_α is Lipschitz continuous with parameter α^{-1}

Theorem: Let $f\colon \mathbf{E} \to \overline{\mathbf{R}}$be closed proper convex. Then the envelope f_α is continuously differentiable on $\mathbf{E}$ with gradient

$$\nabla f_\alpha(x) = \alpha^{-1}(x - \mathrm{prox}_{\alpha f}(x)).$$

Consequently ∇f_α is α^{-1}-smooth.

Proof: Take $\alpha = 1$, then

$$\begin{aligned}
z \in \partial f_\alpha(x) \quad &\iff \quad x \in \partial(f \,\square\, \tfrac{1}{2}\|\cdot\|^2)^\star(z) \\
&\iff \quad x \in \partial\left(f^\star + \left(\tfrac{1}{2}\|\cdot\|^2\right)^\star\right)(z) \\
&\iff \quad x \in \partial f^\star(z) + z \\
&\iff \quad 0 \in \partial(f^\star + \tfrac{1}{2}\|\cdot - x\|^2)(z) \\
&\iff \quad z = \mathrm{prox}_{f^\star}(x) \\
&\iff \quad z = x - \mathrm{prox}_f(x),
\end{aligned}$$

For $\alpha \neq 1$, use the identity $\alpha f_\alpha = (\alpha f)_1$.

Baillon-Haddad Theorem

Theorem: A proper, closed, convex function $f\colon \mathbf{E} \to \overline{\mathbf{R}}$ is μ-strongly convex if and only if the conjugate $f^\star$ is μ^{-1}-smooth.

Proof: ($\Longrightarrow$) Suppose that f is μ-strongly convex and define the convex function $g(x) := f(x) - \frac{\mu}{2}\|x\|^2$. We may then write

$$f^\star = \left(g + \tfrac{\mu}{2}\|\cdot\|^2\right)^\star = g^\star \,\square\, \tfrac{1}{2\mu}\|\cdot\|^2.$$

The right-hand-side is simply the Moreau envelope of $g^\star$ with parameter μ, and is therefore μ^{-1}-smooth.

Baillon-Haddad Theorem

($\Longleftarrow$) Suppose f^* is μ^{-1}-smooth, and set $h := f^*$ and $\beta := \mu^{-1}$ so that h is β-smooth. We know that h is β-smooth is equivalent to

$$0 \leq \langle \nabla h(x) - \nabla h(y), x - y \rangle \leq \beta \left\| x - y \right\|^2 .$$

Set $g := \frac{\beta}{2} \left\| \cdot \right\| - h$. Then

$$\langle \nabla g(y) - \nabla g(x), y - x \rangle = \beta \|y - x\|^2 - \langle \nabla h(y) - \nabla h(x), y - x \rangle \geq 0.$$

Hence, g is cvx. Note that

$$h(y) = \frac{\beta}{2} \|y\|^2 - g(y) = \frac{\beta}{2} \|y\|^2 - g^{\star\star}(y) = \frac{\beta}{2} \|y\|^2 - \sup_x \{\langle y, x \rangle - g^\star(x)\}$$

$$= \inf_x \left[\frac{\beta}{2} \|y\|^2 - \langle y, x \rangle + g^\star(x)\right],$$

so

$$\begin{aligned}
h^\star(z) &= \sup_y \ \{\langle z, y \rangle - h(y)\} \\
&= \sup_y \left[\langle z, y \rangle - \inf_x \left\{ \frac{\beta}{2} \|y\|^2 - \langle y, x \rangle + g^\star(x) \right\}\right] \\
&= \sup_x \sup_y \left[\langle z, y \rangle - \frac{\beta}{2} \|y\|^2 + \langle y, x \rangle - g^\star(x)\right] \\
&= \sup_x \left[\sup_y \ \left\{ \langle z + x, y \rangle - \frac{\beta}{2} \|y\|^2 \right\} - g^\star(x)\right] = \sup_x \ \frac{1}{2\beta} \|z + x\|^2 - g^\star(x).
\end{aligned}$$

So $h^\star(z) - \frac{1}{2\beta} \|z\|^2 = \sup_x \ \left[\frac{1}{\beta} \langle z, x \rangle + \frac{1}{2\beta} \|x\|^2 - g^\star(x)\right]$ is cvx.

Subgradient Dominance Theorem

Theorem: Any proper, closed, α-strongly convex function $f\colon \mathbf{E} \to \overline{\mathbf{R}}$ satisfies the subgradient dominance condition:

$$f(x) - \min f \leq \frac{1}{\alpha}\|v\|^2 \qquad \text{for all } x \in \mathbf{E},\ v \in \partial f(x).$$

Proof: Let $\bar{x}$ be a minimizer of f. Fix any $x \in \mathbf{E}$ and $v \in \partial f(x)$. We compute

$$\begin{aligned} f(x) - f(\bar{x}) &\leq \langle v, x - \bar{x}\rangle \leq \|v\| \cdot \|x - \bar{x}\| \\ &= \|v\| \cdot \|\nabla f^\star(v) - \nabla f^\star(0)\| \leq \frac{1}{\alpha}\|v\|^2. \end{aligned}$$

The Normal Cone to the Epigraph

Proposition: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be proper convex. Then, for all $\bar{x} \in \operatorname{dom} \partial f$, $\partial f(\bar{x}) = \{ v \,|\, (v, -1) \in N_{\operatorname{epi} f}(\bar{x}, f(\bar{x})) \}$.

Proof:

$$
\begin{aligned}
(v, -1) \in N_{\operatorname{epi} f}(x, f(x)) &\iff \langle (v, -1), (x, f(x)) - (\bar{x}, f(\bar{x})) \rangle \le 0 \;\forall\, x \in \operatorname{dom} f \\
&\iff f(\bar{x}) + \langle v, x - \bar{x} \rangle \le f(x) \;\forall\, x \in \operatorname{dom} f \\
&\iff f(\bar{x}) + \langle v, x - \bar{x} \rangle \le f(x) \;\forall\, x \in \mathbf{E}\ .
\end{aligned}
$$

Outer Semicontinuity of the Subdifferential

An important property of the subdifferential is that it is *outer semicontinuous.*

Definition: A multivalued mapping $T : \mathbf{X} \rightrightarrows \mathbf{Y}$ is said to be *outer semicontinuous* on its domain, $\operatorname{dom} T := \{x \,|\, T(x) \neq \emptyset\}$, if for every point $(\bar{x}, \bar{y}) \in (\operatorname{dom} T) \times \mathbf{Y}$ and every sequence $\{(x_i, y_i)\} \subset \mathbf{X} \times \mathbf{Y}$ with $(x_i, y_i) \to (\bar{x}, \bar{y})$ with $y_i \in T(x_i)$ for all i it must be the case that $\bar{y} \in T(\bar{x})$.

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be proper convex. Then ∂f is outer semicontinuous on $\operatorname{dom} \partial f$.

Proof: Let $(\bar{x}, \bar{y}) \in (\operatorname{dom} \partial f) \times \mathbf{E}$ and $\{(x_i, y_i)\} \subset (\operatorname{dom} \partial f) \times \mathbf{E}$ be such that $(x_i, y_i) \to (\bar{x}, \bar{y})$ with $y_i \in \partial f(x_i)$ for all i. We must show $\bar{y} \in \partial f(\bar{x})$. By construction,

$$\operatorname{cl} f(x_i) + \langle y_i,\, x - x_i \rangle \leq f(x) \quad \forall\, x \in \mathbf{E}\ .$$

Hence, given $x \in \mathbf{E}$, using the lower semicontinuity of $\operatorname{cl} f$, we may take the limit in this inequality to find that

$$\operatorname{cl} f(\bar{x}) + \langle \bar{y},\, x - \bar{x} \rangle \leq f(x) \quad \forall\, x \in \mathbf{E}\ .$$

Hence, $\bar{y} \in \partial(\operatorname{cl} f)(\bar{x}) = \partial f(\bar{x})$, where the equality follows since $\bar{x} \in \operatorname{dom} \partial f$.